Journal of the Institute of Mathematics of Jussieu

Additional services for *Journal of the Institute of Mathematics of Jussieu:*

Potential level-lowering for GSp(4)

Claus M. Sorensen

Link to this article: http://journals.cambridge.org/abstract_S147474800900005X

How to cite this article:
Claus M. Sorensen (2009). Potential level-lowering for GSp(4). Journal of the Institute of Mathematics of Jussieu, 8, pp 595-622 doi:10.1017/S147474800900005X

Request Permissions : Click here

POTENTIAL LEVEL-LOWERING FOR GSp(4)

CLAUS M. SORENSEN

*Department of Mathematics, Princeton University, Fine Hall, Washington Road,
Princeton, NJ 08544-1000, USA* (claus@princeton.edu)

Abstract   In this article, we explore a beautiful idea of Skinner and Wiles in the context of GSp(4) over a totally real field. The main result provides congruences between automorphic forms which are Iwahori-spherical at a certain place w, and forms with a tamely ramified principal series at w. Thus, after base change to a finite solvable totally real extension, one can often lower the level at w. For the proof, we first establish an analogue of the Jacquet–Langlands correspondence, using the stable trace formula. The congruences are then obtained on inner forms, which are compact at infinity modulo the centre, and split at all the finite places. The crucial ingredient allowing us to do so, is an important result of Roche on types for principal series representations of split reductive groups.

1. Introduction

Throughout, we fix embeddings $i_\infty : \bar{\mathbb{Q}} \to \mathbb{C}$ and $i_\ell : \bar{\mathbb{Q}} \to \bar{\mathbb{Q}}_\ell$ for each prime ℓ.

In Wiles's proof of Fermat's Last Theorem [**41**], it is shown that certain elliptic curves E over $\mathbb{Q}$ are *modular*: for some/every prime ℓ, the Galois representation

$$\rho_{E,\ell} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(V_\ell(E)), \quad V_\ell(E) = \mathbb{Q}_\ell \otimes_{\mathbb{Z}_\ell} \varprojlim E[\ell^n]$$

on the Tate module is isomorphic to $\rho_{f,\ell}$ for some weight 2 newform f on $\Gamma_0(N)$. Here $\rho_{f,\ell}$ is the ℓ-adic representation attached to f, constructed by Eichler and Shimura. It can be realized on the Tate module of a certain quotient of the Jacobian of the modular curve $X_0(N)$. By an ingenious trick, for the proof of modularity, one may assume that $\ell = 3$ and that the reduction $\bar{\rho}_{E,3}$ is irreducible. This is done by the $3 \leftrightarrow 5$ switch. Suppose $\bar{\rho}_{E,3}$ is reducible. Then the mod 5 representation $\bar{\rho}_{E,5}$ must be irreducible since there are no Galois invariant subgroups in E of order 15. One can then construct another elliptic curve $\mathcal{E}$, with the *same* mod 5 representation, such that $\bar{\rho}_{\mathcal{E},3}$ *is* irreducible. Assuming $\bar{\rho}_{E,3}$ is irreducible, the proof of modularity is divided into two steps.

Residual modularity: $\bar{\rho}_{E,3}$ is modular.

Modularity lifting: $\bar{\rho}_{E,\ell}$ modular $\implies \rho_{E,\ell}$ modular.

The first step is essentially a theorem of Langlands and Tunnell. The point is that $\bar{\rho}_{E,3}$ has solvable image, so one can use the theory of base change for GL_2 to get the result. The second step is more involved. We assume $\bar{\rho}_{E,\ell}$ is irreducible, and denote it by $\bar{\rho}$. For an arbitrary finite set of primes Σ, we introduce the universal deformation ring R_Σ classifying *all* deformations of $\bar{\rho}$, which are well behaved outside Σ. In particular, such deformations are unramified at primes outside Σ where $\bar{\rho}$ is unramified. In addition we have $\mathbb{T}_\Sigma$, the universal modular deformation ring. It is essentially a localization of some Hecke algebra, classifying all *modular* deformations of $\bar{\rho}$ well behaved outside Σ. It is well defined, at least for large Σ, exactly because we assume $\bar{\rho}$ has *some* modular deformation. By universality, there is a natural surjection from R_Σ onto $\mathbb{T}_\Sigma$. The key to step two is to show that this map is an isomorphism:

$$R_\Sigma \xrightarrow{\sim} \mathbb{T}_\Sigma.$$

Wiles gave a *numerical* criterion for this to be an isomorphism: a certain Selmer group and a certain congruence module should have the same size. Using ideas of Ribet [**26**] on level-*raising*, one shows that the sizes of these objects increase in exactly the same way when we add a prime to Σ. Consequently, it suffices to get an isomorphism in the *minimal* case where $\Sigma = \varnothing$. However, it is not clear at all that $\mathbb{T}_\varnothing$ is well defined: does $\bar{\rho}$ have a *minimal* modular deformation? What is needed is level-*lowering* congruences. For example, suppose we have

$$f \in S_2(\Gamma_0(Np)), \qquad \bar{\rho}_{f,\ell} \text{ unramified at } p,$$

where $p \nmid N\ell$. Then there is a form $\tilde{f} \equiv f \mod \ell$ of level N. The crucial case of this result was established by Ribet by a complicated geometric analysis of certain Shimura curves [**27**]. Once $\mathbb{T}_\varnothing$ is known to exist, the minimal case of the isomorphism is finally settled by constructing so-called Taylor–Wiles systems.

Several obstacles occur when one tries to extend the above to higher-dimensional Galois representations. For one thing, it is unknown how to generalize the level-raising part. In [**5**], the first installment of the proof of the Sato–Tate conjecture for elliptic curves with non-integral j-invariant, the expected results are deduced from a conjectural analogue of Ihara's lemma [**15**]. For GSp_4, we offer *some* results in this direction in [**33**]. However, inspired by ideas of Kisin, Taylor found a way to bypass Ihara's lemma and complete the proof of Sato–Tate in [**36**]. Second, one does not know in general how to extend Ribet's level-lowering argument. For $U(2,1)$ though, there has been some recent progress by Helm [**14**]. Instead, in the paper [**5**] for example, one adapts the following elegant idea of Skinner and Wiles: the original paper [**32**] was written in the context of Hilbert modular forms, but for this introduction let us for simplicity assume f is the form of level Np above. An argument *much* simpler than Ribet's, and more amenable to generalization, then shows the existence of a finite solvable totally real extension F over $\mathbb{Q}$, in which all the primes above ℓ split, such that

- $\bar{\rho}_{f,\ell}$ remains irreducible when restricted to $\mathrm{Gal}(\bar{\mathbb{Q}}/F)$,

- there is a form $\tilde{f} \equiv f_F \pmod{\ell}$ over F of level prime-to-(p).

Here f_F denotes the base change of f to F. This is of course weaker than Ribet's result quoted above, but it has the strong advantage of being easier to generalize to other situations. Moreover, this weaker version is often enough in order to prove modularity results. Indeed, by the theory of base change for GL_2, it is sufficient to obtain modularity over the solvable extension F.

In this paper, we explore the Skinner–Wiles trick for GSp_4 over a totally real field F. Our main result is Theorem C below. We first move the initial automorphic representation to an inner form G, which is compact at infinity and split at *all* the finite places. This requires F to have *even* degree over $\mathbb{Q}$. We accomplish this in Theorem B below, as a consequence of the stable trace formula. Then, for G, we obtain the desired congruences in Theorem A. Here, an important ingredient is a result of Roche on types for principal series representations. Using Theorem B again, we then move back to GSp_4. Base change, however, is not yet in such a good shape for GSp_4. Conditional results are available though, due to work of Labesse. Basically, if an automorphic representation has at least two Steinberg components, it admits a weak cyclic base change. We are hopeful that Ngo's recent progress on the fundamental lemma might allow one to weaken the Steinberg condition. Deformations of Galois representations, in the context of GSp_4, has been studied previously by other people. For example, Genestier and Tilouine construct Taylor–Wiles systems in [**8**]. We hope that our main result will be of some interest in this area. In the remainder of this introduction, we state precisely our three theorems mentioned just above.

For now, let F be an arbitrary totally real field of degree $d = [F : \mathbb{Q}]$, and let

$$G_0 = \mathrm{GSp}(2n)_{/F}.$$

Take G to be an arbitrary inner form of G_0 over F such that G_∞ is compact modulo its centre. To fix ideas, one can think of the unitary similitude group:

$$G(R) = \{g \in \mathrm{GL}_n(D \otimes_F R) : {}^t\bar{g}g = \mu(g)I\},$$

for any F-algebra R. Here D is a totally definite quaternion F-algebra. Its centre Z_G is canonically isomorphic to $\mathbb{G}_m$. Let us choose a Hecke character ω, and view it as a central character. Also, we fix an irreducible representation

$$\xi : G_\infty \to \mathrm{GL}(V_\xi).$$

Its complexification, an algebraic representation of $G_\mathbb{C}^d$, is characterized by its highest weight. In our situation, it corresponds to a tuple $a = (a_v)_{v|\infty}$ where

$$a_v = (a_{v,1}, \ldots, a_{v,n}, \tilde{a}_v) \in \mathbb{Z}^{n+1}, \quad a_{v,1} \geqslant \cdots \geqslant a_{v,n} \geqslant 0.$$

Our main result for G is the following analogue of Lemma 4.3.3 in [**5**].

Theorem A. *Let $\ell > 2n$ be a prime. Let π be an automorphic representation of G with central character ω and infinity type ξ. Suppose w is a finite place of F, where the group G splits, such that the following two conditions are satisfied:*

- *$\pi_w^{I_w} \neq 0$, where I_w is an Iwahori subgroup in G_w,*

- *$N(\mathfrak{p}_w) \equiv 1 \pmod{\ell}$.*

For each finite $v \neq w$ fix a compact open subgroup K_v in G_v such that $\pi_v^{K_v} \neq 0$. Then there exists an automorphic representation $\tilde{\pi}$ of G, with central character ω and infinity type ξ, satisfying the following three conditions:

- *$\tilde{\pi} \equiv \pi \pmod{\ell}$,*

- *$\tilde{\pi}_w = \tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$ for at most tamely ramified characters $\{\tilde{\chi}_i, \tilde{\chi}\}$,*

- *$\tilde{\pi}_v^{K_v} \neq 0$ for each finite place $v \neq w$.*

In this theorem, we use the notation from [**35**] for the principal series representations of G_0. Moreover, we use the notation $\tilde{\pi} \equiv \pi \pmod{\ell}$ to signify that the Hecke eigensystems are congruent modulo ℓ, via the fixed embeddings i_∞ and i_ℓ. More concretely, for almost all $v \nmid \ell$ different from w, we have the congruence

$$\eta_{\tilde{\pi}}(\phi) \equiv \eta_\pi(\phi) \pmod{\ell}, \quad \phi \in \mathcal{H}_\mathbb{Z}(G_v, K_v).$$

Here $\eta_\pi(\phi)$ gives the action of ϕ on the K_v-invariant of π_v, which is a one-dimensional subspace when K_v is hyperspecial. This is true for almost all v. One crucial ingredient of the proof of Theorem A is a result of Roche on types for principal series [**29**]. What we use is the following special case. Let χ_i and χ be $n + 1$ characters of $\mathbb{F}_w^*$, inflated to tamely ramified characters of $\mathcal{O}_w^*$. Then, for an irreducible admissible representation π of G_w, we consider the subspace of I_1-invariant vectors transforming according to the characters $\{\chi_i, \chi\}$ under the I-action. Then, this space is *non-zero* if and only if π is a subquotient of

$$\tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$$

for *some* quasi-characters $\tilde{\chi}_i$ extending the χ_i, and some $\tilde{\chi}$ extending χ. Now, the trick is to choose the χ_i wisely such that the above principal series is irreducible for all possible extensions. Here we use an irreducibility criterion due to Tadić, and the assumptions on ℓ and w. We can even take the χ_i and χ to be *trivial* modulo ℓ, and this is used later on in the proof in a crucial way.

To get congruences, as in Theorem A, between automorphic forms on G_0, we need to be able to transfer automorphic representations from G to G_0 and vice versa. That is, we need an analogue in higher rank of the Jacquet–Langlands correspondence for GL_2 and its inner forms ($n = 1$). For this, we assume that

$$n = 2, \qquad d = [F : \mathbb{Q}] \text{ is even.}$$

The first condition seems necessary, in order to apply the trace formula machinery currently available. The second condition ensures the existence (and uniqueness) of a totally definite quaternion algebra D over F, split at all finite places. As above, we consider the unitary similitude group. For F-algebras R,

$$G(R) = \{g \in \mathrm{GL}_2(D \otimes_F R) : {}^t\bar{g}g = \mu(g)I\}.$$

Then G is the inner form of G_0, which is compact at infinity mod centre, and split at *all* the finite places. Throughout the paper, we fix an inner twisting

$$\psi : G \to G_0.$$

That is, an isomorphism over $\bar{\mathbb{Q}}$ such that the Galois actions are intertwined up to G_0-conjugation. Our next theorem provides an analogue of the Jacquet–Langlands correspondence for the groups G and G_0. To state it precisely, we need to introduce some notation and terminology. First, we fix a central character ω and an infinity type ξ as above. Correspondingly, we have an L-packet

$$\Pi_\xi = \otimes_{v|\infty} \Pi_{\xi_v}, \quad \Pi_{\xi_v} = \{\Pi_{\xi_v}^{\mathrm{g}}, \Pi_{\xi_v}^{\mathrm{h}}\}$$

for $G_{0,\infty}$. Here $\Pi_{\xi_v}^{\mathrm{g}}$ and $\Pi_{\xi_v}^{\mathrm{h}}$ are the generic and holomorphic discrete series representations, respectively, of $\mathrm{GSp}_4(\mathbb{R})$ having the same infinitesimal character, and the same central character, as ξ_v. The theorem we are presenting below, only applies to *stable tempered* automorphic representations. Here, by *tempered*, we mean that almost all local components are tempered. This implies cuspidality by Langlands's theory of Eisenstein series [19]: the discrete non-cuspidal spectrum is given by residues of Eisenstein series, and those are non-tempered everywhere, see Proposition 4.5.4 in [18]. Our definition of *stable* is a bit ad hoc: we take it to mean *non-endoscopic* and *non*-CAP (CAP forms are defined below in Definition 1.2), according to the following definition.

Definition 1.1. A cuspidal automorphic representation Π of G_0, or G, is said to be endoscopic if and only if there exists two cuspidal automorphic representations ρ_1 and ρ_2 of GL_2, with the same central character, such that for almost all v:

$$L_v(s, \Pi, \mathrm{spin}) = L_v(s, \rho_1) L_v(s, \rho_2).$$

Here the left-hand side is the Euler factor of the degree-four *spin* L-function of Π. The above definition is from [40]: see the definition of weak endoscopic lift on p. 15, and the pertaining remarks. Let us recall the definition of CAP.

Definition 1.2. A cuspidal automorphic representation Π of G_0, or G, is said to be CAP (cuspidal associated to parabolic) relative to the parabolic P in G_0 if and only if there exists a cuspidal automorphic representations ρ of a Levi factor M_P such that Π is weakly equivalent to the constituents of $\mathrm{Ind}_P^{G_0}(\rho)$.

For GSp_4, these are completely understood due to the work of Piatetski-Shapiro [23] and Soudry [34]. One can construct them by θ-correspondence. Let us mention that Arthur's conjecture predicts that non-CAP *implies* tempered, but this is not known over F. We can now state our main result on functoriality.

Theorem B. *Fix an arbitrary member Π_ξ^+ of the L-packet Π_ξ. Then there is a natural multiplicity-preserving one-to-one correspondence between the sets:*

{stable tempered automorphic π of G with $\omega_\pi = \omega$ and $\pi_\infty = \xi$}

$\overset{1:1}{\longleftrightarrow}$ *{stable tempered cuspidal automorphic Π of G_0 with $\omega_\Pi = \omega$ and $\Pi_\infty = \Pi_\xi^+$}.*

The correspondence takes $\pi \mapsto \Pi_\xi^+ \otimes \pi_f$ and $\Pi \mapsto \xi \otimes \Pi_f$.

The proof of this theorem is an exercise based on deep results of Arthur, Hales, Langlands, Shelstad, Waldspurger and Weissauer. In [**33**] we proved an analogous result over $\mathbb{Q}$. However, in that case, the inner form G is necessarily non-split at some finite place. This makes it impossible to move *from G_0 to G*, as there are no known character relations à la Shelstad at this finite place.

Combining the two preceding theorems, we obtain congruences for GSp_4 à la Skinner and Wiles [**32**] in the case of GL_2. Thus, we continue to assume that $n = 2$ and d is even. The strategy is clear. We start out with a Π on G_0. By Theorem B, it corresponds to a π on G with the same finite part. Theorem A then yields a $\tilde{\pi} \equiv \pi \pmod{\ell}$. Now, we wish to apply Theorem B to $\tilde{\pi}$ and find a corresponding representation $\tilde{\Pi}$ on G_0. For that we need to make sure $\tilde{\pi}$ is stable and tempered. Hence, we assume the original Π admits a Galois representation $\rho_{\Pi,\ell}$ and that its mod ℓ reduction $\bar{\rho}_{\Pi,\ell}$ is irreducible. Since $\tilde{\pi}$ is congruent to π mod ℓ, this ensures that $\tilde{\pi}$ is stable (endoscopic forms and CAP forms have reducible Galois representations). Conjecturally, it is automatically tempered, as mentioned above. One only expects to attach Galois representations to *algebraic* automorphic representations of G_0.

Definition 1.3. A cuspidal automorphic representation Π of G_0 is called algebraic if and only if the following holds. For each archimedean place v of F consider the L-parameter of Π_v, say σ_{Π_v} [**21**]. This is an admissible homomorphism

$$\sigma_{\Pi_v} : W_{F_v} \to {}^L G_0 = \mathrm{GSp}_4(\mathbb{C}) \times W_{F_v},$$

where W_{F_v} is the Weil group. When we restrict σ_{Π_v} to $\mathbb{C}^*$, the corresponding four-dimensional representation should be a sum of characters of the form

$$z \mapsto z^p \bar{z}^q, \quad \text{for } p, q \in \mathbb{Z}.$$

Algebraicity gives a restriction on the possible infinity types ξ. For example, in the Hilbert modular case $n = 1$, the classical weights must have the same parity. By a principle propounded by Langlands, algebraic cusp forms on GL_n should be associated with pure motives. See [**25**] for a precise statement. Here we will only need the ℓ-adic Galois representation $\rho_{\Pi,\ell}$. We admit the following.

Conjecture. *Let Π be a cuspidal automorphic representation of G_0 over F, with infinity type in some Π_ξ, and assume the twist $\Pi \otimes |\det|^{w/2}$ is algebraic for some integer w. Let S_Π denote the set of ramified places of the representation Π. Then there exists a four-dimensional continuous semisimple Galois representation*

$$\rho_{\Pi,\ell} : \mathrm{Gal}(\bar{\mathbb{Q}}/F) \to \mathrm{GL}_4(\bar{\mathbb{Q}}_\ell),$$

which is unramified at all $v \notin S_\Pi$ not dividing ℓ. Moreover, for such places v,

$$L_v(s - w/2, \Pi, \mathrm{spin}) = \det(1 - \rho_{\Pi,\ell}(\mathrm{Frob}_v)N(\mathfrak{p}_v)^{-s})^{-1}.$$

Here Frob_v is the geometric Frobenius. If Π is not CAP, the representation $\rho_{\Pi,\ell}$ is pure of weight w. That is, the eigenvalues of $\rho_{\Pi,\ell}(\mathrm{Frob}_v)$ all have complex absolute value $N(\mathfrak{p}_v)^{w/2}$ for v as above. In other words, each Π_v is tempered.

This was proved over $\mathbb{Q}$ by Weissauer [**38**]. Therefore, if Π is a base change from $\mathbb{Q}$, the conjecture holds. In fact, cyclic base change is *known* for G_0 in many cases. Suppose F is a cyclic extension of $\mathbb{Q}$, and let Π be a cuspidal automorphic representation of G_0 over $\mathbb{Q}$, with at least *two* Steinberg components. Then Π admits a weak cuspidal base change Π_F on G_0 over F. This is due to Labesse in great generality. See Theorem 4.6.2 in [**18**] for a more precise statement.

Theorem C. *Let $\ell > 3$ be a prime. Let Π be a cuspidal automorphic representation of G_0 with central character ω and infinity type Π_ξ^+. We assume Π admits a Galois representation $\rho_{\Pi,\ell}$ as above, and that its mod ℓ reduction $\bar{\rho}_{\Pi,\ell}$ is irreducible. Suppose w is a finite place of F, satisfying the two conditions:*

- $\Pi_w^{I_w} \neq 0$, *where I_w is an Iwahori subgroup in $G_{0,w}$,*

- $N(\mathfrak{p}_w) \equiv 1 \pmod{\ell}$.

For each finite $v \neq w$ fix a compact open subgroup K_v in G_v such that $\Pi_v^{K_v} \neq 0$. Then there exists a cuspidal automorphic representation $\tilde{\Pi}$ of G_0, with central character ω and infinity type Π_ξ^+, satisfying the following three conditions:

- $\tilde{\Pi} \equiv \Pi \pmod{\ell}$,

- $\tilde{\Pi}_w = \tilde{\chi}_1 \times \tilde{\chi}_2 \rtimes \tilde{\chi}$ *for at most tamely ramified characters $\{\tilde{\chi}_1, \tilde{\chi}_2, \tilde{\chi}\}$,*

- $\tilde{\Pi}_v^{K_v} \neq 0$ *for each finite place $v \neq w$.*

What keeps us from getting a precise analogue of the main result in [**32**] is that *strong* cyclic base change is unavailable for GSp(4). Gan and Takeda have recently shown the local Langlands conjecture in this case [**6**], so locally one can base change L-packets. What we need is a global lift compatible with the local lift at *all* places. In the following informal discussion, we assume we have such a strong lift. We take F to be *any* totally real field, and let Π be a cuspidal automorphic representation of $G_{0/F}$ with a Galois representation $\rho_{\Pi,\ell}$. Here $\ell > 3$ is a prime such that the reduction $\bar{\rho}_{\Pi,\ell}$ is irreducible. First, one can base change to a situation where Theorem C applies. Namely, there is a finite solvable totally real extension E/F of even degree with the following properties:

- $\bar{\rho}_{\Pi,\ell}|_{\mathrm{Gal}(\bar{\mathbb{Q}}/E)}$ is irreducible,

- $\mathrm{BC}_{E_{\tilde{w}}/F_w}(\Pi_w)$ is Iwahori-spherical* for all $\tilde{w} \mid w$,

- $N(\mathfrak{p}_{\tilde{w}}) \equiv 1 \pmod{\ell}$ for all $\tilde{w} \mid w$.

Here $w \nmid \ell$ is a fixed finite place of F such that Π_w is ramified. Moreover, to retain some control above ℓ, we may assume that all the places above ℓ split completely in E. The existence of E follows from a well-known fact, also used in [**32**], that one can always find a totally real cyclic extension with prescribed splitting and ramification at a finite set of places. Now, Theorem C applies to any base change of Π. One deduces the existence of a cuspidal $\tilde{\Pi}$ on $G_{0/E}$ with

---

* To be precise, the base change of the L-packet of Π_w is an Iwahori-spherical singleton.

- $\omega_{\tilde{\Pi}} = \omega_{\Pi} \circ N_{E/F}$ and $\tilde{\Pi}_{\infty} = \Pi_{\infty}^{\otimes[E:F]}$,

- $\bar{\rho}_{\tilde{\Pi},\ell} \sim \bar{\rho}_{\Pi,\ell}|_{\mathrm{Gal}(\bar{\mathbb{Q}}/E)}$,

- $\tilde{\Pi}_{\tilde{w}}$ is a tamely ramified principal series for all $\tilde{w} \mid w$.

Furthermore, we may arrange that $\tilde{\Pi}_{\tilde{v}}$ is unramified at all $\tilde{v} \mid v$ such that Π_v is unramified. Recall that the places above ℓ split in E, so we may also arrange for $\tilde{\Pi}$ to have the *same* level as Π at such places. Finally, after another solvable base change, we may assume $\tilde{\Pi}_{\tilde{w}}$ is an *unramified* principal series for all $\tilde{w} \mid w$. That is, we can lower the level by passing to a solvable extension. Using this result, an inductive argument shows the existence of a finite solvable extension E/F such that $\bar{\rho}_{\Pi,\ell}|_{\mathrm{Gal}(\bar{\mathbb{Q}}/E)}$ admits an automorphic deformation $\rho_{\tilde{\Pi},\ell}$ with *minimal* ramification. However, we stress that our discussion in this paragraph relies heavily on the availability of *strong* cyclic base change for GSp(4).

2. Principal series representations of GSp($2n$)

In this section we review two results on principal series needed later. In this section only, F denotes a finite extension of $\mathbb{Q}_p$. We use ν to signify its normalized absolute value $|\cdot|_F$, that is, the modulus character of F. We let $\mathfrak{p}$ denote the maximal ideal in the valuation ring $\mathcal{O}$, and choose a uniformizer ϖ. Also,

$$G_0 = \mathrm{GSp}(2n),$$

viewed as an algebraic group over F. Here the similitude group is defined using

$$\mathcal{J} = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ -1 & & \end{pmatrix}.$$

The group G_0 comes with the similitude character μ. So, for any F-algebra R,

$$G_0(R) = \{g \in \mathrm{GL}_{2n}(R) : {}^t g \mathcal{J} g = \mu(g)\mathcal{J}\}.$$

The subset of upper triangular matrices in G_0 form a Borel subgroup B, and the diagonal matrices form a split maximal torus T of dimension $n+1$. Indeed,

$$T(R) = \left\{ t = \begin{pmatrix} t_1 & & & & & \\ & \cdot^{\cdot^{\cdot}} & & & & \\ & & t_n & & & \\ & & & \mu t_n^{-1} & & \\ & & & & \cdot^{\cdot^{\cdot}} & \\ & & & & & \mu t_1^{-1} \end{pmatrix} \right\},$$

for any R as above. Note that n is the rank of the derived group G_0^{der}. To define principal series, consider quasi-characters $\chi_1, \ldots, \chi_n$ and χ of F^*, and look at

$$\chi_1 \otimes \cdots \otimes \chi_n \otimes \chi : t \mapsto \chi_1(t_1) \cdots \chi_n(t_n)\chi(\mu).$$

We view it as a character of $B(F)$, and induce it to a representation of $G_0(F)$:

$$\chi_1 \times \cdots \times \chi_n \rtimes \chi = \mathrm{Ind}_B^{G_0}(\chi_1 \otimes \cdots \otimes \chi_n \otimes \chi).$$

The induction is normalized by $\delta_B^{1/2}$ in the standard way, where δ_B is the modulus character of $B(F)$. Here we use the notation introduced by Tadić in [**35**].

2.1. An irreducibility criterion of Tadić

For simplicity, we now assume p is *odd*. Then F^* has exactly three characters of order two, and this simplifies the irreducibility criterion we are about to describe. The reducibility of unitary principal series for symplectic groups was first studied by Winarsky [**42**], and then extended by Keys in [**16**]. Using their results, Tadić obtained complete results for similitude groups. We need the following theorem.

Theorem 2.1. *The principal series $\chi_1 \times \cdots \times \chi_n \rtimes \chi$ is irreducible if and only if*

- *there are at most two distinct χ_i of order two,*

- $\chi_i \neq \nu^{\pm 1}$, *for all i,*

- $\chi_i \neq \nu^{\pm 1}\chi_j^{\pm 1}$, *for all $i < j$.*

Proof. This is Theorem 7.9 combined with Remark 7.10 in [**35**]. □

In the last condition, all four combinations of signs are allowed. If all the χ_i are unitary, the last two conditions are clearly satisfied. In particular, for the group GSp(4), unitary principal series are always irreducible. For arbitrary n, the length is at most two: see Corollary 7.8 in [**35**]. Now, let us fix a prime

$$\ell > 2n \text{ such that } N(\mathfrak{p}) \equiv 1 \ (\mathrm{mod}\ \ell).$$

Here $N(\mathfrak{p})$ is the cardinality of the residue field $\mathbb{F}$. We assume such an ℓ exists.

Corollary 2.2. *There exists n tamely ramified characters χ_i of $\mathcal{O}^*$ such that*

$$\tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$$

is irreducible for all quasi-characters $\tilde{\chi}_i$ with $\tilde{\chi}_i|_{\mathcal{O}^*} = \chi_i$, and all $\tilde{\chi}$. Moreover,

$$\chi_i \equiv 1 \ (\mathrm{mod}\ \ell).$$

Proof. We choose a character χ of $\mathbb{F}^*$ of order ℓ, and inflate it to a character of the unit group $\mathcal{O}^*$. We then take χ_i to be χ^i for each i. By assumption on ℓ,

$$\{\chi_1, \ldots, \chi_n\} \cap \{\chi_1^{-1}, \ldots, \chi_n^{-1}\} = \varnothing.$$

If $\tilde{\chi}_i$ are quasi-characters of F^* extending the χ_i, the three irreducibility conditions above are satisfied: none of the tamely ramified $\tilde{\chi}_i$ have order two,

$$\tilde{\chi}_i = \nu^{\pm 1}\tilde{\chi}_j^{\pm 1} \implies \chi_i = \chi_j^{\pm 1}.$$

This contradicts the fact that the above intersection is empty. To show χ_i is trivial mod ℓ, note that its reduction maps into the ℓth roots of unity in $\overline{\mathbb{F}}_\ell$. □

2.2. A result of Roche on types

We continue to assume p is odd. We review a result of Roche, based on an explicit construction of types (in the sense of Bushnell and Kutzko) for principal series representations of split reductive groups. Thus, we let I denote the Iwahori subgroup. That is, the inverse image of $B(\mathbb{F})$ under the reduction map:

$$I = \{g \text{ with } g \ (\mathrm{mod} \ \mathfrak{p}) \in B(\mathbb{F})\} \subset G_0(\mathcal{O}).$$

Also, we introduce the normal subgroup I_1, which is the inverse image of $B_u(\mathbb{F})$:

$$I_1 = \{g \text{ with } g \ (\mathrm{mod} \ \mathfrak{p}) \in B_u(\mathbb{F})\} \subset I.$$

The quotient I/I_1 is naturally identified with $T(\mathbb{F})$, by reducing the diagonal. If π is an admissible representation of G_0, the finite abelian group $T(\mathbb{F})$ acts on the I_1-invariants. Therefore, this space has a decomposition into eigenspaces:

$$\pi^{I_1} = \bigoplus_{\chi_i, \chi} \pi^{I:\chi_1 \otimes \cdots \otimes \chi_n \otimes \chi}.$$

Here the χ_i and χ run over all characters of $\mathbb{F}^*$, and the corresponding direct summand is the subspace of vectors transforming according to the associated character of I. We will need the following consequence of Roche's paper [**29**].

Theorem 2.3. *Let π be an irreducible admissible representation of $G_0(F)$. Then*

$$\pi^{I:\chi_1 \otimes \cdots \otimes \chi_n \otimes \chi} \neq 0$$

if and only if π is a subquotient of a principal series representation of the form

$$\tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$$

for some quasi-characters $\tilde{\chi}_i$ extending the χ_i, and some $\tilde{\chi}$ extending χ.

Proof. By inflation, we view χ_i and χ as characters of $\mathcal{O}^*$ trivial on the 1-units $1 + \mathfrak{p}$. This defines an inertial equivalence class in the Bernstein spectrum,

$$\mathfrak{s} = [T, \tilde{\chi}_1 \otimes \cdots \otimes \tilde{\chi}_n \otimes \tilde{\chi}].$$

Here the $\tilde{\chi}_i$ and $\tilde{\chi}$ are arbitrary extensions of χ_i and χ, respectively. By the construction on p. 367 in [**29**], we get an $\mathfrak{s}$-type, according to Theorem 7.7 in [**29**]. This theorem applies since p is odd. By tameness, see Remark 4.2 in Roche's paper, the compact open subgroup of the $\mathfrak{s}$-type is simply the Iwahori subgroup I. Moreover, the character of I is the one coming from the χ_i and χ as above. By the very definition of a type, $\pi|_I$ contains this character if and only if $\mathfrak{s}$ is the inertial class of π. That is, π is a subquotient of a principal series

$$\tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi},$$

where we may have to modify the characters by unramified twists. Done. $\square$

It follows that π^{I_1} is non-zero if and only if π is a subquotient of a tamely ramified principal series (possibly unramified). As a continuation of Corollary 2.2, we get the following.

Corollary 2.4. *With π as above, and the χ_i as in Corollary 2.2, we have for all χ:*

$$\pi^{I:\chi_1\otimes\cdots\otimes\chi_n\otimes\chi} \neq 0$$

if and only if $\pi = \tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$ for some extensions $\tilde{\chi}_i$ and $\tilde{\chi}$ as above.

3. Modular forms and Hecke algebras

For now, let us fix an arbitrary totally real number field F, of degree $d = [F : \mathbb{Q}]$. We view G_0 as an algebraic F-group, and fix an inner form G over F such that

$$G_\infty = G(F \otimes_\mathbb{Q} \mathbb{R}) = \prod_{v|\infty} G(F_v)$$

is *compact* modulo its centre. The centre Z_G is isomorphic to $\mathbb{G}_m$ over F. The groups G_0 and G are isomorphic over the algebraic closure $\bar{\mathbb{Q}}$, and throughout we fix an inner twisting ψ. By definition, this is an isomorphism over $\bar{\mathbb{Q}}$,

$$\psi : G \to G_0,$$

such that $\sigma\psi \circ \psi^{-1}$ is an inner automorphism of G_0 for all $\sigma \in \text{Gal}(\bar{\mathbb{Q}}/F)$. In this fashion, the equivalence class of G corresponds to a cohomology class in

$$H^1(F, G_0^{\text{ad}}).$$

Here G_0^{ad} is the adjoint group, and H^1 is non-abelian Galois cohomology. More concretely, the group G is the unitary similitude group of an n-dimensional left D-vector space equipped with a non-degenerate hermitian form, where D is a quaternion F-algebra [**24**]. By choosing a basis, we can be even more concrete, and realize G as an algebraic subgroup of $\text{GL}_n(D)$. For any F-algebra R,

$$G(R) = \{g \in \text{GL}_n(D \otimes_F R) : {}^*g\Phi g = \mu(g)\Phi\}.$$

Here ${}^*g = {}^\text{t}\bar{g}$, and Φ is non-degenerate and hermitian. That is, ${}^*\Phi = \Phi$. By an explicit calculation, it is not hard to show that the groups G and G_0 are conjugate in GL_{2n} over $\bar{\mathbb{Q}}$. We refer the reader to Lemma 2.3 in [**9**] for the details. If v is a non-archimedean place of F, there is a unique non-split inner form of G_0 over F_v (the reduced norm on the division quaternion F_v-algebra is surjective, so all hermitian forms are equivalent). This inner form has rank $[n/2]$ modulo its centre. If v is archimedean, the hermitian forms are classified by their signature. Globally, there is always an *even* number of places v such that G is non-split over F_v. At those places, the inner form can be prescribed.

Remark. In the special case $n = 2$, we may alternatively view G as the similitude spin group of a totally definite quadratic form in five variables over F.

3.1. Algebraic modular forms on G

We wish to define modular forms on G as in [**10**]. Thus, we first fix a central character ω. This is naturally identified with a Hecke character of F, not necessarily unitary. Let $\mathcal{A}(\omega)$ be the space of ω-central automorphic forms:

$$f : G(F)\backslash G(\mathbb{A}_F) \to \mathbb{C}, \quad f(zg) = \omega(z)f(g), \quad \forall z \in Z_G(\mathbb{A}_F).$$

These functions are assumed to be G_∞-finite under right translation, and locally constant when restricted to the group of finite adeles $G(\mathbb{A}_F^f)$. In what follows, we simply say f is smooth. The group $G(\mathbb{A}_F)$ acts on the space $\mathcal{A}(\omega)$ by right translation, and it decomposes as a direct sum of irreducible automorphic representations $\pi = \pi_\infty \otimes \pi_f$ with $\omega_\pi = \omega$, each occurring with finite multiplicity. We now fix an infinity type. That is, we fix an irreducible complex representation

$$\xi : G_\infty \to \mathrm{GL}(V_\xi), \quad \xi = \otimes_{v|\infty} \xi_v.$$

We assume it has central character ω_∞. This compatibility is necessary for the following spaces of modular forms to be non-trivial. First, let us introduce

$$\mathcal{A}_\xi(\omega) = \mathrm{Hom}_{G_\infty}(\xi, \mathcal{A}(\omega)) = (\check{\xi} \otimes \mathcal{A}(\omega))^{G_\infty}.$$

It carries an induced action of $G(\mathbb{A}_F^f)$, which is admissible, as we will see below. The space $\mathcal{A}_\xi(\omega)$ is a direct sum of the π_f such that $\xi \otimes \pi_f$ is automorphic, with central character ω. Hence, the ξ-isotypic subspace of $\mathcal{A}(\omega)$ may be written as

$$\mathcal{A}(\omega)[\xi] = \xi \otimes \mathcal{A}_\xi(\omega).$$

We now want to realize $\mathcal{A}_\xi(\omega)$ as a space of vector-valued automorphic forms.

Lemma 3.1. *As a $G(\mathbb{A}_F^f)$-representation, $\mathcal{A}_\xi(\omega)$ can be identified with the space of ω-central $\check{V}_\xi$-valued smooth functions on the quotient $G(F)\backslash G(\mathbb{A}_F)$ such that*

$$\mathcal{F}(gg_\infty^{-1}) = \check{\xi}(g_\infty)\mathcal{F}(g), \quad \forall g_\infty \in G_\infty.$$

Proof. To such an $\mathcal{F}$, we associate the map taking $u \in V_\xi$ to $u \circ \mathcal{F}$. $\square$

By restriction to the finite adeles, we get that the space $\mathcal{A}_\xi(\omega)$ can be identified with the space of ω_f-central $\check{V}_\xi$-valued smooth functions $\mathcal{F}$ on $G(\mathbb{A}_F^f)$ such that

$$\mathcal{F}(\gamma_f g) = \check{\xi}(\gamma_\infty)\mathcal{F}(g), \quad \forall \gamma \in G(F).$$

Let K be a compact open subgroup of $G(\mathbb{A}_F^f)$, and look at the K-invariants:

$$\mathcal{A}_\xi(K, \omega) \simeq \bigoplus_{\pi:\pi_\infty=\xi} a_{\mathrm{disc}}^G(\pi)\pi_f^K.$$

Such forms are said to have level K. Their union, as K varies, is all of $\mathcal{A}_\xi(\omega)$. Moreover, the space above is finite dimensional, thereby proving the aforementioned admissibility. Indeed, if we choose a set of representatives $g_1, \ldots, g_h$ for

$$G(F)\backslash G(\mathbb{A}_F^f)/K,$$

the map taking $\mathcal{F}$ to the h-tuple $(\mathcal{F}(g_1), \ldots, \mathcal{F}(g_h))$ defines an isomorphism

$$\mathcal{A}_\xi(K, \omega) \simeq \bigoplus_{i=1}^{\mathrm{h}} \check{V}_\xi^{\Gamma_i}, \quad \Gamma_i = G(F) \cap g_i G_\infty K g_i^{-1}.$$

Note that each Γ_i is a finite subgroup of G_∞. In fact, by shrinking K, one can often assume they are all trivial. For example, this is true if the projection of K onto some G_v does not contain any non-trivial elements of finite order. In this case, we say K is sufficiently small. We will often make this assumption.

Remark. We always assume K is so small that ω_f is trivial on $K \cap Z_G(\mathbb{A}_F^f)$.

3.2. Hecke algebras for G

We let $\mathcal{H}(G)$ denote the Hecke algebra of $G(\mathbb{A}_F^f)$. That is, the smooth functions with compact support, equipped with the convolution product associated to some fixed Haar measure $\mathrm{d}x$. It acts on $\mathcal{A}_\xi(\omega)$ in the usual way. That is,

$$R(\phi)\mathcal{F}(g) = \int_{G(\mathbb{A}_F^f)} \phi(x)\mathcal{F}(gx)\,\mathrm{d}x.$$

Let K be a compact open subgroup of $G(\mathbb{A}_F^f)$. If we normalize $\mathrm{d}x$ such that K has volume one, the characteristic function e_K is an idempotent in $\mathcal{H}(G)$. Then

$$\mathcal{H}(G, K) = e_K * \mathcal{H}(G) * e_K$$

is the algebra of K-biinvariant compactly supported functions. Clearly, it preserves the level K forms. If we fix a complete set of representatives $\{t\}$ for

$$K \backslash G(\mathbb{A}_F^f)/K,$$

the characteristic functions of the double cosets KtK form a basis for $\mathcal{H}(G, K)$. In fact, the $\mathbb{Z}$-submodule they generate is a subring $\mathcal{H}_\mathbb{Z}(G, K)$. We introduce:

$$\mathbb{T}_\xi(K, \omega) = \mathrm{Im}\{R : \mathcal{H}_\mathbb{Z}(G, K) \to \mathrm{End}_\mathbb{C}\,\mathcal{A}_\xi(K, \omega)\}.$$

More generally, let S be a finite set of finite places, and assume K can be written as $K_S K^S$ where the factor K^S is a product of hyperspecial subgroups $G(\mathcal{O}_v)$. Then the Hecke algebra $\mathcal{H}_\mathbb{Z}(G, K)$ factors correspondingly as a tensor product,

$$\mathcal{H}_\mathbb{Z}(G, K) \simeq \mathcal{H}_\mathbb{Z}(G_S, K_S) \otimes \mathcal{H}_\mathbb{Z}^S(G, K),$$

and we let $\mathbb{T}_\xi^S(K, \omega)$ denote the image of the prime-to-S part $\mathcal{H}_\mathbb{Z}^S(G, K)$ in the above endomorphism algebra. This factor of the Hecke algebra is commutative.

3.3. Arithmetic models

Next, we define models of the spaces $\mathcal{A}_\xi(K, \omega)$ over number fields and finite fields. Thus, we assume our inner twisting ψ is defined over the number field L inside $\bar{\mathbb{Q}}$. By enlarging L, we may assume ω_f maps into L^*, and that L contains F. The complexification of ξ is an algebraic representation of the complex group

$$G(F \otimes_{\mathbb{Q}} \mathbb{C}) = \prod_{v | \infty} G(F_v \otimes_{\mathbb{R}} \mathbb{C}).$$

Such an irreducible representation $\otimes_{v | \infty} \xi_v$ is characterized by its highest weight. This is a dominant character of T. In our case, since G is isomorphic to G_0 over $\mathbb{C}$, the highest weight corresponds to a tuple $a = (a_v)_{v | \infty}$ where for each v:

$$a_v = (a_{v,1}, \ldots, a_{v,n}, \tilde{a}_v) \in \mathbb{Z}^{n+1}, \quad a_{v,1} \geqslant \cdots \geqslant a_{v,n} \geqslant 0.$$

Again, by enlarging L, we may assume this algebraic representation is defined over L, viewed as a subfield of $\mathbb{C}$ via our fixed embedding i_∞. We denote it by

$$\xi : \tilde{G}_{/L} \to \mathrm{GL}(V_\xi(L)), \quad \tilde{G} = \mathrm{Res}_{F/\mathbb{Q}}(G).$$

In particular, the group $G(F)$ acts naturally on $V_\xi(L)$: It is a subgroup of $\tilde{G}(L)$ via the diagonal embedding. This allows us to introduce an L-rational structure on the spaces of modular forms. Indeed, we make the following definition.

Definition 3.2. We define the L-rational modular forms $\mathcal{A}_\xi(K, \omega, L)$ to be the space of ω_f-central $\check{V}_\xi(L)$-valued K-invariant functions $\mathcal{F}$ on $G(\mathbb{A}_F^f)$ such that

$$\mathcal{F}(\gamma_f g) = \check{\xi}(\gamma) \mathcal{F}(g), \quad \forall \gamma \in G(F).$$

It carries an action of the Hecke algebra $\mathbb{T}_\xi(K, \omega)$. We now complete the above space ℓ-adically, for a fixed prime ℓ. Thus, let λ be the place of L determined by the embedding i_ℓ, and let L_λ be the corresponding completion of L. We let

$$\mathcal{A}_\xi(K, \omega, L_\lambda) = L_\lambda \otimes_L \mathcal{A}_\xi(K, \omega, L).$$

Of course, this can be identified with smooth $\check{V}_\xi(L_\lambda)$-valued functions on $G(\mathbb{A}_F^f)$ as above. However, for the purpose of defining integral models and modular forms mod ℓ, it is better to work with a different model. Namely, the following.

Lemma 3.3. *As a $\mathbb{T}_\xi(K, \omega)$-module, the space $\mathcal{A}_\xi(K, \omega, L_\lambda)$ can be identified with the space of $\check{V}_\xi(L_\lambda)$-valued functions $\mathcal{F}$ on the quotient $G(F) \backslash G(\mathbb{A}_F^f)$ such that*

$$\mathcal{F}(gk) = \check{\xi}(k_\lambda^{-1}) \mathcal{F}(g), \quad \forall k \in K,$$

and such that the prime-to-λ part of $Z_G(\mathbb{A}_F^f)$ acts on $\mathcal{F}$ by the character ω_f^λ.

Remark. We abuse notation, and use λ also to denote the place of F defined by i_ℓ. The Hecke action in this model, is given by the following explicit formula:

$$R(e_{KtK}) \mathcal{F}(g) = \sum_{x \in KtK/K} \check{\xi}(x_\lambda) \mathcal{F}(gx).$$

Proof. For a function $\mathcal{F}$ in the former model, the function $g \mapsto \check{\xi}(g_\lambda^{-1})\mathcal{F}(g)$ is left-invariant under $G(F)$, and is easily seen to belong to the latter model—and vice versa: the compatibility between ω and the central character of ξ ensures that F_λ^* acts on the form $\mathcal{F}$ by the central character ω_λ as desired. $\qquad\square$

Now, to define integral models, over the valuation ring $\mathcal{O}_\lambda$, we need a lattice preserved by all k_λ for $k \in K$. We denote the group of such elements by K_λ. For this, we assume $\tilde{G}$ extends to a smooth affine group scheme over $\mathcal{O}_\lambda$, and

$$\xi : \tilde{G}_{/\mathcal{O}_\lambda} \to \mathrm{GL}(V_\xi(\mathcal{O}_\lambda))$$

is an extension of the previous ξ. We assume $\tilde{G}(\mathcal{O}_\lambda)$ contains K_λ. Consequently,

$$\mathcal{A}_\xi(K, \omega, L_\lambda) = L_\lambda \otimes_{\mathcal{O}_\lambda} \mathcal{A}_\xi(K, \omega, \mathcal{O}_\lambda),$$

where $\mathcal{A}_\xi(K, \omega, \mathcal{O}_\lambda)$ is the lattice of $\check{V}_\xi(\mathcal{O}_\lambda)$-valued functions $\mathcal{F}$ as above. We immediately observe that this lattice is not preserved the full Hecke algebra $\mathbb{T}_\xi(K, \omega)$, unless ξ is trivial. Only the prime-to-λ part $\mathbb{T}_\xi^\lambda(K, \omega)$ preserves integrality. Finally, to define modular forms mod ℓ, we tensor by $\mathbb{F}_\lambda$. That is,

$$\mathcal{A}_\xi(K, \omega, \mathbb{F}_\lambda) = \mathbb{F}_\lambda \otimes_{\mathcal{O}_\lambda} \mathcal{A}_\xi(K, \omega, \mathcal{O}_\lambda).$$

This can of course be canonically identified with the space of $\check{V}_\xi(\mathbb{F}_\lambda)$-valued functions $\mathcal{F}$ satisfying the transformation properties in the previous lemma.

4. Congruences

In this section, we will prove Theorem A in the introduction. Let us briefly review the setup. Start with an automorphic representation π of $G(\mathbb{A}_F)$ with

$$\omega_\pi = \omega, \qquad \pi_\infty = \xi, \qquad \pi_f^I \neq 0,$$

where I is a compact open subgroup of $G(\mathbb{A}_F^f)$ of the form $I_w K^w$ for a fixed finite place w of F where G splits. Here I_w is an Iwahori subgroup of $G(F_w)$, and K^w is some compact open subgroup prime-to-w. Moreover, we assume that

$$N(\mathfrak{p}_w) \equiv 1 \pmod{\ell}, \quad \ell > 2n.$$

Let us furthermore assume K^w is factorizable of the form $\prod_{v \neq w} K_v$, where K_v is hyperspecial for $v \notin S_0$. We let S_ℓ denote the places of F above ℓ, and consider

$$S = \{w\} \cup S_0 \cup S_\ell.$$

Then the Hecke algebra $\mathbb{T}_\xi^S(I, \omega)$ acts on π_f^I via an algebra character, denoted

$$\eta_\pi : \mathbb{T}_\xi^S(I, \omega) \to \mathbb{C}.$$

It maps into $\bar{\mathbb{Q}}$, since the Hecke algebra preserves $\mathcal{A}_\xi(I, \omega, L)$. In fact, via i_ℓ, it maps into $\bar{\mathbb{Z}}_\ell$. Indeed, the previously defined integral structures are preserved. It therefore makes sense to reduce η_π mod ℓ. We denote this $\bar{\mathbb{F}}_\ell$-valued character by $\bar{\eta}_\pi$, and let $\mathfrak{m}$ be its kernel. This is a maximal ideal of the algebra $\mathbb{T}_\xi^S(I, \omega)$.

Goal. Find an automorphic representation $\tilde{\pi} \equiv \pi \pmod{\ell}$ of $G(\mathbb{A}_F)$ such that

$$\omega_{\tilde{\pi}} = \omega, \qquad \tilde{\pi}_\infty = \xi, \qquad \tilde{\pi}_f^{K^w} \neq 0,$$

and such that $\tilde{\pi}_w$ is a tamely ramified principal series representation of $G(F_w)$.

Here, the notation $\tilde{\pi} \equiv \pi \pmod{\ell}$ means that $\bar{\eta}_{\tilde{\pi}} = \bar{\eta}_\pi$, where the character

$$\eta_{\tilde{\pi}} : \mathbb{T}_\xi^S(I, \omega) \to \mathbb{C}$$

gives the action of $\mathbb{T}_\xi^S(I, \omega)$ on the K^w-invariants of $\tilde{\pi}$. To reach the goal, we use the characters χ_i of the unit group $\mathcal{O}_w^*$ from Corollary 2.2. Moreover, fix an arbitrary $\chi \equiv 1 \pmod{\ell}$. Having fixed the $\{\chi_i, \chi\}$, we introduce the subspace

$$\mathcal{A}_\xi^{\{\chi_i, \chi\}}(I, \omega)$$

of forms $\mathcal{F}$ invariant under $I_{1w} K^w$, on which I_w acts by $\otimes \chi_i \otimes \chi$. We need

$$\mathcal{A}_\xi^{\{\chi_i, \chi\}}(I, \omega)_\mathfrak{m} \neq 0.$$

Indeed, if this is the case, by Corollary 2.4 there is a $\tilde{\pi}$ with $\bar{\eta}_{\tilde{\pi}} = \bar{\eta}_\pi$ such that

$$\tilde{\pi}_w = \tilde{\chi}_1 \times \cdots \times \tilde{\chi}_n \rtimes \tilde{\chi}$$

for some extensions $\tilde{\chi}_i$ and $\tilde{\chi}$. It suffices to show non-vanishing at $\mathfrak{m}$ mod ℓ:

$$\mathcal{A}_\xi^{\{\chi_i, \chi\}}(I, \omega, \mathbb{F}_\lambda)_\mathfrak{m} \simeq \mathbb{F}_\lambda \otimes_{\mathcal{O}_\lambda} \mathcal{A}_\xi^{\{\chi_i, \chi\}}(I, \omega, \mathcal{O}_\lambda)_\mathfrak{m}.$$

Now recall from Corollary 2.2 that all the χ_i are trivial mod ℓ. Therefore,

$$\mathcal{A}_\xi^{\{\chi_i, \chi\}}(I, \omega, \mathbb{F}_\lambda)_\mathfrak{m} = \mathcal{A}_\xi^{\{1,1\}}(I, \omega, \mathbb{F}_\lambda)_\mathfrak{m},$$

and this is simply the space of I-invariants. Here $\mathfrak{m}$ does occur. We are done.

5. Jacquet–Langlands transfer for $n = 2$

To get congruences between automorphic forms on G_0, we need to establish certain cases of functoriality, analogous to the Jacquet–Langlands correspondence for GL_2 and its inner forms. For this, we need to impose the two conditions

$$n = 2, \qquad d = [F : \mathbb{Q}] \text{ is even.}$$

The first condition is dictated by the present state of the trace formula for symplectic groups. The second condition guarantees the existence of a totally definite quaternion algebra D over F, which is split at all the finite places:

$$D \otimes_\mathbb{Q} \mathbb{R} \simeq \mathbb{H}^d, \qquad D_v \simeq M_2(F_v),$$

for all finite v. Here $\mathbb{H}$ denotes the Hamilton quaternions over $\mathbb{R}$. These conditions determine D uniquely up to isomorphism. We take the hermitian form Φ to be the identity, and look at the unitary similitude group. For F-algebras R,

$$G(R) = \{g \in \mathrm{GL}_2(D \otimes_F R) : {}^*gg = \mu(g)I\}.$$

Then G_∞ is compact modulo its centre, and G is split over F_v for all finite v.

5.1. The trace formula for G

The group G is anisotropic modulo its centre Z_G, so the trace formula takes its simplest form. We keep the central character ω fixed, and look at the space

$$C_c^\infty(G(\mathbb{A}_F), \check{\omega})$$

of smooth $\check{\omega}$-central test functions ϕ on $G(\mathbb{A}_F)$, assumed to be compactly supported modulo the centre. Each such function ϕ defines a compact operator

$$R(\phi)f(g) = \int_{G^{\mathrm{ad}}(\mathbb{A}_F)} \phi(x)f(gx)\,\mathrm{d}x$$

on $\mathcal{A}(\omega)$. It turns out to be of trace class, and the trace formula computes its trace in two different ways. First, one obviously has the spectral expansion:

$$I_{\mathrm{disc}}^G(\phi) \overset{\mathrm{df}}{=} \mathrm{tr}\, R(\phi) = \sum_{\pi:\omega_\pi=\omega} a_{\mathrm{disc}}^G(\pi)\,\mathrm{tr}\,\pi(\phi),$$

in terms of the automorphic spectrum of G. One compares it with the geometric expansion of the trace. First, observe that the operator $R(\phi)$ is given by a kernel,

$$R(\phi)f(g) = \int_{Z_G(\mathbb{A}_F)G(F)\backslash G(\mathbb{A}_F)} K_\phi(g,x)f(x)\,\mathrm{d}x,$$

where we have introduced

$$K_\phi(g,x) = \sum_{\gamma \in G^{\mathrm{ad}}(F)} \phi(g^{-1}\gamma x).$$

To get the trace, we integrate K_ϕ over the diagonal. A short calculation shows that

$$\mathrm{tr}\, R(\phi) = \sum_{\{\gamma\}} \mathrm{vol}(Z_G(\mathbb{A}_F)G_\gamma(F)\backslash G_\gamma(\mathbb{A}_F))\, \mathrm{O}_\gamma(\phi).$$

Here $\{\gamma\}$ is a set of representatives for the conjugacy classes in $G^{\mathrm{ad}}(F)$, and G_γ is the centralizer of γ in G. Moreover, O_γ denotes the orbital integral. That is,

$$\mathrm{O}_\gamma(\phi) = \int_{G_\gamma(\mathbb{A}_F)\backslash G(\mathbb{A}_F)} \phi(x^{-1}\gamma x)\,\mathrm{d}x.$$

5.2. Stable orbital integrals

One would like to rewrite the geometric side in terms of *stable* orbital integrals. To define these, let us first recall the notion of stable conjugacy. Eventually we want to compare the geometric side of the trace formula for G with that of the (much more complicated) trace formula for G_0, in order to compare automorphic spectra. However, the two groups are only isomorphic over $\bar{\mathbb{Q}}$, so one can only compare conjugacy classes over $\bar{\mathbb{Q}}$. We work locally at a fixed place v of F.

Definition 5.1. Elements γ and γ' in $G(F_v)$ are *stably* conjugate if and only if

$$\exists g \in G(\bar{F}_v) : \gamma' = g^{-1}\gamma g \pmod{\text{centre}}.$$

In the case of GL_n, this turns out to coincide with usual conjugacy. However, for G_0 it is coarser. We have used the fact that the derived group G^{der} is simply connected; otherwise, the definition has to be altered slightly. A stable conjugacy class contains only finitely many conjugacy classes, parametrized by

$$\ker\{H^1(F_v, G_\gamma) \to H^1(F_v, G)\}.$$

Again, H^1 is non-abelian Galois cohomology. The parametrization is not difficult; one looks at the cohomology class of the cocycle $\sigma(g)g^{-1}$. In general, for arbitrary semisimple γ, the centralizer G_γ is a connected reductive group. If γ is regular, G_γ is a torus. It is connected, again since G^{der} is simply connected. One can show that the above kernel is a finite abelian subgroup.

Definition 5.2. Suppose γ' is stably conjugate to γ. Then $G_{\gamma'}$ is the inner form of G_γ corresponding to the class in $H^1(F_v, G_\gamma^{\mathrm{ad}})$ given by the conjugacy class of γ' in the above parametrization. Let $e(G_{\gamma'})$ be the Kottwitz sign [**17**]. Then

$$\mathrm{SO}_\gamma(\phi) = \sum_{\gamma'} e(G_{\gamma'}) \, \mathrm{O}_{\gamma'}(\phi)$$

is the stable orbital integral. Here $\gamma' \sim \gamma$ runs over the conjugacy classes.

Globally, stable orbital integrals are defined as products of local ones. More generally, one defines κ-orbital integrals for certain characters κ. The stable case corresponds to $\kappa = 1$. A distribution I on $G(\mathbb{A}_F)$ is called *stable* if it is supported on the stable orbital integrals. That is, for all test functions ϕ as above,

$$\mathrm{SO}_\gamma(\phi) = 0, \quad \forall \text{ semisimple } \gamma \in G(F) \implies I(\phi) = 0.$$

5.3. Elliptic endoscopic triples for G

Unfortunately, one cannot express the geometric side of the trace formula for G purely in terms of stable orbital integrals on G. The distribution I_{disc}^G turns out to be unstable. The solution to this problem is to write I_{disc}^G as a sum of stable distributions on the so-called elliptic endoscopic groups. Actually, to be precise, one looks at equivalence classes of triples (H, s, ζ), consisting of data:

- H is a quasi-split connected reductive group over F,

- s is a semisimple element in the dual group $\hat{G}$,

- $\zeta : \hat{H} \hookrightarrow \hat{G}$ is a homomorphism with image $Z_{\hat{G}}(s)^0$.

Such a triple must satisfy two conditions that we will not state precisely (see p. 20 in [**13**]). First, ζ must be Galois-equivariant up to $\hat{G}$-conjugation. Second, s must satisfy some local triviality condition. Finally, the triple is *elliptic* if

$$\zeta(Z(\hat{H})^{\Gamma})Z(\hat{G})/Z(\hat{G}), \quad \Gamma = \mathrm{Gal}(\bar{\mathbb{Q}}/F),$$

is finite. There is a natural equivalence relation on such triples (see Definition 5.2 in [**13**]). In our case, G admits precisely two elliptic endoscopic triples up to equivalence. The first one being the quasi-split inner form G_0 equipped with $s = 1$ and $\zeta = \mathrm{Id}$. The connected reductive group of the second triple is

$$H = [\mathrm{GL}(2) \times \mathrm{GL}(2)]/\mathbb{G}_m.$$

Here $\mathbb{G}_m$ is a subgroup of the centre via the embedding $z \mapsto (z, z^{-1})$. Moreover,

$$s = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \in \hat{G} = \mathrm{GSp}_4(\mathbb{C}).$$

The dual group $\hat{H}$ is the subgroup of $\mathrm{GL}_2(\mathbb{C}) \times \mathrm{GL}_2(\mathbb{C})$ consisting of pairs of matrices with the same determinant. We may take the homomorphism ζ to be

$$\zeta : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix} \mapsto \begin{pmatrix} \tilde{a} & & & \tilde{b} \\ & a & b & \\ & c & d & \\ \tilde{c} & & & \tilde{d} \end{pmatrix}.$$

It is straightforward to check that its image is indeed the centralizer of s in $\hat{G}$.

5.4. Transfer and the fundamental lemma

In this section, in the beginning, we fix a place v and let $G = G_v$ and so on. We wish to make precise, how one transfers a given test function ϕ on G, or G_0, to families of functions on the endoscopic groups G_0 and H. Let us first explain how to go from G to G_0. The set of (semisimple) stable conjugacy classes in G embeds into the set of stable conjugacy classes in G_0. Namely, let $\gamma \in G$ be semisimple, and look at the G_0-conjugacy class of $\psi(\gamma)$ over $\bar{F}_v$. Recall that we have fixed an inner twisting ψ. By a theorem of Kottwitz and Steinberg, since G_0 is quasi-split and G_0^{der} is simply connected, the conjugacy class of $\psi(\gamma)$ intersects the F_v-rational points G_0 in a stable conjugacy class. That is,

$$\exists \gamma_0 \in G_0 \text{ such that } \gamma_0 \sim_{G_0(\bar{F}_v)} \psi(\gamma).$$

We write $\gamma \leftrightarrow \gamma_0$ in this case, and say that γ_0 is an *image* of γ. For a given γ_0, when such a γ exists, we say that γ_0 *occurs* in G. We need the following result.

Theorem 5.3. *For every $\phi \in C_c^\infty(G, \breve{\omega}_v)$ there exists a $\phi^{G_0} \in C_c^\infty(G_0, \breve{\omega}_v)$ with*

$$\mathrm{SO}_{\gamma_0}(\phi^{G_0}) = \mathrm{SO}_\gamma(\phi)$$

when $\gamma \leftrightarrow \gamma_0$, and such that $\mathrm{SO}_{\gamma_0}(\phi^{G_0}) = 0$ whenever γ_0 does not occur in G.

Proof. In the archimedean case, this was proved by Shelstad in [**30**]. In the non-archimedean case, it is a deep theorem of Waldspurger, proved in [**37**]. □

In this situation, we say that ϕ and ϕ^{G_0} have *matching* orbital integrals. Of course, ϕ^{G_0} is not unique, but $I(\phi^{G_0})$ is well defined for stable distributions I. Next, let us explain how to go from G to H. The discussion with G replaced by G_0 is similar, but simpler. As above, take a semisimple element $\gamma \in G$. We first define what it means for a semisimple $\gamma_H \in H$ to be an *image* of γ. Choose a maximal torus T_H in the centralizer H_{γ_H}, containing γ_H. There is a canonical G-conjugacy class of embeddings of T_H into G defined over $\bar{F}_v$. Choose one, say

$$j : T_H \hookrightarrow G,$$

and assume it maps into the maximal torus T in G. In analogy with the above discussion, we consider the stable conjugacy class of $j(\gamma_H)$. If it contains γ, we say that γ_H is an image of γ. Alternative terminology; say γ *comes* from γ_H.

Definition 5.4. The roots Φ of G relative to T define characters of T_H by pulling back via j. We say that a semisimple element $\gamma_H \in H$ is (G, H)-regular if $\alpha(\gamma_H) \neq 1$ for all $\alpha \in \Phi$ such that the j-restriction α_{T_H} is not a root of H.

The following is the Langlands–Shelstad transfer conjecture, which is known in our setup.

Theorem 5.5. *For every $\phi \in C_c^\infty(G, \breve{\omega}_v)$ there exists a $\phi^H \in C_c^\infty(H, \breve{\omega}_v)$ with*

$$\mathrm{SO}_{\gamma_H}(\phi^H) = \sum_\gamma \Delta_{G,H}(\gamma_H, \gamma) e(G_\gamma) \, \mathrm{O}_\gamma(\phi)$$

for all (G, H)-regular γ_H. The sum is over γ coming from γ_H, up to conjugacy.

Proof. Again, the archimedean case was worked out completely for all groups by Shelstad in [**30**] and [**31**]. The non-archimedean case follows from work of Hales [**11**, **12**] combined with either [**22**] or [**37**]. □

Remark. Once we have normalized the Langlands–Shelstad transfer factors $\Delta_{G,H}$, compatible measures must be used on both sides. Fortunately, we will not need a precise definition of $\Delta_{G,H}$, so we will not give it here. The reader is referred to Chapter 7 in [**13**] for a nice review of the basic properties. We also note that, since H has no endoscopy, the left-hand side of the transfer identity is simply the standard orbital integral $\mathrm{O}_{\gamma_H}(\phi^H)$, up to a sign at most.

Globally, the test functions are spanned by pure tensors $\phi = \otimes \phi_v$, where ϕ_v is the characteristic function of a hyperspecial subgroup K_v for almost all v; they are all conjugate. We would like to transfer ϕ to a function ϕ^H by the formula

$$\phi^H \stackrel{\mathrm{df}}{=} \otimes_v \phi_v^{H_v}.$$

However, for this to make sense, we need to know that we can take $\phi_v^{H_v}$ to be the characteristic function of a hyperspecial subgroup K_v^H for almost all v. This is the so-called standard fundamental lemma, proved by Hales in the case $n = 2$.

Theorem 5.6. $e_{K_v}^{H_v} = e_{K_v^H}$ *(up to a non-zero constant depending on the measures).*

Proof. This is the main result of [**12**]. □

Remark. We mention an amazing observation of Waldspurger. For general groups, the fundamental lemma in fact implies the transfer conjecture [**37**].

To recapitulate, we can transfer a *global* function ϕ on $G(\mathbb{A}_F)$ to a function ϕ^H on $H(\mathbb{A}_F)$, locally with matching orbital integrals, which is well defined on stable distributions. Similarly for functions on the group $G_0(\mathbb{A}_F)$. Furthermore,

$$\phi^{G_0} \stackrel{\mathrm{df}}{=} \otimes_v \phi_v^{G_0,v}$$

defines the transfer of ϕ to $G_0(\mathbb{A}_F)$, well defined since $G = G_0$ at almost all v.

5.5. Character relations for the endoscopic lift

We go back to the local notation from the previous section. Thus, we fix a place v and let $G = G_v$ and so on. In this section, we discuss the local functorial lift pertaining to the L-homomorphism extending the map ζ. The lift takes an irreducible admissible representation ρ of H, and associates to it an L-packet of representations of G_0. If a certain relevancy condition is satisfied, this packet descends to G. The representation ρ is of the form $\rho_1 \otimes \rho_2$, where the ρ_i are representations of GL_2 with the same central character. The distribution $\mathrm{tr}\,\rho$ is stable, since H has no endoscopy. By general results of Arthur [**2**] in the non-archimedean case, and of Shelstad [**30**] in the archimedean case,

$$\mathrm{tr}\,\rho(\phi^H) = \sum_\pi \Delta_{G,H}(\rho, \pi)\,\mathrm{tr}\,\pi(\phi), \quad \forall \phi \in C_c^\infty(G, \check{\omega}_v),$$

for some coefficients $\Delta_{G,H}(\rho, \pi)$ one should think of as spectral analogues of the transfer factors. It is zero for almost all irreducible admissible representations π of G. Similar expansions hold when G is replaced by G_0. In the non-archimedean case, the above expansion has been made very precise by Weissauer in the two preprints [**39**] and [**40**]: for example, when ρ is in the *discrete* series,

$$\mathrm{tr}\,\rho(\phi^H) = \mathrm{tr}\,\pi_+(\rho)(\phi) - \mathrm{tr}\,\pi_-(\rho)(\phi).$$

Here $\{\pi_\pm(\rho)\}$ is the associated L-packet of representations of G (recall that G and G_0 are isomorphic at *all* finite places). The representation $\pi_+(\rho)$ is always generic, in the sense

that it has a Whittaker model, whereas the other member $\pi_-(\rho)$ is always non-generic. One has the following qualitative description.

- $\rho_1 \neq \rho_2$ (ρ *regular*): $\pi_+(\rho)$ is a discrete series, $\pi_-(\rho)$ is supercuspidal.

- $\rho_1 = \rho_2$ (ρ *invariant*): $\{\pi_\pm(\rho)\}$ are limits of discrete series ($\Rightarrow$ tempered).

We will not need it in this paper, but when the residue characteristic of F_v is *odd*, the $\pi_\pm(\rho)$ have a nice description in terms of θ-correspondence for similitude groups. This comes about, by identifying H and its inner form $\breve{H}$ with special orthogonal similitude groups:

$$H \simeq \mathrm{GSO}(V_{\mathrm{split}}), \qquad \breve{H} = [D^* \times D^*]/\mathbb{G}_m \simeq \mathrm{GSO}(V_{\mathrm{aniso}}).$$

Here V_{split} and V_{aniso} are the split and anistropic quaternary quadratic spaces of discriminant one, respectively. Since ρ is in the discrete series, it corresponds to a unique representation $\breve{\rho}$ of $\breve{H}$ under the Jacquet–Langlands correspondence for GL_2. By [**28**], both ρ and $\breve{\rho}$ have unique extensions to the corresponding GOs participating in the θ-correspondence with GSp_4. We denote them by ρ^+ and $\breve{\rho}^+$. In the regular case, they are simply the induced representations. Then

$$\pi_+(\rho) = \theta(\rho^+), \qquad \pi_-(\rho) = \theta(\breve{\rho}^+).$$

For this, see Proposition 1 on p. 14 in [**40**]. In the case of *even* residue characteristic, one can can still get an interpretation in terms of θ-correspondence by looking at the *globally relevant* subset. Finally, when ρ is a *principal* series,

$$\mathrm{tr}\,\rho(\phi^H) = \mathrm{tr}\,\pi_+(\rho)(\phi),$$

for a generic representation $\pi_+(\rho)$ (tempered if ρ is tempered). When one of the ρ_i are one dimensional, the support of the expansion of $\mathrm{tr}\,\rho$ does not define L-packets, but A-packets. Their description is given explicitly in [**39**]. One should mention that, the above results of Weissauer are complete analogues of the results known in the archimedean case, due to the work of Shelstad [**30**]. Globally, one immediately deduces exact analogues of the above character expansions: if $\rho = \otimes\rho_v$ is an automorphic representation of H, then we have

$$\mathrm{tr}\,\rho(\phi^H) = \sum_\pi \Delta_{G,H}(\rho,\pi)\,\mathrm{tr}\,\pi(\phi), \quad \forall\phi \in C_c^\infty(G(\mathbb{A}_F),\breve{\omega}),$$

where $\Delta_{G,H}(\rho,\pi)$ is the product of all the local coefficients $\Delta_{G_v,H_v}(\rho_v,\pi_v)$. Here π runs over the global L-packet for G associated with ρ, that is, the tensor product of all the local L-packets. It is possibly empty for some ρ.

5.6. Stabilization of the trace formula

A key observation is that H itself has *no* endoscopy; stable conjugacy is the same as conjugacy. Indeed, H is essentially just a product of two copies of GL_2. Nevertheless, the trace formula for H is more complicated than that for G, due to the existence of parabolic subgroups defined over F. One has to build the continuous spectrum from

Eisenstein series, and subtract this contribution to get the trace of $R(\phi)$ on the discrete spectrum. In the case of GL_2 this is done in detail in [**7**], and Arthur dealt with arbitrary rank one groups in [**1**]. In addition to the trace of $R(\phi)$ on the discrete spectrum, there are certain so-called scattering terms occurring discretely on the spectral side of the trace formula. We collect all these terms, and denote the resulting distribution by I_{disc}^H. From what we have just said, this is a stable distribution. It can be written as follows:

$$I_{\mathrm{disc}}^H(\phi) = \sum_{\rho:\omega_\rho=\omega} a_{\mathrm{disc}}^H(\rho)\operatorname{tr}\rho(\phi), \quad \forall \phi \in C_c^\infty(H(\mathbb{A}_F), \check\omega).$$

Here ρ runs over all discrete automorphic representations of H with $\omega_\rho = \omega$. It is of the form $\rho_1 \otimes \rho_2$, where the ρ_i are automorphic representations of GL_2 with central character ω. In general, the complex coefficient $a_{\mathrm{disc}}^H(\rho)$ may not be equal to the multiplicity of ρ. However, in the generic case where no ρ_i is a character, they *are* equal. The trace formula for G_0 is even more complicated. There are now three proper parabolic subgroups, up to conjugacy, each contributing to the continuous spectrum. Again, besides the trace of $R(\phi)$ on the discrete spectrum, we get additional discretely occurring terms on the spectral side. They are what Arthur refers to as surviving remnants of Eisenstein series. As before, we construct a distribution $I_{\mathrm{disc}}^{G_0}$ out of these terms, and write it as

$$I_{\mathrm{disc}}^{G_0}(\phi) = \sum_{\Pi:\omega_\Pi=\omega} a_{\mathrm{disc}}^{G_0}(\Pi)\operatorname{tr}\Pi(\phi), \quad \forall \phi \in C_c^\infty(G_0(\mathbb{A}_F), \check\omega).$$

Here Π varies over all discrete automorphic representations of G_0 with central character ω. The coefficient $a_{\mathrm{disc}}^{G_0}(\Pi)$ equals the multiplicity of Π, if Π is cuspidal but not CAP. Recall that a cuspidal automorphic representation Π is CAP with respect to a parabolic P, if it is nearly equivalent to the constituents of a representation induced from a cusp form on a Levi factor M of P. For GL_n, this notion is vacuous by a result of Shalika. On the other hand, for G_0, the CAP forms have been constructed via theta correspondence by Piatetski-Shapiro [**23**] and Soudry [**34**]. As with I_{disc}^G, the distribution $I_{\mathrm{disc}}^{G_0}$ is also unstable. However, if we subtract suitable endoscopic error terms, we can *make* it stable.

Theorem 5.7. *The distribution* $S_{\mathrm{disc}}^{G_0}(\phi) \stackrel{\mathrm{df}}{=} I_{\mathrm{disc}}^{G_0}(\phi) - \frac{1}{4} I_{\mathrm{disc}}^H(\phi^H)$ *is stable.*

Proof. This is a special case of a theorem of Arthur announced in [**3**]. □

We thus arrive at the stable trace formula for G.

Theorem 5.8. $I_{\mathrm{disc}}^G(\phi) = S_{\mathrm{disc}}^{G_0}(\phi^{G_0}) + \frac{1}{4} I_{\mathrm{disc}}^H(\phi^H)$.

Proof. This was first proved by Kottwitz [**17**] and Langlands [**20**], but can be viewed also as a special case of the aforementioned work of Arthur [**3**]. □

5.7. A trace identity at infinity

Finally, we are ready to compare the automorphic spectra of G and G_0. The link is the stable distribution $S_{\mathrm{disc}}^{G_0}$. We insert the character expansions of the $\operatorname{tr}\rho$ into I_{disc}^H, in order

to express the contribution from H to the stable trace formula in terms of the spectrum of G. For all test functions ϕ on G, we get

$$S_{\text{disc}}^{G_0}(\phi^{G_0}) = I_{\text{disc}}^G(\phi) - \tfrac{1}{4} I_{\text{disc}}^H(\phi^H) = \sum_{\pi:\omega_\pi=\omega} \tilde{a}_{\text{disc}}^G(\pi) \operatorname{tr} \pi(\phi),$$

where

$$\tilde{a}_{\text{disc}}^G(\pi) \overset{\text{df}}{=} a_{\text{disc}}^G(\pi) - \frac{1}{4} \sum_{\rho:\omega_\rho=\omega} a_{\text{disc}}^H(\rho) \Delta_{G,H}(\rho,\pi).$$

Similarly, for any test function ϕ_0 on G_0, we get the the analogous expansion:

$$S_{\text{disc}}^{G_0}(\phi_0) = I_{\text{disc}}^{G_0}(\phi_0) - \tfrac{1}{4} I_{\text{disc}}^H(\phi_0^H) = \sum_{\Pi:\omega_\Pi=\omega} \tilde{a}_{\text{disc}}^{G_0}(\Pi) \operatorname{tr} \Pi(\phi_0),$$

where

$$\tilde{a}_{\text{disc}}^{G_0}(\Pi) \overset{\text{df}}{=} a_{\text{disc}}^{G_0}(\Pi) - \frac{1}{4} \sum_{\rho:\omega_\rho=\omega} a_{\text{disc}}^H(\rho) \Delta_{G_0,H}(\rho,\Pi).$$

Taken together, for any pair of matching functions $\phi \mapsto \phi^{G_0}$, we get the relation

$$\sum_{\Pi:\omega_\Pi=\omega} \tilde{a}_{\text{disc}}^{G_0}(\Pi) \operatorname{tr} \Pi(\phi^{G_0}) = \sum_{\pi:\omega_\pi=\omega} \tilde{a}_{\text{disc}}^G(\pi) \operatorname{tr} \pi(\phi).$$

Now, the groups G and G_0 are isomorphic at all finite places, so by linear independence of characters on $G(\mathbb{A}_F^f)$ we may isolate the finite part. That is, for every irreducible admissible representation τ_f of $G(\mathbb{A}_F^f)$ we have the identity

$$\sum_{\Pi_\infty:\omega_{\Pi_\infty}=\omega_\infty} \tilde{a}_{\text{disc}}^{G_0}(\Pi_\infty \otimes \tau_f) \operatorname{tr} \Pi_\infty(\phi_\infty^{G_0}) = \sum_{\pi_\infty:\omega_{\pi_\infty}=\omega_\infty} \tilde{a}_{\text{disc}}^G(\pi_\infty \otimes \tau_f) \operatorname{tr} \pi_\infty(\phi_\infty),$$

for any pair of matching functions $\phi_\infty \mapsto \phi_\infty^{G_0}$. Of course, we assume that the fixed representation τ_f has central character ω_f. In order to compare the coefficients of both sides, we invoke the character relations of Shelstad. However, first we remove the tildes by assuming τ_f is *non-endoscopic*, in the sense that

$$\forall \rho \; \exists v \nmid \infty : \Delta_{G_v,H_v}(\rho_v,\tau_v) = 0.$$

Furthermore, we assume τ_f is *tempered*, meaning that τ_v is tempered for almost all v. Then, by Proposition 6.3 on p. 556 in [**4**] (which assumes the congruence relation on p. 555), the representation $\Pi_\infty \otimes \tau_f$ can only be automorphic for tempered Π_∞; once we know Π_∞ is cohomological for some ξ. It must be, if it contributes to the left-hand side of the above trace identity.

5.8. Shelstad's character relations at infinity

Recall our standing assumption: τ_f is a non-endoscopic tempered representation of $G(\mathbb{A}_F^f)$. We wish to compare the trace identity from the previous section with the character relations of Shelstad [**30**]. In the archimedean case, the local Langlands correspon-

dence is completely known [**21**]. Thus, the irreducible admissible representations of $G_{0,\infty}$ are partitioned into finite subsets Π_μ, the L-packets, parametrized by L-parameters μ. Such a μ corresponds to a tuple

$$\mu = (\mu_v)_{v|\infty}, \quad \mu_v : W_{\mathbb{R}} \to {}^L G_{0,v}.$$

Here $W_{\mathbb{R}}$ is the Weil group of $\mathbb{R}$, a non-trivial extension of $\{\pm 1\}$ by $\mathbb{C}^*$, and μ_v is a so-called admissible homomorphism. For each v, associated with μ_v, there is an L-packet Π_{μ_v} of representations of $G_{0,v}$. Then Π_μ is the tensor product $\otimes_{v|\infty} \Pi_{\mu_v}$. The group G_∞ is compact mod centre, so its L-packets are singletons, parametrized by the subset of *discrete* parameters μ; this means no μ_v maps into a parabolic subgroup. We denote the corresponding representation of G_∞ by π_μ. When μ_v is discrete, the corresponding packet for $G_{0,v}$ consists of two discrete series representations. Precisely one of them is generic (that is, it has a Whittaker model), whereas the other is holomorphic. We will use the notation

$$\Pi_{\mu_v} = \{\Pi_{\mu_v}^{\mathrm{g}}, \Pi_{\mu_v}^{\mathrm{h}}\}.$$

Thus, the L-packet Π_μ consists of exactly 2^d discrete series representations.

Lemma 5.9. *For a fixed μ, the coefficient $a_{\mathrm{disc}}^{G_0}(\Pi_\infty \otimes \tau_f)$ is constant for $\Pi_\infty \in \Pi_\mu$.*

Proof. For a fixed τ_f, we introduce the following distribution on $G_{0,\infty}$,

$$T \overset{\mathrm{df}}{=} \sum_{\Pi_\infty : \omega_{\Pi_\infty} = \omega_\infty} a_{\mathrm{disc}}^{G_0}(\Pi_\infty \otimes \tau_f) \operatorname{tr} \Pi_\infty.$$

Then T is a *stable* distribution. Suppose ϕ is a function on $G_{0,\infty}$ with vanishing stable orbital integrals. Then it matches the zero function on G_∞. By the trace identity from the previous section, $T(\phi) = 0$. Therefore, we can expand T as

$$T = \sum_\mu c_\mu \operatorname{tr} \Pi_\mu, \quad \operatorname{tr} \Pi_\mu \overset{\mathrm{df}}{=} \sum_{\Pi_\infty \in \Pi_\mu} \operatorname{tr} \Pi_\infty.$$

Since τ_f is tempered, only the tempered Π_∞ can contribute to this sum. So, we are summing over the tempered parameters μ. That is, those whose image is bounded when projected onto the dual group. By linear independence of characters for $G_{0,\infty}$, we can match the coefficients and get what we aimed for,

$$a_{\mathrm{disc}}^{G_0}(\Pi_\infty \otimes \tau_f) = c_\mu$$

for all $\Pi_\infty \in \Pi_\mu$. Here μ is an arbitrary (tempered) L-parameter. $\square$

Now we can group together the terms on the left-hand side of the trace identity into packets, and invoke the following character relation due to Shelstad [**30**]:

$$\operatorname{tr} \Pi_\mu(\phi_\infty^{G_0}) = \operatorname{tr} \pi_\mu(\phi_\infty), \quad \mu \text{ discrete}.$$

Moreover, $\operatorname{tr} \Pi_\mu(\phi_\infty^{G_0}) = 0$ when μ is *not* discrete. Again, using linear independence of characters on the group G_∞, we deduce the equality of multiplicities:

$$a_{\mathrm{disc}}^G(\pi_\mu \otimes \tau_f) = a_{\mathrm{disc}}^{G_0}(\Pi_\infty \otimes \tau_f),$$

for all $\Pi_\infty \in \Pi_\mu$ and an arbitrary discrete L-parameter μ. This equality immediately enables us to transfer automorphic representations from G to G_0, and vice versa! The proof of Theorem B on Jacquet–Langlands transfer is complete.

References

1. J. ARTHUR, The Selberg trace formula for groups of F-rank one, *Annals Math. (2)* **100** (1974), 326–385.
2. J. ARTHUR, On local character relations, *Selecta Math.* **2**(4) (1996), 501–579.
3. J. ARTHUR, Towards a stable trace formula, in *Proc. Int. Congress of Mathematicians, Volume 2, Berlin, 1998*, Documenta Mathematica, Extra Volume II, pp. 507–517 (1998).
4. D. BLASIUS AND J. ROGAWSKI, Zeta functions of Shimura varieties, in *Motives, Seattle, WA, 1991*, Proceedings of Symposia in Pure Mathematics, Volume 55, Part 2, pp. 525–571 (American Mathematical Society, Providence, RI, 1994).
5. L. CLOZEL, M. HARRIS AND R. TAYLOR, Automorphy for some l-adic lifts of automorphic mod l representations, preprint (available at www.math.harvard.edu/~rtaylor/).
6. W. T. GAN AND S. TAKEDA, The local Langlands conjecture for GSp(4), preprint (available at www.math.ucsd.edu/~wgan/).
7. S. GELBART AND H. JACQUET, Forms of GL$_2$ from the analytic point of view, in *Automorphic forms, representations and L-functions*, Proceedings of Symposia in Pure Mathematics, Volume 33, Part 1, pp. 213–251 (American Mathematical Society, Providence, RI, 1979).
8. A. GENESTIER AND J. TILOUINE, Systemes de Taylor–Wiles pour GSp(4), in *Formes automorphes, II, Le cas du group* GSp(4), Astérisque, Volume 302, pp. 177–290 (Société Mathématiques de France, Paris, 2005).
9. A. GHITZA, Hecke eigenvalues of Siegel modular forms (mod p) and of algebraic modular forms, *J. Number Theory* **106**(2) (2004), 345–384.
10. B. GROSS, Algebraic modular forms, *Israel J. Math.* **113** (1999), 61–93.
11. T. HALES, Shalika germs on GSp$_4$, Orbites unipotentes et representations, II, *Astérisque* **171–172** (1989), 195–256.
12. T. HALES, The fundamental lemma for Sp(4), *Proc. Am. Math. Soc.* **125**(1) (1997), 301–308.
13. M. HARRIS, Introduction to *The stable trace formula, Shimura varieties, and arithmetic applications*, Volume 1 (available at www.math.jussieu.fr/~harris/), forthcoming.
14. D. HELM, Mazur's principle for $U(2,1)$ Shimura varieties, preprint (available at www.math.harvard.edu/~dhelm/).
15. Y. IHARA, On modular curves over finite fields, in *Discrete Subgroups of Lie Groups and Applications to Moduli, International Colloquium, Bombay, 1973*, pp. 161–202 (Oxford University Press, 1975).
16. D. KEYS, On the decomposition of reducible principal series representations of p-adic Chevalley groups, *Pac. J. Math.* **101**(2) (1982), 351–388.

17. R. Kottwitz, Stable trace formula: elliptic singular terms, *Math. Annalen* **275**(3) (1986), 365–399.

18. J.-P. Labesse, *Cohomologie, stabilisation et changement de base* (with two appendices: Appendix A by L. Clozel and J.-P. Labesse; Appendix B by L. Breen), Astérisque, Volume 257 (Société Mathématiques de France, Paris, 1999).

19. R. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics, Volume 544 (Springer, 1976).

20. R. Langlands, *Les debuts d'une formule des traces stable*, Publications Mathématiques de l'Université Paris VII, Volume 13 (Universite de Paris VII, UER de Mathématiques, Paris, 1983).

21. R. Langlands, On the classification of irreducible representations of real algebraic groups, in *Representation theory and harmonic analysis on semisimple Lie groups*, Mathematical Surveys and Monographs, Volume 31, pp. 101–170 (American Mathematical Society, Providence, RI, 1989).

22. R. Langlands and D. Shelstad, Descent for transfer factors, in *The Grothendieck Festschrift*, Volume II, pp. 485–563, Progress in Mathematics, Volume 87 (Birkhäuser, Boston, MA, 1990).

23. I. Piatetski-Shapiro, On the Saito–Kurokawa lifting, *Invent. Math.* **71**(2) (1983), 309–338.

24. V. Platonov and A. Rapinchuk, *Algebraic groups and number theory* (translated from the 1991 Russian original by R. Rowen), Pure and Applied Mathematics, Volume 139 (Academic Press, 1994).

25. D. Ramakrishnan, Pure motives and automorphic forms, in *Motives, Seattle, WA, 1991*, Proceedings of Symposia in Pure Mathematics, Volume 55, Part 2, pp. 411–446 (American Mathematical Society, Providence, RI, 1994).

26. K. Ribet, Congruence relations between modular forms, in *Proc. Int. Congress of Mathematicians, Volume 1, Warsaw, 1983*, pp. 503–514 (PWN, Warsaw, 1984).

27. K. Ribet, Report on mod l representations of Gal($\bar{\mathbb{Q}}/\mathbb{Q}$), in *Motives, Seattle, WA, 1991*, Proceedings of Symposia in Pure Mathematics, Volume 55, Part 2, pp. 639–676 (American Mathematical Society, Providence, RI, 1994).

28. B. Roberts, The non-Archimedean theta correspondence for GSp(2) and GO(4), *Trans. Am. Math. Soc.* **351**(2) (1999), 781–811.

29. A. Roche, Types and Hecke algebras for principal series representations of split reductive p-adic groups, *Annales Scient. Éc. Norm. Sup.* **31**(3) (1998), 361–413.

30. D. Shelstad, Characters and inner forms of a quasi-split group over $\mathbb{R}$, *Compositio Math.* **39**(1) (1979), 11–45.

31. D. Shelstad, L-indistinguishability for real groups, *Math. Annalen* **259**(3) (1982), 385–430.

32. C. Skinner and A. Wiles, Base change and a problem of Serre, *Duke Math. J.* **107**(1) (2001), 15–25.

33. C. Sorensen, Level-raising for Saito–Kurokawa forms, in preparation.

34. D. Soudry, The CAP representations of GSp(4, $\mathbb{A}$), *J. Reine Angew. Math.* **383** (1988), 87–108.

35. M. Tadic, Representations of p-adic symplectic groups, *Compositio Math.* **90**(2) (1994), 123–181.

36. R. Taylor, Automorphy for some l-adic lifts of automorphic mod l representations, II, preprint (available at www.math.harvard.edu/~rtaylor/).

37. J.-L. Waldspurger, Le lemme fondamental implique le transfert, *Compositio Math.* **105**(2) (1997), 153–236.

38. R. Weissauer, Four dimensional Galois representations, in *Formes automorphes, II, Le cas du group* GSp(4), Astérisque, Volume 302, pp. 67–150 (Société Mathématiques de France, Paris, 2005).

39. R. WEISSAUER, Character identities and Galois representations related to the group GSp(4), preprint (available at www.mathi.uni-heidelberg.de/~weissaue/papers.html).
40. R. WEISSAUER, Endoscopy for GSp(4), preprint (available at www.mathi.uni-heidelberg.de/~weissaue/papers.html).
41. A. WILES, Modular elliptic curves and Fermat's last theorem, *Annals Math.* **141**(3) (1995), 443–551.
42. N. WINARSKY, Reducibility of principal series representations of p-adic Chevalley groups, *Am. J. Math.* **100**(5) (1978), 941–956.