CONSTRUCTION OF AUTOMORPHIC GALOIS REPRESENTATIONS: THE SELF-DUAL CASE

SUG WOO SHIN

This survey paper grew out of the author's lecture notes for the Clay Summer School in 2009. Proofs are often omitted or sketched, but a number of references are given to help the reader find more details if he or she wishes. The article aims to be a guide to the outline of argument for constructing Galois representations from the so-called regular algebraic conjugate self-dual automorphic representations, implementing recent improvements in the trace formula and endoscopy.

Warning. Although it is important to keep track of various constants (e.g. sign matters!), twists by characters, etc in proving representation-theoretic identities, I often ignore them in favor of simpler notation and formulas. Thus many identities and statements in this article should be taken with a grain of salt and not cited in academic work. Correct and precise statements can be found in original papers. I should also point out that significant progress has been made, especially in the case of *non-self-dual* automorphic representations ([HLTT16], [Sch15]) since the original version of these notes was written. I have kept updates minimal (mostly taking place in introduction, adding references to some recent developments) and do not discuss these exciting new results here. The interested readers are referred to survey papers such as [Wei16, Mor16, Scha, Sch16, Car].

Acknowledgments. I heartily thank the Clay Mathematics Institute for its generous support toward the 2009 summer school and choosing an exciting location as Hawaii. I am grateful to the participants for their interest and wonderful questions.

1. Introduction

1.1. **Reading list.** This incomplete list of survey articles is only intended to suggest some starting points for further learning. The reader should not feel compelled to go through too many of them before starting to read the present article.

- [Tay04] is a nice survey of various topics on Galois representations.
- local and global Weil groups, Weil-Deligne groups [Tat79]; L-groups, morphisms of L-groups (L-morphisms) and Satake isomorphisms [Bor79].
- representation theory of GL_n over p-adic fields or archimedean fields with a view toward the local Langlands correspondence [Kud94], [Kna94]
- [Art05] may be a good place where one can start to learn the Arthur-Selberg trace formula. Arthur also wrote several short survey papers on the trace formula, which one may find very helpful.
- (conjectural formulation of) endoscopy for unitary groups: [Rog92, §2], [Mok15]
- [BR94] would be helpful in that it reviews numerous concepts that constantly show up in the study of cohomology of Shimura varieties.

- Surveys on various topics rotating around the trace formula and endoscopy can be found in the Paris book project volume I [CHLN11].
- We do not do enough justice to the history of the subject though we have a few remarks in §1.5 below. The best source is papers by Langlands on the subject, found at:
  `http://publications.ias.edu/rpl/section/26`

The following are research papers and books where one can seriously learn some of the advanced topics that are important to this article. I do not claim by any means that this list is even nearly complete but let me add that there has been recent progress on extending some key constructions like Newton stratification, Igusa varieties, and Rapoport-Zink spaces to the setup for Shimura varieties of Hodge type (and sometimes abelian type) by Hamacher, Howard-Pappas, and W. Kim among others.

- Base change for unitary groups: [Lab11], [Mok15]; for GL_n: [AC89]
- PEL-type Shimura varieties: [Kot92b] (esp. §5), [HT01, Ch III], [Mil05] (esp. §8), [Lan13]
- Newton stratification (in the case of interest): [HT01, III.4], [Man]
- Igusa varieties: [HT01, Ch IV], [Man05]
- Rapoport-Zink spaces: [RZ96], [Far04]
- Stabilization of (elliptic part of) the trace formula: [Kot86], [Kot90]

Finally, §1.5 contains some research papers that are directly related to the main theorem to be discussed.

1.2. **Notation.**

- $\mathbb{A}$ is the adèle ring over $\mathbb{Q}$. If S is a finite set of places of $\mathbb{Q}$ then $\mathbb{A}^S$ is the restricted product of $\mathbb{Q}_v$ over $v \notin S$. In particular, $\mathbb{A}^\infty$ is the ring of finite adèles. If F is a finite extension of $\mathbb{Q}$, $\mathbb{A}_F := \mathbb{A} \otimes_{\mathbb{Q}} F$.
- $\mathrm{Irr}(G(K))$ is the set of isomorphism classes of irreducible smooth representations of $G(K)$, when G is a connected reductive group over a p-adic field K.
- $\mathrm{Irr}(G(\mathbb{A}_F))$ is the set of isomorphism classes of irreducible admissible representations of $G(\mathbb{A})$, when G is a connected reductive group over a number field F. Similarly $\mathrm{Irr}(G(\mathbb{A}_F^S))$ is defined.
- $\mathrm{Groth}(G)$ is the Grothendieck group of admissible and/or continuous representations of G, where G can be a p-adic Lie group, a finite adélic group, a Galois group, and so on. The precise definition is found in [HT01, pp.23-25].
- $\boxplus$ signifies the irreducible parabolic induction either for smooth representations of a p-adic group or for automorphic representations.

1.3. **Main theorem.** Let F be a number field. There is a famous conjecture of Langlands, complemented by an observation of Clozel, Fontaine and Mazur, which goes as follows. For the notion of a compatible family of Galois representations (for varying primes l and field isomorphisms $\overline{\mathbb{Q}}_l \simeq \mathbb{C}$), see [Tay04, §1] or [BLGGT14, §5].

Conjecture 1.1. *There is a bijection between the following two sets consisting of isomorphism classes.*

$$\left\{ \begin{array}{c} \text{cuspidal automorphic} \\ \text{reps } \Pi \text{ of } GL_m(\mathbb{A}_F) \\ \text{algebraic at } \infty \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{compatible systems of} \\ \text{irred. continuous reps} \\ \rho_{l,\iota_l} : \mathrm{Gal}(\overline{F}/F) \to GL_m(\overline{\mathbb{Q}}_l) \\ \text{unram. at all but fin. many places,} \\ \text{de Rham at } l \end{array} \right\}$$

such that

$$\mathrm{WD}(\rho_{l,\iota_l}(\Pi)|_{\mathrm{Gal}(\overline{F}_v/F_v)})^{\mathrm{F-ss}} \simeq \iota_l^{-1} \mathscr{L}_{F_v}(\Pi_v) \tag{1.1}$$

at every finite place v of F.

The algebraicity at ∞ is reviewed in Definition 1.7 below. The "de Rham" property is a technical condition from l-adic Hodge theory. For our purpose it suffices to remark that it is the counterpart of the algebraicity condition for Π at ∞. The functor $\mathrm{WD}(\cdot)$ assigns a Weil-Deligne representation to an l-adic local Galois representation, and the superscript "F-ss" means the Frobenius semisimplification of a Weil-Deligne representation. (See [Tay04, §1] for definitions.) The notation $\mathscr{L}_{F_v}$ denotes a "geometric" normalization (e.g. [Shi11, §2.3]) of the local Langlands correspondence for $GL_m(F_v)$, which was established by Harris-Taylor ([HT01]) and Henniart ([Hen00]) about 10 years ago.

Remark 1.2. In order to uniquely determine the bijection, it suffices to require (1.1) at all but finitely many places v by the strong multiplicity one theorem (on the automorphic side) and the Cebotarev density theorem (on the Galois side).

It is customary to call each direction of the arrow in Conjecture 1.1 as

$\longrightarrow$ construction of Galois representations

$\longleftarrow$ modularity (or automorphy) of Galois representations.

When $m = 1$, Conjecture 1.1 is a consequence of class field theory. The case $m = 2$ with totally real F is discussed in Tilouine's lectures and will not be discussed in my lectures. (This case is separated because Hilbert modular varieties and Shimura curves are used when $m = 2$ while unitary Shimura varieties are used when $m > 2$.)

When $m \geq 3$, the best known case of the above conjecture is the following theorem due to various people. (See §1.5 below for major contributions.) The analogue over totally real fields can be deduced from this theorem (cf. [BLGHT11, Thm 1.1], [BLGGT14, Thm 2.1.1]).

Theorem 1.3. *Assume $m \geq 3$. If*

- *F is a CM field,*
- *Π is a cuspidal automorphic representation of $GL_m(\mathbb{A}_F)$,*
- *$\Pi^\vee \simeq \Pi \circ c$, and*
- *Π_∞ is regular and algebraic,*

then for each prime l and ι_l there exists a semisimple continuous representation

$$\rho_{l,\iota_l}(\Pi) : \mathrm{Gal}(\overline{F}/F) \to GL_m(\overline{\mathbb{Q}}_l),$$

which is unramified at all but finitely many places and de Rham at l, such that (1.1) *holds at every finite place v of F.*

Remark 1.4. It is clear that $\rho_{l,\iota_l}(\Pi)$ is unique up to isomorphism by Cebotarev and Brauer-Nesbitt theorems. We do not know whether $\rho_{l,\iota_l}(\Pi)$ is irreducible in general even though it is expected, unless Π is square-integrable at a finite prime ([TY07, Cor 1.3]). See [BLGGT14, Thm 5.5.2] and [PT15, Thm D] for some partial results. (The dictionary is that the cuspidality of Π should correspond to the irreducibility of $\rho_{l,\iota_l}(\Pi)$.)

Remark 1.5. Roughly speaking, the conditions on Π in the theorem mean that Π comes from a cohomological automorphic representation of a unitary group via quadratic base change.

Remark 1.6. The information about Π at infinite places is encoded by $\rho_{l,\iota_l}(\Pi)$ in the image of complex conjugation (if F has a real place) and the Hodge-Tate numbers at l-adic places. There is also a sign for ρ_{l,ι_l}. See [BLGGT14, Thm 2.1.1] for complete statements (also [CLH16] and [BC11] for complex conjugation and sign; [CLH16] builds on earlier results by Taylor and Taïbi).

My goal is to explain the ideas for the proof of Theorem 1.3.

1.4. **Conditions on Π_∞.** Let us recall some basic terminology regarding Π_∞. The skimming reader should feel free to skip this subsection.

Let F be any number field. Let $\Pi_\infty = \prod_{v|\infty} \Pi_v$ be a representation of $GL_m(F \otimes_{\mathbb{Q}} \mathbb{R}) = \prod_{v|\infty} GL_m(F_v)$. Let $\phi_v : W_{F_v} \to GL_m(\mathbb{C})$ be the L-parameter for Π_v, where $W_{\mathbb{C}} = \mathbb{C}^\times$ and $W_{\mathbb{R}}$ contains $W_{\mathbb{C}}$ as an index two subgroup. Thus we can write

$$\phi_v|_{W_{\mathbb{C}}} \simeq \chi_{v,1} \oplus \cdots \oplus \chi_{v,m}$$

for characters $\chi_{v,i} : \mathbb{C}^\times \to \mathbb{C}^\times$.

Definition 1.7. ([Clo90, Def 1.8]) We say Π_∞ is **algebraic** if there exist $a_{v,i}, b_{v,i} \in \mathbb{Z}$ for all $v|\infty$ and $1 \le i \le m$ such that for all $z \in \mathbb{C}$,

$$\chi_{v,i}(z) = z^{a_{v,i} + \frac{m-1}{2}} \bar{z}^{b_{v,i} + \frac{m-1}{2}}.$$

Remark 1.8. Buzzard and Gee [BG14] defines C-algebraic and L-algebraic representations. The former coincides with algebraic representations above. To define L-algebraic ones, one simply removes $\frac{m-1}{2}$ from the exponent in the above definition.

For simplicity, we restrict ourselves to a CM field F. (See Clozel's article for the general case.) Assume that Π_∞ is algebraic, and reorder indices so that $a_{v,1} \ge \cdots \ge a_{v,m}$ for each v.

Definition 1.9. ([Clo90, Def 3.12], [Shi11, §7.1]) An algebraic representation Π_∞ is said to be **regular algebraic** if $a_{v,1} > \cdots > a_{v,m}$ (or equivalently if Π_∞ has the same infinitesimal character as an algebraic representation Ξ of $(R_{F/\mathbb{Q}} GL_m) \times_{\mathbb{Q}} \mathbb{R}$). A regular algebraic Π_∞ is **slightly regular** if there exist $v|\infty$ and an odd i such that

$$a_{v,i} > a_{v,i+1} + 1.$$

(This may be an unfortunate name. The name indicates that the highest weight of Ξ is slightly regular.)

1.5. **Methods to construct Galois representations.** A natural abundant source of Galois representations is the l-adic cohomology of varieties over number fields. To make a connection with automorphic representations, it is the best to work with Shimura varieties, which come with canonical models over number fields as well as Hecke correspondences which are also defined over the same number fields. In any successful method (except the $m = 1$ case covered by class field theory), the basic starting point is to look for the desired Galois representation attached to a given automorphic representation in the l-adic cohomology of a well-chosen Shimura variety, possibly under some technical conditions. Then one tries to relax those conditions by various methods.

The fundamental idea goes back to 1970s and is due to Langlands, who laid out the program to study the zeta function and cohomology of Shimura varieties by describing the mod p points of Shimura varieties and then comparing the Grothendieck-Lefschetz fixed point formula with the Arthur-Selberg trace formula. Langlands understood the role of endoscopy in this context early on; when the group G is sandwiched between SL_2 and GL_2, he noticed in particular that the zeta function of a Shimura variety was factorized into the L-functions pertaining to not only G but a smaller "endoscopic" group, cf. [Lan79]. For the early history, the reader is referred to Langlands's papers in 1970s; see the last paragraph of [Lan77] for the author's guidance to some of them. (The interested reader may also read his commentary to [Lan77] in 1995 on his IAS website.) Kottwitz substantially contributed to flesh out Langlands's ideas and carried out the program for many PEL-type Shimura varieties in [Kot90, Kot92b, Kot92a], to name a few. Essentially any construction of automorphic Galois representations ultimately relies on this method initiated by Langlands and furthered by Kottwitz, which often goes by the "Langlands–Kottwitz method".

To discuss variants of this method and fine technical points, we list some key considerations in the construction of Galois representations.

(1) good reduction of Shimura varieties
(1′) bad reduction of Shimura varieties (involving nearby cycles, relation with a moduli space of p-divisible groups, etc)
(2) counting points on the special fibers of Shimura varieties at good primes
(2′) counting points on Igusa varieties at primes of bad reduction
(3) Arthur-Selberg trace formula (including twisted trace formula) when endoscopy is trivial
(3′) stable trace formula and nontrivial endoscopy
(4) compactification, boundary contributions in the counting point formula and Arthur-Selberg trace formula
(5) congruences; p-adic approximation

Let us remark on various approaches to Theorem 1.3 when $m \geq 3$.[1] As for (1) and (1)′, every work makes use of certain unitary PEL-type Shimura varieties. Naturally it is essential to consider the case of unitary (similitude) groups in the trace formula method alluded to above.

(i) Collaboration of many people ([LR92]): $m = 3$. (1), (2), (3), (4).
(ii) Clozel ([Clo91]), Kottwitz ([Kot92a], [Kot92b]): (1), (2), (3).

[1]We apologize for suppressing the rich history when $m = 2$ for Galois representations associated to (Hilbert) modular forms. Also when $m = 4$, there has been much earlier work coming from Shimura varieties for GSp_4.

(iii) Harris-Taylor ([HT01]), Taylor-Yoshida ([TY07]): $(1)'$, $(2)'$, (3).
(iv) Morel ([Mor10]): (1), (2), $(3)'$, (4).
(v) Clozel-Harris-Labesse ([CHL11]): (1), (2), $(3)'$.
(vi) Shin ([Shi11]): $(1)'$, $(2)'$, $(3)'$.
(vii) Chenevier-Harris ([CH13]): (5).

Note that only (iii) and (vi) prove (1.1) at ramified places. There $(1)'$ and $(2)'$ are crucial. In (ii) and (iii), Theorem 1.3 was proved under the additional assumption that Π_w is square-integrable at some place w. Later (iv), (v) and (vi) established the theorem without the square-integrability assumption but under a mild condition on Π_∞ when m is even (see Hypothesis 4.9). The last condition was removed by (vii) at the cost of proving (1.1) only up to semisimplification (thus losing the precise information on the monodromy operator). The full version of (1.1) was established by Caraiani [Car12, Car14].

It is worth noting that only special cases of the fundamental lemma were available when (i), (ii) and (iii) were carried out. The improvements (iv)-(vii) have become possible largely thanks to the proof of the fundamental lemma by Laumon, Ngô, Waldspurger and others.

Finally we mention that there is another approach to the bad reduction of Shimura varieties and the counting point formula, shedding some new light on the bad reduction of Shimura varieties and its interaction with representation theory. It may not be unreasonable to say that the flavor of this approach is somewhere between (1)+(2) and $(1)'+(2)'$, if compared with the cited work above. See nice surveys by Rapoport [Rap05] and Haines [Hai05] as well as some research papers [HR12, Sch13a, Sch13b, SS13]. In particular the last two papers simplified the proof (by Harris-Taylor and Henniart) of the local Langlands correspondence for general linear groups and the proof of Theorem 1.3. See [Schb] for a survey on some of these ideas.

From the next section I will mainly follow the approach of (vi), which improves on the method of (iii) especially in the aspect of the counting point formula, its comparison with the Arthur-Selberg trace formula via the stable trace formula. The reader is strongly encouraged to refer to other references as well.

1.6. **The Ramanujan-Petersson conjecture.** Here is a vast generalization of the conjecture (proved by Deligne) that Ramanujan's τ-function satisfies the bound $|\tau(p)| \le 2p^{11/2}$ for all primes p. (This conjecture is a special case of the conjecture below when $m = 2$, $F = \mathbb{Q}$ and Π corresponds to the cuspform Δ of weight 12 and level 1.)

Conjecture 1.10. *(Ramanujan-Petersson) Let $m \ge 1$, F be a number field and Π be a cuspidal automorphic representation of $GL_m(\mathbb{A}_F)$. Then Π_v is tempered at every finite place v of F.*

Remark 1.11. In fact, it would be reasonable to expect something stronger when Π is algebraic. For instance, one can conjecture that if Π is algebraic then for every $\sigma \in \mathrm{Aut}(\mathbb{C})$ and every finite place v, the σ-twist Π_v^σ is tempered. In other words, Π_v should be "absolutely tempered".

By combining Theorem 1.3 and some facts in the representation theory of GL_m over p-adic fields, we obtain

Theorem 1.12. *Conjecture 1.10 is true for F and Π as in Theorem 1.3.*

To be precise, the theorem is a corollary of Theorem 1.3 (with a semisimplified version of (1.1)) when m is even and Π_∞ is slightly regular (or when Π is square-integrable at a finite prime). Several mathematicians have made contributions to this. Namely, whenever the authors in (i)-(vi) of §1.5 proved the local-global compatibility at v for constructed Galois representations, they obtained the corresponding case of Theorem 1.12. Finally in the missing case where m is even and Π_∞ is not slightly regular, Caraiani [Car12] proved Theorem 1.12 in the course of strengthening (1.1), improving on Clozel's result [Clo13] in the unramified case by a different method.

2. Base change for unitary groups

Throughout the article we assume $n \geq 3$. In this case there are no Shimura varieties attached to GL_n (or its inner forms). The next best thing is to study Shimura varieties associated to unitary (similitude) groups since a unitary group becomes isomorphic to GL_n after quadratic extension of the base field. To analyze the cohomology of those Shimura varieties, it is essential to understand automorphic representations of unitary groups, especially in connection with those of GL_n because the representations of GL_n are better understood and also because we would like to prove a theorem about automorphic representations of GL_n rather than a unitary group. This can be achieved by (automorphic) base change for unitary groups, which will be discussed in this section. For the case of unitary similitude groups, see §3.4.

2.1. **Unramified local Langlands correspondence.** This subsection is a general background needed for §2 and §3. Details may be found in [Bor79] and [Min11] for instance. Let G be an unramified connected reductive group over a p-adic field F. Recall that G is said to be unramified over F if G has a smooth integral model over $\mathcal{O}_F$. We will fix such a model, thus also a compact subgroup $G(\mathcal{O}_F) \subset G(F)$, often called "hyperspecial". Denote by $\mathscr{H}^{\mathrm{ur}}(G(F)) := C_c^\infty(G(\mathcal{O}_F)\backslash G(F)/G(\mathcal{O}_F))$ the unramified Hecke algebra equipped with the convolution product. Then the following sets are in canonical bijection with each other (e.g. [Min11, 2.6]). This is a consequence of the Satake isomorphism.

(1) $\mathbb{C}$-algebra morphisms $\chi : \mathscr{H}^{\mathrm{ur}}(G(F)) \to \mathbb{C}$.
(2) isomorphism classes of unramified L-parameters $\varphi : W_F \to {}^L G$.
(3) isomorphism classes of unramified (irreducible admissible) representations of $G(F)$.

By definition φ is unramified if $\varphi(I_F) = (1)$, and $\pi \in \mathrm{Irr}(G(F))$ is unramified if π has a nonzero fixed vector under $G(\mathcal{O}_F)$.

One is invited to define canonical maps between them and prove that they are bijections in case G is a torus. The proof in the general case can be reduced, with some work, to the case of tori by considering a maximal torus of G.

It is an extremely important fact that the bijections among (1), (2) and (3) are *functorial* in G. This means the following: If H, G are connected reductive groups over F and $\widetilde{\eta} : {}^L H \to {}^L G$ is an L-morphism, then there is a canonical map (e.g. [Min11, 2.7])

$$\widetilde{\eta}^* : \mathscr{H}^{\mathrm{ur}}(G(F)) \to \mathscr{H}^{\mathrm{ur}}(H(F))$$

such that $\chi_H : \mathscr{H}^{\mathrm{ur}}(H(F)) \to \mathbb{C}$ and $\varphi_H : W_F \to {}^L H$ correspond if and only if $\chi_H \circ \widetilde{\eta}^*$ and $\widetilde{\eta} \circ \varphi_H$ correspond. The map

$$\widetilde{\eta}_* : \mathrm{Irr}^{\mathrm{ur}}(H(F)) \to \mathrm{Irr}^{\mathrm{ur}}(G(F)) \tag{2.1}$$

corresponding to $\varphi_H \mapsto \widetilde{\eta} \circ \varphi_H$ is basically the "transfer of unramified representations" with respect to $\widetilde{\eta}$. A quick characterization of the map $\widetilde{\eta}_*$ is possible using the following identity: for every $f \in \mathscr{H}^{\mathrm{ur}}(G(F))$,

$$\mathrm{tr}\, \widetilde{\eta}_* \pi_H(f) = \mathrm{tr}\, \pi_H(\widetilde{\eta}^* f). \tag{2.2}$$

2.2. **Setup.** Now we restrict our attention to unitary groups and general linear groups. The following notation will be used.

- F is a CM field with complex conjugation c. Set $F^+ := F^{c=1}$.
- $\vec{n} = (n_1, \ldots, n_r)$, $n_i, r \in \mathbb{Z}_{>0}$, $\sum_{i=1}^r n_i = n$.
- $i_{\vec{n}} : GL_{\vec{n}} \hookrightarrow GL_N$ $(N = \sum_i n_i)$ is the embedding
$$(A_1, \ldots, A_r) \mapsto \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & A_r \end{pmatrix}.$$
- $GL_{\vec{n}} = \prod_{n=1}^r GL_{n_i}$ (over any base).
- U_n is a quasi-split unitary group in n variables over F^+ (for an n-dimensional F-vector space with a Hermitian pairing). Define
$$U_{\vec{n}} := \prod_{1 \le i \le r} U_{n_i}.$$

We will work under the following hypothesis. (This is not a vacuous condition. For instance, $F = \mathbb{Q}(\zeta_5)$ does not satisfy it.)

Hypothesis 2.1. $F = EF^+$ for an imaginary quadratic field $E \subset F$.

Define
$$G^1_{\vec{n}} := R_{F^+/\mathbb{Q}} U_{\vec{n}}, \qquad \widetilde{G}^1_{\vec{n}} := R_{E/\mathbb{Q}}(G^1_{\vec{n}} \times_{\mathbb{Q}} E)$$
where $R_{(\cdot)}$ denotes the Weil restriction of scalars. When $\vec{n} = (n)$, we prefer to write G^1_n and $\mathbb{G}^1_n$ for $G^1_{\vec{n}}$ and $\mathbb{G}^1_{\vec{n}}$, respectively. (These ugly notations anticipate their similitude cousins, which will show up in §3.4. The superscript means that the similitude factor equals 1.) The connected reductive groups $G^1_{\vec{n}}$ and $\mathbb{G}^1_{\vec{n}}$ are defined over $\mathbb{Q}$ and $G^1_{\vec{n}}(\mathbb{Q}) = U_{\vec{n}}(F^+)$, $\mathbb{G}^1_{\vec{n}}(\mathbb{Q}) = G^1_{\vec{n}}(E) = U_{\vec{n}}(F) \simeq GL_{\vec{n}}(F)$. Also note that

$$G^1_{\vec{n}}(\mathbb{A}) = U_{\vec{n}}(\mathbb{A}_{F^+}), \qquad \mathbb{G}^1_{\vec{n}}(\mathbb{A}) = G^1_{\vec{n}}(\mathbb{A}_E) = U_{\vec{n}}(\mathbb{A}_F) \simeq GL_{\vec{n}}(\mathbb{A}_F). \tag{2.3}$$

Let θ denote the action on $\mathbb{G}^1_{\vec{n}}$ induced by $1 \times c$ on $G^1_{\vec{n}} \times_{\mathbb{Q}} E$. Under the above isomorphism, θ is transported to the action $(g_1, \ldots, g_r) \mapsto ({}^t g_1^{-c}, \ldots, {}^t g_r^{-c})$ on $GL_{\vec{n}}(\mathbb{A}_F)$ up to conjugation by $GL_{\vec{n}}(\mathbb{A}_F)$.

Consider the L-groups ${}^L G^1_{\vec{n}} := \widehat{G}_{\vec{n}} \rtimes W_{\mathbb{Q}}$ and ${}^L \mathbb{G}^1_{\vec{n}} := \widehat{\mathbb{G}}^1_{\vec{n}} \rtimes W_{\mathbb{Q}}$. The dual groups may be identified as

$$\widehat{G}^1_{\vec{n}} = GL_{\vec{n}}(\mathbb{C})^{\mathrm{Hom}(F^+,\mathbb{C})}, \qquad \widehat{\mathbb{G}}^1_{\vec{n}} = GL_{\vec{n}}(\mathbb{C})^{\mathrm{Hom}(F,\mathbb{C})}, \tag{2.4}$$

equipped with $W_{\mathbb{Q}}$-actions. There is a natural L-morphism

$$\mathrm{BC}_{\vec{n}} : {}^L G^1_{\vec{n}} \to {}^L \mathbb{G}^1_{\vec{n}}$$

which extends the diagonal embedding on the dual group. On the level of dual groups, $\mathrm{BC}_{\vec{n}}$ maps $(g_\sigma) \mapsto (h_\sigma)$ so that $h_\sigma = h_{\sigma^c} = g_\sigma$ for every $\sigma \in \mathrm{Hom}(F^+, \mathbb{C})$.

2.3. **Local base change.** Fix a finite place v of $\mathbb{Q}$. Consider

$$\pi_v \in \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v)) = \mathrm{Irr}(\prod_{w|v} U_{\vec{n}}(F_w)).$$

(i) For any v, when π_v is unramified, let $\phi(\pi_v) : W_{\mathbb{Q}_v} \to {}^L G^1_{\vec{n}}$ be the corresponding unramified parameter. Define $\mathrm{BC}_{\vec{n},*}(\pi_v) \in \mathrm{Irr}^{\mathrm{ur}}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v))$ corresponding to the unramified parameter $\mathrm{BC}_{\vec{n}} \circ \phi(\pi_v)$. Thus obtain a map (whose image sits inside the set of θ-stable representations).

$$\mathrm{BC}_{\vec{n},*} : \mathrm{Irr}^{\mathrm{ur}}(G^1_{\vec{n}}(\mathbb{Q}_v)) \to \mathrm{Irr}^{\mathrm{ur}}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v)). \tag{2.5}$$

(ii) If v *splits in* E, say u and u^c are primes of E above v. There is an isomorphism $\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v) = G^1_{\vec{n}}(E_v) \simeq G^1_{\vec{n}}(E_u) \times G^1_{\vec{n}}(E_{u^c})$ where θ acts as $(g_1, g_2) \mapsto (g_2^c, g_1^c)$. (Note that $c : E \overset{\sim}{\to} E$ induces $E_u \overset{\sim}{\to} E_{u^c}$.) Define

$$\mathrm{BC}_{\vec{n},*} : \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v)) \to \mathrm{Irr}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v)) \tag{2.6}$$

by $\pi \mapsto \pi \otimes \pi$. The image is clearly θ-stable. It is an exercise to check that the map (2.6) restricts to the map (2.5) defined above if π_v is unramified.

(iii) When $v = \infty$, the base change of any L-packet of $G^1_{\vec{n}}(\mathbb{R})$ can be defined as a representation of $\mathbb{G}^1_{\vec{n}}(R) = G^1_{\vec{n}}(\mathbb{C})$, but the details will be omitted. The fact that this is possible should not be surprising as the local Langlands correspondence is known for any real or complex group due to Langlands ([Lan89]). We will be interested in only those L-packets consisting of discrete series representations of $G^1_{\vec{n}}(\mathbb{R})$ and their base change.

Although the "explicit" base change in the above list does not exhaust all cases, it suffices for constructing Galois representations. (To our knowledge this observation was first made by Harris.) This is desirable since the base change for unitary groups is not established in full generality yet.

We will often write BC or $\mathrm{BC}_{\vec{n}}$ for the map $\mathrm{BC}_{\vec{n},*}$ in (2.5) or (2.6) if there is no danger of confusion.

In general the base change along a finite cyclic extension is characterized by a trace identity with respect to the transfer of test functions. The existence of transfer for any endoscopy (twisted or untwisted) is a consequence of the recent proof of the fundamental lemma, thanks to Laumon, Ngô, Waldspurger and others. In particular the transfer of test functions is known for any cyclic base change, which is an instance of twisted endoscopy. (However see Remark 2.4 below.)

Proposition 2.2. (1) *In case (i) and (ii) above, let* $f_v \in C_c^\infty(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v))$*. Then there exists* $\phi_v \in C_c^\infty(G^1_{\vec{n}}(\mathbb{Q}_v))$ *with matching orbital integrals. In case (ii), if* $f_v \in \mathscr{H}^{\mathrm{ur}}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v))$ *then one can take* $\phi_v = \widetilde{BC}_{\vec{n}}^{\,*}(f_v)$ *in the notation of §2.1.*

(2) *The function* ϕ_v *in part (1) satisfies the following in case (ii).*

$$\widetilde{\mathrm{tr}}\, BC(\pi_v)(f_v) = \mathrm{tr}\, \pi_v(\phi_v), \qquad \forall \pi_v \in \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v))$$

In case (i), the same identity holds for all unramified $\pi_v \in \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v))$.

Sketch of proof. Part (1) is a consequence of the fundamental lemma as explained above. As for Part (2), (2.2) is the desired identity in case (i). In case (ii) it is shown by a direct computation (which is not hard). □

Remark 2.3. We apologize for the imprecision in the statement of (1). We chose not to write out the precise identity for matching of orbital integrals. Moreover the experienced reader must have noticed an abuse of language. Orbital integrals there really mean stable (twisted) orbital integrals.

Remark 2.4. The second assertion of (1) is usually refereed to as the base change fundamental lemma and due to Kottwitz, Clozel and Labesse and was proved almost 20 years ago.

Remark 2.5. You may have noticed that no properties of unitary groups are used in §2.3. Indeed, local base change can be defined in the same manner as above when

- the quadratic extension $E/\mathbb{Q}$ is replaced with any finite cyclic extension E'/E'',
- $G^1_{\vec{n}}$ is replaced with any connected reductive group H over E'',
- $\mathbb{G}^1_{\vec{n}}$ is replaced with $\mathbb{H} := R_{E'/E''}(H \times_{E''} E')$ and
- v runs over places of E''.

What is nice about the quadratic base change from $G^1_{\vec{n}}$ to $\mathbb{G}^1_{\vec{n}}$ is that the representation theory of $\mathbb{G}^1_{\vec{n}}$ (which is a general linear group) is much more complete than that of other groups. As such, this base change enables us to study representations of $G^1_{\vec{n}}$ through those of $\mathbb{G}^1_{\vec{n}}$. (For instance, one may try to define a local or global L-packet for $G^1_{\vec{n}}$ as the set of those representations of $G^1_{\vec{n}}$ whose base change images are isomorphic.)

2.4. **Weak base change (global).**

Definition 2.6. Let $\pi \in \mathrm{Irr}(G^1_n(\mathbb{A}))$ and Π be an automorphic representation of $\mathbb{G}^1_n(\mathbb{A})$. We say that Π is a *weak base change* of π and write $\Pi = \mathrm{WBC}(\pi)$ if $\mathrm{BC}(\pi_v) \simeq \Pi_v$ for all but finitely many v. (This makes sense as π_v is unramified and so $\mathrm{BC}(\pi_v)$ is defined for almost all v.)

By strong multiplicity one theorem for GL_n, the weak base change $\mathrm{WBC}(\pi)$ is unique (up to isomorphism) if it exists (cf. (2.3)). There is a fairly general existence result by Clozel and Labesse.

Proposition 2.7. *(*[Lab11, Cor 5.3]*) If π_∞ is a discrete series whose infinitesimal character is sufficiently regular, then* $\mathrm{WBC}(\pi)$ *exists. In other words, the conclusion is that there exists an automorphic representation Π of $\mathbb{G}^1_n(\mathbb{A})$ such that* $\mathrm{BC}(\pi_v) \simeq \Pi_v$ *for almost all v.*

Remark 2.8. Actually, in the construction of Galois representations, it is not necessary to use this proposition until the last stage where the p-adic approximation argument ([CH13]) is used. The reverse of (weak) base change, or "descent", often turns out to be more essential. Labesse ([Lab11]) also proves many instances of descent.

We will be mostly interested in the case when

$$\mathrm{WBC}(\pi) = \Pi \simeq \mathrm{Ind}(\Pi_1 \otimes \cdots \otimes \Pi_r) \tag{2.7}$$

such that each Π_i is *cuspidal* and $\Pi_i^\theta \simeq \Pi_i$. (In general, it can happen that some Π_i is discrete but not cuspidal and also that for $i \neq j$, $\Pi_i^\theta \simeq \Pi_j$ and $\Pi_j^\theta \simeq \Pi_i$.) It is built into assumption that the parabolic induction of (2.7) is irreducible. The Π as above can be thought of as coming from the discrete tempered spectrum for $G^1(\mathbb{A}) = U(\mathbb{A}_{F^+})$. For our application, the case $r \leq 2$ is the most important.

Definition 2.9. Suppose that $\mathrm{WBC}(\pi)$ exists and has the form as in (2.7). We say π is *stable* if $r = 1$ and *endoscopic* if $r > 1$.

3. Endoscopy for unitary groups

3.1. **Endoscopic groups.** Define $\mathscr{E}(G^1_n)$ (resp. $\mathscr{E}^{\mathrm{ell}}(G^1_n)$) to be the set of $G^1_{\vec{n}}$ such that $\sum_{i=1}^r n_i = n$ (resp. $\sum_{i=1}^r n_i = n$ and $r \le 2$). In other words,

$$\mathscr{E}^{\mathrm{ell}}(G^1) = \{G^1_n\} \cup \{G^1_{n_1,n_2} \,|\, n_1 + n_2 = n,\ n_1, n_2 > 0\}.$$

The elements of $\mathscr{E}^{\mathrm{ell}}(G^1)$ are called the elliptic endoscopic groups for G^1_n and play a fundamental role in the trace formula for G^1_n. For each $G^1_{\vec{n}} \in \mathscr{E}^{\mathrm{ell}}(G^1)$, fix an L-embedding

$$\widetilde{\eta}^1_{\vec{n}} : {}^L G^1_{\vec{n}} \to {}^L G^1_n \tag{3.1}$$

extending $\widehat{G}^1_{\vec{n}} \to \widehat{G}^1_n$ given by the block diagonal embedding $i_{\vec{n}}$. The choice of $\widetilde{\eta}^1_{\vec{n}}$ is not unique and there is no canonical choice in general. An explicit form of $\widetilde{\eta}^1_{\vec{n}}$ will be omitted but the interested reader may find it as an exercise or look up [Rog92, §1.2], cf. [Shi11, §3.2].

3.2. **Endoscopic transfer.** Let $G^1_{\vec{n}} \in \mathscr{E}(G^1_n)$. (In fact, $G^1_{\vec{n}} \in \mathscr{E}^{\mathrm{ell}}(G^1_n)$ suffices for our later discussion.) There should be a transfer of representations from $G^1_{\vec{n}}$ to G^1_n corresponding to $\widetilde{\eta}^1_{\vec{n}}$. Such a transfer is called an *endoscopic transfer* and an instance of the Langlands functoriality. Although the Langlands functoriality is still largely open, which should involve the description of L-packets among other things, there are some favorable cases (cf. §2.3), as we will see below, where the local transfer of representations can be explicitly worked out.

(i) For any finite v,

$$\widetilde{\eta}^1_{\vec{n},*} : \mathrm{Irr}^{\mathrm{ur}}(G^1_{\vec{n}}(\mathbb{Q}_v)) \to \mathrm{Irr}^{\mathrm{ur}}(G^1_n(\mathbb{Q}_v))$$

is defined as in (2.1).

(ii) If v splits in E,

$$\widetilde{\eta}^1_{\vec{n},*} : \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v)) \to \mathrm{Irr}(G^1_n(\mathbb{Q}_v))$$

is defined as a character twist of the parabolic induction, noting that $G^1_{\vec{n}}(\mathbb{Q}_v)$ and $G^1_n(\mathbb{Q}_v)$ are isomorphic to $\prod_w GL_{\vec{n}}(F_w)$ and $\prod_w GL_n(F_w)$, respectively, where w runs over the half of the places dividing v. The character twist depends on $\widetilde{\eta}^1_{\vec{n},*}$ at v, which is not the identity map on $W_{\mathbb{Q}_v}$ (inside ${}^L G^1_{\vec{n}}$) in general.

(iii) If $v = \infty$, the transfer of representations is defined by the Langlands correspondence for $G^1_{\vec{n}}(\mathbb{R})$ and $G^1_n(\mathbb{R})$. (We are especially interested in the discrete series representations.)

We have the analogue of Proposition 2.2 characterizing the endoscopic transfer described above with respect to the transfer of test functions. As explained in the paragraph above Proposition 2.2, the existence of transfer is known.

Proposition 3.1. (1) *In case (i) and (ii) above, let $f_v \in C^\infty_c(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v))$. Then there exists $\phi_v \in C^\infty_c(G^1_{\vec{n}}(\mathbb{Q}_v))$ with matching orbital integrals. Moreover in case (ii), if $f_v \in \mathscr{H}^{\mathrm{ur}}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v))$ then one can take $\phi_v = (\widetilde{\eta}^1_{\vec{n}})^*(f_v)$ in the notation of §2.1.*

(2) *The function ϕ_v in part (1) satisfies the following in case (ii).*

$$\widetilde{\mathrm{tr}}\,\left(\widetilde{\eta}^1_{\vec{n},*}(\pi_{H,v})\right)(f_v) = \mathrm{tr}\,\pi_{H,v}(\phi_v), \qquad \forall \pi_{H,v} \in \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v))$$

In case (i), the same identity holds for all unramified $\pi_{H,v} \in \mathrm{Irr}(G^1_{\vec{n}}(\mathbb{Q}_v))$.

Sketch of proof. The same remarks in the proof of Proposition 2.2 apply. Remark 2.3 is valid here as well. □

Remark 3.2. The article [CHL11] would be an excellent place to learn details about the material of this subsection.

3.3. **Base change and endoscopic transfer.** Let $G^1_{\vec{n}} \in \mathscr{E}(G^1_n)$ as before. We would like to understand the interplay between the local base change and the local endoscopic transfer. For this let us look at the following diagram of L-morphisms. The three maps $\widetilde{\eta}_{\vec{n}}$, $\mathrm{BC}_{\vec{n}}$ and BC_n are already defined. It is easy to choose $\widetilde{i}_{\vec{n}}$ explicitly so that the diagram commutes.

$$\begin{tikzcd} {}^L G^1_{\vec{n}} \arrow[r, "\widetilde{\eta}^1_{\vec{n}}"] \arrow[d, "\mathrm{BC}_{\vec{n}}"'] & {}^L G^1_n \arrow[d, "\mathrm{BC}_n"] \\ {}^L \mathbb{G}^1_{\vec{n}} \arrow[r, "\widetilde{i}_{\vec{n}}"'] & {}^L \mathbb{G}^1_n \end{tikzcd} \tag{3.2}$$

By §2.1, we immediately obtain a commutative diagram at each finite v.

$$\begin{tikzcd} \mathrm{Irr}^{\mathrm{ur}}(G^1_{\vec{n}}(\mathbb{Q}_v)) \arrow[r, "\widetilde{\eta}^1_{\vec{n},*}"] \arrow[d, "\mathrm{BC}_{\vec{n},*}"'] & \mathrm{Irr}^{\mathrm{ur}}(G^1_n(\mathbb{Q}_v)) \arrow[d, "\mathrm{BC}_{n,*}"] \\ \mathrm{Irr}^{\mathrm{ur}}(\mathbb{G}^1_{\vec{n}}(\mathbb{Q}_v)) \arrow[r, "\widetilde{i}_{\vec{n},*}"'] & \mathrm{Irr}^{\mathrm{ur}}(\mathbb{G}^1_n(\mathbb{Q}_v)) \end{tikzcd} \tag{3.3}$$

If v splits in E or $v = \infty$, there is a similar diagram with $\mathrm{Irr}(\cdot)$ in place of $\mathrm{Irr}^{\mathrm{ur}}(\cdot)$.

The functoriality with respect to $\widetilde{i}_{\vec{n}}$ is simply a parabolic induction up to a twist by a character, noting that $\mathbb{G}^1_{\vec{n}}$ and $\mathbb{G}^1_n$ are essentially $GL_{\vec{n}}$ and GL_n (2.3). (This statement can be made precise and proved.) This leads to the following interesting observation. Suppose that $\pi \in \mathscr{A}(G^1_n(\mathbb{A}))$ is the endoscopic transfer with respect to $\widetilde{\eta}^1_{\vec{n}}$ of $\pi_0 \in \mathscr{A}(G^1_{\vec{n}}(\mathbb{A}))$ at almost all places, where $\vec{n} \neq (n)$. Assume the sufficient regularity of Proposition 2.7 for π and π_0. Then $\mathrm{WBC}(\pi) \in \mathscr{A}(\mathbb{G}^1_n(\mathbb{A}))$ and $\mathrm{WBC}(\pi_0) \in \mathscr{A}(\mathbb{G}^1_{\vec{n}}(\mathbb{A}))$ exist. The commutativity of (3.3) implies that

$$\mathrm{WBC}(\pi) \simeq \mathrm{Ind}^{G^1_n}_{G^1_{\vec{n}}}(\mathrm{WBC}(\pi_0) \otimes \chi) \tag{3.4}$$

for a certain character χ (determined by $\widetilde{\eta}^1_{\vec{n}}$). We have shown that if π is the image of endoscopic transfer from an endoscopic group $G^1_{\vec{n}}$ different from G^1_n, then $\mathrm{WBC}(\pi)$ is endoscopic, namely it is induced from a representation of a proper parabolic subgroup (if it is also the case that $\mathrm{WBC}(\pi)$ is induced from a cuspidal representation). This justifies the terminology of Definition 2.9. The converse is expected to be true and proved in some cases. (The converse says that any endoscopic π arises from some $G^1_{\vec{n}} \neq G^1_n$ via endoscopic transfer.)

3.4. **Unitary similitude groups.** The results of §2 and §3 carry over with very minor changes to unitary similitude groups $G_{\vec{n}}$ (including the case $\vec{n} = (n)$) sitting inside the following exact sequence where $G_{\vec{n}} \to \mathbb{G}_m$ is the multiplier map.

$$1 \to G^1_{\vec{n}} \to G_{\vec{n}} \to \mathbb{G}_m \to 1$$

Define $\mathbb{G}_{\vec{n}} := R_{E/\mathbb{Q}}(G_{\vec{n}} \times_{\mathbb{Q}} E)$. Observe that

$$\mathbb{G}_{\vec{n}}(\mathbb{A}) = G(\mathbb{A}_E) \simeq GL_1(\mathbb{A}_E) \times GL_{\vec{n}}(\mathbb{A}_F).$$

The extra factor $GL_1(\mathbb{A}_E)$, which did not exist in $\mathbb{G}^1_{\vec{n}}(\mathbb{A})$, is a nuisance but does not increase technical difficulty.

It is the groups $G_{\vec{n}}$ and $\mathbb{G}_{\vec{n}}$ that constantly show up later on. Why do we care about $G_{\vec{n}}$ and $\mathbb{G}_{\vec{n}}$ when it looks simpler to work with $G^1_{\vec{n}}$ and $\mathbb{G}^1_{\vec{n}}$? The main reason is that $G_{\vec{n}}$ naturally occurs in the context of PEL-type Shimura varieties, whose cohomology is an essential input in the construction of Galois representations. Although it is possible to understand the representations of $G_{\vec{n}}$ through those of $G^1_{\vec{n}}$, it seems more satisfactory to deal with $G_{\vec{n}}$ directly. Nevertheless, in the first reading of the subject, it may be harmless to ignore the similitude part and pretend that you are working with unitary groups.

4. Shimura varieties

We keep the previous notation. In particular, F is a CM field, F^+ is the maximal totally real subfield of F, and $F = EF^+$ for some imaginary quadratic field $E \subset F$.

4.1. **Choice of unitary group.** From here on, we will fix an *odd* integer $n \in \mathbb{Z}_{\geq 3}$. Let G be an inner form of the quasi-split group G_n over $\mathbb{Q}$. Such a G is also a unitary similitude group and fits into an exact sequence

$$1 \to G^1 \to G \to \mathbb{G}_m \to 1$$

so that G^1 is an inner form of G^1_n. When n is odd, there is no obstruction in finding G such that

- G^1 is quasi-split at all finite places and
- $G^1(\mathbb{R}) \simeq U(1, n-1) \times U(0, n)^{[F^+:\mathbb{Q}]-1}$.

(In general, there is a cohomological obstruction for finding a unitary group with prescribed local conditions at all places. The obstruction always vanishes if n is odd, which is not the case if n is even. See [Clo91, §2] for a detailed computation of cohomological obstructions.) The main reason for choosing G^1 to be quasi-finite at all finite places is that we want to see in the cohomology of Shimura varieties as many Galois and automorphic representations as possible.[2] In case you wonder why we choose $G^1(\mathbb{R})$ as above, see Remark 4.4 and 5.2.

Remark 4.1. In work of Clozel, Kottwitz and Harris-Taylor, they had an assumption which implies that G^1 is *not* quasi-finite at some finite place (but they had the same $G^1(\mathbb{R})$). That assumption was imposed mainly because some techniques in the trace formula were available only in limited cases at that time (but there was another good reason for Harris-Taylor when they proved a counting point formula). This is why Π was assumed to be square-integrable at a finite place in their proof of Theorem 1.3. The whole point of recent work by several people is to remove the last assumption on Π, so it would be reasonable to appreciate the quasi-split condition on G^1 at finite places.

[2] A general principle is that a quasi-split group has more representations than its non quasi-split inner forms, either locally or globally. For an example, think of the Jacquet-Langlands correspondence between GL_n and its inner form coming from a division algebra.

4.2. **l-adic cohomology.** Let $\mathrm{Sh} = \{\mathrm{Sh}_U\}$ be (the projective system of) Shimura varieties associated to G. (To be precise we also have to choose an $\mathbb{R}$-morphism $R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \to G$ but in our case there is a natural choice for this morphism, which we suppress.) Each Sh_U is a smooth quasi-projective variety over F of dimension $n-1$ (if U is sufficiently small). From now on, we make another

Hypothesis 4.2. $F^+ \neq \mathbb{Q}$

so that $G^1(\mathbb{R})$ has at least one compact factor $U(0,n)$. Then it can be shown that Sh_U is projective. Postponing the moduli problem for Sh_U to §5.1, we would like to explain the rough structure of the l-adic étale cohomology of Sh.

Let ξ be an irreducible algebraic representation of G over $\overline{\mathbb{Q}}_l$. The ξ gives rise to a compatible system of smooth $\overline{\mathbb{Q}}_l$-sheaves $\mathscr{L}_\xi$ on Sh_U. (We have seen this in Tilouine's lectures. For a precise construction, refer to [HT01, III.2].) The $\overline{\mathbb{Q}}_l$-vector space

$$H^k(\mathrm{Sh}, \mathscr{L}_\xi) := \varinjlim_U H^k(\mathrm{Sh}_U \times_F \overline{F}, \mathscr{L}_\xi)$$

is equipped with a smooth action of $G(\mathbb{A}^\infty) \times \mathrm{Gal}(\overline{F}/F)$. (Smoothness means that each vector has an open stabilizer in $G(\mathbb{A}^\infty)$.) There is a direct sum decomposition

$$H^k(\mathrm{Sh}, \mathscr{L}_\xi) = \bigoplus_{\pi^\infty} \pi^\infty \otimes R_l^k(\pi^\infty)$$

where π^∞ runs over irreducible admissible representations of $G(\mathbb{A}^\infty)$ and $R_l^k(\pi^\infty)$ is a finite dimensional representation of $\mathrm{Gal}(\overline{F}/F)$. (Comparison with Matsushima's formula tells us that the above formula is indeed a direct sum.)

Often it is convenient to consider virtual representations (integral combinations of representations with possibly negative coefficients)

$$H(\mathrm{Sh}, \mathscr{L}_\xi) = \sum_{k\geq 0} (-1)^k H^k(\mathrm{Sh}, \mathscr{L}_\xi), \qquad R_l^k(\pi^\infty) = \sum_{k\geq 0} (-1)^k R_l^k(\pi^\infty).$$

Let $\Pi = \chi \otimes \Pi^1$ be an automorphic representation (not necessarily cuspidal) of $\mathbb{G}(\mathbb{A}) = G(\mathbb{A}_E) \simeq GL_1(\mathbb{A}_E) \times GL_n(\mathbb{A}_F)$. Define a finite sum

$$R_l(\Pi) := \sum_{\mathrm{WBC}(\pi^\infty) = \Pi^\infty} R_l(\pi^\infty).$$

In order to understand $H^k(\mathrm{Sh}, \mathscr{L}_\xi)$ (for instance if one is interested in the L-functions for Sh_U), it would be ideal to describe $R_l(\pi^\infty)$ for each π^∞, or even $R_l^k(\pi^\infty)$ for each $k \in \mathbb{Z}_{\geq 0}$. For the construction of Galois representations, it suffices to describe $R_l(\Pi)$.

Let $\Pi^1 = \boxplus_{i=1}^r \Pi_i^1$ such that each Π_i^1 is θ-stable (cf. (2.7)). We assume that $R_l(\Pi) \neq 0$. (Roughly speaking, this is the case if Π_∞ is the base change of the discrete series representation "determined by" ξ.) Then we expect that

(1) If Π^1 is cuspidal ($r = 1$) then $R_l(\Pi) = 0$ or $R_l(\Pi)$ should correspond to Π^1 (in the sense of Conjecture 1.1) up to a character twist.
(2) In general, there should exist some i such that $R_l(\Pi)$ corresponds to Π_i^1.
(3) Among $H^k(\mathrm{Sh}, \mathscr{L}_\xi)$, only H^{n-1} contributes to $R_l(\Pi)$. (This is not expected to be true if some θ-stable representation Π_i^1 is discrete but not cuspidal.) So $R_l(\Pi)$ should come with the sign $(-1)^{n-1}$.

Remark 4.3. The same assertion should hold for $R_l(\pi^\infty)$ in place of $R_l(\Pi)$ if $\mathrm{WBC}(\pi^\infty) = \Pi^\infty$. In particular, we should allow a nonzero integral multiplicity for $R_l(\Pi)$ (which is harmless) in general.

Remark 4.4. The above expectation is valid only under the condition on $G^1(\mathbb{R})$ as in §4.1. Let us briefly discuss what happens to (1) and (2) if that condition is given up. If $U(1, n-1)$ is replaced with $U(a, b)$ with $a + b = n$ then $R_l(\Pi)$ should be roughly the a-th exterior power of the Galois representation associated to Π^1 if Π^1 is cuspidal. (Depending on the convention, you get the b-th exterior power, which is dual to the a-th exterior power up to twist.) In the $U(a, b)$ case, if Π^1 is not cuspidal, the description of $R_l(\Pi)$ is more complicated. In fact there is a concrete recipe for predicting Galois representations in the cohomology of PEL-type Shimura varieties which are of unitary or symplectic type in terms of Arthur's parameters ([Kot90, §10]). The above expectation for $R_l(\Pi)$ should be viewed as a special case of the general principle.

4.3. **Finding a candidate Galois representation.** Let us go back to the setting of Theorem 1.3. Changing notation slightly from the theorem, let Π^1 denote a cuspidal automorphic representation of $GL_m(\mathbb{A}_F)$ which is θ-stable. A big question is where to look for $\rho_l(\Pi^1)$ corresponding to Π^1. One can learn the following recipe from some experience.

<u>m: odd</u>

Take $n := m$ and choose $\psi : \mathbb{A}_E^\times / E^\times \to \mathbb{C}^\times$ so that $\Pi := \psi \otimes \Pi^1$ is a θ-stable representation of $\mathbb{G}_n(\mathbb{A})$. Then $\rho_l(\Pi^1) := R_l(\Pi)$ should work for the main theorem (up to a nonzero multiplicity and a character twist).

<u>m: even</u>

Take $n := m + 1$ and set $\Pi_1 := \Pi^1$. Choose $\psi : \mathbb{A}_E^\times / E^\times \to \mathbb{C}^\times$ and $\Pi_2 : \mathbb{A}_F^\times / F^\times \to \mathbb{C}^\times$ so that Π defined as below is a θ-stable representation of $\mathbb{G}_n(\mathbb{A})$. (In particular, $\Pi_2^\vee = \Pi_2^{-1}$ should hold.)

$$\Pi_M := \psi \otimes \Pi_1 \otimes \Pi_2, \qquad \Pi := \mathrm{Ind}_{\mathbb{G}_{m,1}}^{\mathbb{G}_n}(\Pi_M) = \psi \otimes (\Pi_1 \boxplus \Pi_2) \tag{4.1}$$

Then $\rho_l(\Pi_1) := R_l(\Pi)$ should work (again up to a nonzero multiplicity and a character twist), as long as $R_l(\Pi)$ corresponds to not Π_2 but Π_1. (See the paragraph preceding Remark 4.3.) If Π_1 is slightly regular, then Π_2 (and ψ) can be chosen such that $R_l(\Pi)$ corresponds to Π_1. (The proof involves an explicit sign computation in endoscopy theory, especially for real L-parameters. As this is not very intuitive at first sight, you may regard this as a black box. In case you are curious, the relevant result is Lemma 7.3 of [Shi11].)

The problem is more tractable if we control the set of primes where ramification occurs:

Hypothesis 4.5. If v is a finite place of $\mathbb{Q}$ such that either v is ramified in F or Π_v is ramified, then v splits in E.

If v splits in E then $G(\mathbb{Q}_v)$ is isomorphic to a product of general linear groups (cf. (5.1)) in which case almost everything is better understood than other groups, so the ramification causes less difficulty. (If v is not split in E then $G(\mathbb{Q}_v)$ is a v-adic unitary similitude group.)

4.4. **What should we do?** To prove Theorem 1.3, we must show that $\rho_l(\Pi^1)$ defined as in §4.3 does correspond to Π^1 at each finite place (excluding those dividing l) via the local Langlands correspondence. This amounts to analyzing $R_l(\Pi)$ in the two cases depending on the parity of m. Indeed, the main theorem will essentially follow from

Theorem 4.6. *Keep the notation of §4.3. At each finite place w not dividing l,*

$$R_l(\Pi)|_{W_{F_w}} \sim \begin{cases} \mathscr{L}_{F_w}(\Pi^1_w), & \text{if } m \text{ is odd,} \\ \mathscr{L}_{F_w}(\Pi_{1,w}) \text{ or } \mathscr{L}_{F_w}(\Pi_{2,w}), & \text{if } m \text{ is even} \end{cases}$$

where $\sim$ means an isomorphism up to a nonzero multiplicity, a character twist and semisimplification.

In fact we prove the theorem under one more hypothesis:

Hypothesis 4.7. $p := w|_{\mathbb{Q}}$ splits in E.

That is to say, we compute $R_l(\Pi)$ only at those places where G is a product of general linear groups. This may seem like a defect, but it turns out that once the theorem is shown under Hypothesis 4.7, the latter can be removed without much difficulty (cf. §4.5).

Remark 4.8. The ambiguity in $\sim$ raises nontrivial issues, but they are resolved after all. The possibility that the multiplicity could be greater than one is handled by Taylor's trick ([HT01, Prop VII.1.8]). The character twist is not a problem as one can always twist back. As for the last issue, Taylor and Yoshida ([TY07]) removed semisimplification by proving the purity of $WD(R_l(\Pi)|_{W_{F_w}})$ by studying Shimura varieties with Iwahori level structure. See their article for further detail.

If m is odd, it is not hard to see that Theorem 4.6 implies Theorem 1.3 (cf. Remark 4.8 above). In case m is even, if

$$R_l(\Pi)|_{W_{F_w}} \sim \mathscr{L}_{F_w}(\Pi_{1,w}) \tag{4.2}$$

(resp. $R_l(\Pi)|_{W_{F_w}} \sim \mathscr{L}_{F_w}(\Pi_{2,w})$) at one w then it is so at every other $w \nmid l$. We would be happy if (4.2) is the case. Unfortunately we are unable to tell whether $\mathscr{L}_{F_w}(\Pi_{1,w})$ or $\mathscr{L}_{F_w}(\Pi_{2,w})$ occurs. (This problem is linked with the computation of the sign e_2 in (6.8).) Nevertheless, if $\mathscr{L}_{F_w}(\Pi_{2,w})$ occurs in the formula for a given $\Pi = \chi \otimes (\Pi_1 \boxplus \Pi_2)$, we can show the following key fact ([Shi11, Lem 7.3]) by some technical sign computation in endoscopy: if we suppose

Hypothesis 4.9. Π_1 is assumed to be slightly regular when m is even

then there exists Π'_2 such that if we set $\Pi' := \chi \otimes (\Pi_1 \boxplus \Pi'_2)$, then $R_l(\Pi')|_{W_{F_w}} \sim \mathscr{L}_{F_w}(\Pi_{1,w})$ for all $w \nmid l$. This suffices for the purpose of deducing Theorem 1.3 under the running hypotheses, as $R_l(\Pi')$ is essentially the desired Galois representation of that theorem (cf. Remark 4.8).

In summary, if Theorem 4.6 is known to be valid under Hypotheses 2.1, 4.2, 4.5 and 4.7, then Theorem 1.3 can be proved under Hypotheses 2.1, 4.2, 4.5, 4.7 and 4.9. In §4.5 below, we briefly explain how to obtain Theorem 1.3 by removing all the hypotheses. Starting from §5, our focus will be how to tackle Theorem 4.6 under Hypotheses 2.1, 4.2, 4.5 and 4.7.

4.5. **Removal of hypotheses.** Suppose that Theorem 1.3 is shown under Hypotheses 2.1, 4.2, 4.5, 4.7 and 4.9. In other words, we assume that the desired Galois representations are constructed under these hypotheses. Using them as initial seeds, we can remove all the hypotheses except Hypothesis 4.9 by various tricks. Among key ingredients are Arthur-Clozel's base change for general linear groups ([AC89]) and the so-called patching lemma (due to Blasius-Ramakrishnan [BR89] and generalized by [Sor]) among others. Rather than delving into detail, we refer the reader to [CH13, 3.1] or the proof of Theorem VII.1.9 in [HT01] for this type of argument. (The corresponding part in [Shi11] appears in the proof of Proposition 7.4 and Theorem 7.5.)

Finally Hypothesis 4.9 needs to go away. Chenevier and Harris constructed Galois representations without Hypothesis 4.9 (but assuming the other hypotheses and then removing them again) by p-adic congruences. Namely they made use of p-adic families of Galois representations on eigenvarieties for definite unitary groups. They derived various expected properties for $\rho_{l,\iota_l}(\Pi)$, including l-adic Hodge theoretic properties, but established a weaker form of (1.1). Caraiani [Car12, **?**] obtained (1.1) in general (without Hypothesis 4.9). A main point in her work is to show that the Galois representations associated to Π in the cohomology of $U(1,n-1)\times U(1,n-1)$-Shimura varieties are pure in a spirit similar to [TY07].

5. LOCAL GEOMETRY OF SHIMURA VARIETIES

As we commented at the end of §4.4, our aim in the rest of the article is to sketch the idea of proof of Theorem 1.3 under Hypotheses 2.1, 4.2, 4.5, 4.7 and 4.9 assuming $v\nmid l$.

In §5.1 we briefly discuss integral models for Shimura varieties. We omit the detail on the integral models with bad reduction (defined in terms of Drinfeld level structure at p) but note that these play a crucial role in establishing the first basic identity in §5.5. It is worth noting at the outset that the contents of §5.2-§5.5 are not needed in the analysis of good reduction (namely the methods (1) and (2) of §1.5). Indeed, in the case of good reduction modulo p, Kottwitz ([Kot92b]) derived a nice counting point formula for the whole special fiber of Shimura varieties (generalizing earlier work of Ihara and Langlands) rather than for an individual Newton stratum. Kottwitz's formula describes the Hecke action outside a prime p and the Frobenius action at p on the cohomology of Shimura varieties. However, the strategy should be modified quite a bit in the case of bad reduction and our aim is to introduce some of the new tools, which are largely due to Harris and Taylor in the setting of unitary Shimura varieties. In the case of GL_2 the tools were developed earlier by Deligne and Carayol.

5.1. **Moduli definition of integral models.** Keep the notation and hypotheses from §4.4. It is convenient to define a place $u:=w|_E$ of E (as a restriction of w to E).

Let $U=U^p\times U_p$ be an open compact subgroup of $G(\mathbb{A}^\infty)$. We will assume that U_p is a maximal compact subgroup of $G(\mathbb{Q}_p)$. (It may not be hyperspecial as p may ramify in F.) Consider the following moduli problem which associates to a connected locally noetherian scheme S the set of equivalence classes $\{(A,\lambda,i,\bar{\eta})\}/\sim$ where

- A is an abelian scheme over S.
- $\lambda: A\to A^\vee$ is a prime-to-p polarization.
- $i:\mathcal{O}_F\hookrightarrow \mathrm{End}(A)\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}$ such that $\lambda\circ i(f)=i(f^c)^\vee\circ\lambda$, $\forall f\in\mathcal{O}_F$.

- $\bar{\eta}^p$ is a $\pi_1(S,s)$-invariant U^p-orbit of isomorphisms of $F \otimes_{\mathbb{Q}} \mathbb{A}^{\infty,p}$-modules $\eta^p : V \otimes_{\mathbb{Q}} \mathbb{A}^{\infty,p} \stackrel{\sim}{\to} V^p A_s$ which take the pairing $\langle \cdot, \cdot \rangle$ to the λ-Weil pairing up to $(\mathbb{A}^{\infty,p})^\times$-multiples. (Here s is any geometric point of S. The map $\bar{\eta}^p$ for any two choices of s can be identified.)
- $\mathrm{Lie}\, A$ with the induced $\mathcal{O}_F$-action satisfies a "determinant" condition ([Kot92b, §5]).
- Two quadruples $(A_1, \lambda_1, i_1, \bar{\eta}^p_1)$ and $(A_2, \lambda_2, i_2, \bar{\eta}^p_2)$ are equivalent if there is a prime-to-p isogeny $A_1 \to A_2$ taking $\lambda_1, i_1, \bar{\eta}^p_1$ to $\gamma\lambda_2, i_2, \bar{\eta}^p_2$ for some $\gamma \in \mathbb{Z}^\times_{(p)}$.

If U^p is sufficiently small, the above functor is representable by a smooth projective $\mathcal{O}_{F_w}$-scheme, which is denoted Sh_{U^p}. (Recall that U_p is maximal, which means we are in prime-to-p level.)

If U_p is not maximal but a certain congruence subgroup of $G(\mathbb{Q}_p)$, the integral model for Sh_U can be constructed by adding Drinfeld level structure at p to the moduli problem ([HT01, Ch II.2, III.4]). In general Sh_U has bad reduction mod p, although Sh_U is smooth over $\mathcal{O}_{F_w}$ if U_p is maximal compact. Although we will not discuss this further, the integral model with bad reduction plays a crucial role in the proof of Theorem 5.11.

5.2. **Newton stratification.** The prototype for Newton stratification is seen on the mod p fiber $\overline{Y}$ of an affine elliptic modular curve Y. Let k be a field of characteristic p. A k-point on $\overline{Y}$ corresponds to an elliptic curves over k (with level structure). Thus there is a set-theoretic decomposition

$$\overline{Y} = \overline{Y}^{\mathrm{ord}} \coprod \overline{Y}^{\mathrm{ss}}$$

such that the points of $\overline{Y}^{\mathrm{ord}}$ (resp. $\overline{Y}^{\mathrm{ss}}$) correspond to ordinary (resp. supersingular) elliptic curves. Note that $\overline{Y}^{\mathrm{ord}}$ (resp. $\overline{Y}^{\mathrm{ss}}$) is Zariski open (resp. closed) in $\overline{Y}$.

There is an analogous construction for the mod w fiber $\overline{\mathrm{Sh}}_{U^p}$ of Sh_{U^p} (and also for the mod w fiber of integral models with Drinfeld level structure at p). To do this, we need a little preparation. Let $(\mathscr{A}^{\mathrm{univ}}, \lambda^{\mathrm{univ}}, i^{\mathrm{univ}})$ denote the universal abelian scheme over $\overline{\mathrm{Sh}}_{U^p}$ with polarization and endomorphism structure. Let s be an $\overline{\mathbb{F}}_p$-point of $\overline{\mathrm{Sh}}_{U^p}$. Then $\mathscr{A}^{\mathrm{univ}}_s$, denoting the fiber of $\mathscr{A}^{\mathrm{univ}}$ at s, is an abelian variety over $\overline{\mathbb{F}}_p$. Since its p-divisible group $\mathscr{A}^{\mathrm{univ}}_s[p^\infty]$ is equipped with an action of $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p \simeq \prod_{x|p} \mathcal{O}_{F_x}$ (and a polarization), it is decomposed as

$$\mathscr{A}^{\mathrm{univ}}_s[p^\infty] = \bigoplus_{x|p} \mathscr{A}^{\mathrm{univ}}_s[x^\infty]$$

with respect to that action. The $\mathcal{O}_{F_w}$-height of each $\mathscr{A}^{\mathrm{univ}}_s[x^\infty]$ is n. The determinant condition in the moduli problem implies that the dimension of $\mathscr{A}^{\mathrm{univ}}_s[x^\infty]$, which is the same as $\dim_{\overline{\mathbb{F}}_p} \mathrm{Lie}\, \mathscr{A}^{\mathrm{univ}}_s[x^\infty]$, equals 1 if $x = w$ and 0 if $x \neq w$ and $x|_E = u$. In other words, whereas in the latter case, $\dim_{\overline{\mathbb{F}}_p} \mathrm{Lie}\, \mathscr{A}^{\mathrm{univ}}_s[x^\infty]$ is an étale p-divisible group. (If $x|_E \neq w|_E$ then $x^c|_E = w|_E$ and $\mathscr{A}^{\mathrm{univ}}_s[x^\infty]$ is isomorphic to $\mathscr{A}^{\mathrm{univ}}_s[(x^c)^\infty]^\vee$ via the prime-to-p polarization induced by λ^{univ}.) The upshot is that there is a stratification of $\overline{\mathrm{Sh}}_{U^p}$ into locally closed subsets

$$\overline{\mathrm{Sh}}_{U^p} = \coprod_{0 \le h \le n-1} \overline{\mathrm{Sh}}^{(h)}_{U^p}$$

such that $s \in \overline{\mathrm{Sh}}_{U^p}(\overline{\mathbb{F}}_p)$ lands in $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ if and only if the maximal étale quotient of $\mathscr{A}^{\mathrm{univ}}_s[w^\infty]$ has $\mathcal{O}_{F_w}$-height h. For any $0 \le h' \le n-1$, the union of $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ for all $0 \le h \le h'$ is Zariski closed in $\overline{\mathrm{Sh}}_{U^p}$. The last fact reflects the principle that the Newton polygon of a p-divisible group can only go up under specialization on the base scheme. Each locally closed subset $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ is viewed as a scheme with the reduced subscheme structure. It turns out that each $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ has dimension h (but it is a nontrivial fact that every stratum is nonempty).

Remark 5.1. Recall that Newton polygons may be used to classify isogeny classes of p-divisible groups over $\overline{\mathbb{F}}_p$ in the context of Dieudonné theory. In general the Newton stratification for Shimura varieties is defined according to isogeny classes of p-divisible groups with polarization and endomorphism structure. In our case it can be shown that the latter isogeny classes are classified by the integer h introduced above (which amounts to the Newton polygon of height n, dimension 1 and "étale height" h with nonnegative slopes).

Remark 5.2. Our Shimura varieties are nice in that the study of the universal p-divisible group $\mathscr{A}^{\mathrm{univ}}[p^\infty]$ over $\overline{\mathrm{Sh}}_{U^p}$ (with additional structure) is essentially reduced to the study of p-divisible groups of dimension 1. This dimension is linked to the signature of $G^1(\mathbb{R})$ in §4.1 via the determinant condition in the moduli problem. (For a different choice of signatures, the dimension is higher than 1 in general.) In the dimension 1 case, the deformation theory of p-divisible groups works nicely and the Drinfeld level structure for the integral models behaves well. This gives another reason why our choice of G is favorable (cf. §4.1, Remark 4.3).

5.3. **Igusa varieties.** Igusa varieties in the case of elliptic modular curves were studied in [Igu59] and [KM85]. They also appear in the context of p-adic automorphic forms as in Hida's work. (See [Hid04] for instance.) In work of Harris and Taylor, Igusa varieties and Rapoport-Zink spaces are used to study the bad reduction of Shimura varieties (Theorem 5.5). We will use the same strategy here. In this subsection we give basic definitions for Igusa varieties.

Temporarily fix $0 \le h \le n-1$. Let $\Sigma^{(h)} := \prod_{x|p} \Sigma_x$ and $\Sigma_w = \Sigma^0_w \times \Sigma^{\mathrm{et}}_w$ where each Σ_x denotes the p-divisible group over $\overline{\mathbb{F}}_p$ with an $\mathcal{O}_{F_x}$-action and

- Σ^0_w is connected of $\mathcal{O}_{F_w}$-height $n-h$ and dimension 1,
- Σ^{et}_w is étale of $\mathcal{O}_{F_w}$-height h and
- Σ_x is étale of $\mathcal{O}_{F_x}$-height n for $x \ne w$, $x|u$.
- $\Sigma_{x^c} = \Sigma^\vee_x$ for $x|u$.

The Newton stratification is defined so that the geometric fibers of $\mathscr{A}^{\mathrm{univ}}[x^\infty]$ are isogenous to Σ_x as p-divisible groups with $\mathcal{O}_{F_x}$-action on $\overline{\mathrm{Sh}}^{(h)}_{U^p}$. (In fact, a little more is true in our particular case. Namely, $\mathscr{A}^{\mathrm{univ}}[p^\infty]$ is fiberwise isogenous to $\Sigma^{(h)}$ with $\mathcal{O}_F \otimes \mathbb{Z}_p$-action and polarization.)

Define

$$J^{(h)}_w(\mathbb{Q}_p) := \mathrm{Aut}_{\mathcal{O}_{F_w}}(\Sigma_h) \simeq D^\times_{n-h,F_w} \times GL_n(F_w)$$

and

$$J^{(h)}(\mathbb{Q}_p) := \mathbb{Q}_p^\times \times J^{(h)}_w(\mathbb{Q}_p) \times \prod_{x|u,\ x\ne w} GL_n(F_x)$$

where D_{n-h} is a central division algebra over $\mathcal{O}_{F_w}$ with Hasse invariant $\frac{1}{n-h}$.

Remark 5.3. The group $J^{(h)}(\mathbb{Q}_p)$ may be naturally identified with the automorphism group of $\Sigma^{(h)}$ with $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$-action and polarization (the latter preserved up to $\mathbb{Q}_p^\times$-multiples) in the isogeny category. We will often ignore the $\mathbb{Q}_p^\times$-part to simplify exposition, and this allows us to view $J^{(h)}(\mathbb{Q}_p)$ loosely as the automorphism group of $\prod_{x|u} \Sigma_x$ with $\prod_{x|u} \mathcal{O}_{F_x}$-action in the isogeny category.

Note that

$$G(\mathbb{Q}_p) \simeq \mathbb{Q}_p^\times \times \prod_{x|u} GL_n(F_x). \tag{5.1}$$

Thus $J^{(h)}$ is an inner form of a Levi subgroup of (a parabolic subgroup of) G over $\mathbb{Q}_p$.

Example 5.4. In the case of elliptic modular curves, the analogue of $J^{(h)}(\mathbb{Q}_p)$ is $\mathbb{Q}_p^\times \times \mathbb{Q}_p^\times$ (resp. $D^\times$) for the ordinary (resp. supersingular) stratum, where D is a central division algebra over $\mathbb{Q}_p$ of degree 4.

By abuse of notation, the pullback of $\mathscr{A}^{\mathrm{univ}}$ from $\overline{\mathrm{Sh}}_{U^p}$ to $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ will still be denoted by $\mathscr{A}^{\mathrm{univ}}$. Let $\mathscr{G}^0$ (resp. $\mathscr{G}^{\mathrm{et}}$) be the maximal connected sub p-divisible group (resp. maximal quotient p-divisible group) of $\mathscr{A}^{\mathrm{univ}}[p^\infty]$.

We will introduce Igusa varieties and Rapoport-Zink spaces which are closely related to the stratum $\overline{\mathrm{Sh}}^{(h)}_{U^p}$. Let $m \geq 1$. The first object $\mathrm{Ig}^{(h)}_{U^p,m}$ is the moduli space over $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ which associates to a scheme S over $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ the quadruple of isomorphisms $j = (j_{p,0}, j^0_w, j^{\mathrm{et}}_w, \{j_x\})$ where

- $j_{p,0} : \mathbb{Z}_p^\times \xrightarrow{\sim} \mathbb{Z}_p(1)^\times$.
- $j^0_w : \Sigma^0_h[w^m] \times_{\overline{\mathbb{F}}_p} S \xrightarrow{\sim} \mathscr{G}^0[w^m] \times_{\overline{\mathrm{Sh}}^{(h)}_{U^p}} S$ which is compatible with $\mathcal{O}_{F_w}$-actions,
- $j^{\mathrm{et}}_w : \Sigma^{\mathrm{et}}_h[w^m] \times_{\overline{\mathbb{F}}_p} S \xrightarrow{\sim} \mathscr{G}^{\mathrm{et}}[w^m] \times_{\overline{\mathrm{Sh}}^{(h)}_{U^p}} S$ which is compatible with $\mathcal{O}_{F_w}$-actions,
- $j_x : \Sigma_x[x^m] \times_{\overline{\mathbb{F}}_p} S \xrightarrow{\sim} \mathscr{G}[x^m] \times_{\overline{\mathrm{Sh}}^{(h)}_{U^p}} S$ for $x|u$ and $x \neq w$ which is compatible with $\mathcal{O}_{F_x}$-actions and

To be precise, there is a technical condition imposed on the data that $(j_{p,0}, j^0_w, j^{\mathrm{et}}_w, \{j_x\})$ should be liftable to the level of p-divisible groups.

Remark 5.5. An equivalent formulation is that $\mathrm{Ig}^{(h)}_{U^p,m}$ parametrizes graded isomorphisms $j : \Sigma^{(h)}[p^m] \xrightarrow{\sim} \mathrm{gr}(\mathscr{G}^0)[p^m]$ compatible with $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$-actions and polarizations, the latter up to $(\mathbb{Z}/p^m\mathbb{Z})^\times$-multiple, where the grading is given by slope filtration. (See [Man05, Def 3].) We have chosen to give a more down-to-earth moduli problem above.

It turns out that each $\mathrm{Ig}^{(h)}_{U^p,m}$ is a smooth variety over $\overline{\mathbb{F}}_p$ and finite Galois over $\overline{\mathrm{Sh}}^{(h)}_{U^p}$. By abuse of notation, $\mathscr{L}_\xi$ will also denote its pullback from $\overline{\mathrm{Sh}}^{(h)}_{U^p}$ to each $\mathrm{Ig}^{(h)}_{U^p,m}$. Define

$$H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi) := \sum_{k \geq 0} (-1)^k \varinjlim_{U^p,m} H^k_c(\mathrm{Ig}^{(h)}_{U^p,m}, \mathscr{L}_\xi)$$

which is naturally a virtual representation of $G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p)$. Indeed, the action of $G(\mathbb{A}^{\infty,p})$ is inherited from the Hecke action on the tower $\overline{\mathrm{Sh}}^{(h)}_{U^p}$. (Note that the tower of $\overline{\mathrm{Sh}}^{(h)}_{U^p}$

for varying U^p is invariant under the Hecke action of $G(\mathbb{A}^{\infty,p})$.) The action of $J^{(h)}(\mathbb{Q}_p)$ is defined by extending the natural action of

$$\mathbb{Z}_p^\times \times (\mathcal{O}_{D_{n-h,F_w}} \times GL_h(\mathcal{O}_{F_w})) \times \prod_{x|u, x\neq w} GL_n(\mathcal{O}_{F_w})$$

on the quadruple $(j_{p,0}, j_w^0, j_w^{\mathrm{et}}, \{j_x\})$. The computation of $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)$, which we will focus on in §6, is the most innovative part of [Shi11] and a vital input for the computation of $H(\mathrm{Sh}, \mathscr{L}_\xi)$.

5.4. **Rapoport-Zink spaces and the map** $\mathrm{Mant}^{(h)}$**.** The Rapoport-Zink spaces are local analogues of PEL Shimura varieties. Indeed, the former are moduli spaces of p-divisible groups with additional structure whereas the latter are moduli spaces of abelian schemes with additional (PEL) structure. Just as a Shimura variety is constructed from a reductive group over $\mathbb{Q}$ and other data, a Rapoport-Zink space is associated with a reductive group over $\mathbb{Q}_p$ and some other data. Recall that our main strategy is based on the philosophy that the cohomology of Shimura varieties is closely related to the global Langlands correspondence. Similarly it is believed that the cohomology of Rapoport-Zink spaces has a lot to do with the local Langlands correspondence. Indeed, this was one of the main motivations to study these spaces. For some precise conjectures, see Remark 5.10 below.

Rapoport-Zink spaces were introduced by Rapoport and Zink in an attempt to generalize the non-abelian Lubin-Tate spaces and the Drinfeld spaces. The latter two spaces are associated with general linear groups and the unit groups of division algebras, and had been studied the most in connection with the local Langlands correspondence for GL_n (and in fact the Jacquet-Langlands correspondence as well). When $n = 1$, this reduces to the well-known relationship between the classical Lubin-Tate theory and local class field theory.

In our setting, the relevant Rapoport-Zink spaces are associated with $G(\mathbb{Q}_p)$ and isomorphic to (products of) the so-called non-abelian Lubin-Tate spaces. To be more concrete, the Rapoport-Zink space with no level structure, denoted $\mathrm{RZ}_w^{(h)}$, represents (as a formal scheme) the moduli problem

$$RZ_w^{(h)}(S) = \{(H, i, \beta)\}/\simeq$$

from the category of $\mathcal{O}_{F_w^{\mathrm{ur}}}$-scheme in which p is locally nilpotent to the category of sets, where

- H is a p-divisible group over S,
- $i : \mathcal{O}_{F_w}\mathbb{Q}_p \hookrightarrow \mathrm{End}(H)$ is a $\mathbb{Q}_p$-algebra morphism,
- $\beta : \Sigma_w^{(h)} \times_{\overline{\mathbb{F}}_p} \overline{S} \to H \times_S \overline{S}$ is a quasi-isogeny compatible with $\mathcal{O}_{F_w}$-actions, where $\overline{S}$ is the closed subscheme of S defined by the ideal sheaf $p\mathcal{O}_S$.
- The determinant condition as in [RZ96, 3.23.(a)].

Then $\mathrm{RZ}_w^{(h)}$ is a Rapoport-Zink space associated to $R_{F_w/\mathbb{Q}_p}GL_n$. The Rapoport-Zink space $\mathrm{RZ}_0^{(h)}$ (without level structure) associated to G can be defined as the product of $\mathrm{RZ}_w^{(h)}$ and the zero-dimensional Rapoport-Zink spaces accounting for $\mathbb{Q}_p^\times$ and $\prod_{x|u,x\neq w} GL_n$ in the decomposition 5.1.

The formal scheme $\mathrm{RZ}_0^{(h)}$ gives rise to a rigid analytic space. With this space at the bottom, one can throw in level structure to construct a projective system of rigid analytic spaces $RZ^{(h)} = \{RZ^{(h)}_{U_p}\}$ indexed by open compact subgroups U_p of $G(\mathbb{Q}_p)$. They are (not of finite type but) locally of finite type over $\widehat{F}^{\mathrm{ur}}_w$.

We will not give detail, but the l-adic cohomology of $\mathrm{RZ}^{(h)}$ comes equipped with a natural commuting action of $J^{(h)}(\mathbb{Q}_p)$, $G(\mathbb{Q}_p)$ and W_{F_w}. (The group $J^{(h)}(\mathbb{Q}_p)$ acts on the deformation datum for each fixed level U_p. On the other hand, $G(\mathbb{Q}_p)$ acts on the tower $\mathrm{RZ}^{(h)}$ in the style of Hecke correspondences.) To study its cohomology effectively, especially in connection with the cohomology of Shimura varieties, it is useful to define the following map

$$\mathrm{Mant}^{(h)}_n : \mathrm{Groth}(J^{(h)}(\mathbb{Q}_p)) \to \mathrm{Groth}(G(\mathbb{Q}_p) \times W_{F_w})$$

as

$$\mathrm{Mant}^{(h)}_n(\rho) := \sum_{i,j\geq 0} (-1)^{i+j} \varinjlim_{U_p} \mathrm{Ext}^i_{J^{(h)}(\mathbb{Q}_p)}(H^j_c(\mathrm{RZ}^{(h)}_{U_p}, \overline{\mathbb{Q}}_l), \rho).$$

In the case under consideration, a complete description of $\mathrm{Mant}^{(h)}$ was given by Harris-Taylor. (See [Shi11, Prop 2.2] for a summary of results.) The case $h = 0$ turns out to be the most interesting. To show the flavor we state a result in the supercuspidal case, which was proved by Carayol ([Car86]) for $n = 2$ and Harris-Taylor for any n. (The analogue of the Mant map for Drinfeld spaces was computed in [Car90] for $n = 2$ and in [Dat07] for any n.) When $n = 1$, the formula essentially follows from the classical Lubin-Tate theory.

Theorem 5.6. *(*[HT01, Thm VII.1.3]*) Let $\rho \in \mathrm{Irr}(J^{(h)}(\mathbb{Q}_p))$ be such that $JL(\rho)$ is supercuspidal. Then*

$$\mathrm{Mant}^{(0)}_n(\rho) = (-1)^{n-1} \cdot JL(\rho) \otimes \mathscr{L}_{F_w}(JL(\rho)).$$

The known proofs of the theorem are global in nature in that a key input comes from the cohomology of Shimura varieties and its interaction with $\mathrm{Mant}^{(h)}$ (as in Theorem 5.11 below). The proof of the local Langlands correspondence for GL_n over p-adic fields, either by Henniart or by Harris-Taylor, is also global as it relies on a result on the cohomology of Shimura varieties.

For $h > 0$ we have an induction formula, which is implicit in [HT01].

Theorem 5.7. $\mathrm{Mant}^{(h)}_n(\rho_1 \otimes \rho_2) = \mathrm{Ind}^{GL_n}_{GL_{n-h,h}}(\mathrm{Mant}^{(0)}_{n-h}(\rho_1) \otimes \rho_2)$.

As the notation suggests,

$$\mathrm{Mant}^{(0)}_{n-h} : \mathrm{Groth}(\mathbb{Q}_p^\times \times D^\times_{n-h}) \to \mathrm{Groth}(\mathbb{Q}_p^\times \times GL_{n-h}(F_w) \times W_{F_w})$$

is built from the Rapoport-Zink spaces for GL_{n-h}, corresponding to the Newton polygon height $n-h$ and dimension 1. As we already remarked, $\mathrm{Mant}^{(0)}_{n-h}$ can be explicitly described thanks to Harris and Taylor. Therefore Theorem 5.7 enables us to compute $\mathrm{Mant}^{(h)}_n$ for all $1 \leq h \leq n-1$ as well.

Remark 5.8. It would natural and interesting to figure out $H^j_c(RZ^{(h)}_{U_p}, \overline{\mathbb{Q}}_l)$ directly, without taking the alternating sum over j. In the case at hand, the result was obtained by Boyer ([Pas09]).

Remark 5.9. In the history of class field theory, local class field theory was first proved by using its global counterpart, but later established by purely local methods. Thus it would be desirable to find a purely local proof of Theorem 5.6 and the local Langlands correspondence.

Let us mention some partial results (which are not meant to be exhaustive by any means). For the first problem, see [Str05]. As for the second problem, Bushnell and Henniart ([BH05a], [BH05b]) explicitly constructed the Langlands correspondence in the "essentially tame" case.

Remark 5.10. Theorem 5.6 is concerned with the Rapoport-Zink spaces for GL_n corresponding to the "basic" Newton polygon of pure slope $1/n$. It is basic in the sense that it lies above the other Newton polygons with the same end points. A natural generalization of Theorem 5.6 would be a description of the analogue of $\mathrm{Mant}^{(h)}$ for other Rapoport-Zink spaces. In this direction of research, the most prominent conjectures seem to be the following, which are very precise but stated here only loosely.

- Kottwitz's conjecture ([Rap95, Conj 5.1], cf. [Har01, Conj 5.3, 5.4]) - the Mant map for a basic Newton polygon is described in terms of discrete L-parameters.
- Harris's conjecture ([Har01, Conj 5.2]) - the Mant map for a non-basic Newton polygon is obtained by an induction formula.

In fact Theorem 5.7 is a special case of Harris's conjecture, but the general case is a wide open question. For a progress toward the first (resp. second) conjecture, see [Far04] (resp. [Man08]). The paper [Shi12] provides a little extra information in the case of Rapoport-Zink spaces for GL_n.

5.5. **The first basic identity.** The cohomology of each Newton stratum in the special fiber of Shimura varieties can be related to the cohomology of Igusa varieties and Rapoport-Zink spaces via a very neat formula. For our Shimura varieties, the result is due to Harris and Taylor. (Although their Shimura varieties are attached to an inner form of our G, their proof carries over to our case.) The following is a reformulation in the style of Mantovan's theorem. See the remark below Theorem 5.11.

Theorem 5.11. *(Harris-Taylor) The following holds in* $\mathrm{Groth}(G(\mathbb{A}^\infty) \times W_{F_w})$. *(We regard* $\mathrm{Mant}^{(h)}$ *as the identity map on the space of* $G(\mathbb{A}^{\infty,p})$*-representations.)*

$$H(\mathrm{Sh}, \mathscr{L}_\xi) = \sum_{0 \le h \le n-1} \mathrm{Mant}^{(h)}(H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi))$$

Sketch of proof. For simplicity, we assume $\mathscr{L}_\xi = \overline{\mathbb{Q}}_l$. Let $R\Psi_{\mathrm{Sh}}$ denote the nearby cycle complex on $\overline{\mathrm{Sh}}$ associated with $\overline{\mathbb{Q}}_l$. Then

$$H(\mathrm{Sh}, \overline{\mathbb{Q}}_l) = H(\overline{\mathrm{Sh}}, R\Psi_{\mathrm{Sh}}) = \sum_{0 \le h \le n-1} H(\overline{\mathrm{Sh}}^{(h)}, R\Psi_{\mathrm{Sh}})$$

in $\mathrm{Groth}(G(\mathbb{A}^\infty) \times W_{F_w})$, where $R\Psi_{\mathrm{Sh}}$ also denotes the sheaves on $\overline{\mathrm{Sh}}^{(h)}$ by abuse of notation.

Let us pretend that $m : \overline{\mathrm{RZ}}^{(h)} \times_{\overline{\mathbb{F}}_p} \mathrm{Ig}^{(h)} \to \overline{\mathrm{Sh}}^{(h)}$ is a Galois covering with Galois group $J^{(h)}(\mathbb{Q}_p)$. Although this is not literally true, we hope that this helps the reader to grasp some core ideas more easily.

Let $p_1 : \overline{\mathrm{RZ}}^{(h)} \times_{\overline{\mathbb{F}}_p} \mathrm{Ig}^{(h)} \to \overline{\mathrm{RZ}}^{(h)}$ denote the projection map. Berkovich's theory provides $R\Psi_{\mathrm{RZ}}$ on $\overline{\mathrm{RZ}}^{(h)}$, which is the analogue for formal schemes of nearby cycle complexes. Harris and Taylor proved (and Mantovan generalized) a deep fact that

$$m^* R\Psi_{\mathrm{Sh}}^{(h)} \simeq p_1^* R\Psi_{\mathrm{RZ}},$$

which roughly asserts that Shimura varieties and Rapoport-Zink spaces present the same kind of singularities along the Newton stratum for h. Then the following is morally true. The second last equality uses the fact that $H_c(\overline{\mathrm{RZ}}^{(h)}, R\Psi_{\mathrm{RZ}})$ is dual to $H_c(\mathrm{RZ}^{(h)}, \overline{\mathbb{Q}}_l)$ via Berkovich's theory.

$$\begin{aligned} H(\overline{\mathrm{Sh}}^{(h)}, R\Psi_{\mathrm{Sh}}) &= H_*(J^{(h)}(\mathbb{Q}_p), H_c(\overline{\mathrm{RZ}}^{(h)} \times \mathrm{Ig}^{(h)}, m^* R\Psi_{\mathrm{Sh}}) \\ &= H_*(J^{(h)}(\mathbb{Q}_p), H_c(\overline{\mathrm{RZ}}^{(h)} \times \mathrm{Ig}^{(h)}, p_1^* R\Psi_{\mathrm{RZ}}) \\ &= \mathrm{Tor}_{C_c^\infty(J^{(h)}(\mathbb{Q}_p))}(H_c(\overline{\mathrm{RZ}}^{(h)}, R\Psi_{\mathrm{RZ}}), H_c(\mathrm{Ig}^{(h)}, \overline{\mathbb{Q}}_l)) \\ &= \mathrm{Ext}_{J^{(h)}(\mathbb{Q}_p)}(H_c(\mathrm{RZ}^{(h)}, \overline{\mathbb{Q}}_l), H_c(\mathrm{Ig}^{(h)}, \overline{\mathbb{Q}}_l)) \\ &= \mathrm{Mant}^{(h)}(H_c(\mathrm{Ig}^{(h)}, \overline{\mathbb{Q}}_l)). \end{aligned}$$

□

Remark 5.12. Theorem 5.11 was extended by Mantovan ([Man05], [Man11]) to PEL-type Shimura varieties of unitary or symplectic type when the PEL datum is "unramified" at p (which amounts to the running assumption in [Kot92b]) and Kottwitz's integral model of Sh (with good reduction) is proper over $\mathcal{O}_{F_w}$. Fargues obtained a similar formula ([Far04, Cor 4.6.3]) when restricted to the basic (cf. Remark 5.10) stratum.

Remark 5.13. In the problem of understanding $H(\mathrm{Sh}, \mathscr{L}_\xi)$ as a virtual representation of $G(\mathbb{A}^\infty) \times W_{F_w}$, two sources of difficulty are bad reduction (or ramified Galois action) at w and global endoscopy for G. Theorem 5.11 enables us to separate the two kinds of difficulty. Namely, the information of bad reduction is mostly contained in $\mathrm{Mant}^{(h)}$ (which arises from a purely local geometric object) while the global endoscopy is captured by $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)$. It is worth noting that $\mathrm{Ig}^{(h)}$ is global in nature whereas $\mathrm{RZ}^{(h)}$ is a purely local object which can be defined independently of Shimura varieties.

Write $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi) = \sum_{i \in I} n_i [\pi_i^{\infty,p}][\rho_i]$ where I is an index set, $\pi^{\infty,p} \in \mathrm{Irr}(G(\mathbb{A}^{\infty,p}))$ and $\rho_p \in \mathrm{Irr}(J^{(h)}(\mathbb{Q}_p))$. Define the "$\Pi^{\infty,p}$-part" as

$$H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\} := \sum_{\mathrm{WBC}(\pi_i^{\infty,p}) = \Pi^{\infty,p}} n_i[\rho_i] \ \in \mathrm{Groth}(J^{(h)}(\mathbb{Q}_p)). \tag{5.2}$$

This is independent of the expansion of $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)$. We can prove

Corollary 5.14. *In* $\mathrm{Groth}(G(\mathbb{Q}_p) \times W_{F_w})$*,*

$$\pi_p \otimes R_l(\Pi) = \sum_{0 \le h \le n-1} \mathrm{Mant}^{(h)}(H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\}).$$

Proof. The corollary is basically obtained by taking the $\{\Pi^{\infty,p}\}$-part of Theorem 5.11. □

Since we know how $\mathrm{Mant}^{(h)}$ works, the proof of Theorem 4.6 is reduced to the problem of understanding $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\}$ as a virtual representation of $J^{(h)}(\mathbb{Q}_p)$. This brings us to the next section.

6. Cohomology of Igusa varieties

6.1. **Counting point formula.** The action of $G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p)$ on $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\}$ is basically given by "Hecke correspondences". (The $G(\mathbb{A}^{\infty,p})$-action is indeed compatible with the Hecke action on Shimura varieties but the $J^{(h)}(\mathbb{Q}_p)$-action is more subtle. We will ignore the subtlety here, though.) The trace of the Hecke action on $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)$ can be computed in terms of fixed points of $\mathrm{Ig}^{(h)}(\overline{\mathbb{F}}_p)$ under algebraic correspondences thanks to Fujiwara and Varshavsky, who proved Deligne's conjecture ([Fuj97, Cor 5.4.5], [Var07, Thm 2.3.2]). The latter is a version of the Grothendieck-Lefschetz trace formula and needed here as $\mathrm{Ig}^{(h)}$ is usually non-proper over $\overline{\mathbb{F}}_p$. Roughly speaking, Deligne's conjecture says that the Grothendieck-Lefschetz trace formula holds for an algebraic correspondence on a non-proper variety if that correspondence is twisted by a large enough power of Frobenius.

Since $\mathrm{Ig}^{(h)}$ is a moduli space for $(A, \lambda, i, \bar{\eta}^p)$ as well as certain isomorphisms of p-divisible groups, the fixed points under a correspondence on $\mathrm{Ig}^{(h)}$ are naturally described in terms of the moduli data. Our hope is to extract some automorphic information from the fixed point formula. The best way might be to relate the fixed point formula to an analogous[3] formula in automorphic representation theory, such as the Arthur-Selberg trace formula. But the latter formula has obviously no reference to abelian varieties or their structures. So the main problem is to massage the fixed-point formula for $\mathrm{Ig}^{(h)}$ to obtain a trace formula for the Hecke action on $\mathrm{Ig}^{(h)}$ which resembles the geometric side of the trace formula for G. (So to speak, it is about the passage from (6.2) to the statement of Theorem 6.3 below.) In the context of unitary Shimura varieties, this was carried out by Harris and Taylor for a certain $U(1, n-1)$-type unitary group with no endoscopy. A trace formula for Igusa varieties was proved ([Shi10]) for any PEL-type Shimura varieties associated to unitary or symplectic groups (possibly with nontrivial endoscopy), in the spirit of Langlands-Kottwitz's formula ([Kot92b]) for Shimura varieties with good reduction.

Before stating the result, we define the notion of Kottwitz triples in our context (which are somewhat different from those for Shimura varieties with good reduction as in [Kot90, §2]).

Definition 6.1. By an **effective Kottwitz triple** (of type $0 \le h \le n-1$), we mean a triple $(\gamma_0; \gamma, \delta)$ where

- $\gamma_0 \in G(\mathbb{Q})$ is semisimple, and elliptic in $G(\mathbb{R})$
- $\gamma \in G(\mathbb{A}^{\infty,p})$ and $\gamma_0 \sim_{\overline{\mathbb{A}}^{\infty,p}} \gamma$.
- $\delta \in J^{(h)}(\mathbb{Q}_p)$ is acceptable (to be explained) and $\gamma_0 \sim_{\overline{\mathbb{Q}}_p} \delta$ in $G(\overline{\mathbb{Q}}_p)$ via a natural embedding $J^{(h)}(\overline{\mathbb{Q}}_p) \hookrightarrow G(\overline{\mathbb{Q}}_p)$ (natural up to $G(\overline{\mathbb{Q}}_p)$-conjugacy).
- a certain Galois cohomology invariant $\alpha(\gamma_0; \gamma, \delta)$ vanishes.

Two Kottwitz triples $(\gamma_0; \gamma, \delta) \sim (\gamma_0'; \gamma', \delta')$ are said to be equivalent if $\gamma_0 \sim_{st} \gamma_0'$, $\gamma \sim_{\mathbb{A}^{\infty,p}} \gamma'$, and $\delta \sim \delta'$. For each $0 \le h \le n-1$, we define $KT^{(h),\mathrm{eff}}$ to be the set of equivalence classes of all effective Kottwitz triples of type h.

Remark 6.2. The word "effective" refers to the last condition, which is closely tied with the phenomenon of endoscopy. The analogous fact is that an element $\gamma \in G(\mathbb{A})$ is not always $G(\overline{\mathbb{A}})$-conjugate to an element of $G(\mathbb{Q})$. The failure is detected by the nonvanishing of a similar Galois cohomology invariant. When we say that a group G over $\mathbb{Q}$ has "no endoscopy", it

[3] Only remotely analogous, a priori.

indicates that this failure does not occur. (In that case no endoscopic groups other than the quasi-split inner form of G will contribute to the stable trace formula for G.)

The counting point formula is stated below. The terminology "acceptable" will not be defined but morally means that "twisted by enough power of Frobenius" in a suitable sense.

Theorem 6.3. *(*[Shi09, Thm 13.1]*)*
If $\varphi \in C_c^\infty(G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p))$ is acceptable, then

$$\mathrm{tr}\,(\varphi | H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)) = \sum_{(\gamma_0;\gamma,\delta)\in KT^{(h),\mathrm{eff}}} \mathrm{vol}(I_\infty(\mathbb{R})^1)^{-1} |A(I_0)|\ \mathrm{tr}\,\xi(\gamma_0) \cdot O_{(\gamma,\delta)}^{G(\mathbb{A}^{\infty,p})\times J^{(h)}(\mathbb{Q}_p)}(\varphi)$$

Sketch of proof. For simplicity, assume ξ is trivial so that $\mathscr{L}_\xi = \overline{\mathbb{Q}}_l$. Let $U_p(m) \subset J^{(h)}(\mathbb{Q}_p)$ denote the kernel of $\mathrm{Aut}(\Sigma^{(h)}) \to \mathrm{Aut}(\Sigma^{(h)}[p^m])$. It is enough to treat the case where $\varphi = \varphi^{\infty,p}\varphi'_p$ with $\varphi^{\infty,p} = \mathrm{char}_{U^p g^p U^p}$ and $\varphi'_p = \mathrm{char}_{U_p(m) g_p U_p(m)}$, as the general case is obtained by taking linear combinations. Then the left hand side is identified with

$$\mathrm{tr}\,([U^p g^p U^p] \times [U_p(m) g_p U_p(m)] | H_c(\mathrm{Ig}^{(h)}_{U^p,m}, \overline{\mathbb{Q}}_l)) \tag{6.1}$$

where $[U^p g^p U^p]$ and $[U_p(m) g_p U_p(m)]$ are Hecke correspondences for $\mathrm{Ig}^{(h)}_{U^p,m}$. The solution of Deligne's conjecture allows us to evaluate (6.1) as the number of fixed points on $\mathrm{Ig}^{(h)}(\overline{\mathbb{F}}_p)$ under the product correspondence. Recall from §5.3 that

$$\mathrm{Ig}^{(h)}(\overline{\mathbb{F}}_p) = \{(A, \lambda, i, \overline{\eta}^p, j)\}/\simeq$$

where A is an abelian variety over $\overline{\mathbb{F}}_p$ equipped with additional structure. It is not difficult to show that the number of fixed points corresponding to a given (A, λ, i) equals a sum of orbital integral of φ. To summarize the situation more precisely, we have

$$\mathrm{tr}\,(\varphi | H_c(\mathrm{Ig}^{(h)}, \overline{\mathbb{Q}}_l)) = \sum_{\{(A,\lambda,i)\}/\simeq} \left(\sum_{a\in \mathrm{Aut}^0(A,\lambda,i)} (\mathrm{const.}) \cdot O_a^{G(\mathbb{A}^{\infty,p})\times J^{(h)}(\mathbb{Q}_p)}(\varphi) \right) \tag{6.2}$$

where Aut^0 is the automorphism in the isogeny category (with additional structure). Then the proof is essentially completed by proving a natural bijection between $KT^{(h),\mathrm{eff}}$ and the set of A, λ, i, a in the sum. This is the core of the argument and involves a refinement of Honda-Tate theory, CM-lifting of abelian varieties from characteristic p to characteristic 0, Galois cohomology computations, theory of isocrystals over $\overline{\mathbb{F}}_p$ and others. $\square$

We do not lose generality by restricting ourselves to acceptable functions. More precisely,

Lemma 6.4. *(*[Shi09, Lem 6.4]*) For $\Pi_1, \Pi_2 \in \mathrm{Groth}(G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p))$,*

$$\mathrm{tr}\,\Pi_1(\varphi) = \mathrm{tr}\,\Pi_2(\varphi)$$

for all acceptable functions $\varphi \in C_c^\infty(G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p))$ if and only if $\Pi_1 = \Pi_2$ in the Grothendieck group.

Remark 6.5. The lemma was used in [HT01] without stating it as a lemma.

6.2. **Stabilization.** We intend to use Theorem 6.3 to understand $H_c(\mathrm{Ig}_b, \mathscr{L}_\xi)$ in terms of automorphic representations of G. It is very common (e.g. in the trace formula approach to the Langlands functoriality) that the trace formula should be stabilized to have interesting applications. Thus it is natural to attempt to stabilize the right hand side of Theorem 6.3. In other words, we want to rewrite the sum of orbital integrals as a sum of stable orbital integrals on (G and its endoscopic groups). As far as elliptic conjugacy classes are concerned, the stabilization process has been well-known thanks to Langlands and Kottwitz. In fact it has been conditional on the fundamental lemma, but the latter is recently established by Laumon, Ngô, Waldspurger and others.

However, there is an immediate obstacle due to the peculiarity of our trace formula. First of all, $G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p)$ is a strange topological group in that it is not the set of $\mathbb{A}^\infty$-points of any reductive group over $\mathbb{Q}$. So it is not a priori clear how to adapt the stabilization process and make sense of the Langlands-Shelstad transfer at p. This problem is successfully solved in [Shi10] for PEL-Shimura varieties of unitary or symplectic type. The result has the following form.

Theorem 6.6. *(*[Shi10, Thm 7.2]*) Let $\varphi \in C_c^\infty(G(\mathbb{A}^{\infty,p}) \times J^{(h)}(\mathbb{Q}_p))$ be an acceptable function. For each $G_{\vec{n}} \in \mathscr{E}^{\mathrm{ell}}(G)$, one can construct a function $\phi^{\vec{n}}$ from φ such that*

$$\mathrm{tr}\,(\varphi | H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)) = |\ker^1(\mathbb{Q}, G)| \sum_{G_{\vec{n}} \in \mathscr{E}^{\mathrm{ell}}(G)} \iota(G, G_{\vec{n}}) ST_e^{G_{\vec{n}}}(\phi^{\vec{n}})$$

(See the end of §6.2 for nice properties enjoyed by $\phi^{\vec{n}}$.)

Sketch of idea. It suffices to handle the case $\varphi = (\prod_{v \neq p,\infty} \varphi_v) \times \varphi'_p$. Away from p and ∞, the function $\phi_v^{\vec{n}}$ is the Langlands-Shelstad transfer of φ_v via $\widetilde{\eta}_{\vec{n}}$. At $v = \infty$ one has an explicit construction using Shelstad's real endoscopy and Clozel-Delorme's pseudo-coefficients for discrete series. The most interesting and important for applications is the case of $v = p$. Here the construction of $\phi_p^{\vec{n}}$ from φ'_p has no analogue in the usual trace formula business, as we must find a natural transfer from $J^{(h)}(\mathbb{Q}_p)$ (not $G(\mathbb{Q}_p)$) to $G_{\vec{n}}(\mathbb{Q}_p)$ which is an endoscopic group of $G(\mathbb{Q}_p)$ (but typically not of $J^{(h)}(\mathbb{Q}_p)$).

The idea is that there is a natural finite set of groups $\{M_H\}$ where each M_H is simultaneously an endoscopic group of $J^{(h)}$ and a Levi subgroup of $G_{\vec{n}}$. For each M_H, φ'_H transfers to M_H by the Langlands-Shelstad transfer and then to $G_{\vec{n}}$ by a certain non-standard transfer (which makes sense if φ' is acceptable in the same sense as in Theorem 6.3). Then φ'_p is constructed as the signed sum of these transfers over the set of all M_H which intervene. □

After all the stable trace formula above will be used to extract spectral information. Thanks to a concrete description of $\phi^{\vec{n}}$ as sketched above (via the Langlands-Shelstad transfer or other means), the following is known (when φ and $\phi^{\vec{n}}$ admit product decompositions). The notation $\mathrm{Red}^{(h)}_{\vec{n}}$ will be defined in the next subsection when $\vec{n} = (n)$ and $\vec{n} = (n-1, 1)$. These are the only cases which concern us.

Proposition 6.7. (1) *When $v \neq p, \infty$, the identity in Proposition 3.1 holds (whenever case (i) or (ii) applies).*

(2) *(*[Shi11, Lem 5.10]*) When $v = p$, we have for all $\pi_{H,v} \in \mathrm{Irr}(G_{\vec{n}}(\mathbb{Q}_p))$,*

$$\mathrm{tr}\,\left(\mathrm{Red}^{(h)}_{\vec{n}}(\pi_{H,v})\right)(\varphi'_p) = \mathrm{tr}\,\pi_{H,v}(\phi_p^{\vec{n}}), .$$

(3) *When $v=\infty$, the trace of $\phi^{\vec{n}}_\infty$ on any discrete series of $G_{\vec{n}}(\mathbb{R})$ explicitly.*

6.3. **The maps $\mathrm{Red}^{(h)}_n$ and $\mathrm{Red}^{(h)}_{n-1,1}$.** The main reference for this subsection is [Shi11, §5.5]. Recall that

$$J^{(h)}(\mathbb{Q}_p)\simeq \mathbb{Q}_p^\times\times(D^\times_{n-h,F_w}\times GL_h(F_w))\times\prod_{x|u,\ x\neq w} GL_n(F_x)$$

$$G(\mathbb{Q}_p)\simeq \mathbb{Q}_p^\times\times\prod_{x|u} GL_n(F_x).$$

Denote by

$$LJ_{n-h,h}:\mathrm{Groth}(GL_{n-h}(F_w)\times GL_h(F_w)\to \mathrm{Groth}(D^\times_{n-h,F_w}\times GL_h(F_w))$$

Badulescu's Jacquet-Langlands map ([Bad07]) on the first factor and the identity map on the second. For $\pi_w\in\mathrm{Irr}(GL_n(F_w))$, set

$$\mathrm{Red}^{(h)}_{n;w}(\pi_w):=LJ_{n-h,h}(\mathrm{Jac}_{n-h,h}(\pi_w))$$

where $\mathrm{Jac}_{n-h,h}$ is the Jacquet module from GL_n to $GL_{n-h}\times GL_h$. Define

$$\mathrm{Red}^{(h)}_n:\mathrm{Groth}(G(\mathbb{Q}_p))\to\mathrm{Groth}(J^{(h)}(\mathbb{Q}_p))$$

so that for irreducible $\pi_p=\pi_{p,0}\otimes(\otimes_{x|u}\pi_x)$,

$$\mathrm{Red}^{(h)}_n(\pi_p):=\pi_{p,0}\otimes\mathrm{Red}^{(h)}_{n;w}(\pi_w)\otimes\left(\bigotimes_{x|u,x\neq w}\pi_x\right).$$

Remark 6.8. We are not being precise about normalization, for instance that of the Jacquet modules and parabolic inductions in this subsection (and other places). Also we dropped the sign appearing in [Shi11, §5.5] as the final result will be stated up to sign. See [Shi11, §5.5] for precise normalizations and signs.

The definition of $\mathrm{Red}^{(h)}_{n-1,1}$ is more technical as it is supposed to account for endoscopic terms. Recall

$$G(\mathbb{Q}_p)\simeq \mathbb{Q}_p^\times\times\prod_{x|u}(GL_{n-1}(F_x)\times GL_1(F_x)).$$

Let us define its w-part

$$\mathrm{Red}^{(h)}_{n-1,1;w}:\mathrm{Groth}(GL_{n-1}(F_w)\times GL_1(F_w))\to\mathrm{Groth}(D^\times_{n-h,F_w}\times GL_h(F_w))$$

$$\pi_{w,1}\otimes\pi_{w,2}\ \mapsto\ \begin{cases}0, & \text{if}\ \ h=0,\\ \mathrm{Ind}^{GL_{n-h,h}}_{GL_{n-h,h-1,1}}\left(\mathrm{Jac}_{n-h,h-1}(\pi_{w,1})\otimes\pi_{w,2}\right), & \text{if}\ \ 0<h<n-1,\\ \mathrm{Ind}^{GL_{n-h,h}}_{GL_{n-h,h-1,1}}\left(\mathrm{Jac}_{n-h,h-1}(\pi_{w,1})\otimes\pi_{w,2}\right)-\pi_{w,2}\otimes\pi_{w,1}, & \text{if}\ \ h=n-1.\end{cases}$$

The notation $\mathrm{Ind}^{GL_{n-h,h}}_{GL_{n-h,h-1,1}}$ means the obvious parabolic induction from $GL_{n-h}\times GL_{h-1}\times GL_1$ to $GL_{n-h}\times GL_h$. Again we avoid the issue of precise sign and normalization. Finally define

$$\mathrm{Red}^{(h)}_{n-1,1}:\mathrm{Groth}(G_{n-1,1}(\mathbb{Q}_p))\to\mathrm{Groth}(J^{(h)}(\mathbb{Q}_p))$$

for irreducible $\pi_{H,p} = \pi_{H,p,0} \otimes \pi_{H,w} \otimes (\otimes_{x|u,x\neq w} \pi_{H,x,1} \otimes \pi_{H,x,2})$ by

$$\mathrm{Red}^{(h)}_{n-1,1}(\pi_{H,p}) := \mathrm{Red}^{(h)}_{n-1,1;w}(\pi_{H,w}) \otimes \bigotimes_{x|u,x\neq w} \mathrm{Ind}^{GL_n}_{GL_{n-1,1}}(\pi_{H,x,1} \otimes \pi_{H,x,2}). \tag{6.3}$$

Example 6.9. Let $\pi_{w,1} \in \mathrm{Irr}(GL_{n-1}(F_w))$, $\pi_{w,2} \in \mathrm{Irr}(GL_1(F_w))$ be supercuspidal representations. Let $\pi_w := \pi_{w,1} \boxplus \pi_{w,2}$. (The induction is always irreducible as $n \geq 3$.) Then the above formulas yield

$$\mathrm{Red}^{(h)}_{n-1,1;w}(\pi_w) = \begin{cases} 0, & \text{if} \quad h \neq 1, n-1, \\ LJ(\pi_{w,1}) \otimes \pi_{w,2}, & \text{if} \quad h = 1, \\ \pi_{w,2} \otimes \pi_{w,1}, & \text{if} \quad h = n-1. \end{cases} \tag{6.4}$$

$$\mathrm{Red}^{(h)}_{n-1,1;w}(\pi_{w,1} \otimes \pi_{w,2}) = \begin{cases} 0, & \text{if} \quad h \neq 1, n-1, \\ LJ(\pi_{w,1}) \otimes \pi_{w,2}, & \text{if} \quad h = 1, \\ -\pi_{w,2} \otimes \pi_{w,1}, & \text{if} \quad h = n-1. \end{cases} \tag{6.5}$$

Remark 6.10. It may appear that $\mathrm{Red}^{(h)}_{n-1,1}$ is a very unnatural map, but it is not. It is the signed sum of two functorial transfers from $G_{n-1,1}$ to $J^{(h)}$ represented by the L-morphisms which occur naturally in the stabilization problem of §6.2. Since only inner forms of general linear groups are involved, the transfers may be made explicit, and thereby $\mathrm{Red}^{(h)}_{n-1,1}$ was obtained in [Shi11].

6.4. **Application of the twisted trace formula.** Denote by $\pi_p \in \mathrm{Irr}(G(\mathbb{Q}_p))$ a representation such that $\mathrm{BC}(\pi_p) \simeq \Pi_p$. (Such a π_p is unique up to isomorphism as p splits in E.) When m is even, define $\pi_{H,p} \in \mathrm{Irr}(G_{n-1,1}(\mathbb{Q}_p))$ such that $\mathrm{BC}(\pi_{H,p}) \simeq \psi_p \otimes \Pi_{1,p} \otimes \Pi_{2,p}$ in the notation of §4.3. In the latter case we recall (7.5), which says in particular

$$\Pi^1_w = \mathrm{Ind}(\Pi_{1,w} \otimes \Pi_{2,w}) \tag{6.6}$$

(The parabolic induction can be shown to be irreducible.) Recall that $H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\}$ was defined in (5.2).

Theorem 6.11. *For each $0 \leq h \leq n-1$, the following equalities hold in* $\mathrm{Groth}(J^{(h)}(\mathbb{Q}_p))$, *where the sign e_2 is independent of h. The constants in the formulas are some explicit positive integers.*

(1) *(m odd)*

$$H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi))\{\Pi^{\infty,p}\} = (\mathrm{const.}) \cdot [\mathrm{Red}^{(h)}_n(\pi_p)]. \tag{6.7}$$

(2) *(m even)*

$$H_c(\mathrm{Ig}^{(h)}, \mathscr{L}_\xi)\{\Pi^{\infty,p}\} = (\mathrm{const.}) \cdot \frac{1}{2}\left[(\mathrm{Red}^{(h)}_n(\pi_p) + e_2 \mathrm{Red}^{(h)}_{n-1,1}(\pi_{H,p}))\right]. \tag{6.8}$$

Sketch of proof. Suppose that φ admits a product decomposition and that each $\phi^{\vec{n}}$ in Theorem 6.6 is a transfer of $f^{\vec{n}}$ in base change (§2.3). Then the following identity is essentially due to Labesse:

$$ST^{G_{\vec{n}}}_e(\phi^{\vec{n}}) = \widetilde{T}^{\mathbb{G}_{\vec{n}}}(f^{\vec{n}})$$

where the right hand side is the twisted trace formula for $\mathbb{G}_{\vec{n}}$ (with respect to θ in §2.2). Thus Theorem 6.6 implies that

$$\begin{aligned}\mathrm{tr}\,(\varphi|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)) &= |\ker^1(\mathbb{Q},G)|\sum_{\vec{n}}\iota(G,G_{\vec{n}})\widetilde{T}^{\mathbb{G}_{\vec{n}}}(f^{\vec{n}}) \qquad (6.9)\\ &\sim \widetilde{T}^{\mathbb{G}_n}(f^n)+\frac{1}{2}\widetilde{T}^{\mathbb{G}_{n-1,1}}(f^{n-1,1})+(\text{other terms}).\end{aligned}$$

The notation $\sim$ indicates that $|\ker^1(\mathbb{Q},G)|$ is ignored. The spectral expansion of the twisted trace formula looks like

$$\widetilde{T}^{\mathbb{G}_n}(f^n)=\sum_{\Pi'}\widetilde{\mathrm{tr}}\,\Pi'(f^n)+\frac{1}{2}\sum_{\Pi'_M}\widetilde{\mathrm{tr}}\,(\mathrm{Ind}(\Pi'_M))(f^n)+(\text{other terms}).$$

The first (resp. second) sum runs over θ-stable automorphic representations of Π' of $GL_n(\mathbb{A}_F)$ (resp. Π'_M of $(GL_{n-1}\times GL_1)(\mathbb{A}_F)$). The twisted trace with respect to θ is denoted by $\widetilde{\mathrm{tr}}$. When $\vec{n}=(n-1,1)$,

$$\widetilde{T}^{\mathbb{G}_{n-1,1}}(f^{n-1,1})=\sum_{\Pi'_H}\widetilde{\mathrm{tr}}\,\Pi'_H(f^{n-1,1})+(\text{other terms}).$$

What (3.4) means for us is essentially (ignoring the character twist there)

$$\widetilde{\mathrm{tr}}\,(\Pi'_H)^{\infty,p}((f^{n-1,1})^{\infty,p})=\widetilde{\mathrm{tr}}\,\mathrm{Ind}((\Pi'_H)^{\infty,p})((f^n)^{\infty,p}).$$

The identity holds outside p and ∞ because everything is the usual transfer along the way outside p,∞ but there are some deviations from the usual transfer at p and ∞, as we have seen in the stabilization process.

On the other hand, the base change identities in Proposition 2.2 allows to rewrite the left hand side of (6.9) as

$$\mathrm{tr}\,(\varphi|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi))=\mathrm{tr}\,((f^n)^{\infty,p}\varphi'_p|BC^p(H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi))) \qquad (6.10)$$

where BC^p is the map applying local base change away from p and ∞.

Now we look back at the situation of §4.3 and suppose that m is odd. We can separate the $\Pi^{\infty,p}$-part from the two sides of (6.9) (with (6.10) applied to the left hand side) by varying test functions outside p and ∞. Then only the $\Pi'=\Pi$ term survives on the right hand side. (For this we appeal to the strong multiplicity one theorem of Jacquet and Shalika.) Hence

$$\mathrm{tr}\,(\varphi'_p|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)\{\Pi^{\infty,p}\})\sim\widetilde{\mathrm{tr}}\,\Pi_p(f^n_p)\cdot\widetilde{\mathrm{tr}}\,\Pi_\infty(f^n_\infty).$$

Propositions 2.2.(2) and 6.7.(2) tell us that

$$\widetilde{\mathrm{tr}}\,\Pi_p(f^n_p)=\mathrm{tr}\,\pi_p(\phi^n_p)=\mathrm{tr}\,\mathrm{Red}^{(h)}_n(\varphi'_p) \qquad (6.11)$$

and $\widetilde{\mathrm{tr}}\,\Pi_\infty(f^n_\infty)$ turns out to be a constant (depending only on ξ). Hence we obtain

$$\mathrm{tr}\,(\varphi'_p|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)\{\Pi^{\infty,p}\})\sim\mathrm{tr}\,\mathrm{Red}^{(h)}_n(\varphi'_p).$$

Since φ'_p can be chosen to be an arbitrary acceptable function, Lemma 6.4 concludes the proof.

It remains to treat the case when m is even. Again we can separate the $\Pi^{\infty,p}$-part from the two sides of (6.9) and notice that only the terms for $\Pi'_M = \Pi_M$ and $\Pi'_H = \Pi_M$ survive. Hence

$$\begin{aligned}\mathrm{tr}\,(\varphi'_p|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)\{\Pi^{\infty,p}\}) \quad \sim \quad & \frac{1}{2}\widetilde{\mathrm{tr}}\,\mathrm{Ind}(\Pi_{M,p})(f_p^n)\cdot\widetilde{\mathrm{tr}}\,\mathrm{Ind}(\Pi_{M,\infty})(f_\infty^n)\\ & +\frac{1}{2}\widetilde{\mathrm{tr}}\,\Pi_{M,p}(f_p^{n-1,1})\cdot\widetilde{\mathrm{tr}}\,\Pi_{M,\infty}(f_\infty^{n-1,1})\end{aligned}$$

By applying Propositions 2.2.(2) and 6.7.(2) to the second term at p, we obtain

$$\widetilde{\mathrm{tr}}\,\Pi_{M,p}(f_p^{n-1,1}) = \mathrm{tr}\,\pi_{H,p}(\phi_p^{n-1,1}) = \mathrm{tr}\,\mathrm{Red}^{(h)}_{n-1,1}(\varphi'_p).$$

One can compute that $\widetilde{\mathrm{tr}}\,\mathrm{Ind}(\Pi_{M,\infty})(f_\infty^n)$ and $\widetilde{\mathrm{tr}}\,\Pi_{M,\infty}(f_\infty^{n-1,1})$ are constants which coincide up to ± 1. Assign $e_2 = 1$ if they are the same and $e_2 = -1$ otherwise. By the above identity and (6.11),

$$\mathrm{tr}\,(\varphi'_p|H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)\{\Pi^{\infty,p}\}) \sim \frac{1}{2}\mathrm{tr}\,\mathrm{Red}^{(h)}_n(\varphi'_p) + \frac{1}{2}e_2\cdot\mathrm{tr}\,\mathrm{Red}^{(h)}_{n-1,1}(\varphi'_p).$$

The proof is finished by Lemma 6.4 as in the case when m is odd.

□

Remark 6.12. In work of Harris-Taylor, where no endoscopy arises, Theorem [HT01, Thm V.5.4] corresponds to part (1) of Theorem 6.11. In fact their theorem takes the form

$$H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi) = (\mathrm{const.})\cdot\mathrm{Red}^{(h)}_n H(\mathrm{Sh},\mathscr{L}_\xi) \qquad (6.12)$$

and is justified by the comparison of the trace formulas for $H_c(\mathrm{Ig}^{(h)},\mathscr{L}_\xi)$ and $H(\mathrm{Sh},\mathscr{L}_\xi)$. (For the latter, the Arthur-Selberg trace formula suffices as the Galois action is to be forgotten in the identity.) As long as no endoscopy occurs, (6.12) generalizes ([Shi12, Thm 6.7]). In the presence of endoscopy, a simple identity like (6.12) is not expected and the type of argument in the above proof seems to be more effective.

7. Proof of Theorem 4.6 under hypotheses

In the remainder we finish the proof of Theorem 4.6 under running hypotheses, namely Hypotheses 2.1, 4.2, 4.5, 4.7 and 4.9). This achieves the goal of the article: it was already explained how Theorem 4.6 implies Theorem 1.3 under the same running hypotheses, and §4.5 gave a short outline and references for a strategy to get rid of those assumptions.

The main ingredients of this section are the computation of $\mathrm{Mant}^{(h)}_n$ (§5.4), Corollary 5.14 and Theorem 6.11. As in the previous sections the argument will be sketched while overlooking some delicate points. A precise treatment can be found in section 6.2 of [Shi11], especially the proof of Theorem 6.4 there.

7.1. In case m is odd.

Lemma 7.1. *([Shi11, Prop 2.3]) For every $\pi_p \in \mathrm{Irr}(G(\mathbb{Q}_p))$, the following holds in* $\mathrm{Groth}(G(\mathbb{Q}_p)\times W_{F_w})$.

$$\sum_{h=0}^{n-1}\mathrm{Mant}^{(h)}_n(\mathrm{Red}^{(h)}_n(\pi_p)) = \pi_p\otimes\mathscr{L}_{F_w}(\Pi^1_w) \qquad (7.1)$$

Idea of proof. By taking the explicit description of $\mathrm{Mant}_n^{(h)}$ and $\mathrm{Red}_n^{(h)}$ (§5.4, §6.3) as inputs, one proves the lemma by computations with representations of p-adic general linear groups. We omit the detail, but see Example 7.2 below. □

Corollary 5.14 and Theorem 6.11 show that

$$\pi_p \otimes R(\Pi)|_{W_{F_w}} \sim \sum_{h=0}^{n-1} \mathrm{Mant}_n^{(h)}(\mathrm{Red}_n^{(h)}(\Pi_w^1)).$$

The above lemma tells us that the latter is equal to $\pi_p \otimes \mathscr{L}_{F_w}(\Pi_w^1)$. Hence the first part of Theorem 4.6 holds.

Example 7.2. We illustrate the proof of Lemma 7.1 in a particular case. Let

$$\pi_p = \pi_{p,0} \otimes (\otimes_{x|u} \pi_x) \tag{7.2}$$

where $\pi_w = \pi_{w,1} \oplus \pi_{w,2}$ is as in Example 6.9. According to Theorems 5.6 and 5.7,

$$\begin{aligned} \mathrm{Mant}_n^{(1)}(LJ(\pi_{w,1}) \otimes \pi_{w,2}) &= \mathrm{Ind}_{GL_{n-1,1}}^{GL_n}(\mathrm{Mant}_{n-1}^{(0)}(\pi_{w,1}) \otimes \pi_{w,2}) \\ &= (\pi_{w,1} \boxplus \pi_{w,2}) \otimes \mathscr{L}_{F_w}(\pi_{w,1}). \end{aligned} \tag{7.3}$$

Similarly

$$\mathrm{Mant}_n^{(n-1)}(\pi_{w,2} \otimes \pi_{w,1}) = (\pi_{w,1} \boxplus \pi_{w,2}) \otimes \mathscr{L}_{F_w}(\pi_{w,2}). \tag{7.4}$$

In light of (6.4), the left hand side of (7.1) is computed as

$$\begin{aligned} & \pi_{p,0} \otimes \Big(\mathrm{Mant}_n^{(1)}(LJ(\pi_{w,1}) \otimes \pi_{w,2}) + \mathrm{Mant}_n^{(n-1)}(\pi_{w,2} \otimes \pi_{w,1})\Big) \otimes (\otimes_{x|u} \pi_x) \\ = \;& \pi_p \otimes (\mathscr{L}_{F_w}(\pi_{w,1}) + \mathscr{L}_{F_w}(\pi_{w,2})) \;=\; \pi_p \otimes \mathscr{L}_{F_w}(\pi_{w,1} \boxplus \pi_{w,2}) \\ = \;& \pi_p \otimes \mathscr{L}_{F_w}(\pi_w) \;=\; \pi_p \otimes \mathscr{L}_{F_w}(\Pi_w^1). \end{aligned}$$

The last identity uses $\Pi_w^1 \simeq \pi_w$, which follows from the fact that $\Pi_p = BC(\pi_p)$ (cf. §2.3.(ii)).

7.2. **In case m is even.**

Lemma 7.3. *For every $\pi_{H,p} \in \mathrm{Irr}(G_{n-1,1}(\mathbb{Q}_p))$, the following holds in* $\mathrm{Groth}(G(\mathbb{Q}_p) \times W_{F_w})$.

$$\sum_{h=0}^{n-1} \mathrm{Mant}_n^{(h)}(\mathrm{Red}_{n-1,1}^{(h)}(\pi_{H,p})) = \pi_p \otimes (\mathscr{L}_{F_w}(\Pi_{1,w}) - \mathscr{L}_{F_w}(\Pi_{2,w})) \tag{7.5}$$

Proof. The proof is contained in the proof of [Shi11, Thm 6.4.(ii)]. Also see Example 7.5 below. □

Remark 7.4. This is an amazing identity. Lemmas 7.1 and 7.3 demonstrate how the representations in different Newton polygon strata add up to the expected Galois representation, even in the endoscopic case.

Again by Corollary 5.14 and Theorem 6.11,

$$\pi_p \otimes R(\Pi)|_{W_{F_w}} \sim \frac{1}{2} \sum_{h=0}^{n-1} \mathrm{Mant}_n^{(h)} \left(\mathrm{Red}_n^{(h)}(\pi_p) \pm \mathrm{Red}_{n-1,1}^{(h)}(\pi_{H,p})\right),$$

where the sign depends on e_1 and e_2 in the cited theorem. The equality (6.6) and Lemmas 7.1 and 7.3 identify the right hand side with

$$\begin{aligned}
& \pi_p \otimes \left(\frac{1}{2} \left(\mathscr{L}_{F_w}(\Pi_{1,w} \boxplus \Pi_{2,w}) \pm (\mathscr{L}_{F_w}(\Pi_{1,w}) - \mathscr{L}_{F_w}(\Pi_{2,w})) \right) \right) \\
= \; & \pi_p \otimes \left(\frac{1}{2} \left(\mathscr{L}_{F_w}(\Pi_{1,w}) + \mathscr{L}_{F_w}(\Pi_{2,w}) \pm (\mathscr{L}_{F_w}(\Pi_{1,w}) - \mathscr{L}_{F_w}(\Pi_{2,w})) \right) \right) \\
= \; & \pi_p \otimes \mathscr{L}_{F_w}(\Pi_{1,w}) \quad \text{or} \quad \pi_p \otimes \mathscr{L}_{F_w}(\Pi_{2,w})
\end{aligned}$$

depending on the sign. This concludes the proof of Theorem 4.6 in the second case.

Example 7.5. Lemma 7.3 can be shown without pain in a particular case as follows. Take π_p as in (7.2) where π_w is as in the setting of Example 6.9. Recall that $\pi_{H,p}$ was given at the start of §6.4. If we write

$$\pi_{H,p} = \pi_{H,p,0} \otimes \pi_{H,w} \otimes (\otimes_{x|u, x\neq w} \pi_{H,x,1} \otimes \pi_{H,x,2})$$

then

$$\pi_x = \pi_{H,x,1} \boxplus \pi_{H,x,2}, \quad \forall x|u.$$

Now the left hand side of (7.5) is identified with the following, with help of (6.3), (6.5), (7.3) and (7.4).

$$\begin{aligned}
& \pi_{p,0} \otimes \left(\mathrm{Mant}_n^{(1)}(LJ(\pi_{w,1}) \otimes \pi_{w,2}) - \mathrm{Mant}_n^{(n-1)}(\pi_{w,2} \otimes \pi_{w,1}) \right) \otimes (\otimes_{x|u} \pi_x) \\
= \; & \pi_p \otimes (\mathscr{L}_{F_w}(\pi_{w,1}) - \mathscr{L}_{F_w}(\pi_{w,2})) \,.
\end{aligned}$$

References

[AC89] James Arthur and Laurent Clozel, *Simple algebras, base change, and the advanced theory of the trace formula*, Annals of Mathematics Studies, vol. 120, Princeton University Press, Princeton, NJ, 1989. MR 1007299

[Art05] James Arthur, *An introduction to the trace formula*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 1–263. MR 2192011

[Bad07] Alexandru Ioan Badulescu, *Jacquet-Langlands et unitarisabilité*, J. Inst. Math. Jussieu **6** (2007), no. 3, 349–379. MR 2329758

[BC11] Joël Bellaïche and Gaëtan Chenevier, *The sign of Galois representations attached to automorphic forms for unitary groups*, Compos. Math. **147** (2011), no. 5, 1337–1352. MR 2834723

[BG14] Kevin Buzzard and Toby Gee, *The conjectural connections between automorphic representations and Galois representations*, Automorphic forms and Galois representations. Vol. 1, London Math. Soc. Lecture Note Ser., vol. 414, Cambridge Univ. Press, Cambridge, 2014, pp. 135–187. MR 3444225

[BH05a] Colin J. Bushnell and Guy Henniart, *The essentially tame local Langlands correspondence. I*, J. Amer. Math. Soc. **18** (2005), no. 3, 685–710. MR 2138141

[BH05b] ———, *The essentially tame local Langlands correspondence. II. Totally ramified representations*, Compos. Math. **141** (2005), no. 4, 979–1011. MR 2148193

[BLGGT14] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor, *Potential automorphy and change of weight*, Ann. of Math. (2) **179** (2014), no. 2, 501–609. MR 3152941

[BLGHT11] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, *A family of Calabi-Yau varieties and potential automorphy II*, P.R.I.M.S. **98** (2011), 29–98.

[Bor79] A. Borel, *Automorphic L-functions*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 27–61. MR 546608 (81m:10056)

[BR89] Don Blasius and Dinakar Ramakrishnan, *Maass forms and Galois representations*, Galois groups over **Q** (Berkeley, CA, 1987), Math. Sci. Res. Inst. Publ., vol. 16, Springer, New York, 1989, pp. 33–77. MR 1012167

[BR94] Don Blasius and Jonathan D. Rogawski, *Zeta functions of Shimura varieties*, Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, pp. 525–571. MR 1265563 (95e:11051)

[Car] Ana Caraiani, *Lecture notes on perfectoid Shimura varieties*, `http://swc.math.arizona.edu/aws/2017/2017CaraianiNotes.pdf`.

[Car86] Henri Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. (4) **19** (1986), no. 3, 409–468. MR 870690 (89c:11083)

[Car90] H. Carayol, *Nonabelian Lubin-Tate theory*, Automorphic forms, Shimura varieties, and L-functions, Vol. II (Ann Arbor, MI, 1988), Perspect. Math., vol. 11, Academic Press, Boston, MA, 1990, pp. 15–39. MR 1044827

[Car12] Ana Caraiani, *Local-global compatibility and the action of monodromy on nearby cycles*, Duke Math. J. **161** (2012), no. 12, 2311–2413. MR 2972460

[Car14] ———, *Monodromy and local-global compatibility for $l = p$*, Algebra Number Theory **8** (2014), no. 7, 1597–1646. MR 3272276

[CH13] Gaëtan Chenevier and Michael Harris, *Construction of automorphic Galois representations, II*, Camb. J. Math. **1** (2013), no. 1, 53–73. MR 3272052

[CHL11] Laurent Clozel, Michael Harris, and Jean-Pierre Labesse, *Construction of automorphic Galois representations, I*, On the stabilization of the trace formula, Stab. Trace Formula Shimura Var. Arith. Appl., vol. 1, Int. Press, Somerville, MA, 2011, pp. 497–527. MR 2856383

[CHLN11] Laurent Clozel, Michael Harris, Jean-Pierre Labesse, and Bao-Châu Ngô (eds.), *On the stabilization of the trace formula*, Stabilization of the Trace Formula, Shimura Varieties, and Arithmetic Applications, vol. 1, International Press, Somerville, MA, 2011. MR 2742611

[CLH16] Ana Caraiani and Bao V. Le Hung, *On the image of complex conjugation in certain Galois representations*, Compos. Math. **152** (2016), no. 7, 1476–1488. MR 3530448

[Clo90] Laurent Clozel, *Motifs et formes automorphes: applications du principe de fonctorialité*, Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 77–159. MR 1044819 (91k:11042)

[Clo91] ———, *Représentations galoisiennes associées aux représentations automorphes autoduales de* $\mathrm{GL}(n)$, Inst. Hautes Études Sci. Publ. Math. (1991), no. 73, 97–145. MR 1114211 (92i:11055)

[Clo13] ———, *Purity reigns supreme*, Int. Math. Res. Not. IMRN (2013), no. 2, 328–346. MR 3010691

[Dat07] J.-F. Dat, *Théorie de Lubin-Tate non-abélienne et représentations elliptiques*, Invent. Math. **169** (2007), no. 1, 75–152. MR 2308851 (2008g:22022)

[Far04] Laurent Fargues, *Cohomologie des espaces de modules de groupes p-divisibles et correspondances de Langlands locales*, Astérisque (2004), no. 291, 1–199, Variétés de Shimura, espaces de Rapoport-Zink et correspondances de Langlands locales. MR 2074714

[Fuj97] Kazuhiro Fujiwara, *Rigid geometry, Lefschetz-Verdier trace formula and Deligne's conjecture*, Invent. Math. **127** (1997), no. 3, 489–533. MR 1431137

[Hai05] Thomas J. Haines, *Introduction to Shimura varieties with bad reduction of parahoric type*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 583–642. MR 2192017

[Har01] Michael Harris, *Local Langlands correspondences and vanishing cycles on Shimura varieties*, European Congress of Mathematics, Vol. I (Barcelona, 2000), Progr. Math., vol. 201, Birkhäuser, Basel, 2001, pp. 407–427. MR 1905332

[Hen00] Guy Henniart, *Une preuve simple des conjectures de Langlands pour* $\mathrm{GL}(n)$ *sur un corps p-adique*, Invent. Math. **139** (2000), no. 2, 439–455. MR 1738446

[Hid04] Haruzo Hida, *p-adic automorphic forms on Shimura varieties*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2004. MR 2055355

[HLTT16] Michael Harris, Kai-Wen Lan, Richard Taylor, and Jack Thorne, *On the rigid cohomology of certain Shimura varieties*, Res. Math. Sci. **3** (2016), Paper No. 37, 308. MR 3565594

[HR12] Thomas J. Haines and Michael Rapoport, *Shimura varieties with $\Gamma_1(p)$-level via Hecke algebra isomorphisms: the Drinfeld case*, Ann. Sci. Éc. Norm. Supér. (4) **45** (2012), no. 5, 719–785 (2013). MR 3053008

[HT01] Michael Harris and Richard Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, NJ, 2001, With an appendix by Vladimir G. Berkovich. MR 1876802 (2002m:11050)

[Igu59] Jun-ichi Igusa, *Kroneckerian model of fields of elliptic modular functions*, Amer. J. Math. **81** (1959), 561–577. MR 0108498

[KM85] Nicholas M. Katz and Barry Mazur, *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies, vol. 108, Princeton University Press, Princeton, NJ, 1985. MR 772569

[Kna94] A. W. Knapp, *Local Langlands correspondence: the Archimedean case*, Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, pp. 393–410. MR 1265560

[Kot86] Robert E. Kottwitz, *Stable trace formula: elliptic singular terms*, Math. Ann. **275** (1986), no. 3, 365–399. MR 858284 (88d:22027)

[Kot90] ———, *Shimura varieties and λ-adic representations*, Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988), Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, pp. 161–209. MR 1044820

[Kot92a] ———, *On the λ-adic representations associated to some simple Shimura varieties*, Invent. Math. **108** (1992), no. 3, 653–665. MR 1163241

[Kot92b] ———, *Points on some Shimura varieties over finite fields*, J. Amer. Math. Soc. **5** (1992), no. 2, 373–444. MR 1124982

[Kud94] Stephen S. Kudla, *The local Langlands correspondence: the non-Archimedean case*, Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, pp. 365–391. MR 1265559

[Lab11] J.-P. Labesse, *Changement de base CM et séries discrètes*, On the stabilization of the trace formula, Stab. Trace Formula Shimura Var. Arith. Appl., vol. 1, Int. Press, Somerville, MA, 2011, pp. 429–470. MR 2856380

[Lan77] R. P. Langlands, *Shimura varieties and the Selberg trace formula*, Canad. J. Math. **29** (1977), no. 6, 1292–1299. MR 0498400

[Lan79] ———, *On the zeta functions of some simple Shimura varieties*, Canad. J. Math. **31** (1979), no. 6, 1121–1216. MR 553157

[Lan89] ———, *On the classification of irreducible representations of real algebraic groups*, Representation theory and harmonic analysis on semisimple Lie groups, Math. Surveys Monogr., vol. 31, Amer. Math. Soc., Providence, RI, 1989, pp. 101–170. MR 1011897

[Lan13] Kai-Wen Lan, *Arithmetic compactifications of PEL-type Shimura varieties*, London Mathematical Society Monographs Series, vol. 36, Princeton University Press, Princeton, NJ, 2013. MR 3186092

[LR92] Robert P. Langlands and Dinakar Ramakrishnan (eds.), *The zeta functions of Picard modular surfaces*, Université de Montréal, Centre de Recherches Mathématiques, Montreal, QC, 1992. MR 1155223

[Man] Elena Mantovan, *The Newton stratification*, in this volume.

[Man05] ———, *On the cohomology of certain PEL-type Shimura varieties*, Duke Math. J. **129** (2005), no. 3, 573–610. MR 2169874

[Man08] ———, *On non-basic Rapoport-Zink spaces*, Ann. Sci. Éc. Norm. Supér. (4) **41** (2008), no. 5, 671–716. MR 2504431

[Man11] ———, *l-adic étale cohomology of PEL type Shimura varieties with non-trivial coefficients*, WIN—women in numbers, Fields Inst. Commun., vol. 60, Amer. Math. Soc., Providence, RI, 2011, pp. 61–83. MR 2777800

[Mil05] J. S. Milne, *Introduction to Shimura varieties*, Harmonic analysis, the trace formula, and Shimura varieties, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 265–378. MR 2192012

[Min11] Alberto Minguez, *Unramified representations of unitary groups*, On the stabilization of the trace formula, Stab. Trace Formula Shimura Var. Arith. Appl., vol. 1, Int. Press, Somerville, MA, 2011, pp. 389–410. MR 2856377

[Mok15] Chung Pang Mok, *Endoscopic classification of representations of quasi-split unitary groups*, Mem. Amer. Math. Soc. **235** (2015), no. 1108, vi+248. MR 3338302

[Mor10] Sophie Morel, *On the cohomology of certain noncompact Shimura varieties*, Annals of Mathematics Studies, vol. 173, Princeton University Press, Princeton, NJ, 2010, With an appendix by Robert Kottwitz. MR 2567740 (2011b:11073)

[Mor16] ———, *Construction de représentations galoisiennes de torsion [d'après Peter Scholze]*, Astérisque (2016), no. 380, Séminaire Bourbaki. Vol. 2014/2015, Exp. No. 1102, 449–473. MR 3522182

[Pas09] Boyer Pascal, *Monodromie du faisceau pervers des cycles évanescents de quelques variétés de Shimura simples*, Invent. Math. **177** (2009), no. 2, 239–280. MR 2511742

[PT15] Stefan Patrikis and Richard Taylor, *Automorphy and irreducibility of some l-adic representations*, Compos. Math. **151** (2015), no. 2, 207–229. MR 3314824

[Rap95] Michael Rapoport, *Non-Archimedean period domains*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 423–434. MR 1403942

[Rap05] ———, *A guide to the reduction modulo p of Shimura varieties*, 2005, Automorphic forms. I, pp. 271–318. MR 2141705 (2006c:11071)

[Rog92] Jonathan D. Rogawski, *Analytic expression for the number of points mod p*, The zeta functions of Picard modular surfaces, Univ. Montréal, Montreal, QC, 1992, pp. 65–109. MR 1155227

[RZ96] M. Rapoport and Th. Zink, *Period spaces for p-divisible groups*, Annals of Mathematics Studies, vol. 141, Princeton University Press, Princeton, NJ, 1996. MR 1393439

[Scha] Peter Scholze, *Perfectoid spaces and their applications*, Proceedings of the ICM 2014.

[Schb] ———, *The local Langlands correspondence for GL_n over p-adic fields, and the cohomology of compact unitary Shimura varieties*, in this volume.

[Sch13a] ———, *The Langlands-Kottwitz approach for some simple Shimura varieties*, Invent. Math. **192** (2013), no. 3, 627–661. MR 3049931

[Sch13b] ———, *The local Langlands correspondence for GL_n over p-adic fields*, Invent. Math. **192** (2013), no. 3, 663–715. MR 3049932

[Sch15] ———, *On torsion in the cohomology of locally symmetric varieties*, Ann. of Math. (2) **182** (2015), no. 3, 945–1066. MR 3418533

[Sch16] ———, *Perfectoid Shimura varieties*, Jpn. J. Math. **11** (2016), no. 1, 15–32. MR 3510678

[Shi09] Sug Woo Shin, *Counting points on Igusa varieties*, Duke Math. J. **146** (2009), no. 3, 509–568. MR 2484281

[Shi10] ———, *A stable trace formula for Igusa varieties*, J. Inst. Math. Jussieu **9** (2010), no. 4, 847–895. MR 2684263

[Shi11] S. W. Shin, *Galois representations arising from some compact Shimura varieties*, Ann. of Math. (2) **173** (2011), no. 3, 1645–1741. MR 2800722

[Shi12] Sug Woo Shin, *On the cohomology of Rapoport-Zink spaces of EL-type*, Amer. J. Math. **134** (2012), no. 2, 407–452. MR 2905002

[Sor] Claus Sorensen, *A patching lemma*, Stabilization of the trace formula, Shimura varieties, and arithmetic applications, to appear.

[SS13] Peter Scholze and Sug Woo Shin, *On the cohomology of compact unitary group Shimura varieties at ramified split places*, J. Amer. Math. Soc. **26** (2013), no. 1, 261–294. MR 2983012

[Str05] Matthias Strauch, *On the Jacquet-Langlands correspondence in the cohomology of the Lubin-Tate deformation tower*, Astérisque (2005), no. 298, 391–410, Automorphic forms. I. MR 2141708

[Tat79] J. Tate, *Number theoretic background*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 3–26. MR 546607 (80m:12009)

[Tay04] Richard Taylor, *Galois representations*, Ann. Fac. Sci. Toulouse Math. (6) **13** (2004), no. 1, 73–119. MR 2060030

[TY07] Richard Taylor and Teruyoshi Yoshida, *Compatibility of local and global Langlands correspondences*, J. Amer. Math. Soc. **20** (2007), no. 2, 467–493. MR 2276777 (2007k:11193)

[Var07] Yakov Varshavsky, *Lefschetz-Verdier trace formula and a generalization of a theorem of Fujiwara*, Geom. Funct. Anal. **17** (2007), no. 1, 271–319. MR 2306659

[Wei16] Jared Weinstein, *Reciprocity laws and Galois representations: recent breakthroughs*, Bull. Amer. Math. Soc. (N.S.) **53** (2016), no. 1, 1–39. MR 3403079