# INTRODUCTION TO THE LANGLANDS–KOTTWITZ METHOD

YIHANG ZHU

## 1. Introduction

This chapter is a survey of Kottwitz's two papers [Kot92] and [Kot90]. The first step towards understanding the (étale) cohomology of Shimura varieties is to study the trace of a Hecke operator away from $p$ composed with a power of the Frobenius at $p$, where $p$ is a prime of good reduction. In the paper [Kot92], Kottwitz obtains a formula for this trace, for PEL Shimura varieties (which are assumed to be compact and not of type D). Kottwitz's formula involves orbital integrals away from $p$ and twisted orbital integrals at $p$. We shall refer to it as the *unstabilized Langlands–Kottwitz formula.* The shape of this formula is reminiscent of another formula, which stems from automorphic representation theory rather than Shimura varieties, namely the elliptic part of the Arthur–Selberg trace formula. In fact, the main idea of the so-called Langlands–Kottwitz method is that the Langlands–Kottwitz formula could be compared with the Arthur–Selberg trace formula, and the comparison should eventually reflect the reciprocity between Galois representations and automorphic representations.

However, apart from very special situations, the unstabilized Langlands–Kottwitz formula is not directly comparable with the Arthur–Selberg trace formula. One has to *stabilize* both the formulas in order to compare them. Recall from [Har11] that the elliptic part of the Arthur–Selberg trace formula can be stabilized, in the sense that it can be rewritten in terms of stable orbital integrals on the endoscopic groups. In a similar sense, the unstabilized Langlands–Kottwitz formula could also be stabilized, and this is the main result of the paper [Kot90]. The stabilization procedure for the Langlands–Kottwitz formula in [Kot90] is analogous to the stabilization procedure for the elliptic part of the Arthur–Selberg trace formula, to which Kottwitz also made major contributions (see [Kot84] [Kot86], and also the exposition in [Har11]). At the same time there are also important differences between the two stabilization procedures, which mainly happen at the places $p$ and $\infty$.

After the stabilization, the Langlands–Kottwitz formula is expressed in terms of the elliptic parts of the stable Arthur–Selberg trace formulas for the endoscopic groups. See Theorem 6.1.2 for the precise statement. In the special case of unitary Shimura varieties, the stabilized formula is stated in [CHL11, (4.3.5)], which is among the main inputs to the chapter [CHL11]. The present chapter, to some extent, could be viewed as an extended explanation of the formula [CHL11, (4.3.5)].

In our exposition of the unstabilized Langlands–Kottwitz formula, we shall specialize to the case of unitary PEL Shimura varieties. More precisely, we only consider the unitary similitude group of a global Hermitian space over a field, as opposed to over a more general division algebra. This is already enough for the chapter [CHL11], and our main purpose for restricting to this special case is to gain some concreteness and to simplify certain proofs. We hope that nothing essential in the general PEL case is lost. Our exposition of the stabilization procedure is, however, for a general Shimura datum, under some simplifying assumptions of a group theoretic nature.

In the paper [Kot90], Kottwitz assumes several hypotheses in local harmonic analysis, including the Fundamental Lemma for Endoscopy, the Fundamental Lemma for Unstable Base Change, and the Langlands–Shelstad Transfer Conjecture. These hypotheses are all proved theorems now, and we shall point out the references when they are invoked in the argument.

Here is the organization of this chapter. In §2 we describe the problem. In §3 - §5 we discuss the unstabilized Langlands–Kottwitz formula in [Kot92]. In §6 we discuss the stabilization procedure in [Kot90]. In §7 we make some brief remarks on developments following Kottwitz's work.

Last updated: February 18, 2020.

**Acknowledgment** I would like to thank the editors of this volume for inviting me to contribute, and the anonymous referee for many helpful remarks, suggestions, and corrections. I would like to thank Mark Kisin for introducing me to Shimura varieties. I also thank Tasho Kaletha, Kai-Wen Lan, Sophie Morel, Sug Woo Shin, and Wei Zhang for explaining to me many subtle details about Shimura varieties and endoscopy over the years. Last but not least, I would like to thank the authors and editors of the first volume of this book series, which served as a great guide for me, when I was entering the field of automorphic forms and trace formula as a student.

## 2. Description of the problem

2.1. **The unitary PEL and Shimura data.** We follow the setting of [CHL11] but work slightly more generally. Let $F$ be a totally real field and let $\mathcal{K}/F$ be a totally imaginary quadratic extension.[1] Denote by $\mathbf{c}$ the nontrivial element in $\mathrm{Gal}(\mathcal{K}/F)$. Let $(V,[\cdot,\cdot])$ be a non-degenerate Hermitian space for $\mathcal{K}/F$, i.e., $V$ is a finite dimensional $\mathcal{K}$-vector space and $[\cdot,\cdot]$ is a $\mathbf{c}$-sesquilinear form satisfying $[x,y]=\mathbf{c}([y,x])$ for all $x,y\in V$. We assume that $(V,[\cdot,\cdot])$ is *anisotropic*, which means $[x,x]\neq 0$ for all $x\in V-\{0\}$. For instance, the Hermitian spaces $V_1$ and $V_2$ considered in [CHL11, §3] are both anisotropic.

Fix an element $\epsilon\in\mathcal{K}$ such that $\mathbf{c}(\epsilon)=-\epsilon$. We define

$$\langle\cdot,\cdot\rangle := \mathrm{Tr}_{\mathcal{K}/\mathbb{Q}}(\epsilon[\cdot,\cdot]). \tag{2.1.1}$$

Then $\langle\cdot,\cdot\rangle$ is a $\mathbb{Q}$-bilinear $\mathbb{Q}$-valued non-degenerate alternating form on $V$. It satisfies

$$\langle ax,y\rangle = \langle x,\mathbf{c}(a)y\rangle,\ \forall x,y\in V, a\in\mathcal{K}.$$

It can be easily checked that $[\cdot,\cdot]$ is in turn uniquely determined by $\langle\cdot,\cdot\rangle$, via (2.1.1).

Finally we choose $h:\mathbb{C}\to\mathrm{End}_{\mathbb{R},\mathcal{K}}(V\otimes_{\mathbb{Q}}\mathbb{R})$ to be an $\mathbb{R}$-algebra morphism such that $\langle h(z)x,y\rangle=\langle x,h(\bar z)y\rangle$ for all $x,y\in V\otimes_{\mathbb{Q}}\mathbb{R}$ and such that the symmetric bilinear form $(x,y)\mapsto\langle x,h(i)y\rangle$ on $V\otimes_{\mathbb{Q}}\mathbb{R}$ is positive definite. Then in the terminology of [GN, Definition 3.4.1], the datum

$$(B:=\mathcal{K},*:=\mathbf{c},V,\langle\cdot,\cdot\rangle,h)$$

is a *rational PE-structure*. Our datum is in fact a special case of the unitary PEL data (in a broader sense) considered in [Nic], and of course they are all special cases of the PEL data considered in [GN], following [Kot92].

We write $C$ for the algebra $\mathrm{End}_{\mathcal{K}}(V)$ and write $*$ for the involution on $C$ given by the alternating form $\langle\cdot,\cdot\rangle$, characterized by

$$\langle cx,y\rangle=\langle x,c^*y\rangle,\ \forall x,y\in V,c\in C.$$

The reductive group $G$ over $\mathbb{Q}$ associated to the above data as in [GN, §3.5] and [Roz, §1.1] is the same as the unitary similitude group $\mathrm{GU}(V,[\cdot,\cdot])$ defined in [CHL11, §1.1]. By definition, for any $\mathbb{Q}$-algebra $R$, we have

$$\begin{aligned} G(R) &= \left\{g\in\mathrm{GL}_{\mathcal{K}}(V\otimes_{\mathbb{Q}}R)|\exists\nu(g)\in R^\times\ \forall x,y\in V,\ [gx,gy]=\nu(g)[x,y]\right\}\\ &= \left\{g\in\mathrm{GL}_{\mathcal{K}}(V\otimes_{\mathbb{Q}}R)|\exists\nu(g)\in R^\times\ \forall x,y\in V,\ \langle gx,gy\rangle=\nu(g)\langle x,y\rangle\right\}\\ &= \left\{g\in(C\otimes_{\mathbb{Q}}R)^\times|\exists\nu(g)\in R^\times, g^*g=\nu(g)\right\}. \end{aligned}$$

Moreover, the factors of similitude $\nu(g)$ in the above three descriptions are the same, and we have a $\mathbb{Q}$-morphism $\nu:G\to\mathbb{G}_m$.

Let $E$ be the reflex field of the PE-structure (see [GN, §3.4]).[2] The $\mathbb{R}$-algebra morphism $h$ induces a homomorphism between $\mathbb{R}$-algebraic groups

$$\mathbb{S}=\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\,\mathbb{G}_m\longrightarrow G_{\mathbb{R}},$$

which we still denote by $h$. Let $X$ be the $G(\mathbb{R})$-conjugacy class of

$$h^{-1}:\mathbb{S}\longrightarrow G_{\mathbb{R}}.$$

[1] Our $\mathcal{K}$ plays the same role as the simple $\mathbb{Q}$-algebra $B$ in [Kot92]. In [Kot92], the center of $B$ is denoted by $F$, which for us is of course still $B=\mathcal{K}$. Note our different usage of the notation $F$.

[2] This notation is not to be confused with the imaginary quadratic field $E$ in [CHL11].

Then $(G, X)$ is a Shimura datum whose reflex field is $E$.[3]

*Remark* 2.1.1. We record two important facts about the reductive group $G$, which are easy to check and simplify the theory greatly.

(1) The derived subgroup $G^{\mathrm{der}}$ is simply connected.
(2) The center $Z_G$ satisfies the condition that its maximal $\mathbb{R}$-split subtorus is $\mathbb{Q}$-split. (This subtorus is $\mathbb{G}_m$.)

The following lemma also plays an important role in [Kot92]:

**Lemma 2.1.2** ([Kot92, Lemma 7.1])**.** *Let $Q$ be any algebraically closed field of characteristic zero. Then two elements in $G(Q)$ are conjugate if and only if they have the same image under $\nu$ and their images in $\mathrm{GL}_{\mathcal{K}}(V \otimes_{\mathbb{Q}} Q) \cong \mathrm{GL}_{\dim_{\mathcal{K}} V}(Q) \times \mathrm{GL}_{\dim_{\mathcal{K}} V}(Q)$ are conjugate.* □

We next fix *p-integral data* as in [Roz, §1.2]. Let $p$ be a rational prime that is unramified in $\mathcal{K}$. We need to choose a $\mathbb{Z}_{(p)}$-order $\mathcal{O}_B$ in $B$ and a $\mathbb{Z}_p$-lattice $\Lambda$ in $V_{\mathbb{Q}_p}$ such that $\Lambda$ is stable under $\mathcal{O}_B$ and self-dual under $\langle\cdot,\cdot\rangle$. We take $\mathcal{O}_B := \mathcal{O}_{\mathcal{K}} \otimes \mathbb{Z}_{(p)}$, and fix the choice of $\Lambda$.

Of course the existence of $\Lambda$ is a non-trivial assumption on the prime $p$, and in particular implies that $G_{\mathbb{Q}_p}$ is unramified. More precisely, as in [Roz], the choice of $\Lambda$ determines a reductive model $\mathcal{G}$ over $\mathbb{Z}_p$ of $G_{\mathbb{Q}_p}$ and a hyperspecial subgroup $K_p := \mathcal{G}(\mathbb{Z}_p)$ of $G(\mathbb{Q}_p)$.

2.2. **The integral models.** Fix a prime $\mathfrak{p}$ of $E$ above $p$. Given the above data, we obtain a smooth quasi-projective scheme $S_{K^p}$ over $\mathcal{O}_{E,(\mathfrak{p})}$ for each small enough compact open subgroup $K^p$ of $G(\mathbb{A}_f^p)$. (In fact, the choice of $\Lambda$ is not needed for the construction of $S_{K^p}$.) Roughly speaking, $S_{K^p}$ parametrizes $\mathbb{Z}_{(p)}$-polarized projective abelian schemes with an action of $\mathcal{O}_B$ respecting the determinant condition, equipped with $K^p$-level structure, up to prime-to-$p$ isogeny. This is the topic of [Roz]. Moreover, our assumption that $(V, [\cdot,\cdot])$ is anisotropic implies that $G$ is anisotropic modulo center, which further implies that $S_{K^p}$ is projective over $\mathcal{O}_{E,(\mathfrak{p})}$. See the Introduction of this volume for a general discussion of the issue of projectivity of the integral models. In the following, all compact open subgroups $K^p$ of $G(\mathbb{A}_f^p)$ that appear are understood to be small enough.

When $K^p$ varies, we obtain a natural tower $\varprojlim_{K^p} S_{K^p,E}$ with finite étale transition maps. The group $G(\mathbb{A}_f^p)$ acts on the tower by acting on the level structures in the moduli problem. Thus for any $g \in G(\mathbb{A}_f^p)$ and any compact open subgroups $K_1^p, K_2^p \subset G(\mathbb{A}_f^p)$ satisfying $g^{-1}K_1^p g \subset K_2^p$, we have a finite étale morphism

$$[\cdot g] = [\cdot g]_{K_1^p,K_2^p} : S_{K_1^p} \longrightarrow S_{K_2^p}.$$

These morphisms are called the *Hecke operators*. In particular the transition maps in the tower $\varprojlim_{K^p} S_{K^p,E}$ are the Hecke operators $[\cdot 1]_{K_1^p,K_2^p}$.

For any reductive group $H$ over $\mathbb{Q}$, denote by $\ker^1(\mathbb{Q}, H)$ the pointed set

$$\ker(\mathbf{H}^1(\mathbb{Q}, H) \to \mathbf{H}^1(\mathbb{A}, H)).$$

Recall from [Roz, §7] that the generic fiber of the tower $\varprojlim_{K^p} S_{K^p,E}$, as a tower of schemes over $E$, is identified with the disjoint union of $\left|\ker^1(\mathbb{Q}, G)\right|$ copies of the tower $\varprojlim_{K^p} \mathrm{Sh}_{K^pK_p}(G, X)$ of canonical models for the Shimura datum $(G, X) = (G, h^{-1})$.[4] The reader is referred to [GN] or [Mil05] for the theory of canonical models. This identification is equivariant with respect to the $G(\mathbb{A}_f^p)$-actions on both sides. On $\mathbb{C}$-points, the Hecke operator

$$[\cdot g] : \mathrm{Sh}_{K_1^pK_p}(\mathbb{C}) \longrightarrow \mathrm{Sh}_{K_2^pK_p}(\mathbb{C})$$

where $g, K_1^p, K_2^p$ are as above, is the map

$$G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K_1^pK_p) \longrightarrow G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K_2^pK_p)$$
$$(x, h) \mapsto (x, hg).$$

[3]To be more precise, when $G(\mathbb{R})$ is compact modulo center, which is possible in our setting, the pair $(G, X)$ would violate the axiom [Del79, (2.1.1.3)] for a Shimura datum. However for our purpose this does not matter too much. One may actually use the PEL datum and the associated moduli problem to replace the theory of Shimura varieties in such a case.

[4]When $G(\mathbb{R})$ is compact modulo center, the pair $(G, X)$ is not a Shimura datum, and we understand $\mathrm{Sh}_{K^pK_p}(G, X)$ simply as the (zero-dimensional) locally symmetric space $G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K^pK_p)$. See footnote 3.

See [Nic] for more discussion on the $\mathbb{C}$-points of unitary Shimura varieties and Hecke operators between them.

*Remark* 2.2.1. There are different conventions in the literature on the notion of canonical models, and they are related by replacing $h$ by $h^{-1}$. Our convention here, which is the same as that in [CHL11], follows [Kot92]. Note that there is a mistake in [Kot90] which switches $h$ and $h^{-1}$, see the remark at the end of [Kot92], cf. also [CHL11, §3].

2.3. **$\lambda$-adic automorphic sheaves.** Fix a finite dimensional representation $\xi$ of $G(\mathbb{C})$. It descends to an algebraic representation over a number field $L(\xi)$. Let $\lambda$ be a place of $L(\xi)$ coprime to $p$. Let $\ell := \lambda|_{\mathbb{Q}}$. Then $\xi$ induces a $\lambda$-adic representation of $G(\mathbb{Q}_\ell)$. Using the action of $G(\mathbb{Q}_\ell) \subset G(\mathbb{A}_f^p)$ on the tower $\varprojlim_{K^p} S_{K^p}$, we obtain a $\lambda$-*adic automorphic sheaf on the tower.* More precisely, we obtain the following datum:

(1) For each $K^p$ (small enough), we have a $\lambda$-adic sheaf $\mathscr{F}_{K^p}$ on $S_{K^p}$.
(2) For each Hecke operator $[\cdot g] : S_{K_1^p} \to S_{K_2^p}$, where $K_1^p, K_2^p$ are as in (1) and $g \in G(\mathbb{A}_f^p)$ is such that $g^{-1} K_1^p g \subset K_2^p$, we have a canonical identification $\eta : [\cdot g]^* \mathscr{F}_{K_2^p} \xrightarrow{\sim} \mathscr{F}_{K_1^p}$, satisfying certain compatibility conditions.

For more details of the construction see [Kot92, §6], cf. also [Pin92a, §5.1], [HT01, §III.2], [Car86, §2.1], in which the constructions can be easily adapted to the $G(\mathbb{A}_f^p)$-tower of integral models here.

*Remark* 2.3.1. In order to construct $\mathscr{F}_{K^p}$ for a fixed $K^p$, it is crucial that the sub-tower $\varprojlim_{K'^p, K'^p \subset K^p} S_{K'^p}$ over $S_{K^p}$ has Galois group $K^p$, or equivalently, that for any open normal subgroup $K'^p$ of $K^p$ the Galois covering $S_{K'^p} \to S_{K^p}$ has Galois group $K^p/K'^p$. This is indeed guaranteed by property (2) in Remark 2.1.1, when $K^p$ is small enough. This point is not emphasized explicitly in [Kot92], but in fact all the PEL Shimura data of types A and C treated in [Kot92] satisfy property (2) in Remark 2.1.1. (This holds even more generally for all Hodge-type Shimura data, but may fail for abelian-type Shimura data.)

2.4. **The goal.** Let $\mathbb{F}_q = \mathbb{F}_{p^r}$ be the residue field of $E$ at $\mathfrak{p}$. Fix a small enough compact open subgroup $K^p \subset G(\mathbb{A}_f^p)$. Let the symbol $k$ denote either $\overline{\mathbb{F}}_q$ or $\overline{E}$. Consider the $\lambda$-adic cohomology groups

$$\mathbf{H}^*_{k,K^p} := \mathbf{H}^*(S_{K^p,k}, \mathscr{F}_{K^p}).$$

Of course $\mathbf{H}^*_{\overline{E},K^p}$ is the direct sum of $\left|\ker^1(\mathbb{Q}, G)\right|$ copies of the analogous cohomology group of $\mathrm{Sh}_{K^pK_p}(G,X)_{\overline{E}}$. The $G(\mathbb{A}_f^p)$ action on the tower $\varprojlim_{K'^p, K'^p \subset K^p} S_{K'^p}$ endows $\mathbf{H}^*_{k,K^p}$ with the structure of a module over the Hecke algebra

$$\mathcal{H}_{K^p} := \mathcal{H}(G(\mathbb{A}_f^p)//K^p, L(\xi))$$

consisting of $L(\xi)$-valued $K^p$-biinvariant locally constant compactly supported functions on $G(\mathbb{A}_f^p)$. See §2.5 below for a more precise description of the Hecke action.[5] Moreover $\mathrm{Gal}(\overline{E}/E)$ (respectively $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$) acts on $\mathbf{H}^*_{\overline{E},K^p}$ (respectively $\mathbf{H}^*_{\overline{\mathbb{F}}_q,K^p}$) by transport of structure, and this action commutes with the action of $\mathcal{H}_{K^p}$.

Since $S_{K^p}$ is proper smooth over $\mathcal{O}_{E,(\mathfrak{p})}$, we know that the $\mathrm{Gal}(\overline{E}/E)$-action on $\mathbf{H}^*_{\overline{E},K^p}$ is unramified at $\mathfrak{p}$, and we have a canonical isomorphism

$$\mathbf{H}^*_{\overline{E},K^p} \cong \mathbf{H}^*_{\overline{\mathbb{F}}_q,K^p}$$

as $(\mathcal{H}_{K^p} \times \mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q))$-modules.

The main goal of this chapter is to understand the Euler characteristic

$$\begin{aligned} T(j, f^{p,\infty}) &:= \sum_i (-1)^i \operatorname{Tr}(\Phi_q^j \times f^{p,\infty}, \mathbf{H}^i_{\overline{E},K^p}) \\ &= \sum_i (-1)^i \operatorname{Tr}(\Phi_q^j \times f^{p,\infty}, \mathbf{H}^i_{\overline{\mathbb{F}}_q,K^p}) \in L(\xi)_\lambda, \end{aligned}$$

where

- $\Phi_q$ is the **geometric** $q$-Frobenius in $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$

[5] Strictly speaking one also needs to choose a Haar measure on $G(\mathbb{A}_f^p)$ to pin down the action of $\mathcal{H}_{K^p}$. We always choose the one giving volume 1 to $K^p$.

- $f^{p,\infty} \in \mathcal{H}_{K^p}$
- $j \in \mathbb{N}$ is large enough.

Without loss of generality, we may assume $f^{p,\infty}$ is the characteristic function of $K^p g^{-1} K^p$ for some $g \in G(\mathbb{A}_f^p)$, as such functions span $\mathcal{H}_{K^p}$. We write $T(j,g)$ for $T(j, f^{p,\infty})$ in this case.

2.5. **Applying the Lefschetz–Verdier trace formula.** We shall apply the Lefschetz–Verdier trace formula (or rather its simplified version conjectured by Deligne) to compute $T(j,g)$. Denote

$$K_g^p := K^p \cap gK^pg^{-1}.$$

Consider the cohomological correspondence

$$u(j,g) : [\cdot 1]^*(\Phi^j)^*\mathscr{F}_{K^p} \cong [\cdot 1]^*\mathscr{F}_{K^p} \xrightarrow[\eta]{\sim} \mathscr{F}_{K_g^p} \xrightarrow[\eta^{-1}]{\sim} [\cdot g]^*\mathscr{F}_{K^p}$$

supported on the geometric correspondence

(2.5.1)
$$\begin{tikzcd} & S_{K_g^p,\overline{\mathbb{F}}_q} \arrow[dl, "a:={[\cdot g]}"'] \arrow[dr, "c:=\Phi^j\circ{[\cdot 1]}_{K_g^p,K^p}"] & \\ S_{K^p,\overline{\mathbb{F}}_q} & & S_{K^p,\overline{\mathbb{F}}_q} \end{tikzcd}$$

Here $\Phi$ denotes the $q$-Frobenius endomorphism of the variety $S_{K^p,\overline{\mathbb{F}}_q}$, and we have used the natural isomorphism $\mathscr{F}_{K^p} \cong (\Phi^j)^*\mathscr{F}_{K^p}$ in the above definition of $u(j,g)$. We have an induced endomorphism $u(j,g)_*$ on $\mathbf{H}^*_{\overline{\mathbb{F}}_q,K^p}$. See [Pin92b, §1] for more details on this formalism.[6]

By definition, when $j=0$, the endomorphism $u(j,g)_*$ gives the action of $f^{p,\infty} = 1_{K^pg^{-1}K^p} \in \mathcal{H}_{K^p}$ on $\mathbf{H}^*_{\overline{\mathbb{F}}_q,K^p}$.

**Lemma 2.5.1.** *$T(j,g)$ is equal to the Euler characteristic*

$$\operatorname{Tr} u(j,g)_* := \sum_i (-1)^i \operatorname{Tr}\left(u(j,g)_*, \mathbf{H}^*_{\overline{\mathbb{F}}_q,K^p}\right)$$

*of $u(j,g)_*$.*

*Proof.* This follows from the definition of the $\mathcal{H}_{K^p}$-module structure on $\mathbf{H}^*(\overline{\mathbb{F}}_q, K^p)$ mentioned above, and the compatibility between the geometric Frobenius $\Phi_q \in \operatorname{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$ and the Frobenius endomorphism $\Phi$ of $S_{K^p,\overline{\mathbb{F}}_q}$ acting on $\lambda$-adic cohomology. □

To compute $\operatorname{Tr} u(j,g)_*$, let $\operatorname{Fix}(j,g)$ denote the (geometric) fixed points of (2.5.1), namely

$$\operatorname{Fix}(j,g) = \left\{x' \in S_{K_g^p}(\overline{\mathbb{F}}_q) :\ a(x') = c(x')\right\}.$$

By general consideration (see [Pin92b, Lemma 7.1.2]) we know that $\operatorname{Fix}(j,g)$ is finite if $j$ is large enough. Recall that the general Lefschetz–Verdier trace formula asserts:

$$\operatorname{Tr} u(j,g)_* = \sum_{x'\in\operatorname{Fix}(j,g)} \operatorname{LT}_{x'},$$

where $\operatorname{LT}_{x'}$ are abstractly defined *local terms* and are in general difficult to compute. According to a conjecture of Deligne, when $j$ is large enough, we have a similar formula:

$$\operatorname{Tr} u(j,g)_* = \sum_{x'\in\operatorname{Fix}(j,g)} \operatorname{LT}^{naive}_{x'}, \tag{2.5.2}$$

where $\operatorname{LT}^{naive}_{x'}$ are the so-called *naive local terms*, which are much easier to compute. Deligne's conjecture has been proved by Pink [Pin92b] in special cases (which already suffices for us) and by Fujiwara [Fuj97] in general (cf. also [Var05]). It is this version (2.5.2) of the Lefschetz–Verdier trace formula that we shall use.

[6] Note that in [Pin92b] $a$ would be denoted by a right arrow and $c$ would be denoted by a left arrow.

In our case, for $x' \in \mathrm{Fix}(j,g)$ and $x := a(x')$, the naive local term $\mathrm{LT}^{naive}_{x'}$ is given by the trace on the stalk $\mathscr{F}_{K^p,x}$ of the endomorphism

$$\mathscr{F}_{K^p,x} \cong (c^*\mathscr{F}_{K^p})_{x'} \xrightarrow{u(j,g)} (a^*\mathscr{F}_{K^p})_{x'} \cong \mathscr{F}_{K^p,x}.$$

See [Kot92, §16] and [Pin92b, §1.5]. This turns out to be quite easy to understand once we have a good parametrization of $\mathrm{Fix}(j,g)$. Hence the main problem is to understand $\mathrm{Fix}(j,g)$, or to "count its points".

*Example* 2.5.2. When $g = 1$, we have $\mathrm{Fix}(j,g) = S_{K^p}(\mathbb{F}_{q^j})$.

2.6. **The strategy for understanding** $\mathrm{Fix}(j,g)$**.** Recall from [Roz, §3.1] or [Kot92, §5] that the $\overline{\mathbb{F}}_q$-points on $S_{K^p}$ correspond to isomorphism classes in the $\mathbb{Z}_{(p)}$-linear category $\mathscr{C}_{\overline{\mathbb{F}}_q}$, where

- the objects in $\mathscr{C}_{\overline{\mathbb{F}}_q}$ are quadruples
$$(\bar{A}, \lambda, \iota, \bar{\eta})$$
where $(\bar{A}, \lambda, \iota)$ is a $\mathbb{Z}_{(p)}$-polarized abelian variety over $\overline{\mathbb{F}}_q$ with an action of $\mathcal{O}_B = \mathcal{O}_\mathcal{K} \otimes \mathbb{Z}_{(p)}$ satisfying the compatibility between $\mathbf{c}$ and the Rosati involution for $\lambda$, and satisfying the determinant condition with respect to $h$, and $\bar{\eta}$ is a $K^p$-level structure.
- the morphisms from $(\bar{A}, \lambda, \iota, \bar{\eta})$ to $(\bar{A}_1, \lambda_1, \iota_1, \bar{\eta}_1)$ are the elements of $\mathrm{Hom}(A, B) \otimes_\mathbb{Z} \mathbb{Z}_{(p)}$ that preserve the additional structures (where we only require the polarizations to be preserved up to a scalar).

To compare with [Roz, §3.1], the groupoid $\mathcal{F}_{K^p}(\overline{\mathbb{F}}_q)$ in that chapter is obtained by considering all objects in $\mathscr{C}_{\overline{\mathbb{F}}_q}$ and only isomorphisms between them.

Write $q^j$ as $p^n$. For a moment we suppose $g = 1$. Then by Example 2.5.2, we know that an object $(\bar{A}, \lambda, \iota, \bar{\eta}) \in \mathscr{C}_{\overline{\mathbb{F}}_q}$ represents a point in $\mathrm{Fix}(j,g)$ if and only if $(\bar{A}, \lambda, \iota, \bar{\eta})$ is "defined over $\mathbb{F}_{p^n}$" in a suitable sense. In this case, by forgetting $\bar{\eta}$ we obtain a triple $(\bar{A}, \lambda, \iota)$, which should be "defined over $\mathbb{F}_{p^n}$" in a suitable sense.

Now let $g$ be general. If $(\bar{A}, \lambda, \iota, \bar{\eta})$ represents a point in $\mathrm{Fix}(j,g)$, we can still equip $(\bar{A}, \lambda, \iota)$ with a "virtual $\mathbb{F}_{p^n}$-structure". (The precise definition is to be given later.) We will define a $\mathbb{Q}$-linear category $\mathscr{V}_{n,\mathcal{K},pol}$ of such triples $(\bar{A}, \lambda, \iota)$ equipped with a virtual $\mathbb{F}_{p^n}$-structure. Thus we get a well-defined map $\mathbf{f}$ from $\mathrm{Fix}(j,g)$ to the set $\mathscr{I}$ of isomorphism classes in $\mathscr{V}_{n,\mathcal{K},pol}$. In order to understand $\mathrm{Fix}(j,g)$, it suffices to understand the subset $\mathrm{im}(\mathbf{f})$ of $\mathscr{I}$, and to study the fibers of $\mathbf{f}$.

We carry out the strategy outlined above in §3 - §5, arriving at the unstabilized Langlands–Kottwitz formula for $T(j,g)$ in Theorem 5.3.1.

## 3. Honda-Tate theory for virtual abelian varieties

### 3.1. Virtual abelian varieties with a polarizable condition.

**Definition 3.1.1.** Let $n \in \mathbb{N}$ and let $\sigma$ be the arithmetic $p$-Frobenius. Let $\mathscr{A}_{\overline{\mathbb{F}}_p}$ be the category of abelian varieties over $\overline{\mathbb{F}}_p$. Define the $\mathbb{Z}_{(p)}$-linear category $\mathscr{V}^p_n$ of *virtual abelian varieties over* $\mathbb{F}_{p^n}$ *up to prime-to-p isogeny* as follows:

- the objects are pairs $A = (\bar{A}, u)$, where $\bar{A}$ is an object of $\mathscr{A}_{\overline{\mathbb{F}}_p} \otimes \mathbb{Z}_{(p)}$ and $u$ is an isomorphism $\sigma^n(\bar{A}) \to \bar{A}$ in $\mathscr{A}_{\overline{\mathbb{F}}_p} \otimes \mathbb{Z}_{(p)}$. Here $\sigma^n(\bar{A})$ denotes the base change of $\bar{A}$ along the map $\sigma^n : \overline{\mathbb{F}}_p \to \overline{\mathbb{F}}_p$.
- the morphisms from $(A_1, u_1)$ to $(A_2, u_2)$ are morphisms $f : A_1 \to A_2$ in $\mathscr{A}_{\overline{\mathbb{F}}_p} \otimes \mathbb{Z}_{(p)}$ such that $fu_1 = u_2\sigma^n(f)$.

For $A = (\bar{A}, u) \in \mathscr{V}^p_n$, define its *Frobenius endomorphism*

$$\pi_A \in \mathrm{End}_{\mathscr{A}_{\overline{\mathbb{F}}_p} \otimes \mathbb{Z}_{(p)}}(\bar{A}) = \mathrm{End}(\bar{A}) \otimes \mathbb{Z}_{(p)}$$

to be the composition of the $p^n$-Frobenius homomorphism $\bar{A} \to \sigma^n(\bar{A})$ with $u$. In particular, $\mathrm{End}_{\mathscr{V}^p_n}(A)$ is the centralizer of $\pi_A$ in $\mathrm{End}(\bar{A}) \otimes \mathbb{Z}_{(p)}$.

**Definition 3.1.2.** Let $\mathscr{V}_n$ be the isogeny category $\mathscr{V}^p_n \otimes_{\mathbb{Z}_{(p)}} \mathbb{Q}$. For $A, B \in \mathscr{V}_n$, we write $\mathrm{End}(A)_\mathbb{Q}$ and $\mathrm{Hom}(A,B)_\mathbb{Q}$ for the $\mathbb{Q}$-vector spaces $\mathrm{End}_{\mathscr{V}_n}(A)$ and $\mathrm{Hom}_{\mathscr{V}_n}(A,B)$ respectively. Denote by $\mathbb{Q}[\pi_A]$ the $\mathbb{Q}$-subalgebra of $\mathrm{End}(A)_\mathbb{Q}$ generated by $\pi_A$.

**Definition 3.1.3.** Let $c \in \mathbb{Q}$ and $A \in \mathscr{V}_n$. By a *$c$-polarization of $A$*, we mean a $\mathbb{Q}$-polarization $\lambda$ of $\bar{A}$ such that $\pi_A^* \lambda = c\lambda$. This only exists when $c > 0$. If there exists a $c$-polarization of $A$ we say $A$ is *$c$-polarizable.* Let $\mathscr{V}_{n,c}$ be the full subcategory of $\mathscr{V}_n$ consisting of $c$-polarizable objects.

*Remark* 3.1.4. Let $A$ be an abelian variety over $\mathbb{F}_{p^n}$. Then $A$ defines an object of $\mathscr{V}_n$, by taking $\bar{A}$ to be $A_{\overline{\mathbb{F}}_p}$ and $u$ to be the canonical isomorphism. For such $A$, a $p^n$-polarization is the same as a polarization $A \to A^\vee$ defined over $\mathbb{F}_{p^n}$.

**Definition 3.1.5.** Let $c \in \mathbb{Q}_{>0}$. By a *$c$-number* we mean an element $\pi \in \overline{\mathbb{Q}}$ whose images under all embeddings $\overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ have non-negative valuations and whose images under all embeddings $\overline{\mathbb{Q}} \to \mathbb{C}$ have absolute value $c^{1/2}$.

*Example* 3.1.6. Suppose $c = p^n$. Recall that a *Weil $p^n$-number* is by definition an **algebraic integer** all of whose complex embeddings have absolute value $p^{n/2}$. Hence a $p^n$-number $\pi$ in the sense of Definition 3.1.5 is a Weil $p^n$-number if and only if $\pi$ is a unit at all places of $\mathbb{Q}(\pi)$ coprime to $p$. (To check this equivalence, use the product formula.) For example, let $p = 7$. Then $7(1+2i)/(1-2i)$ is a $7^2$-number but not a Weil $7^2$-number.

In the following theorem we summarize results about $\mathscr{V}_{n,c}$ proved in [Kot92, §10]. The theorem can be viewed as a generalization of Honda-Tate theory, and its proof also heavily relies on the latter.

**Theorem 3.1.7** (Kottwitz). *For $c \in \mathbb{Q}$, the category $\mathscr{V}_{n,c}$ is nonempty only if $c > 0$ and $v_p(c) = n$. Assume this is the case. Then $\mathscr{V}_{n,c}$ is a semi-simple abelian category with all objects having finite length. For any simple object $A$ of $\mathscr{V}_{n,c}$, the $\mathbb{Q}$-algebra $\mathrm{End}(A)_{\mathbb{Q}}$ is a division algebra with center $\mathbb{Q}[\pi_A]$ (which is a field). The Hasse invariant of $\mathrm{End}(A)_{\mathbb{Q}}$ at all places of $\mathbb{Q}[\pi_A]$ are given by explicit formulas, see* [Kot92, Lemma 10.11]. *The image of $\pi_A$ under any embedding $\mathbb{Q}[\pi_A] \to \overline{\mathbb{Q}}$ is a $c$-number. Moreover $A \mapsto \pi_A$ induces a bijection from the set of isomorphism classes of simple objects in $\mathscr{V}_{n,c}$ to the set of Galois orbits of $c$-numbers.* □

3.2. **The endomorphism structure.**

**Definition 3.2.1.** Let $\mathscr{V}_{n,c,\mathcal{K}}$ be the category whose objects are pairs $(A, \iota)$ with $A \in \mathscr{V}_{n,c}$ and $\iota$ a $\mathbb{Q}$-algebra map $\mathcal{K} \to \mathrm{End}(A)_{\mathbb{Q}}$, and whose morphisms are morphisms in $\mathscr{V}_{n,c}$ that are compatible with the $\mathcal{K}$-actions. In the following we also use the simpler notations $\mathrm{End}_{\mathcal{K}}(A,\iota)_{\mathbb{Q}}$ or $\mathrm{End}_{\mathcal{K}}(A)_{\mathbb{Q}}$ to denote $\mathrm{End}_{\mathscr{V}_{n,c,\mathcal{K}}}(A,\iota)$.

As shown in [Kot92, Lemmas 3.2, 3.3], we can formally deduce the structure of $\mathscr{V}_{n,c,\mathcal{K}}$ from that of $\mathscr{V}_{n,c}$, and obtain the following corollary.

**Corollary 3.2.2** ([Kot92, Lemma 10.13])**.** *Let $c \in \mathbb{Q}$ such that $c > 0$ and $v_p(c) = n$. Then $\mathscr{V}_{n,c,\mathcal{K}}$ is a semi-simple abelian category with all objects having finite length. Fix an embedding $\mathcal{K} \to \overline{\mathbb{Q}}$. The simple objects of $\mathscr{V}_{n,c,\mathcal{K}}$ are classified by $c$-numbers in $\overline{\mathbb{Q}}$ up to conjugacy over $\mathcal{K}$.*

*Sketch of proof.* We only give the definition of the simple object $X_\pi$ in $\mathscr{V}_{n,c,\mathcal{K}}$ (up to isomorphism) corresponding to a $c$-number $\pi$, in terms of the classification in Theorem 3.1.7. Let $A_\pi$ be the simple object in $\mathscr{V}_{n,c}$ corresponding to $\pi$ as in Theorem 3.1.7. Choose a $\mathbb{Q}$-vector space isomorphism $\mathcal{K} \cong \mathbb{Q}^d$. Then the action of $\mathcal{K}$ on itself by multiplication gives an embedding $\mathcal{K} \hookrightarrow M_d(\mathbb{Q})$. Consider $A_\pi \otimes_{\mathbb{Q}} \mathcal{K} := A_\pi^{\oplus d} \in \mathscr{V}_{n,c}$. Define $\iota$ to be the composition

$$\mathcal{K} \hookrightarrow M_d(\mathbb{Q}) \xrightarrow{\mathbb{Q} \to \mathrm{End}(A_\pi)_{\mathbb{Q}}} M_d(\mathrm{End}(A_\pi)_{\mathbb{Q}}) = \mathrm{End}(A_\pi \otimes_{\mathbb{Q}} \mathcal{K})_{\mathbb{Q}}.$$

Then $X_\pi$ is given by $(A_\pi \otimes_{\mathbb{Q}} \mathcal{K}, \iota)$. □

*Remark* 3.2.3. With the notation in the above proof, we have

$$\mathrm{End}_{\mathcal{K}}(X_\pi)_{\mathbb{Q}} = \mathcal{K} \otimes_{\mathbb{Q}} \mathrm{End}(A_\pi)_{\mathbb{Q}}.$$

## 4. Kottwitz triples

4.1. **The notion of a Kottwitz triple and its Kottwitz invariant.** Denote $\Gamma := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and $\Gamma_v := \mathrm{Gal}(\overline{\mathbb{Q}}_v/\mathbb{Q}_v)$ for all places $v$ of $\mathbb{Q}$. For each $v$ we choose an embedding $\overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_v$, and correspondingly we have an embedding $\Gamma_v \to \Gamma$. For each $n \in \mathbb{N}$, denote $W_n := W(\mathbb{F}_{p^n})$ (the ring of Witt vectors) and $L_n := W_n[1/p] = \mathrm{Frac}(W_n)$. Denote $W := W(\overline{\mathbb{F}}_p)$ and $L := W[1/p] = \mathrm{Frac}(W)$.

**Definition 4.1.1.** Let $n \in \mathbb{N}$. An $\mathbb{F}_{p^n}$*-Kottwitz triple* is a triple $(\gamma_0, \gamma, \delta)$, where $\gamma_0$ is a stable conjugacy class of semi-simple $\mathbb{R}$-elliptic elements in $G(\mathbb{Q})$, $\gamma$ is a conjugacy class in $G(\mathbb{A}_f^p)$, and $\delta$ is a $\sigma$-conjugacy class in $G(L_n)$, satisfying the following conditions:

(1) $\gamma_0$ is stably conjugate to $\gamma$.

(2) $\gamma_0$ (viewed as an element of $G(\mathbb{Q}_p)$) is an $n$-th norm of $\delta$, i.e., $\gamma_0$ is conjugate to $\delta\sigma(\delta)\cdots\sigma^{n-1}(\delta)$ in $G(\overline{\mathbb{Q}}_p)$. Here recall that the $n$-th norm of $\delta$ is well-defined as a stable conjugacy class in $G(\mathbb{Q}_p)$. See [Kot82] for more details.

(3) Let $B(G_{\mathbb{Q}_p})$ be the set of $\sigma$-conjugacy classes in $G(L)$. We require that the image of $\delta$ under the Kottwitz map $B(G_{\mathbb{Q}_p}) \to X^*(Z(\widehat{G})^{\Gamma_p})$ is equal to $-\mu_1$, where $\mu_1$ is the natural image in $X^*(Z(\widehat{G})^{\Gamma_p})$ of the Hodge cocharacter $\mu_h : \mathbb{G}_m \to G_{\mathbb{C}}$ associated to $h : \mathbb{S} \to G_{\mathbb{R}}$.[7] See [Kot85] for fundamental facts about $B(G_{\mathbb{Q}_p})$ and the definition of the Kottwitz map, and see [Kot90, §2] for how $\mu_h$ has a natural image in $X^*(Z(\widehat{G})^{\Gamma_p})$.

*Remark* 4.1.2. Note that $\gamma_0$ (up to stable conjugacy) is uniquely determined by $\gamma$.

*Remark* 4.1.3. In general, if $G$ is a reductive group $\mathbb{Q}$ that has a Shimura datum, then $G_{\mathbb{R}}$ has elliptic maximal tori. For such $G$, the two notions of ellipticity over $\mathbb{R}$ for semi-simple elements of $G(\mathbb{R})$ (p. 392 of [Kot86]) are equivalent. For a semi-simple element $\gamma_0 \in G(\mathbb{Q})$, we say it is $\mathbb{R}$-*elliptic* if $\gamma_0 \in G(\mathbb{R})$ is elliptic over $\mathbb{R}$.

**Definition 4.1.4** (cf. [Kot86], [Har11, §4])**.** Let $\gamma_0 \in G(\mathbb{Q})$ be a semi-simple elliptic element. Let $I_0 := G_{\gamma_0}$. Define $\mathfrak{K}(\gamma_0/\mathbb{Q}) = \mathfrak{K}(I_0/\mathbb{Q})$ to be the finite abelian group consisting of elements in $\pi_0((Z(\widehat{I_0})/Z(\widehat{G}))^{\Gamma}) = (Z(\widehat{I_0})/Z(\widehat{G}))^{\Gamma}$ whose images in $\mathbf{H}^1(\Gamma, Z(\widehat{G}))$ are locally trivial.

In [Kot90, §2], Kottwitz attaches an element $\alpha(\gamma_0, \gamma, \delta) \in \mathfrak{K}(\gamma_0/\mathbb{Q})^D$ (where the superscript $D$ denotes the Pontryagin dual) to each $\mathbb{F}_{p^n}$-Kottwitz triple $(\gamma_0, \gamma, \delta)$. The construction also depends on the Shimura datum. We call $\alpha(\gamma_0, \gamma, \delta)$ the *Kottwitz invariant of* $(\gamma_0, \gamma, \delta)$. We will sketch its construction below. Of course, now we view $\gamma_0$ as an element of $G(\mathbb{Q})$ on the nose as opposed to a stable conjugacy class, so that the group $I_0 := G_{\gamma_0}$ is defined. When $\gamma_0$ varies in the stable conjugacy class, the groups $\mathfrak{K}(\gamma_0/\mathbb{Q})$ are canonically identified with each other and the elements $\alpha(\gamma_0, \gamma, \delta)$ are compatible with respect to the these identifications.

We now sketch the construction of $\alpha(\gamma_0, \gamma, \delta)$. As recalled in [Har11, §3], at each finite place $v \neq p$ the difference between the $v$-component of $\gamma$ and the element $\gamma_0 \in G(\mathbb{Q}_v)$ is measured by a class $\alpha_v$ in

$$\ker(\mathbf{H}^1(\mathbb{Q}_v, I_0) \to \mathbf{H}^1(\mathbb{Q}_v, G)).$$

Recall that Kottwitz's "abelianization map" gives a canonical bijection (see e.g. [Har11, §4]):

$$\mathbf{H}^1(\mathbb{Q}_v, H) \xrightarrow{\sim} \pi_0(Z(\widehat{H})^{\Gamma_v})^D,$$

for $H = I_0$ or $H = G$. Hence $\alpha_v$ can be viewed as a character

$$\alpha_v : Z(\widehat{I_0})^{\Gamma_v} \longrightarrow \mathbb{C}^\times$$

whose restriction to $(Z(\widehat{I_0})^{\Gamma_v})^0 Z(\widehat{G})^{\Gamma_v}$ is trivial.

At $v = p$, by a theorem of Steinberg we can find $c \in G(L)$ such that $c\gamma_0 c^{-1} = \delta\sigma(\delta)\cdots\sigma^{n-1}(\delta)$. Define $b := c^{-1}\delta\sigma(c)$. Then $b$ is a well-defined element in $B(I_{0,\mathbb{Q}_p})$. Define $\alpha_p$ to be the image of $b$ under the Kottwitz map (see [Kot85])

$$B(I_{0,\mathbb{Q}_p}) \longrightarrow X^*(Z(\widehat{I_0})^{\Gamma_p}).$$

At $v = \infty$, Kottwitz uses the homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ to construct an element

$$\alpha_\infty \in X^*(Z(\widehat{I_0})^{\Gamma_\infty}), \tag{4.1.1}$$

which depends only on the $G(\mathbb{R})$-conjugacy class of $h$. Its image in $X^*(Z(\widehat{G})^{\Gamma_\infty})$ is equal to $\mu_1$, see [Kot90, p. 167].

[7]Recall that $h$ is involved in the PEL datum, and the Shimura datum in question is $(G, h^{-1})$. On p. 419 of [Kot92] Kottwitz formulates this condition by requiring $\delta$ to have image equal to $\mu_1$, which should be a mistake. In fact, any $\delta$ that contributes non-trivially to the right hand side of [Kot92, (19.5)] must have image $-\mu_1$.

Thus at every place $v$ of $\mathbb{Q}$, we have constructed a character

$$\alpha_v : Z(\widehat{I_0})^{\Gamma_v} \longrightarrow \mathbb{C}^\times. \tag{4.1.2}$$

By condition (3) in Definition 4.1.1, the $\alpha_v$'s can be extended to characters

$$\beta_v : Z(\widehat{I_0})^{\Gamma_v} Z(\widehat{G}) \longrightarrow \mathbb{C}^\times,$$

such that

$$\beta_v|_{Z(\widehat{G})} = \begin{cases} \mu_1, & v = \infty \\ -\mu_1, & v = p \\ 1, & \text{else.} \end{cases}$$

Recall that our goal is to define a character

$$\alpha(\gamma_0, \gamma, \delta) : \mathfrak{K}(\gamma_0/\mathbb{Q}) \to \mathbb{C}^\times.$$

Let $\kappa \in \mathfrak{K}(\gamma_0/\mathbb{Q})$. Using the Chebotarev density theorem, it is easy to see that

$$\mathfrak{K}(\gamma_0/\mathbb{Q}) = \left[\bigcap_v Z(\widehat{I_0})^{\Gamma_v} Z(\widehat{G})\right] / Z(\widehat{G}).$$

Thus $\kappa$ has lifts $\kappa_v \in Z(\widehat{I_0})^{\Gamma_v} Z(\widehat{G})$ for all $v$. We define

$$\alpha(\gamma_0, \gamma, \delta)(\kappa) := \prod_v \beta_v(\kappa_v).$$

This definition is independent of the choices of $\kappa_v$.

4.2. **Associating Kottwitz triples to virtual abelian varieties with additional structures.** Let $A$ be an abelian variety over $\mathbb{F}_{p^n}$. Let $\bar{A} := A_{\overline{\mathbb{F}}_p}$. We have the following well-known constructions:

(1) We have the $\mathbb{A}_f^p$-Tate module $\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$ of $\bar{A}$, which is a finite free $\mathbb{A}_f^p$-module with a distinguished endomorphism, namely the action of $\pi_A$.
(2) We have the cristalline homology $\bar{\Lambda} = \mathbf{H}_1(\bar{A}/W)$, which is by definition the $W$-linear dual of the usual cristalline cohomology $\mathbf{H}^1_{\text{cris}}(\bar{A}/W)$. Thus $\bar{\Lambda}$ is a finite free $W$-module, and the usual Frobenius on $\mathbf{H}^1_{\text{cris}}(\bar{A}/W)$ induces a $\sigma$-linear automorphism $\Phi$ of $\bar{\Lambda}[1/p]$, still called the Frobenius. We have $\Phi(\bar{\Lambda}) \supset \bar{\Lambda}$. Moreover, the $W_n$-linear dual $\Lambda = \mathbf{H}_1(A/W_n)$ of $\mathbf{H}^1_{\text{cris}}(A/W_n)$ provides a canonical $W_n$-structure for $\bar{\Lambda}$. In other words we have a canonical isomorphism $\Lambda \otimes_{W_n} W \xrightarrow{\sim} \bar{\Lambda}$. The Frobenius $\Phi$ on $\bar{\Lambda}[1/p]$ stabilizes $\Lambda[1/p]$. We refer the reader to [Hai05, §14] for a good introduction to these constructions, including a discussion on the duality.

We would like to generalize the above constructions to all $A \in \mathscr{V}_n$. The generalization of (1) is verbatim. On p. 403 of [Kot92] Kottwitz shows how to generalize (2), which is in fact already needed in the proof of Theorem 3.1.7. More precisely:

**Lemma 4.2.1.** *For all $A = (\bar{A}, u) \in \mathscr{V}_n$, let $\bar{\Lambda} = \mathbf{H}_1(\bar{A}/W)$ be the $W$-linear dual of $\mathbf{H}^1_{\text{cris}}(\bar{A}/W)$. There is a canonical $W_n$-structure for $\bar{\Lambda}$, denoted by $\Lambda = \mathbf{H}_1(A/W_n)$, which is functorial with respect to the morphisms in $\mathscr{V}_n$. The Frobenius $\Phi$ on $\bar{\Lambda}[1/p]$ stabilizes $\Lambda[1/p]$. Moreover, the operator $\Phi^{-n}$ on $\Lambda$ coincides with the endomorphism of $\Lambda$ induced functorially by $\pi_A \in \mathrm{End}(A)_{\mathbb{Q}}$.*

*Sketch of proof.* The key point is that $\bar{\Lambda}$ and $\bar{\Lambda}' := \mathbf{H}_1(\sigma^n(\bar{A})/W)$ are related by a canonical isomorphism

$$\bar{\Lambda} \otimes_{W,\sigma^n} W \cong \bar{\Lambda}'.$$

Hence $u$ induces a $\sigma^n$-linear automorphism $u$ of $\bar{\Lambda}$. We define

$$\Lambda := \left\{x \in \bar{\Lambda} | ux = x\right\}.$$

The rest of the properties are easy to check. ∎

Fix $c \in \mathbb{Q}_{>0}$ with $v_p(c) = n$. Consider $A \in \mathscr{V}_{n,c}$ and a $c$-polarization $\lambda$ of it. Then $\lambda$ induces an $\mathbb{A}_f^p(1)$-valued alternating form on $\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$, denoted by $(\cdot,\cdot)_\lambda$, satisfying

$$(\pi_A v, \pi_A w)_\lambda = c(v,w)_\lambda,\ \forall v,w \in \mathbf{H}_1(\bar{A}, \mathbb{A}_f^p). \tag{4.2.1}$$

Similarly, we have an alternating form $(\cdot,\cdot)_\lambda$ on the $L_n$-vector space $\mathbf{H}_1(A/W_n)[1/p]$, well-defined up to a scalar in $W_n^\times$, satisfying

$$(\Phi v, \Phi w)_\lambda = c'\sigma(v,w),\ \forall v,w \in \mathbf{H}_1(A/W_n)[1/p],$$

for some $c' \in L_n^\times$ whose norm to $\mathbb{Q}_p^\times$ is $c^{-1}$. See [Kot92, p. 404] for details. Now suppose in addition that we have a $\mathbb{Q}$-algebra map $\iota : \mathcal{K} \to \mathrm{End}(A)_\mathbb{Q}$ which is *compatible with* $\lambda$ in the sense that $\iota$ is equivariant with respect to the automorphism $\mathbf{c}$ on $\mathcal{K}$ and the Rosati involution on $\mathrm{End}(A)_\mathbb{Q}$ defined by $\lambda$. Then by functoriality we have actions of $\mathcal{K}$ on $\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$ and on $\mathbf{H}_1(A/W_n)[1/p]$, commuting with $\pi_A$ and $\Phi$ respectively.

We now refine the category $\mathscr{V}_{n,c,\mathcal{K}}$ in §3.2 by including a $c$-polarization as part of the structure.

**Definition 4.2.2.** Define the category $\mathscr{V}_{n,c,\mathcal{K},pol}$ of *$c$-polarized virtual abelian varieties over* $\mathbb{F}_{p^n}$ as follows. The objects are triples $(A,\iota,\lambda)$, where $(A,\iota) \in \mathscr{V}_{n,c,\mathcal{K}}$ and $\lambda$ is a $c$-polarization of $A$ which is compatible with $\iota$ in the sense mentioned above. The morphisms are the morphisms of $\mathscr{V}_{n,c,\mathcal{K}}$ that preserve the polarizations up to a scalar. Define $\mathscr{V}_{n,c}^G$ to be the full subcategory of $\mathscr{V}_{n,c,\mathcal{K},pol}$ consisting of objects $(A,\iota,\lambda)$ satisfying the following two conditions:

(1) Fix an arbitrary $\mathbb{A}_f^p$-linear isomorphism $\mathbb{A}_f^p \cong \mathbb{A}_f^p(1)$. There is an $\mathbb{A}_f^p$-linear isomorphism

$$(\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p), (\cdot,\cdot)_\lambda) \to (V, \langle\cdot,\cdot\rangle) \otimes_\mathbb{Q} \mathbb{A}_f^p$$

which is $\mathcal{K}$-equivariant and preserves the alternating forms up to a scalar.

(2) There is an $L_n$-linear isomorphism

$$(\mathbf{H}_1(A/W_n)[1/p], (\cdot,\cdot)_\lambda) \to (V, \langle\cdot,\cdot\rangle) \otimes_\mathbb{Q} L_n$$

which is $\mathcal{K}$-equivariant and preserves the alternating forms up to a scalar.

**Definition 4.2.3.** Let $(A,\iota,\lambda) \in \mathscr{V}_{n,c}^G$. Choose isomorphisms as in (1) and (2) of Definition 4.2.2. Using these isomorphisms, we translate $\pi_A^{-1}$ to an element $\gamma$ of $\mathrm{GL}(V)(\mathbb{A}_f^p)$, and translate $\Phi$ to a $\sigma$-linear automorphism $\delta'$ of $V \otimes_\mathbb{Q} L_n$. Then it is easy to see that $\gamma \in G(\mathbb{A}_f^p)$ and $\delta' = \delta \circ \sigma$ for some $\delta \in G(L_n)$. Moreover the conjugacy class of $\gamma$ in $G(\mathbb{A}_f^p)$ and the $\sigma$-conjugacy class of $\delta$ in $G(L_n)$ are independent of the choices. We say that $(\gamma,\delta)$ is associated to $(A,\iota,\lambda)$.

*Remark* 4.2.4. From (4.2.1) we know that $\gamma$ in the above definition has factor of similitude $\nu(\gamma) = c^{-1}$.

**Definition 4.2.5.** Let $\mathscr{V}_{n,c}^{(G,h^{-1})}$ be the full subcategory of $\mathscr{V}_{n,c}^G$ consisting of objects whose associated $\delta$ as in Definition 4.2.3 satisfies condition (3) in Definition 4.1.1.

**Proposition 4.2.6** ([Kot92, §14])**.** *Let $(A,\iota,\lambda) \in \mathscr{V}_{n,c}^{(G,h^{-1})}$. Let $(\gamma,\delta)$ be associated to it as in Definition 4.2.3. Then the stable conjugacy class of $\gamma$ contains a $\mathbb{Q}$-point $\gamma_0$, which in particular has factor of similitude $\nu(\gamma_0) = c^{-1}$ (see Remark 4.2.4). Moreover, $(\gamma_0,\gamma,\delta)$ is an $\mathbb{F}_{p^n}$-Kottwitz triple.*

*Sketch of proof.* For any $\mathbb{Q}$-algebra $R$, let $I(R)$ be the automorphism group of $(A,\iota,\lambda)$ in the $R$-linear category $\mathscr{V}_{n,c}^{(G,h^{-1})} \otimes_\mathbb{Q} R$. Then the functor $R \mapsto I(R)$ is representable by a reductive group $I$ over $\mathbb{Q}$. To show the existence of $\gamma_0$, the key claim is that any maximal torus $T$ of $I$ defined over $\mathbb{Q}$ admits an embedding to $G$ defined over $\mathbb{Q}$. Once the claim is proved, noting that $\pi_A$ is a $\mathbb{Q}$-point of the center of $I$, we define $\gamma_0$ to be the image of $\pi_A^{-1}$ under a $\mathbb{Q}$-embedding $T \to G$.

The proof of the claim contains two steps. Let $N$ be the centralizer of $T(\mathbb{Q})$ in $\mathrm{End}_\mathcal{K}(A)_\mathbb{Q}$. Then the Rosati involution $*$ on $\mathrm{End}(A)_\mathbb{Q}$ induced by $\lambda$ stabilizes $N$ and we have $\dim_\mathcal{K} N = \dim_\mathcal{K} V$. In fact it can be shown that $(N, *|_N)$ is a CM algebra (i.e. a product of CM fields). The algebraic group $T$ is recovered from $N$ by

$$T(R) = \left\{ g \in (N \otimes_\mathbb{Q} R)^\times | gg^* \in R^\times \right\}$$

for any $\mathbb{Q}$-algebra $R$. The first step in the proof of the claim is to show the existence of a $\mathcal{K}$-algebra embedding $N \to C = \mathrm{End}_{\mathcal{K}}(V)$. This step is trivial in our case, as $C$ is a matrix algebra over $\mathcal{K}$ of the right dimension. The second step is to show that one can modify the embedding $N \to C$ such that it respects the involutions on $N$ and $C$. By cohomological considerations and the theory of transferring tori, we reduce to checking that certain local obstructions vanish. At finite places $v \neq p$, the local obstructions vanish because of condition (1) in Definition 4.2.2. At $p$ and $\infty$, the obstructions vanish more or less automatically.

Once the existence of $\gamma_0$ is proved, it is not too difficult to check that $(\gamma_0, \gamma, \delta)$ is an $\mathbb{F}_{p^n}$-Kottwitz triple with the help of Lemma 2.1.2. For example, condition (3) in Definition 4.1.1 holds by definition, and condition (2) in that definition boils down to the relation $\Phi^{-n} = \pi_A$ in Lemma 4.2.1. $\square$

**Definition 4.2.7.** In the situation of Proposition 4.2.6, we say that $(\gamma_0, \gamma, \delta)$ is the $\mathbb{F}_{p^n}$-Kottwitz triple associated to $(A, \iota, \lambda) \in \mathscr{V}_{n,c}^{(G,h^{-1})}$.

The following result is arguably the technical core of [Kot92].

**Theorem 4.2.8** ([Kot92, Lemma 18.1])**.** *Let $(\gamma_0, \gamma, \delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple. Then $(\gamma_0, \gamma, \delta)$ is associated to an object $(A, \iota, \lambda)$ in $\mathscr{V}_{n,c}^{(G,h^{-1})}$ if and only if the following conditions hold:*

*(1) The factor of similitude $\nu(\gamma_0)$ of $\gamma_0$ is equal to $c^{-1}$.*
*(2) The Kottwitz invariant $\alpha(\gamma_0, \gamma, \delta)$ is trivial.*
*(3) There is a $W_n$-lattice $\Lambda$ in $V \otimes_{\mathbb{Q}} L_n$ such that $(\delta\sigma)\Lambda \supset \Lambda$.*

4.3. **Sketch of the "only if" part in Theorem 4.2.8.** We only sketch the proof of (2), which is the most difficult part. Firstly Kottwitz observes that by special properties of $G$, the group $\mathfrak{K}(\gamma_0/\mathbb{Q})^D$ always maps injectively into $X^*(Z(\widehat{I_0})^\Gamma)$, where $I_0 := G_{\gamma_0}$. For $(A, \iota, \lambda) \in \mathscr{V}_{n,c}^{(G,h^{-1})}$, denote by $\alpha(A, \iota, \lambda)$ the image in $X^*(Z(\widehat{I_0})^\Gamma)$ of the Kottwitz invariant of the Kottwitz triple associated to $(A, \iota, \lambda)$. It suffices to show that $\alpha(A, \iota, \lambda)$ is trivial.

In fact Kottwitz proves a generalization (see Proposition 4.3.2 below) of the statement $\alpha(A, \iota, \lambda) = 0$, and this generalization is also needed in the "if" part of Theorem 4.2.8. To state it we need an auxiliary set.

**Definition 4.3.1.** Let $\mathscr{V}^{aux}$ be the set of quadruples $(\gamma_0; A, \iota, \lambda)$, where $(A, \iota, \lambda)$ is an object of $\mathscr{V}_{n,c,\mathcal{K},pol}$ and $\gamma_0 \in G(\mathbb{Q})$, satisfying the following conditions:

- $\gamma_0$ is a semi-simple $\mathbb{R}$-elliptic element with $\nu(\gamma_0) = c^{-1}$.
- Let $Y$ be an indeterminate over $\mathcal{K}$. Define a $\mathcal{K}[Y]$-module structure on $V$ (resp. $\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$) by letting $Y$ act as $\gamma_0^{-1}$ (resp. $\pi_A$). We require that there exists a $\mathcal{K}[Y] \otimes_{\mathbb{Q}} \mathbb{A}_f^p$-module isomorphism $\phi : V \otimes_{\mathbb{Q}} \mathbb{A}_f^p \xrightarrow{\sim} \mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$. Here we do **not** require $\phi$ to preserve the alternating forms $\langle \cdot, \cdot \rangle$ on $V \otimes_{\mathbb{Q}} \mathbb{A}_f^p$ and $(\cdot, \cdot)_\lambda$ on $\mathbf{H}_1(\bar{A}, \mathbb{A}_f^p)$ up to a scalar.

Let $(A, \iota, \lambda) \in \mathscr{V}_{n,c}^{(G,h^{-1})}$ and let $(\gamma_0, \gamma, \delta)$ be the associated $\mathbb{F}_{p^n}$-Kottwitz triple. Then $(\gamma_0; A, \iota, \lambda)$ is an element in $\mathscr{V}^{aux}$. However in the definition of $\mathscr{V}^{aux}$ above we only require $(A, \iota, \lambda) \in \mathscr{V}_{n,c,\mathcal{K},pol}$. Kottwitz shows that the construction

$$\mathscr{V}_{n,c}^{(G,h^{-1})} \ni (A, \iota, \lambda) \mapsto \alpha(A, \iota, \lambda) \tag{4.3.1}$$

can be extended to a map

$$\mathscr{V}^{aux} \ni (\gamma_0; A, \iota, \lambda) \mapsto \alpha(\gamma_0; A, \iota, \lambda) \in X^*(Z(\widehat{I_0})^\Gamma), \ I_0 := G_{\gamma_0}. \tag{4.3.2}$$

Kottwitz then proves:

**Proposition 4.3.2** ([Kot92, Lemma 15.2])**.** *The element $\alpha(\gamma_0; A, \iota, \lambda)$ is trivial for all $(\gamma_0; A, \iota, \lambda) \in \mathscr{V}^{aux}$.*

4.3.3. *Sketch of the construction (4.3.2).* Consider a finite place $v \neq p$. Let $\phi$ be as in Definition 4.3.1. The difference between the two alternating forms on $V \otimes_{\mathbb{Q}} \mathbb{Q}_v$, given by $\langle \cdot, \cdot \rangle$ and $\phi^{-1}((\cdot, \cdot)_\lambda)$, is measured by a class in $\mathbf{H}^1(\mathbb{Q}_v, I_0)$. In fact, the $\mathbb{Q}$-algebraic group $I_0$ is the automorphism group of $(V, \langle \cdot, \cdot \rangle)$, as a $\mathcal{K}[Y]$-module equipped with an alternating form up to a scalar. The $\mathcal{K}[Y] \otimes_{\mathbb{Q}} \mathbb{Q}_v$-module $\mathbf{H}_1(\bar{A}, \mathbb{Q}_v)$ is by assumption isomorphic to $V \otimes_{\mathbb{Q}} \mathbb{Q}_v$, and therefore

$$(\mathbf{H}_1(\bar{A}, \mathbb{Q}_v), (\cdot, \cdot)_\lambda) \otimes_{\mathbb{Q}_v} \overline{\mathbb{Q}}_v$$

is isomorphic to

$$(V, \langle\cdot,\cdot\rangle) \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}_v,$$

as $\mathcal{K}[Y] \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}_v$-modules equipped with alternating forms up to a scalar. Thus their difference over $\mathbb{Q}_v$ is measured by a class in $\mathbf{H}^1(\mathbb{Q}_v, I_0)$.

Using this class in $\mathbf{H}^1(\mathbb{Q}_v, I_0)$ we construct, as in the end of §4.1, a character

$$\alpha_v : Z(\widehat{I_0})^{\Gamma_v} \to \mathbb{C}^\times.$$

We also have a character

$$\alpha_\infty : Z(\widehat{I_0})^{\Gamma_\infty} \to \mathbb{C}^\times$$

as in (4.1.1). At $p$, by Steinberg's theorem there exists a $\mathcal{K}(Y) \otimes_{\mathbb{Q}} L$-module isomorphism

$$\mathbf{H}_1(\bar{A}/W)[1/p] \xrightarrow{\sim} V \otimes_{\mathbb{Q}} L$$

that preserves the alternating forms, where $Y$ acts on the left hand side by $\pi_A$ and on the right hand side by $\gamma_0^{-1}$. Then $\Phi$ acting on the left hand side translates to some $b\sigma$ acting on the right hand side, for $b \in I_0(L)$ well-defined up to $\sigma$-conjugacy. We define $\alpha_p \in X^*(Z(\widehat{I_0})^{\Gamma_p})$ to be the image of $b$ under the Kottwitz map for $I_{0,\mathbb{Q}_p}$. Define

$$\alpha(\gamma_0; A, \iota, \lambda) := \prod_v (\alpha_v)|_{Z(\widehat{I_0})^\Gamma}. \tag{4.3.3}$$

The above is a sketch of the definition of (4.3.2). We leave it as an exercise for the reader to check that (4.3.2) indeed extends (4.3.1).

4.3.4. *Sketch of the proof of Proposition 4.3.2.* Our specific PEL datum with $B = \mathcal{K}$ simplifies the proof. Let $I$ be the reductive group over $\mathbb{Q}$ of automorphisms of $(A, \iota, \lambda)$ in $\mathscr{V}_{n,c,\mathcal{K},pol}$, as in the proof of Proposition 4.2.6. Let $M := \operatorname{End}_{\mathcal{K}}(A)_{\mathbb{Q}}$. Then $M$ is equipped with the Rosati involution $*$ and

$$I(R) = \left\{ g \in (M \otimes R)^\times | g^* g \in R^\times \right\}$$

for any $\mathbb{Q}$-algebra $R$. Let $M_0$ be the centralizer of $\gamma_0$ in $C = \operatorname{End}_{\mathcal{K}}(V)$. Then similarly $M_0$ is stable under the involution $*$ on $C$ and we have

$$I_0(R) = \left\{ g \in (M_0 \otimes R)^\times | g^* g \in R^\times \right\}.$$

Consider the $I_0$-torsor of $\mathcal{K}$-linear $*$-isomorphisms from $M_0$ to $M$ that take $\gamma_0 \in M_0$ to $\pi_A^{-1} \in M$. The second condition in Definition 4.3.1 guarantees that this torsor is nonempty, and it follows that there is an inner twisting $i : I \to I_0$ (as algebraic groups over $\mathbb{Q}$) canonical up to $I_0(\overline{\mathbb{Q}})$-conjugation. By the general theory of transferring tori, we find a maximal torus $T$ of $I$ that transfers to $I_0$ along $i$, i.e., there exists $g_0 \in I_0(\overline{\mathbb{Q}})$ such that $(\operatorname{Int}(g_0) \circ i)|_T : T \to I_0$ is defined over $\mathbb{Q}$. Let $N$ be the centralizer of $T$ in $M$ and let $N'$ be the centralizer of $(\operatorname{Int}(g_0) \circ i)(T)$ in $M_0$. Then $N \cong N'$, and it is a CM algebra containing $\mathcal{K}$ with $\mathcal{K}$-dimension equal to $\dim_{\mathcal{K}} V$.

Essentially by replacing $\mathcal{K}$ with $N'$ in the whole discussion, we reduce to the special case where $\dim_{\mathcal{K}} V = 1$. In this case we know that $I = I_0 = G = T$ is a torus, given by

$$T(R) = \left\{ g \in (\mathcal{K} \otimes R)^\times | g \cdot \mathbf{c}(g) \in R^\times \right\}.$$

Moreover $\dim A = [\mathcal{K} : \mathbb{Q}]/2$. The element $\alpha(\gamma_0; A, \iota, \lambda)$, which we would like to show is trivial, now lies in $X^*(\widehat{T}^\Gamma) = X_*(T)_\Gamma$. The vanishing of $\alpha(\gamma_0; A, \iota, \lambda)$ in this case is shown in [Kot92, Lemma 13.2]. We sketch its proof below.

Note that if we replace $\lambda$ by some $\lambda'$ such that $(\gamma_0; A, \iota, \lambda')$ is still in $\mathscr{V}^{aux}$, then $\alpha(\gamma_0; A, \iota, \lambda')$ differs from $\alpha(\gamma_0; A, \iota, \lambda)$ by the image in $X_*(T)_\Gamma$ of a global class in $\mathbf{H}^1(\mathbb{Q}, T)$. By Tate-Nakayama, any such image is trivial. Hence $\alpha(\gamma_0; A, \iota, \lambda)$ is independent of $\lambda$. Moreover $\alpha(\gamma_0; A, \iota, \lambda)$ is independent of the virtual $\mathbb{F}_{p^n}$-structure on the abelian variety $\bar{A}$ over $\overline{\mathbb{F}}_p$, as long as the virtual structure is compatible with the $\mathcal{K}$-action on $\bar{A}$. Hence we may write $\alpha(\gamma_0; \bar{A}, \iota)$ for $\alpha(\gamma_0; A, \iota, \lambda)$. By Tate's CM lifting theorem, the pair $(\bar{A}, \iota)$ (up to $\mathbb{Q}$-isogeny) lifts to a CM abelian variety $(\mathcal{A}, \iota)$ with CM by $\mathcal{K}$ over characteristic zero.

We view $\bar{A}$ simply as the reduction of $\mathcal{A}$, ignoring the "up to $\mathbb{Q}$-isogeny". After suitably choosing a $\mathbb{Q}$-symmetrization on $\mathcal{A}$ and a non-degenerate alternating form on $V$ (which need not be $\langle\cdot,\cdot\rangle$), we obtain an element $b \in B(T_{\mathbb{Q}_p})$ that describes the isocrystal $\mathbf{H}_1(\bar{A}/W)[1/p]$. Moreover these choices can be made in such a convenient way that the vanishing of $\alpha(\gamma_0; \bar{A}, \iota)$ is easily reduced to the following claim.

**Claim** ([Kot92, Lemma 13.1]). The image of $b$ under the Kottwitz map for $T_{\mathbb{Q}_p}$

$$B(T_{\mathbb{Q}_p}) \to X^*(\widehat{T}^{\Gamma_p})$$

(which is a bijection) is equal to the image of $\mu_{(\mathcal{A},\iota)}$ under the natural projection

$$X_*(T) \to X_*(T)_{\Gamma_p} = X^*(\widehat{T}^{\Gamma_p}).$$

Here $\mu_{(\mathcal{A},\iota)} \in X_*(T)$ is the cocharacter corresponding to the CM type of $(\mathcal{A}, \iota)$.

The proof of [Kot92, Lemma 13.1] involves $p$-adic Hodge theory and the Shimura–Taniyama reciprocity, see [Kot92, §§12, 13]. We note that Reimann and Zink proved a stronger form of [Kot92, Lemma 13.1] in [RZ88] (for $p \neq 2$). We refer the reader to the proof of [Hai02, Theorem 7.2] for an explanation of how the result of Reimann–Zink can be interpreted so as to imply [Kot92, Lemma 13.1]. □

4.4. **Sketch of the "if" part in Theorem 4.2.8.** Let $(\gamma_0, \gamma, \delta)$ be such an $\mathbb{F}_{p^n}$-Kottwitz triple. As usual we write $I_0$ for $G_{\gamma_0}$. The first step is to construct an object $X'$ in $\mathscr{V}_{n,c,\mathcal{K},pol}$, such that $(\gamma_0; X') \in \mathscr{V}^{aux}$ (see Definition 4.3.1). Let $f$ be the characteristic polynomial of $\gamma_0^{-1} \in \mathrm{GL}_{\mathcal{K}}(V)$ over $\mathcal{K}$. Let $f = \prod_i f_i$ be the irreducible factorization over $\mathcal{K}$. It can be shown that all the roots of $f_i$ are $c$-numbers, so each $f_i$ gives rise to a simple object $X_i$ in $\mathscr{V}_{n,c,\mathcal{K}}$ via Corollary 3.2.2. Then $X'$ can be constructed by forming the direct sum of the $X_i$'s with suitable multiplicities and choosing a suitable $c$-polarization.

Having constructed $X'$, we recall from (4.3.3) that the invariant $\alpha(\gamma_0; X') \in X^*(Z(\widehat{I_0})^{\Gamma})$ is constructed from local invariants

$$\alpha_v(\gamma_0; X') \in X^*(Z(\widehat{I_0})^{\Gamma_v}).$$

Also recall from (4.1.2) that the invariant $\alpha(\gamma_0, \gamma, \delta)$ is constructed from local invariants

$$\alpha_v(\gamma_0, \gamma, \delta) \in X^*(Z(\widehat{I_0})^{\Gamma_v}).$$

Let

$$\beta_v := \alpha_v(\gamma_0, \gamma, \delta) - \alpha_v(\gamma_0; X').$$

By assumption the invariant $\alpha(\gamma_0, \gamma, \delta)$ is trivial, and by Proposition 4.3.2 the invariant $\alpha(\gamma_0; X')$ is trivial. Therefore

$$\prod_v \beta_v = 0 \in \mathfrak{K}(I_0/\mathbb{Q})^D.$$

Let $I$ be the reductive group over $\mathbb{Q}$ of automorphisms of $X'$ in $\mathscr{V}_{n,c,\mathcal{K},pol}$, as in the proof of Proposition 4.2.6. We have seen in §4.3 that there is an inner twisting $I \to I_0$ which is canonical up to $I_0(\overline{\mathbb{Q}})$-conjugation. Hence we have canonical identifications

$$X^*(Z(\widehat{I_0})^{\Gamma_v}) \cong X^*(Z(\widehat{I})^{\Gamma_v})$$

$$\mathfrak{K}(I_0/\mathbb{Q}) \cong \mathfrak{K}(I/\mathbb{Q}).$$

The elements $\beta_v$ can be viewed as elements of $X^*(Z(\widehat{I})^{\Gamma_v})$ whose product is zero in $\mathfrak{K}(I/\mathbb{Q})^D$. It follows that the $\beta_v$'s are localizations of a common global class $\beta \in \mathbf{H}^1(\mathbb{Q}, I)$. We can then *twist* $X'$ *by* $\beta$, to obtain an object $X = (X')^{\beta} \in \mathscr{V}_{n,c,\mathcal{K},pol}$. Here $X$ is characterized by the property that it is isomorphic to $X'$ in $\mathscr{V}_{n,c,\mathcal{K},pol} \otimes \overline{\mathbb{Q}}$, and such that the functor that assigns to each $\mathbb{Q}$-algebra $R$ the set of isomorphisms $X' \to X$ in $\mathscr{V}_{n,c,\mathcal{K},pol} \otimes R$ is represented by the $I$-torsor corresponding to $\beta$. It can be checked that $X \in \mathscr{V}_{n,c}^{(G,h^{-1})}$ and that $(\gamma_0, \gamma, \delta)$ is its associated Kottwitz triple.

Q.E.D. for Theorem 4.2.8.

## 5. The unstabilized Langlands–Kottwitz formula

5.1. **Definition of orbital integrals.** Let $(\gamma_0,\gamma,\delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple. Assume $\alpha(\gamma_0,\gamma,\delta)$ is trivial. Let $I_0 := G_{\gamma_0}$. It is shown in [Kot90, §3] that there exists an inner form $I$ of $I_0$, such that

- $I(\mathbb{R})$ is compact modulo $Z_G$.
- $I_{\mathbb{Q}_v}$ is isomorphic to $G_{\gamma_v}$ (as inner forms of $I_{0,\mathbb{Q}_v}$), for all finite places $v \neq p$.
- $I_{\mathbb{Q}_p}$ is isomorphic to the $\sigma$-centralizer $G_\delta^\sigma$ of $\delta$ (as inner forms of $I_{0,\mathbb{Q}_p}$).

We refer the reader to [Kot90, §3] for the more precise statement.

Recall that $\mathbb{F}_q = \mathbb{F}_{p^r}$ is the residue field of $E$ at $\mathfrak{p}$. Assume $r|n$. The cocharacter $\mu_h : \mathbb{G}_m \to G_{\mathbb{C}}$ descends to a conjugacy class of cocharacters $\mu_{\overline{\mathbb{Q}}}$ over $\overline{\mathbb{Q}}$, and this conjugacy class is defined over $E$. Since $G_{\mathbb{Q}_p}$ is unramified (with reductive model $\mathcal{G}$ over $\mathbb{Z}_p$), the $G(\overline{\mathbb{Q}}_p)$-conjugacy class of $\mu_{\overline{\mathbb{Q}}}$ contains a cocharacter $\mu : \mathbb{G}_m \to G_{L_r}$ which is defined over $L_r = E_{\mathfrak{p}}$ and factors through a maximal $W_r$-split torus of $\mathcal{G}_{W_r}$. We fix such a cocharacter $\mu$ and define

$$\phi_n : G(L_n) \to \mathbb{Z}$$

to be the characteristic function of $\mathcal{G}(W_n)\mu(p^{-1})\mathcal{G}(W_n)$. Let

$$f^{p,\infty} : G(\mathbb{A}_f^p) \to \mathbb{Z}$$

be the characteristic function of $K^p g^{-1} K^p$, as in the end of §2.4.

Fix Haar measures on $I_0(\mathbb{A}_f^p)$ and $I_0(\mathbb{Q}_p)$, and *transfer* them to $I(\mathbb{A}_f^p) \cong G_\gamma(\mathbb{A}_f^p)$ and $I(\mathbb{Q}_p) \cong G_\delta^\sigma(\mathbb{Q}_p)$ respectively. We refer the reader to [Kot88, p. 631] for the notion of transferring Haar measures between inner forms. Fix the Haar measure on $G(\mathbb{A}_f^p)$ that gives volume 1 to $K^p$. (Compare footnote 5.) Fix the Haar measure on $G(L_n)$ that gives volume 1 to $\mathcal{G}(W_n)$. Define the *orbital integral*

$$O_\gamma(f^{p,\infty}) := \int_{G_\gamma(\mathbb{A}_f^p)\backslash G(\mathbb{A}_f^p)} f^{p,\infty}(x^{-1}\gamma x)dx$$

and the *twisted orbital integral*

$$TO_\delta(\phi_n) := \int_{G_\delta^\sigma(\mathbb{Q}_p)\backslash G(L_n)} \phi_n(y^{-1}\delta\sigma(y))dy.$$

Here

$$G_\delta^\sigma(\mathbb{Q}_p) := \left\{y \in G(L_n) | y^{-1}\delta\sigma(y) = \delta\right\}.$$

Define

$$O(\gamma_0,\gamma,\delta) := \mathrm{vol}(I(\mathbb{Q})\backslash I(\mathbb{A}_f))O_\gamma(f^{p,\infty})TO_\delta(\phi_n) \tag{5.1.1}$$

$$c(\gamma_0,\gamma,\delta) := \mathrm{vol}(I(\mathbb{Q})\backslash I(\mathbb{A}_f)) \left|\ker(\ker^1(\mathbb{Q}, I_0) \to \ker^1(\mathbb{Q}, G))\right|. \tag{5.1.2}$$

Then $O(\gamma_0,\gamma,\delta)$ is independent of the choices of Haar measures on $I_0(\mathbb{A}_f^p)$ and $I_0(\mathbb{Q}_p)$, whereas $c(\gamma_0,\gamma,\delta)$ depends on the product measure on $I_0(\mathbb{A}_f)$.

**Lemma 5.1.1** ([Kot92, p. 441])**.** *Let $(\gamma_0,\gamma,\delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple. Assume $\alpha(\gamma_0,\gamma,\delta)$ is trivial, $r|n$, and $O(\gamma_0,\gamma,\delta) \neq 0$. Then $(\gamma_0,\gamma,\delta)$ satisfies the conditions (1) (3) in Theorem 4.2.8 for some $c \in \mathbb{Q}_{>0}$ with $v_p(c) = n$.* □

5.2. **Counting** $\mathrm{Fix}(j,g)$**.** Let $n := rj$, so that $p^n = q^j$. Consider a point $x' \in \mathrm{Fix}(j,g)$. Then $x'$ is represented by an object $(\bar{A},\lambda,\iota,\bar{\eta}) \in \mathscr{C}_{\overline{\mathbb{F}}_q}$ (see §2.6) such that there is an isomorphism

$$u : (\bar{A},\lambda,\iota,\bar{\eta}\cdot g) \xrightarrow{\sim} \sigma^n(\bar{A},\lambda,\iota,\bar{\eta}) := (\sigma^n(\bar{A}),\sigma^n(\lambda),\sigma^n(\iota),\sigma^n(\bar{\eta}))$$

in the category $\mathscr{C}_{\overline{\mathbb{F}}_q}$. Then $(\bar{A},u,\iota,\lambda)$ defines an object $X$ in $\mathscr{V}_{n,c,\mathcal{K},pol}$ for some $c \in \mathbb{Q}$. The isomorphism class of $X$ only depends on $x'$. Moreover it can be checked that $X \in \mathscr{V}_{n,c}^{(G,h^{-1})}$, see [Kot92, pp. 430-431].

Let $\mathscr{I}_{n,c}$ denote the set of isomorphism classes in $\mathscr{V}_{n,c}^{(G,h^{-1})}$, and let $\mathscr{I}_n := \coprod_{c\in\mathbb{Q}} \mathscr{I}_{n,c}$. Thus we have a well-defined map

$$\mathbf{f} : \mathrm{Fix}(j,g) \to \mathscr{I}_n.$$

We omit the proofs of the following three results.

**Lemma 5.2.1** ([Kot92, §16]). *Let $x' \in \mathrm{Fix}(j,g)$. Let $(\gamma_0,\gamma,\delta)$ be the Kottwitz triple associated to $\mathbf{f}(x')$ as in Definition 4.2.7. Then $\alpha(\gamma_0,\gamma,\delta)$ is trivial by Theorem 4.2.8, and so $O(\gamma_0,\gamma,\delta)$ in (5.1.1) is defined. The fiber of $\mathbf{f}$ containing $x'$ has cardinality $O(\gamma_0,\gamma,\delta)$. Moreover, for any point $x'_1$ in this fiber, the naive local term $\mathrm{LT}^{naive}_{x'_1}$ at $x'_1$ is equal to $\mathrm{Tr}_\xi(\gamma_0)$.* ☐

**Lemma 5.2.2** ([Kot92, §17 and p. 441]). *Let $(\gamma_0,\gamma,\delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple. Then the set of $y \in \mathscr{I}_n$ whose associated Kottwitz triple is $(\gamma_0,\gamma,\delta)$ has cardinality either $0$ or $\left|\ker^1(\mathbb{Q},I_0)\right|$.* ☐

**Lemma 5.2.3** ([Kot92, p. 441]). *Let $(\gamma_0,\gamma,\delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple. If $TO_\delta(\phi_n) \neq 0$, then $\delta$ satisfies condition (3) in Theorem 4.2.8.* ☐

5.3. **The formula.** We now assemble what we have done to obtain the unstabilized Langlands–Kottwitz formula.

**Theorem 5.3.1** ([Kot92, (19.5)]). *Fix $K^p$ and $g \in G(\mathbb{A}_f^p)$. Let $f^{p,\infty}$ be the characteristic function of $K^pg^{-1}K^p$. When $j \in \mathbb{N}$ is large enough, we have*

$$T(j,f^{p,\infty})\left|\ker^1(\mathbb{Q},G)\right|^{-1} = \sum_{(\gamma_0,\gamma,\delta)} c(\gamma_0,\gamma,\delta)O_\gamma(f^p)TO_\delta(\phi_n)\,\mathrm{Tr}_\xi(\gamma_0), \tag{5.3.1}$$

*where the summation is over $\mathbb{F}_{p^n}$-Kottwitz triples whose Kottwitz invariants are trivial, with $n = rj$.*

*Proof.* This follows from Lemma 2.5.1, (2.5.2), Theorem 4.2.8, Lemma 5.1.1, Lemma 5.2.1, Lemma 5.2.2, Lemma 5.2.3, and the following identity:

$$\left|\ker^1(\mathbb{Q},G)\right|\left|\ker(\ker^1(\mathbb{Q},G_{\gamma_0}) \to \ker^1(\mathbb{Q},G))\right| = \ker^1(\mathbb{Q},G_{\gamma_0}) \tag{5.3.2}$$

which is valid for any semisimple element $\gamma_0 \in G(\mathbb{Q})$. We explain the truth of (5.3.2). In fact, the pointed sets $\ker^1(\mathbb{Q},G_{\gamma_0})$ and $\ker^1(\mathbb{Q},G)$ have canonical structures of abelian groups (see e.g. [Kot84]), so (5.3.2) is equivalent to the surjectivity of $\ker^1(\mathbb{Q},G_{\gamma_0}) \to \ker^1(\mathbb{Q},G)$. But $\ker^1(\mathbb{Q},Z_G)$ already maps onto $\ker^1(\mathbb{Q},G)$, as shown in [Kot92, §7]. ☐

*Remark* 5.3.2. The assumption that $j$ is large enough is only needed for the validity of (2.5.2). If we had defined $T(j,g) = T(j,f^{p,\infty})$ to be the right hand side of (2.5.2), then Theorem 5.3.1 would be true for all $j \in \mathbb{N}_{\geq 1}$.

*Remark* 5.3.3. By linearity, Theorem 5.3.1 holds for arbitrary

$$f^{p,\infty} \in \mathcal{H}_{K^p} = \mathcal{H}(G(\mathbb{A}_f^p)//K^p, L(\xi)).$$

Of course then the condition that $j$ is large enough depends on $f^{p,\infty}$.

*Remark* 5.3.4. As noted in the beginning of §2.4, the quantity

$$T(j,f^{p,\infty})\left|\ker^1(\mathbb{Q},G)\right|^{-1}$$

is the trace on the cohomology of $\mathrm{Sh}_{K^pK_p}(G,X)_{\overline{E}}$ analogous to $T(j,g)$.

## 6. Stabilization

6.1. **The goal.** Let $f^{p,\infty} \in \mathcal{H}_{K^p} = \mathcal{H}(G(\mathbb{A}_f^p)//K^p, L(\xi))$. We fix a field embedding $L(\xi) \to \mathbb{C}$ and think of $f^{p,\infty}$ as taking values in $\mathbb{C}$. Denote by $T'(j,f^{p,\infty})$ the right hand side of (5.3.1). Namely

$$T'(j,f^{p,\infty}) := \sum_{(\gamma_0,\gamma,\delta)} c(\gamma_0,\gamma,\delta)O_\gamma(f^p)TO_\delta(\phi_n)\,\mathrm{Tr}_\xi(\gamma_0) \tag{6.1.1}$$

where the summation is over $\mathbb{F}_{p^n}$-Kottwitz triples with trivial Kottwitz invariant. Our next goal is to describe the stabilization of $T'(j,f^{p,\infty})$. We may and shall assume that $\xi$ is irreducible without loss of generality. Denote by $\chi \in X^*(Z_G)$ the central character of $\xi$.

One may think of $T'(j,f^{p,\infty})$ as the Shimura-variety analogue of $T_e$, the elliptic part of Arthur–Selberg trace formula for $G$. The stabilization of $T_e$ is the main topic of [Har11], and is of the following form (see [Har11, Theorem 6.2]):

$$T_e(\phi) = \sum_{(H,s,\eta)\in\mathfrak{E}} i(G,H)ST_e^{H,*}(\phi^H). \tag{6.1.2}$$

Strictly speaking, in the setting of [Har11], $G$ is assumed to be anisotropic, while our $G$ is only anisotropic modulo center. However for this motivational discussion we ignore this point. Here, as well as in what follows, we replace the notation $\xi$ used in [Har11] by $\eta$. The stabilization of $T'(j, f^{p,\infty})$ is structurally very similar to (6.1.2).

We recall briefly the ingredients in (6.1.2), which will also appear in the stabilization of $T'(j, f^{p,\infty})$. The notion of an elliptic endoscopic triple for $G$ over $\mathbb{Q}$ or $\mathbb{Q}_v$ is discussed in [Har11, §5]. We let $\mathfrak{E}$ (resp. $\mathfrak{E}_v$) denote a fixed set of representatives of the isomorphism classes of elliptic endoscopic triples for $G$ over $\mathbb{Q}$ (resp. $\mathbb{Q}_v$). Each element of $\mathfrak{E}$ (resp. $\mathfrak{E}_v$) is a triple $(H, s, \eta)$ consisting of

- a semi-simple element $s \in \widehat{G}$
- a quasi-split reductive group $H$ over $\mathbb{Q}$ (resp. $\mathbb{Q}_v$)
- an embedding $\eta : \widehat{H} \to \widehat{G}$ such that $\eta(\widehat{H})$ centralizes $s$

satisfying extra properties. Since $G^{\mathrm{der}}$ is simply connected, we can further extend $\eta$ to an $L$-embedding ${}^L H \to {}^L G$, by [Lan79, Proposition 1]. We fix such a choice and still denote it by $\eta$.[8] For example, an explicit presentation of $\mathfrak{E}, \mathfrak{E}_v$ and the embeddings ${}^L H \to {}^L G$ in special cases can be found in [Mor10b, §2.3] and [CHL11, §1.3].

*Remark* 6.1.1. For our specific $G = \mathrm{GU}(V, [\cdot, \cdot])$, we know that $H^{\mathrm{der}}$ is simply connected for all $(H, s, \eta)$ in $\mathfrak{E}$ or $\mathfrak{E}_v$. This will greatly simplify the theory.

For each $(H, s, \eta) \in \mathfrak{E}$, the constant $i(G, H)$ is defined as

$$\tau(G)\tau(H)^{-1}\lambda(H, s, \eta),$$

where $\tau(G)$ and $\tau(H)$ are the Tamagawa numbers and $\lambda(H, s, \eta)$ is the order of the outer automorphism group of the triple $(H, s, \eta)$. See [Har11, §5] for details.

Let $(H, s, \eta) \in \mathfrak{E}$. The notions of $(G, H)$-regular semisimple elements in $H(\mathbb{Q})$, and stable orbital integrals of them, are discussed in [Har11, §5]. Recall from [Har11, p. 30] or [Kot90, p. 189] that the *$(G, H)$-regular elliptic part of the stable trace formula for $H$* is the distribution

$$C_c^\infty(H(\mathbb{A})) \ni h \mapsto ST_e^{H,*}(h) := \tau(H) \sum_{\gamma_H} SO_{\gamma_H}(h), \tag{6.1.3}$$

where

- $\gamma_H$ runs through a set of representatives of the stable conjugacy classes in $H(\mathbb{Q})$ that are elliptic and $(G, H)$-regular.
- each $SO_{\gamma_H}(h)$ is the stable orbital integral of $h$ over $\gamma_H$, defined for instance on p. 21 of [Har11].

At this point we have described all the ingredients on the right hand side of (6.1.2) except $\phi^H$. The function $\phi^H : H(\mathbb{A}) \to \mathbb{C}$ is defined to be any *Langlands–Shelstad transfer* of $\phi$.

The desired stabilization of $T'(j, f^{p,\infty})$ is of the same form as (6.1.2), except that the test function on each $H$ is defined differently.

**Theorem 6.1.2** ([Kot90, Theorem 7.2])**.**

$$T'(j, f^{p,\infty}) = \sum_{(H,s,\eta)\in\mathfrak{E}} i(G, H) ST_e^{H,*}(f_H), \tag{6.1.4}$$

*where for each $(H, s, \eta) \in \mathfrak{E}$ the function $f_H : H(\mathbb{A}) \to \mathbb{C}$ is a product $f_H = f_{H,\infty} f_{H,p} f_H^{p,\infty}$. The factor $f_H^{p,\infty} \in C_c^\infty(H(\mathbb{A}_f^p))$ is any Langlands–Shelstad transfer of $f^{p,\infty}$ (with respect to certain normalizations to be specified below). The factors $f_{H,p} \in C_c^\infty(H(\mathbb{Q}_p))$ and $f_{H,\infty} \in C^\infty(H(\mathbb{Q}_p))$ are to be explicitly specified below. The meanings of $i(G, H)$ and $ST_e^{H,*}(\cdot)$ are the same as in (6.1.2).*

*Remark* 6.1.3. The main point here is that the functions $f_{H,p}, f_{H,\infty}$ are specified in an explicit way that is **independent** of $f^{p,\infty}$.

[8] In general, when $G^{\mathrm{der}}$ is not simply connected, the notion of an *endoscopic datum* in the sense of [LS87] refers to a suitable quadruple $(H, \mathcal{H}, s, \eta : \mathcal{H} \to {}^L G)$. Our endoscopic triple $(H, s, \eta)$ equipped with the chosen extension of $\eta$ amounts to an endoscopic datum $(H, \mathcal{H}, s, \eta : \mathcal{H} \to {}^L G)$ equipped with an isomorphism $\mathcal{H} \cong {}^L H$. Such an object is sometimes called an *extended endoscopic datum* in the literature, e.g. [Kal11].

*Remark* 6.1.4. The function $f_{H,\infty}$ is not compactly supported, but it has compact support modulo center and transforms under the split component of the center by a character. In particular the same definition (6.1.3) of $ST_e^{H,*}(f_H)$ still makes sense.

*Remark* 6.1.5. Even if we drop the $(G,H)$-regular condition and allow all elliptic elements in the summation (6.1.3), the equality (6.1.4) still holds. In fact, the function $f_{H,\infty}$ to be specified below will have the property that the resulting $f_H$ satisfies $SO_{\gamma_H}(f_H)=0$ for all elliptic non-$(G,H)$-regular $\gamma_H$. See the remark below [CHL11, (4.3.5)] and Remark 6.3.5.

*Remark* 6.1.6. As we mentioned in the introduction, Kottwitz's proof of Theorem 6.1.2 was conditional on certain hypotheses in local harmonic analysis, including the existence of Langlands–Shelstad transfer for the statement to even make sense. These hypotheses have been proved now, as we shall see later.

The rest of this chapter is devoted to explaining the definition of $f_H$ and the proof of Theorem 6.1.2. For brevity we write $T'$ for $T'(j,f^{p,\infty})$. In the following we do not need to assume that we are in the special case of unitary Shimura varieties. All we need is that $(G,X)=(G,h^{-1})$ is a Shimura datum, satisfying the two simplifying conditions in Remark 2.1.1 and the simplifying condition in Remark 6.1.1. (It is also not necessary to assume that $G$ is anisotropic over $\mathbb{Q}$.) Of course in this generality we do not necessarily know the truth of the unstabilized Langlands–Kottwitz formula,[9] but we may and will take (6.1.1) as the definition of $T'$ and still prove Theorem 6.1.2.

## 6.2. **Preliminary steps.**

**Definition 6.2.1.** For $H$ a reductive group over $\mathbb{Q}$ or $\mathbb{Q}_v$, we denote by $A_H$ the maximal split torus in the center of $H$, over $\mathbb{Q}$ or $\mathbb{Q}_v$ respectively.

In particular, by Remark 2.1.1 we have $(A_G)_{\mathbb{R}}=A_{G_{\mathbb{R}}}$.

**Lemma 6.2.2.** *Let $(\gamma_0,\gamma,\delta)$ be an $\mathbb{F}_{p^n}$-Kottwitz triple such that $\alpha(\gamma_0,\gamma,\delta)$ is trivial. The quantity $c(\gamma_0,\gamma,\delta)$ defined in (5.1.2) is equal to*

$$\tau(G)\,|\mathfrak{K}(\gamma_0/\mathbb{Q})|\,\mathrm{vol}(A_G(\mathbb{R})^0\backslash I_0^{comp}(\mathbb{R}))^{-1}.$$

*Here $I_0:=G_{\gamma_0}$ and $I_0^{comp}$ is the compact-mod-center inner form of $I_{0,\mathbb{R}}$. Note that $I_0^{comp}$ exists because $\gamma_0$ is $\mathbb{R}$-elliptic, and is unique as an isomorphism class of inner forms of $I_{0,\mathbb{R}}$. The Haar measure on $I_0^{comp}(\mathbb{R})$ being used in the definition of $\mathrm{vol}(A_G(\mathbb{R})^0\backslash I_0^{comp}(\mathbb{R}))$ is characterized by its transfer to $I_0(\mathbb{R})$; we normalize the latter such that its product with the chosen Haar measure on $I_0(\mathbb{A}_f)$ is the Tamagawa measure on $I_0(\mathbb{A})$.*

*Proof.* Let $I$ be the inner form of $I_0$ as in §5.1. Then $I_0^{comp}$ can be realized as $I_{\mathbb{R}}$. We have

$$\mathrm{vol}(I(\mathbb{Q})\backslash I(\mathbb{A}_f))=\tau(I)\,\mathrm{vol}(A_G(\mathbb{R})^0\backslash I_0^{comp}(\mathbb{R}))^{-1}.$$

Moreover $\tau(I)=\tau(I_0)$ by [Kot88]. Hence we have

$$c(\gamma_0,\gamma,\delta)=\left|\ker(\ker^1(\mathbb{Q},I_0)\to\ker^1(\mathbb{Q},G))\right|\tau(I_0)\,\mathrm{vol}(A_G(\mathbb{R})^0\backslash I_0^{comp}(\mathbb{R}))^{-1}.$$

By [Kot84] and [Kot88], we have

$$\tau(G)=\left|\ker^1(\mathbb{Q},G)\right|^{-1}\left|\pi_0(Z(\widehat{G})^\Gamma)\right|,$$

and similarly for $G$ replaced by $I_0$. Thus it suffices to prove

$$\frac{\left|\ker(\ker^1(\mathbb{Q},I_0)\to\ker^1(\mathbb{Q},G))\right|\left|\ker^1(\mathbb{Q},G)\right|\left|\pi_0(Z(\widehat{I}_0)^\Gamma)\right|}{\left|\ker^1(\mathbb{Q},I_0)\right|\left|\pi_0(Z(\widehat{G})^\Gamma)\right|\left|\mathfrak{K}(I_0/\mathbb{Q})\right|}=1.$$

But this follows from the exact sequence

$$\ker^1(\mathbb{Q},I_0)\to\ker^1(\mathbb{Q},G)\to\mathfrak{K}(I_0/\mathbb{Q})^D\to X^*(\pi_0(Z(\widehat{I}_0)^\Gamma))\to X^*(\pi_0(Z(\widehat{G})^\Gamma))\to 0,$$

cf. [Kot86, p. 395] for the dual version. The exactness at $X^*(\pi_0(Z(\widehat{G})^\Gamma))$ follows from the assumption that $\gamma_0$ is elliptic. □

[9] See §7 for a discussion of the known cases.

Consequently we have

$$T' = \sum_{\substack{(\gamma_0,\gamma,\delta)\\ \alpha(\gamma_0,\gamma,\delta)=1}} \tau(G)\, |\mathfrak{K}(\gamma_0/\mathbb{Q})| \operatorname{vol}(A_G(\mathbb{R})^0 \backslash I_0^{comp}(\mathbb{R}))^{-1} O_\gamma(f^{p,\infty}) TO_\delta(\phi_n) \operatorname{Tr}_\xi(\gamma_0).$$

We now insert the global Kottwitz sign

$$1 = e(I)$$

into each summand, where $I$ is the inner form of $I_0$ associated to $(\gamma_0, \gamma, \delta)$ as in §5.1. We have a product formula

$$e(I) = \prod_v e(I_{\mathbb{Q}_v})$$

with the local Kottwitz signs (see [Kot83]), where

- The sign $e(I_{\mathbb{R}}) = e(I_0^{comp})$ depends only on $\gamma_0$.
- For a finite place $v \neq p$, the sign $e(I_{\mathbb{Q}_v}) = e(G_{\gamma_v})$ depends only on the $v$-component $\gamma_v$ of $\gamma$. We write $e(\gamma_v)$ for it. Almost all of them are equal to 1. Write $e(\gamma)$ for $\prod_{v \neq p,\infty} e(\gamma_v)$.
- The sign $e(I_{\mathbb{Q}_p}) = e(G_\delta^\sigma)$ depends only on $\delta$. We write $e(\delta)$ for it.

Denote

$$\bar{v}(\gamma_0) = \bar{v}(I_0) := \operatorname{vol}(A_G(\mathbb{R})^0 \backslash I_0^{comp}(\mathbb{R})) e(I_0^{comp}). \tag{6.2.1}$$

Thus we have

$$T' = \sum_{\substack{(\gamma_0,\gamma,\delta)\\ \alpha(\gamma_0,\gamma,\delta)=1}} \tau(G)\, |\mathfrak{K}(\gamma_0/\mathbb{Q})|\, \bar{v}(\gamma_0)^{-1} e(\gamma) e(\delta) O_\gamma(f^{p,\infty}) TO_\delta(\phi_n) \operatorname{Tr}_\xi(\gamma_0).$$

Note that in the above formula, each summand makes sense even if $\alpha(\gamma_0, \gamma, \delta) \neq 1$. Therefore by Fourier inversion on the finite abelian group $\mathfrak{K}(\gamma_0)$, we have

$$T' = \tau(G) \sum_{\gamma_0} \sum_{\kappa \in \mathfrak{K}(\gamma_0/\mathbb{Q})} \bar{v}(\gamma_0)^{-1} \operatorname{Tr}_\xi(\gamma_0) \sum_{(\gamma,\delta)} e(\gamma) e(\delta) O_\gamma(f^{p,\infty}) TO_\delta(\phi_n) \langle \alpha(\gamma_0, \gamma, \delta), \kappa \rangle, \tag{6.2.2}$$

where

- $\gamma_0$ runs through a set of representatives for the $\mathbb{R}$-elliptic stable conjugacy classes in $G(\mathbb{Q})$.
- $\gamma$ runs through a set of representatives for the conjugacy classes in $G(\mathbb{A}_f^p)$ that are stably conjugate to $\gamma_0$.
- $\delta$ runs through a set of representatives for the $\sigma$-conjugacy classes in $G(L_n)$ of which $\gamma_0$ is an $n$-th norm, and which satisfy the condition (3) in Definition 4.1.1.
- In the summation there are only finitely many nonzero terms.

The identity (6.2.2) is [Kot90, (4.2)]. Note the similarity between (6.2.2) and [Har11, (4.8)], but also note that the places $\infty$ and $p$ play very different roles in the two situations.

6.3. **The archimedean place.** Let $(H, s, \eta) \in \mathfrak{E}_\infty$. We assume that *an elliptic maximal torus of* $G_{\mathbb{R}}$ *comes from* $H$. We explain the meaning of this assumption.

Throughout the whole theory of endoscopy for the group $G_{\mathbb{R}}$, we need to fix a quasi-split reductive group $G^*$ over $\mathbb{R}$ together with an inner twisting $\psi : G^* \to G_{\mathbb{R}}$. For any maximal torus $T'$ of $H_{\mathbb{C}}$, there is a canonical $G^*(\mathbb{C})$-conjugacy class $\mathscr{E}_{T'}$ of embeddings $T' \to G^*_{\mathbb{C}}$. Now suppose $T_H \subset H$ is a maximal torus defined over $\mathbb{R}$. By an *admissible embedding* of $T_H$ into $G_{\mathbb{R}}$, we mean an embedding $j : T_H \hookrightarrow G_{\mathbb{R}}$ defined over $\mathbb{R}$, which is the composition of an element in $\mathscr{E}_{T_{H,\mathbb{C}}}$ with $\psi$. In general $T_H$ may not have any admissible embedding into $G_{\mathbb{R}}$. For more details see [LS87, §1] or [Ren11, §2].

We can now explain the meaning of our assumption: We require that there exists a maximal torus $T_H$ of $H$ defined over $\mathbb{R}$, and an admissible embedding $j : T_H \hookrightarrow G_{\mathbb{R}}$, such that $T_G := j(T_H)$ is an **elliptic** maximal torus of $G_{\mathbb{R}}$. In particular, $T_H$ is an elliptic maximal torus of $H$ and $A_H \cong A_{G_{\mathbb{R}}}$. Note that the existence of elliptic maximal tori in $H$ is already a non-trivial assumption on $H$ in general. We fix such a triple $(T_H, T_G, j)$.

The definition of $f_{H,\infty}$ depends on the choice of $(T_H, T_G, j)$ as above, as well as the choice of a Borel $B_{G,H}$ of $G_{\mathbb{C}}$ containing $T_{G,\mathbb{C}}$. Let $(T_H, T_G, j, B_{G,H})$ be fixed. In the following we denote this quadruple simply by $(j, B_{G,H})$. Then $(j, B_{G,H})$ determines an embedding of root systems

$$j_* : R(H_{\mathbb{C}}, T_{H,\mathbb{C}}) \to R(G_{\mathbb{C}}, T_{G,\mathbb{C}}),$$

and determines a Borel $B_H$ of $H_{\mathbb{C}}$ containing $T_H$ by pulling back the $B_{G,H}$-positive roots along $j_*$. Via $j_*$ we view the Weyl group $\Omega_H := \Omega(H_{\mathbb{C}}, T_{H,\mathbb{C}})$ of $H_{\mathbb{C}}$ as a subgroup of the Weyl group $\Omega := \Omega(G_{\mathbb{C}}, T_{G,\mathbb{C}})$ of $G_{\mathbb{C}}$. We also have a subset $\Omega_*$ of $\Omega$, consisting of elements $\omega \in \Omega$ such that $(\omega \circ j, B_{G,H})$ still determines the Borel $B_H$. The multiplication induces a bijection

$$\Omega_H \times \Omega_* \to \Omega$$

known as the *Kostant decomposition*.

Let $\xi^*$ be the contragredient representation of $\xi$. Let $\varphi_{\xi^*}$ be the elliptic Langlands parameter of $G_{\mathbb{R}}$ corresponding to $\xi^*$, so that the L-packet of $\varphi_{\xi^*}$ is the packet of (essentially) discrete series representations of $G(\mathbb{R})$ whose central and infinitesimal characters are the same as those of $\xi^*$. Let $\Phi_H(\varphi_{\xi^*})$ be the set of equivalence classes of elliptic Langlands parameters $\varphi_H$ of $H$ such that $\eta \circ \varphi_H$ is equivalent to $\varphi_{\xi^*}$, where $\eta$ is the fixed L-embedding ${}^L H \to {}^L G$. As on p. 185 of [Kot90], we have a bijection

$$\Phi_H(\varphi_{\xi^*}) \xrightarrow{\sim} \Omega_*$$
$$\varphi_H \mapsto \omega_*(\varphi_H),$$

characterized by the condition that $\varphi_H$ is *aligned* with

$$(\omega_*(\varphi_H)^{-1} \circ j, B_{G,H}, B_H).$$

We refer to [Kot90, pp. 184-185] for the meaning of the last condition. For general expositions of the Langlands–Shelstad theory of L-packets of discrete series representations over $\mathbb{R}$, we refer the reader to [Ada11] or [Taï14, §4.2.1] and the references therein.

For any $\varphi_H \in \Phi_H(\varphi_{\xi^*})$, define the *averaged pseudo-coefficient*

$$f_{\varphi_H} := |\Pi(\varphi_H)|^{-1} \sum_{\pi \in \Pi(\varphi_H)} f_\pi,$$

where $\Pi(\varphi_H)$ is the L-packet of $\varphi_H$, and $f_\pi \in C^\infty(H(\mathbb{R}))$ is a normalized pseudo-coefficient for $\pi$, constructed in [CD85]. Note that the normalization of $f_\pi$ depends on the choice of a Haar measure on $H(\mathbb{R})$. We also know that each $f_\pi$, and hence $f_{\varphi_H}$, transform under $A_H(\mathbb{R})^0 \cong A_{G_{\mathbb{R}}}(\mathbb{R})^0$ by the central character $\chi^{-1}$ of $\xi^*$.

**Definition 6.3.1.** Fix the choice of $(j, B_{G,H})$ as above and fix a Haar measure on $H(\mathbb{R})$. Define

$$f_{H,\infty} := (-1)^{q(G_{\mathbb{R}})} \langle \mu_{T_G}, s\rangle_j \sum_{\varphi_H \in \Phi_H(\varphi_{\xi^*})} \det(\omega_*(\varphi_H)) f_{\varphi_H}.$$

Here $\mu_{T_G}$ is the Hodge cocharacter of any $G(\mathbb{R})$-conjugate of $h$ that factors through $T_G$. The pairing $\langle \mu_{T_G}, s\rangle_j \in \mathbb{C}^\times$ is defined as follows: We use $j$ to view $\mu_{T_G}$ as a cocharacter of $T_H$, and view $s$ as an element of $Z(\widehat{H})$. Then we use the canonical pairing

$$X_*(T_H) \times Z(\widehat{H}) \to X^*(Z(\widehat{H})) \times Z(\widehat{H}) \to \mathbb{C}^\times$$

to define $\langle \mu_{T_G}, s\rangle_j$. The integer $q(G_{\mathbb{R}})$ is half the real dimension of the symmetric space of $G(\mathbb{R})$.

On the other hand, it is a result of Shelstad (reviewed by Kottwitz on p. 184 of [Kot90]) that for the chosen $(j, B_{G,H})$, there is a normalization of the transfer factors between $H$ and $G_{\mathbb{R}}$ (see [Har11]) denoted by $\Delta_{j,B_{G,H}}$, such that for all $(G,H)$-regular elements $\gamma_H \in T_H(\mathbb{R})$ the values $\Delta_{j,B_{G,H}}(\gamma_H, j(\gamma_H))$ are given by an explicit formula. We do not record this explicit formula, but only remark that it involves the Langlands correspondence for tori over $\mathbb{R}$.

*Remark* 6.3.2. In practice, it is sometimes necessary to compare the normalization $\Delta_{j,B_{G,H}}$ with other normalizations. For example, we obtain another normalization by realizing $G_{\mathbb{R}}$ as a pure or rigid inner form of $G^*_{\mathbb{R}}$ and by choosing a Whittaker datum on $G^*_{\mathbb{R}}$, see the introduction of [Kal16]. In principle, such a comparison can be done by comparing the shapes of the associated endoscopic character identities. See [Zhu18] for a special case of such a computation.

Definition 6.3.1 is motivated by the following computation of stable orbital integrals:

**Proposition 6.3.3** ([Kot90, (7.4)])**.** *Let $\gamma_H \in H(\mathbb{R})$ be a semi-simple $(G,H)$-regular element. If $\gamma_H$ is not elliptic (i.e. not in any elliptic maximal torus), then $SO_{\gamma_H}(f_{H,\infty}) = 0$. Suppose $\gamma_H$ is elliptic. Then*

$$SO_{\gamma_H}(f_{H,\infty}) = \langle \mu_{T_G}, s\rangle_j \ \Delta_{j,B_{G,H}}(\gamma_H, \gamma_0)\bar{v}(\gamma_0)^{-1}\operatorname{Tr}_\xi(\gamma_0).$$

*Here $\gamma_0 \in G(\mathbb{R})$ is any element such that $\gamma_H$ is an image of $\gamma_0$ (see* [Har11, §5] *for the notion of "image"). The term $\bar{v}(\gamma_0)$ is defined as in (6.2.1), with respect to the Haar measure transferred from the one on $H_{\gamma_H}(\mathbb{R})$ (where $H_{\gamma_H}$ is an inner form of $I_0^{comp}$) that is used to define $SO_{\gamma_H}$.* □

*Remark* 6.3.4. Note that $\gamma_0$ in the above proposition exists. In fact, since all elliptic maximal tori of $H$ are conjugate under $H(\mathbb{R})$, we may first conjugate $\gamma_H$ into $T_H$ and then map it to $G$ via $j$, and the image is an example of $\gamma_0$.

*Remark* 6.3.5. As a special case of an observation of Morel [Mor11, Remarque 3.2.6], we have $SO_{\gamma_H}(f_{H,\infty}) = 0$ for any semi-simple $\gamma_H \in H(\mathbb{R})$ that is not $(G,H)$-regular.

6.4. **Away from $p,\infty$.** Away from $p,\infty$, the ingredients we need from local harmonic analysis are the same as those in the stabilization of $T_e$ discussed in [Har11], namely the Langlands–Shelstad Transfer Conjecture and the Fundamental Lemma. They are now proved theorems thanks to the work of Ngô [Ngô10], Waldspurger [Wal97] [Wal06], Cluckers-Loeser [CL10], and Hales [Hal95]. See [Har11, §6] for more historical remarks. In the following we state these two proved conjectures and their adelic consequence. The statements get simplified by the fact that $G$ and all its local and global elliptic endoscopic groups $H$ have simply connected derived groups in our specific case.

**Theorem 6.4.1** (Langlands–Shelstad Transfer Conjecture)**.** *Let $v$ be a finite place of $\mathbb{Q}$ and let $(H,s,\eta) \in \mathfrak{E}_v$. Fix a normalization of the transfer factors between $H$ and $G_{\mathbb{Q}_v}$. Fix Haar measures on $G(\mathbb{Q}_v)$ and $H(\mathbb{Q}_v)$. For any $f \in C_c^\infty(G(\mathbb{Q}_v))$, there exists $f^H \in C_c^\infty(H(\mathbb{Q}_v))$, called a* Langlands–Shelstad transfer *of $f$, satisfying the following properties. For any semi-simple $(G,H)$-regular element $\gamma_H \in H(\mathbb{Q}_v)$, we have*

$$SO_{\gamma_H}(f^H) = \begin{cases} 0, & \gamma_H \text{ is not an image from } G(\mathbb{Q}_v) \\ \Delta(\gamma_H,\gamma)O_\gamma^\kappa(f), & \gamma_H \text{ is an image of a semisimple } \gamma \in G(\mathbb{Q}_v) \end{cases} \tag{6.4.1}$$

*where in the second situation,*

- *the element*
$$\kappa \in \mathfrak{K}(G_\gamma/\mathbb{Q}_v) = \pi_0(Z(\widehat{G_\gamma})^{\Gamma_v})/\pi_0(Z(\widehat{G})^{\Gamma_v})$$
*is the natural image of $s$ (see* [Kot86]*). The $\kappa$-orbital integral $O_\gamma^\kappa$ is defined as in* [Kot86, §5] *or* [Har11, (4.9)].
- *$SO_{\gamma_H}(f^H)$ and $O_\gamma^\kappa(f)$ are defined with respect to the fixed Haar measures on $G(\mathbb{Q}_v)$ and $H(\mathbb{Q}_v)$, and compatible Haar measures on $G_\gamma(\mathbb{Q}_v)$ and $H_{\gamma_H}(\mathbb{Q}_v)$.*
- *$\Delta(\gamma_H,\gamma) \in \mathbb{C}$ is the transfer factor.*

□

**Theorem 6.4.2** (Fundamental Lemma)**.** *In the situation of Theorem 6.4.1, assume $G$ is unramified over $\mathbb{Q}_v$. Also assume $(H,s,\eta)$ is unramified, meaning that $H$ is unramified over $\mathbb{Q}_v$ and that $\eta : {}^LH \to {}^LG$ is unramified at $v$, in the sense that the restriction of $\eta$ to the inertia group at $v$ (as a subgroup of ${}^LH$) is trivial. Fix Haar measures on $G(\mathbb{Q}_v)$ and $H(\mathbb{Q}_v)$ giving volume 1 to hyperspecial subgroups. Let $K$ (resp. $K_H$) be an arbitrary hyperspecial subgroup of $G(\mathbb{Q}_v)$ (resp. $H(\mathbb{Q}_v)$). Then $1_{K_H}$ is a Langlands–Shelstad transfer of $1_K$ as in Theorem 6.4.1, for the canonical unramified normalization of transfer factors defined in* [Hal93]. □

We now state an adelic consequence of the above theorems. Let $(H,s,\eta) \in \mathfrak{E}$. Then it is unramified at almost all places. Suppose for almost all places $v$ at which $G$ and $(H,s,\eta)$ are unramified, that the transfer factors at $v$ are normalized under the canonical unramified normalization.

**Corollary 6.4.3.** *For any $f \in C_c^\infty(G(\mathbb{A}_f^p))$, there exists $f^H \in C_c^\infty(H(\mathbb{A}_f^p))$ such that the $\mathbb{A}_f^p$-analogue of (6.4.1) holds. Here the meaning of an adelic $(G,H)$-regular element is as on pp. 178-179 of* [Kot90]*, and all the orbital integrals are defined with respect to adelic Haar measures.* □

*Remark* 6.4.4. The original Langlands–Shelstad Transfer Conjecture is only stated for $G$-regular elements. The more general conjecture with $(G, H)$-regular elements is reduced to that version by Langlands–Shelstad [LS90, §2.4]. The same remark applies to the Fundamental Lemma.

**Definition 6.4.5.** Let $(H, s, \eta) \in \mathfrak{E}$. Assume that an elliptic maximal torus of $G_{\mathbb{R}}$ comes from $H_{\mathbb{R}}$. Also assume that the reductive group $H$ is unramified over $\mathbb{Q}_p$ (but we do not assume that $\eta$ is unramified over $\mathbb{Q}_p$). Fix the Haar measure on $G(\mathbb{A}_f^p)$ giving volume 1 to $K^p$. Fix the Haar measure on $H(\mathbb{A}_f^p)$ such that its product with the chosen Haar measure on $H(\mathbb{R})$ in Definition 6.3.1 and the Haar measure on $H(\mathbb{Q}_p)$ giving volume 1 to hyperspecial subgroups is equal to the Tamagawa measure on $H(\mathbb{A})$. Normalize the transfer factors between $H$ and $G$ at all places away from $p$ and $\infty$ in a way that satisfies the hypothesis in Corollary 6.4.3. We define $f_H^{p,\infty} \in C_c^\infty(H(\mathbb{A}_f^p))$ to be a Langlands–Shelstad transfer of $f^{p,\infty}$ as in Corollary 6.4.3.

6.5. **The place $p$.** At the place $p$, the ingredient needed from local harmonic analysis is a variant of the Fundamental Lemma. It is a special case of the Twisted Fundamental Lemma, in the situation known as the "unstable cyclic base change". We shall state it in the form needed here, after recalling some facts. Again, the statement is simplified by the fact that $G$ and all its endoscopic groups $H$ over $\mathbb{Q}_p$ have simply connected derived groups.

Let $k$ be a finite extension of $\mathbb{Q}_p$. Let $H_1$ be an unramified reductive group over $k$. Then by the Satake isomorphism we know that the Hecke algebras $\mathcal{H}(H_1(k)//K)$, for different hyperspecial subgroups $K$ of $H_1(k)$, are canonically isomorphic to each other. We denote them by the common notation $\mathcal{H}^{\mathrm{ur}}(H_1(k))$, called the *unramified Hecke algebra*. Moreover, let $H_2$ be another such reductive group and let ${}^L H_1 \to {}^L H_2$ be an unramified L-morphism. Then we have an associated algebra map $\mathcal{H}^{\mathrm{ur}}(H_2(k)) \to \mathcal{H}^{\mathrm{ur}}(H_1(k))$, defined using the Satake isomorphism. We refer the reader to [Car79], [Bor79], [HR10], [ST16] for expositions of the Satake isomorphism.

Fix $(H, s, \eta) \in \mathfrak{E}_p$. Assume temporarily the following two conditions:

(1) $(H, s, \eta)$ is unramified.

(2) The element $s$, when viewed as an element of $Z(\widehat{H})$, is fixed by $\Gamma_p$.

Let $R := \mathrm{Res}_{L_n/\mathbb{Q}_p} G$. Then $R$ is unramified over $\mathbb{Q}_p$. On pp. 179-180 of [Kot90], Kottwitz uses $\eta : {}^L H \to {}^L G$ to construct an unramified L-morphism ${}^L H \to {}^L R$, which will serve as a component of a *twisted endoscopic datum.* We then obtain a map

$$b : \mathcal{H}^{\mathrm{ur}}(R(\mathbb{Q}_p)) = \mathcal{H}^{\mathrm{ur}}(G(L_n)) \to \mathcal{H}^{\mathrm{ur}}(H(\mathbb{Q}_p)).$$

**Theorem 6.5.1** (Fundamental Lemma for Unstable Cyclic Base Change)**.** *Let $(H, s, \eta) \in \mathfrak{E}_p$ be an element that satisfies the above two temporary assumptions. For any $f \in \mathcal{H}^{\mathrm{ur}}(G(L_n))$ and for any semi-simple $(G, H)$-regular element $\gamma_H \in H(\mathbb{Q}_p)$, we have*

$$SO_{\gamma_H}(b(f)) = \sum_{\delta} \langle \beta(\gamma_0, \delta), s\rangle \Delta(\gamma_H, \gamma_0) e(\delta) TO_\delta(f), \tag{6.5.1}$$

*where*

- *$\gamma_0 \in G(\mathbb{Q}_p)$ is any semi-simple element such that $\gamma_H$ is an image of it. It always exists because $G_{\mathbb{Q}_p}$ is quasi-split.*
- *$\delta$ runs through a set of representatives for the $\sigma$-conjugacy classes in $G(L_n)$ such that $\gamma_0$ is an $n$-th norm of $\delta$.*
- *$\beta(\gamma_0, \delta)$ is the character $Z(\widehat{G_{\gamma_0}})^{\Gamma_p} Z(\widehat{G}) \to \mathbb{C}^\times$ defined in the same way as $\beta_p$ at the end of §4.1. The pairing $\langle \beta(\gamma_0, \delta), s\rangle$ is defined in the obvious way.*
- *$\Delta(\gamma_H, \gamma_0)$ is the transfer factor under the canonical unramified normalization.*
- *$\mathrm{SO}_{\gamma_H}$ and $TO_\delta$ are defined with respect to the Haar measures on $H(\mathbb{Q}_p)$ and $G(L_n)$ giving volume 1 to hyperspecial subgroups, and compatible Haar measures on $H_{\gamma_H}(\mathbb{Q}_p)$ and $G_\delta^\sigma(\mathbb{Q}_p)$.*

*Proof.* This is [Kot90, (7.3)] (with a more general function $f$ here). In [Kot90] it is shown that [Kot90, (7.3)] is equivalent to [Kot90, (7.2)]. In [Mor10b, Appendix A], Kottwitz reduces [Kot90, (7.2)] to a **special case** of the usual form of the Twisted Fundamental Lemma, and he also reduces the $(G, H)$-regular case to the $G$-regular case. In [Mor10b, §9], Morel shows how the desired statement can be further reduced to the Twisted Fundamental Lemma for the unit element in the unramified Hecke algebra.

The general Twisted Fundamental Lemma is known for the unit element in the unramified Hecke algebra for large $p$, thanks to the work of Ngô [Ngô10], Waldspurger [Wal06] [Wal08], Cluckers-Loeser [CL10] and others. By the recent work of Lemaire, Moeglin, and Waldspurger [LMW15] [LW15], it is also known for general elements of the unramified Hecke algebra, including the case of small $p$. □

**Corollary 6.5.2.** *Let $(H, s, \eta) \in \mathfrak{E}_p$. We assume that the reductive group $H$ is unramified over $\mathbb{Q}_p$, but we drop the temporary assumptions (1) (2) above. Fix an arbitrary normalization of transfer factors between $H$ and $G_{\mathbb{Q}_p}$. For any $f \in \mathcal{H}^{\mathrm{ur}}(G(L_n))$, there exists a function $b'(f) \in \mathcal{H}^{\mathrm{ur}}(H(\mathbb{Q}_p))$, such that the conclusion of Theorem 6.5.1 holds for $b'(f)$.*

*Proof.* This follows from Theorem 6.5.1, see [Kot90, p. 181]. □

**Definition 6.5.3.** Let $(H, s, \eta) \in \mathfrak{E}$. Assume that $H$ is unramified over $\mathbb{Q}_p$ and that an elliptic maximal torus of $G_{\mathbb{R}}$ comes from $H_{\mathbb{R}}$. Take the normalization of transfer factors between $H_{\mathbb{Q}_P}$ and $G_{\mathbb{Q}_p}$ such that its product with the normalization $\Delta_{j,B_{G,H}}$ at $\infty$ and the normalization away from $p, \infty$ in Definition 6.4.5 satisfies the global product formula (see [LS87, §6]). We define $f_{H,p}$ to be $b'(\phi_n)$ as in Corollary 6.5.2, where $\phi_n$ is defined in §5.1.

6.6. **The final step.** Let $(H, s, \eta) \in \mathfrak{E}$. If it does not satisfy the assumptions in Definition 6.5.3, we define $f_H := 0$. Otherwise we define $f_H := f_{H,\infty} f_H^{p,\infty} f_{H,p}$, with $f_{H,\infty}$, $f_H^{p,\infty}$, $f_{H,p}$ defined in Definitions 6.3.1, 6.4.5, 6.5.3 respectively.

Consider a pair $(\gamma_0, \kappa)$ appearing on the right hand side of (6.2.2). Assume that the summand indexed by it is non-zero. Recall from [Kot86, §9] or [Har11, §5] that $(\gamma_0, \kappa)$ determines a set of *endoscopic quadruples* $\pi^{-1}(\gamma_0, \kappa)$. Each element of $\pi^{-1}(\gamma_0, \kappa)$ is an *endoscopic quadruple* of the form $(H, s, \eta, \gamma_H)$, with $(H, s, \eta) \in \mathfrak{E}$ and $\gamma_H \in H(\mathbb{Q})$ a semi-simple $(G, H)$-regular element that is an image of $\gamma_0$. In particular, $\gamma_H$ is $\mathbb{R}$-elliptic, and an elliptic maximal torus of $G_{\mathbb{R}}$ comes from $H_{\mathbb{R}}$. Moreover it can be checked that $H_{\mathbb{Q}_p}$ is necessarily unramified, as on p. 189 of [Kot90]. Assembling Proposition 6.3.3, Corollary 6.4.3, Corollary 6.5.2, and using the global product formula of transfer factors, it is easy to see that $SO_{\gamma_H}(f_H)$ is equal to the summand in (6.2.2) indexed by $(\gamma_0, \kappa)$. Then by the same counting argument as in [Har11, §5] and the following easy observation, we know that the right hand side of (6.2.2) is equal to the right hand side of (6.1.4), and Theorem 6.1.2 is proved. The observation is that if $\delta \in G(L_n)$ is an element that violates condition (3) in Definition 4.1.1, then $TO_\delta(\phi_n) = 0$.

Q.E.D. for Theorem 6.1.2.

## 7. Final remarks

7.1. **What is after stabilization?** The next step after Theorem 6.1.2 may be called "destabilization". This amounts to expanding the right hand side of (6.1.4) spectrally, in terms of automorphic representations, and then rewriting everything in terms of $G$ alone, so that the endoscopic groups $H$'s no longer appear. In this way, one would obtain a description of the cohomology of the Shimura variety in terms of the conjectural relation between Galois representations and automorphic representations. In fact, a suitable implementation of the destabilization step could be used to construct Galois representations attached to automorphic representations.

Kottwitz outlines in [Kot90, Part II] the destabilization, assuming Arthur's multiplicity conjectures and some related hypotheses on the stable trace formula. For the special case of unitary Shimura varieties, a lot has been known with the aid of the base change of automorphic representations from the unitary group to $\mathrm{GL}_n$. We refer the reader to [Mor10b], [Shi11], and [CHL11].

For cases beyond the unitary groups, the hypotheses made in [Kot90, Part II] have been largely verified for quasi-split classical groups, by Arthur [Art13] (for symplectic and special orthogonal groups) and Mok [Mok15] (for unitary groups). Extensions of these results to non-quasi-split inner forms are available in certain cases, see [KMSW14] and [Taï15], but the general case is still open. We remark that for the theory of Shimura varieties it is usually insufficient to understand only quasi-split groups, as the existence of a Shimura datum often forces the group to be non-quasi-split already at infinity. In [Zhu18], the results from [Art13] and [Taï15] are applied to the problem of destabilization for some special cases of orthogonal Shimura varieties, where the special orthogonal groups in question have signature $(n, 2)$ at infinity and are not quasi-split.

7.2. **What is known beyond PEL type?** An important circle of ideas that is closely related to Kottwitz's work [Kot92] is the Langlands–Rapoport conjecture [LR87] (later corrected by Reimann [Rei97]). This conjecture gives a group theoretic description of the reduction of the Shimura variety at hyperspecial primes. It is essentially shown in [LR87] that this conjecture implies the unstabilized Langlands–Kottwitz formula (5.3.1) (cf. Remark 5.3.4). Recently, a modified version of the Langlands–Rapoport conjecture has been proved by Kisin [Kis17] for all abelian-type Shimura varieties. See [Kis17] for more historical remarks and references. In an ongoing project [KSZ], it is expected that the results in [Kis17] can be refined in such a way as to imply the generalization of the unstabilized Langlands–Kottwitz formula for all abelian-type Shimura varieties. Also the stabilization procedure is carried out in [KSZ] in this generality, with the usual group theoretic assumptions removed. Thus the expected main result of [KSZ] will generalize the stabilized formula (6.1.4) from PEL type to abelian type.

7.3. **What is beyond the compact and good reduction case?** Kottwitz's work has also been generalized to some non-compact Shimura varieties, where the intersection cohomology of the Baily–Borel compactification turns out to be the correct object to compare with Arthur–Selberg trace formulas. See [LR92], [Lau97], [Mor10a], [Mor10b], [Mor11], and [Zhu18]. In an orthogonal direction of generalization, the bad reduction of Shimura varieties is an extremely active and fruitful field of study, with profound arithmetic consequences which are not obtainable by studying good reductions only. See e.g. [HT01], [Rap05], [Hai05], [Shi11], [Man], [Shi], [Sch], and the references therein.

## References


[Ada11] Jeffrey Adams. Discrete series and characters of the component group. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 369–387. Int. Press, Somerville, MA, 2011.

[Art13] James Arthur. *The endoscopic classification of representations*, volume 61 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2013. Orthogonal and symplectic groups.

[Bor79] A. Borel. Automorphic $L$-functions. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 27–61. Amer. Math. Soc., Providence, R.I., 1979.

[Car79] P. Cartier. Representations of $p$-adic groups: a survey. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, pages 111–155. Amer. Math. Soc., Providence, R.I., 1979.

[Car86] Henri Carayol. Sur les représentations $l$-adiques associées aux formes modulaires de Hilbert. *Ann. Sci. École Norm. Sup. (4)*, 19(3):409–468, 1986.

[CD85] Laurent Clozel and Patrick Delorme. Pseudo-coefficients et cohomologie des groupes de Lie réductifs réels. *C. R. Acad. Sci. Paris Sér. I Math.*, 300(12):385–387, 1985.

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

[CL10] Raf Cluckers and François Loeser. Constructible exponential functions, motivic Fourier transform and transfer principle. *Ann. of Math. (2)*, 171(2):1011–1065, 2010.

[Del79] Pierre Deligne. Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 247–289. Amer. Math. Soc., Providence, R.I., 1979.

[Fuj97] Kazuhiro Fujiwara. Rigid geometry, Lefschetz-Verdier trace formula and Deligne's conjecture. *Invent. Math.*, 127(3):489–533, 1997.

[GN] Alain Genestier and Bao Châu Ngô. Chapter in vol. II.

[Hai02] Thomas J. Haines. On connected components of Shimura varieties. *Canad. J. Math.*, 54(2):352–395, 2002.

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

[Hal93] Thomas C. Hales. A simple definition of transfer factors for unramified groups. In *Representation theory of groups and algebras*, volume 145 of *Contemp. Math.*, pages 109–134. Amer. Math. Soc., Providence, RI, 1993.

[Hal95] Thomas C. Hales. On the fundamental lemma for standard endoscopy: reduction to unit elements. *Canad. J. Math.*, 47(5):974–994, 1995.

[Har11] Michael Harris. An introduction to the stable trace formula. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 3–47. Int. Press, Somerville, MA, 2011.

[HR10] Thomas J. Haines and Sean Rostami. The Satake isomorphism for special maximal parahoric Hecke algebras. *Represent. Theory*, 14:264–284, 2010.

[HT01] Michael Harris and Richard Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich.

[Kal11] Tasho Kaletha. Endoscopic character identities for depth-zero supercuspidal $L$-packets. *Duke Math. J.*, 158(2):161–224, 2011.

[Kal16] Tasho Kaletha. Rigid inner forms of real and $p$-adic groups. *Ann. of Math. (2)*, 184(2):559–632, 2016.

[Kis17] Mark Kisin. Mod $p$ points on Shimura varieties of abelian type. *J. Amer. Math. Soc.*, 30(3):819–914, 2017.

[KMSW14] T. Kaletha, A. Minguez, S. W. Shin, and P.-J. White. Endoscopic Classification of Representations: Inner Forms of Unitary Groups. *ArXiv e-prints*, September 2014.

[Kot82] Robert E. Kottwitz. Rational conjugacy classes in reductive groups. *Duke Math. J.*, 49(4):785–806, 1982.

[Kot83] Robert E. Kottwitz. Sign changes in harmonic analysis on reductive groups. *Trans. Amer. Math. Soc.*, 278(1):289–297, 1983.

[Kot84] Robert E. Kottwitz. Stable trace formula: cuspidal tempered terms. *Duke Math. J.*, 51(3):611–650, 1984.

[Kot85] Robert E. Kottwitz. Isocrystals with additional structure. *Compositio Math.*, 56(2):201–220, 1985.

[Kot86] Robert E. Kottwitz. Stable trace formula: elliptic singular terms. *Math. Ann.*, 275(3):365–399, 1986.

[Kot88] Robert E. Kottwitz. Tamagawa numbers. *Ann. of Math. (2)*, 127(3):629–646, 1988.

[Kot90] Robert E. Kottwitz. Shimura varieties and $\lambda$-adic representations. In *Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988)*, volume 10 of *Perspect. Math.*, pages 161–209. Academic Press, Boston, MA, 1990.

[Kot92] Robert E. Kottwitz. Points on some Shimura varieties over finite fields. *J. Amer. Math. Soc.*, 5(2):373–444, 1992.

[KSZ] Mark Kisin, Sug Woo Shin, and Yihang Zhu. The stable trace formula for certain Shimura varieties of abelian type. in preparation.

[Lan79] R. P. Langlands. Stable conjugacy: definitions and lemmas. *Canad. J. Math.*, 31(4):700–725, 1979.

[Lau97] Gérard Laumon. Sur la cohomologie à supports compacts des variétés de Shimura pour $\mathrm{GSp}(4)_{\mathbf{Q}}$. *Compositio Math.*, 105(3):267–359, 1997.

[LMW15] B. Lemaire, C. Moeglin, and J.-L. Waldspurger. Le lemme fondamental pour l'endoscopie tordue: réduction aux éléments unités. *ArXiv e-prints*, June 2015.

[LR87] R. P. Langlands and M. Rapoport. Shimuravarietäten und Gerben. *J. Reine Angew. Math.*, 378:113–220, 1987.

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

[LS87] R. P. Langlands and D. Shelstad. On the definition of transfer factors. *Math. Ann.*, 278(1-4):219–271, 1987.

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

[LW15] B. Lemaire and J.-L. Waldspurger. Le lemme fondamental pour l'endoscopie tordue: le cas où le groupe endoscopique non ramifié est un tore. *ArXiv e-prints*, November 2015.

[Man] Elena Mantovan. Chapter in vol. II.

[Mil05] J. S. Milne. Introduction to Shimura varieties. In *Harmonic analysis, the trace formula, and Shimura varieties*, volume 4 of *Clay Math. Proc.*, pages 265–378. Amer. Math. Soc., Providence, RI, 2005.

[Mok15] Chung Pang Mok. Endoscopic classification of representations of quasi-split unitary groups. *Mem. Amer. Math. Soc.*, 235(1108):vi+248, 2015.

[Mor10a] Sophie Morel. The intersection complex as a weight truncation and an application to Shimura varieties. In *Proceedings of the International Congress of Mathematicians. Volume II*, pages 312–334, New Delhi, 2010. Hindustan Book Agency.

[Mor10b] Sophie Morel. *On the cohomology of certain noncompact Shimura varieties*, volume 173 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2010. With an appendix by Robert Kottwitz.

[Mor11] Sophie Morel. Cohomologie d'intersection des variétés modulaires de Siegel, suite. *Compos. Math.*, 147(6):1671–1740, 2011.

[Ngô10] Bao Châu Ngô. Le lemme fondamental pour les algèbres de Lie. *Publ. Math. Inst. Hautes Études Sci.*, (111):1–169, 2010.

[Nic] Marc-Hubert Nicole. Chapter in vol. II.

[Pin92a] Richard Pink. On $l$-adic sheaves on Shimura varieties and their higher direct images in the Baily-Borel compactification. *Math. Ann.*, 292(2):197–240, 1992.

[Pin92b] Richard Pink. On the calculation of local terms in the Lefschetz-Verdier trace formula and its application to a conjecture of Deligne. *Ann. of Math. (2)*, 135(3):483–525, 1992.

[Rap05] Michael Rapoport. A guide to the reduction modulo $p$ of Shimura varieties. *Astérisque*, (298):271–318, 2005. Automorphic forms. I.

[Rei97] Harry Reimann. *The semi-simple zeta function of quaternionic Shimura varieties*, volume 1657 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1997.

[Ren11] David Renard. Endoscopy for real reductive groups. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 95–141. Int. Press, Somerville, MA, 2011.

[Roz] Sandra Rozensztajn. Chapter in vol. II.

[RZ88] Harry Reimann and Thomas Zink. Der Dieudonnémodul einer polarisierten abelschen Mannigfaltigkeit vom CM-Typ. *Ann. of Math. (2)*, 128(3):461–482, 1988.

[Sch] Peter Scholze. Chapter in vol. II.

[Shi] Sug Woo Shin. Chapter in vol. II.

[Shi11] Sug Woo Shin. Galois representations arising from some compact Shimura varieties. *Ann. of Math. (2)*, 173(3):1645–1741, 2011.

[ST16] Sug Woo Shin and Nicolas Templier. Sato-Tate theorem for families and low-lying zeros of automorphic $L$-functions. *Invent. Math.*, 203(1):1–177, 2016. Appendix A by Robert Kottwitz, and Appendix B by Raf Cluckers, Julia Gordon and Immanuel Halupczok.

[Taï14] O. Taïbi. Dimensions of spaces of level one automorphic forms for split classical groups using the trace formula. *ArXiv e-prints*, June 2014.

[Taï15] O. Taïbi. Arthur's multiplicity formula for certain inner forms of special orthogonal and symplectic groups. *ArXiv e-prints*, October 2015.

[Var05] Yakov Varshavsky. A proof of a generalization of Deligne's conjecture. *Electron. Res. Announc. Amer. Math. Soc.*, 11:78–88, 2005.

[Wal97] J.-L. Waldspurger. Le lemme fondamental implique le transfert. *Compositio Math.*, 105(2):153–236, 1997.

[Wal06] J.-L. Waldspurger. Endoscopie et changement de caractéristique. *J. Inst. Math. Jussieu*, 5(3):423–525, 2006.

[Wal08] J.-L. Waldspurger. L'endoscopie tordue n'est pas si tordue. *Mem. Amer. Math. Soc.*, 194(908):x+261, 2008.

[Zhu18] Yihang Zhu. The stabilization of the Frobenius–Hecke traces on the intersection cohomology of orthogonal Shimura varieties. *ArXiv e-prints*, January 2018.