# The Newton stratification of Shimura varieties and the Arthur-Selberg trace formula

Arno Kret

THÈSE DE DOCTORAT

présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS-SUD

Préparée à l'École Doctorale
de Mathématiques de la région Paris-Sud

Spécialité : Mathématiques

par

Arno Kret

# Stratification de Newton des variétés de Shimura et formule des traces d'Arthur-Selberg

Soutenue le 10 Décembre 2012 devant la Commission d'examen :

| | | | |
|---|---|---|---|
| M. | Henri | CARAYOL | (Rapporteur) |
| M. | Laurent | CLOZEL | (Directeur de thèse) |
| M. | Laurent | FARGUES | (Directeur de thèse) |
| M. | Michael | HARRIS | (Examinateur) |
| M. | Guy | HENNIART | (Examinateur) |
| M. | Sug Woo | SHIN | (Rapporteur, absent à la soutenance) |
| M. | Emmanuel | ULLMO | (Examinateur) |

MATHEMATIQUES
ORSAY

Thèse préparée au
**Département de Mathématiques d'Orsay**
Laboratoire de Mathématiques (UMR 8628), Bât. 425
Université Paris-Sud 11
91 405 Orsay CEDEX

**Résumé**

Nous étudions la stratification de Newton des variétés de Shimura de type PEL aux places de bonne réduction.

Nous considérons la strate basique de certaines variétés de Shimura simples de type PEL modulo une place de bonne réduction. Sous des hypothèses simplificatrices nous prouvons une relation entre la cohomologie $\ell$-adique de ce strate basique et la cohomologie de la variété de Shimura complexe. En particulier, nous obtenons des formules explicites pour le nombre de points dans la strate basique sur des corps finis, en termes de représentations automorphes. Nous obtenons les résultats à l'aide de la formule des traces et de la troncature de la formule de Kottwitz pour le nombre de points sur une variété de Shimura sur un corps fini.

Nous montrons, en utilisant la formule des traces, que n'importe quelle strate de Newton d'une variété de Shimura de type PEL de type (A) est non vide en une place de bonne réduction. Ce résultat a déjà été établi par Viehmann-Wedhorn [**104**] ; nous donnons une nouvelle preuve de ce théorème.

Considerons la strate basique des variétés de Shimura associées à certains groupes unitaires dans les cas où cette strate est une variété finie. Alors, nous démontrons un résultat d'équidistribution pour les opérateurs de Hecke agissant sur cette strate. Nous relions le taux de convergence avec celui de la conjecture de Ramanujan. Dans nos formules ne figurent que des représentations automorphes cuspidales sur $\mathrm{GL}_n$ pour lesquelles cette conjecture est connue, et nous obtenons donc des estimations très bonnes sur la vitesse de convergence.

En collaboration avec Erez Lapid nous calculons le module de Jacquet d'une représentation en échelle pour tout sous-groupe parabolique standard du groupe général linéaire sur un corps local non-archimédien.

**Abstract**

We study the Newton stratification of Shimura varieties of PEL type, at the places of good reduction.

We consider the basic stratum of certain simple Shimura varieties of PEL type at a place of good reduction. Under simplifying hypotheses we prove a relation between the $\ell$-adic cohomology of this basic stratum and the cohomology of the complex Shimura variety. In particular we obtain explicit formulas for the number of points in the basic stratum over finite fields, in terms of automorphic representations. We obtain our results using the trace formula and truncation of the formula of Kottwitz for the number of points on a Shimura variety over a finite field.

We prove, using the trace formula that any Newton stratum of a Shimura variety of PEL-type of type (A) is non-empty at a prime of good reduction. This result is already established by Viehmann-Wedhorn [**104**]; we give a new proof of this theorem.

We consider the basic stratum of Shimura varieties associated to certain unitary groups in cases where this stratum is a finite variety. Then, we prove an equidistribution result for Hecke operators acting on the basic stratum. We relate the rate of convergence to the bounds from the Ramanujan conjecture of certain particular cuspidal automorphic representations on $\mathrm{GL}_n$. The Ramanujan conjecture turns out to be known for these automorphic representations, and therefore we obtain very sharp estimates on the rate of convergence.

We prove that any connected reductive group $G$ over a non-Archimedean local field has a cuspidal representation.

Together with Erez Lapid we compute the Jacquet module of a Ladder representation at any standard parabolic subgroup of the general linear group over a non-Archimedean local field.

**Remerciements**

L'achèvement de cette thèse n'aurait pas été possible sans la contribution de nombreuses personnes que je tiens à remercier ici.

D'abord et avant tout, je voudrai remercier mon directeur de thèse Laurent Clozel. Je veux que ce soit clair que Laurent Clozel m'a beaucoup aidé au cours de mon doctorat. Il m'a expliqué la méthode, et donné ses notes sur la courbe modulaire (Chapitre 2). Après que le calcul pour la courbe modulaire a terminé, il m'a beaucoup aidé avec tous les chapitres (excepté l'annexe B que j'ai écrit avec Erez Lapid), me donnant beaucoup d'idées et vérifiant soigneusement les arguments que j'ai écrit (mais je reste responsable de toute erreur restante).

Je remercie aussi mon deuxième directeur Laurent Fargues, qui m'a aidé avec les isocristaux et la géométrie.

Je remercie Erez Lapid pour m'avoir invité en Israël et pour avoir calculé pour moi les modules de Jacquet des représentations en échelle.

Je remercie mes rapporteurs Sug Woo Shin et Henri Carayol pour la relecture de ce travail, ainsi que Michael Harris, Guy Henniart et Emmanuel Ullmo pour leur participation à mon jury de thèse.

En outre, je remercie les mathématiciens suivants pour leur aide et leurs explications : Ioan Badulescu, Bas Edixhoven, Guy Henniart, Florian Herzig, Gérard Laumon, Hendrik Lenstra, Alberto Minguez, Ben Moonen, Frans Oort, Olivier Schifmann, Peter Scholze, Sug Woo Shin, Marko Tadic, Lenny Taelman et Emmanuel Ullmo.

Je remercie mes collegues (et amis) pour toutes les dicussions qu'on a eues au cours de ces dernièrs années. Je pense à : Ramla, Giuseppe, Caroline, Przemyslaw, Miaofen, Maarten, Martino, Hatem, Christian, Wen-Wei,Yongqi, Judith, Marco, Stefano, Bartosz, Martin, Arne, Shenghao, Chunh-Hui, et Haoran. En particulier, je tiens à remercier Ariyan, Giovanni, John et Marco.

Je remercie Caterina, Elodie et Jana pour leur aide dans la rédaction.

Je remercie mes amis pour les très belles soirées à Paris et ailleurs : Amira, Angela, Ariane, Boye, Caterina, Catherine, Clara, Federica, Francesca, Jana, Lexie, Liviana, Lorenzo, Martina, Michiel, Nynke, Robbert, Sanela, Teodolinda, Valentina, et Vito.

Finalement je voudrais remercier ma famille pour leur support constant pendant les années. Ici je pense en particular à mes parents et ma soeur et Jasper, mais aussi à tous mes grand-parents, oncles, tantes et cousins.

# Table des matières

# Introduction

Dans cette thèse, nous étudions la réduction des variétés de Shimura de type PEL [1] modulo des nombres premiers de bonne réduction. Plus précisément, nous étudions la stratification de Newton de ces variétés modulo $p$. Les variétés de Shimura de type PEL sont des espaces de modules de variétés abéliennes avec certaines structures additionnelles de type PEL. La stratification de Newton des variétés de Shimura de type PEL consiste en des lieux où l'isocristal attaché aux variétés abéliennes est constant. Ces strates de Newton sont elles-mêmes des variétés et nous voulons comprendre leur cohomologie $\ell$-adique.

## 1. Histoire et motivation

L'étude des strates de Newton a commencé avec Frans Oort, qui les a définies pour l'espace classique de Siegel. À son tour, il étend le travail de Grothendieck et aussi de Katz qui ont étudié le comportement des cristaux associés à des groupes $p$-divisibles dans les familles.

Pour l'espace de Siegel, Oort a déterminé les strates de Newton non vides, et a calculé les dimensions de ces strates [**85**]. Le premier résultat d'Oort (le fait que les strates sont non vides) est démontré dans [**84**] et a été conjecturé plus tôt par Grothendieck dans [**42**]. Oort a étudié en outre les orbites de Hecke dans les strates de Newton, et a introduit d'autres stratifications différentes de la stratification de Newton (que nous ne considérerons pas dans cette thèse).

La définition de la stratification de Newton a ensuite été étendue à toutes les variétés de Shimura de type PEL par Rapoport et Richarz [**88**]. Leur article est apparu après les travaux de Kottwitz sur les isocristaux avec des structures additionnelles [**55**] (voir aussi [**60**]).

Pour une discussion plus détaillée de l'histoire du sujet nous renvoyons le lecteur à l'article de Rapoport [**87**] ; une autre référence utile est l'article de Mantovan [**74**].

## 2. La stratification de Newton

Avant d'énoncer les résultats de cette thèse, rappelons d'abord plus précisément la définition de la stratification de Newton.

Nous avons déjà expliqué brièvement ci-dessus que l'on étudie les variétés de Shimura de type PEL et que ces variétés ont une interprétation comme espaces de modules de variétés abéliennes avec certaines structures additionnelles de type PEL.

1. **P**olarization, **E**ndomorphisms and **L**evel structure.

Pourquoi est-ce que cette interprétation comme problème des modules est utile ? Nous l'utilisons pour réduire la variété de Shimura modulo $p$ et définir la stratification de Newton : *A priori* une variété de Shimura $S$ est définie seulement sur un certain corps des nombres $E$ (le corps réflex), et donc "réduction modulo $p$" n'a aucun sens. Avant de pouvoir réduire la variété modulo un nombre premier $p$, nous avons besoin d'un modèle de $\mathcal{S}$ de $S$, sur, disons, l'anneau $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$. Bien sûr, les modèles existent, mais ils ne sont pas uniques et leur réduction dépend du modèle que l'on choisit. Mais rappelons que nous avons supposé que $S$ a une interprétation comme problème des modules de type PEL sur $E$, et donc les choses se simplifient. Le problème des modules peut être étendu à un problème des modules sur l'anneau $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$, et le problème étendu est représentable par un champ de Deligne-Mumford [**59**, §5]. Sous des hypothèses naturelles, ce champ est un schéma quasi-projectif lisse sur $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$. Pour avoir la représentabilité par un schéma lisse il faut que le groupe compact ouvert $K \subset G(\mathbb{A}_{\mathrm{f}})$ soit suffisamment petit hors $p$ et hyperspecial à $p$ ; nous supposerons, pour simplifier, que ce soit le cas. Ensuite, nous avons un choix canonique pour le modèle $\mathcal{S}$ de $S$ sur $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$, et nous choisissons ce modèle. On remplace désormais $S$ par son modèle $\mathcal{S}$ sur $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$.

La variété $S \otimes \mathbb{F}_p$ se décompose canoniquement en certaines pièces appelées *strates de Newton*. Pour définir ces strates, on utilise de nouveau l'interprétation de $S$ comme espace de modules : Pour chaque point $x \in S(\overline{\mathbb{F}}_p)$ on peut considérer le module de Dieudonné rationnel $\mathbb{D}(\mathcal{A}_x[p^\infty]) \otimes \mathbb{Q}$ de la variété abélienne $\mathcal{A}_x$ correspondant au point $x$. Ces modules de Dieudonné sont des isocristaux et les structures additionnelles sur $\mathcal{A}_x$ induisent des structures additionnelles sur l'isocristal $\mathbb{D}(\mathcal{A}_x[p^\infty]) \otimes \mathbb{Q}$. Lorsqu'il est équipé de ces structures, l'objet $\mathbb{D}(\mathcal{A}_x[p^\infty]) \otimes \mathbb{Q}$ est un *isocristal avec $G$-structure* (ici, $G$ est le groupe de la donnée de Shimura de $S$). Nous sommes intéressés par cet objet à isomorphisme près. On note $B(G_{\mathbb{Q}_p})$ pour l'ensemble des isocristaux avec des $G$-structures additionelles. Donc $\mathbb{D}(\mathcal{A}_x[p^\infty]) \otimes \mathbb{Q} \in B(G_{\mathbb{Q}_p})$.

Maintenant, pour chaque élément $b \in B(G_{\mathbb{Q}_p})$ on note $S_b(\overline{\mathbb{F}}_p)$ le sous-ensemble de $S(\overline{\mathbb{F}}_p)$ constitué d'éléments $x \in S(\overline{\mathbb{F}}_p)$ tels que $b = \mathbb{D}(\mathcal{A}_x[p^\infty]) \otimes \mathbb{Q} \in B(G_{\mathbb{Q}_p})$. Le sous-ensemble $S_b(\overline{\mathbb{F}}_p) \subset S(\overline{\mathbb{F}}_p)$ provient d'un sous-schéma réduit et localement fermé $S_b$ de $S$ [**88**]. La collection des schémas $\{S_b\}_{b \in B(G_{\mathbb{Q}_p})}$ est la *stratification de Newton de $S$*, et les $S_b$ sont les *strates de Newton*.

Les correspondances de Hecke sur la variété $S(\mathbb{C})$ sont algébriques, et définies sur le corps $E$. Elles s'étendent aussi au modèle de $S$ sur $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$, parce que leur action peut être décrite en termes de l'interprétation de $S$ comme problème des modules. En particulier, nous avons les correspondances de Hecke sur $S \otimes \mathbb{F}_p$. Ces correspondances de Hecke respectent la stratification de Newton, de sorte qu'elles peuvent être restreintes aux différentes strates de Newton. Par conséquent les espaces de cohomologie $\mathrm{H}^i_{\text{ét}}(S_{b,\overline{\mathbb{F}}_p}, \overline{\mathbb{Q}}_\ell)$ (avec $\ell \neq p$) sont des modules sur l'algèbre de Hecke de $G$. Ces espaces de cohomologie portent aussi une action du groupe de Galois $\mathrm{Gal}(\overline{\mathbb{F}}_p/k)$ (où $k$ est un corps résiduel de $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$) qui commute avec l'action de l'algèbre de Hecke.

## 3. Cette thèse

Nous donnons un bref aperçu des nos résultats.

Dans cette thèse, on étudie la stratification de Newton des variétés de Shimura de type PEL, en des places de bonne réduction. Nous introduisons une nouvelle méthode pour étudier les strates de Newton. Notre méthode utilise (la restriction de) la formule de Kottwitz et des formes automorphes. En utilisant cette méthode nous répondrons à certaines questions classiques.

Nous nous posons quatre questions générales sur les strates de Newton $S_b$ (cf. §1) :

(1) Pour quels éléments $b \in B(G_{\mathbb{Q}_p})$, la strate $S_b \subset S$ correspondante est-elle non vide ?

(2) Pour $b \in B(G_{\mathbb{Q}_p})$ donné, peut-on calculer la dimension de la variété $S_b$ ?

(3) Peut-on décrire la fonction zêta de $S_b$ ?

(4) Peut-on décrire le cohomologie $\ell$-adique de $S_b$ en tant que module de Galois/Hecke ?

Nous avons numéroté les questions en difficulté croissante. Souvent, une réponse satisfaisante à la question $(i)$ donne également une réponse satisfaisante à la question $(i-1)$.

Dans cette thèse, nous répondrons partiellement aux quatre questions ci-dessus. Maintenant nous écrivons quelques énoncés imprécis afin de donner une idée des résultats. Nous préciserons nos théorèmes principaux dans la section suivante.

**Question (1).** Kottwitz a introduit l'ensemble des isocristaux $\mu$-admissibles $B(G,\mu) \subset B(G)$, où $\mu$ est défini par la variété de Shimura. Pour tout point $x \in S(\overline{\mathbb{F}}_p)$ l'isocristal associé se trouve dans le sous-ensemble $B(G,\mu) \subset B(G)$ (Rapoport-Richarz). Ainsi, les strates de Newton associées aux isocristaux non-admissibles sont vides. Récemment Wedhorn et Viehmann ont établi, pour les variétés de PEL de type (A) et (C), qu'inversement, pour $b$ un isocristal $\mu$-admissible donné, il existe un point $x \in S(\overline{\mathbb{F}}_p)$ dont l'isocristal est $b$. Nous établissons le résultat de Wedhorn et Viehmann dans le Chapitre 4 pour les variétés de type (A). Même si notre résultat n'est pas nouvau, notre preuve est complètement différente : la formule des traces remplace des arguments délicats de géométrie algébrique.

**Question (2).** Dans le Chapitre 2 on établit une formule pour la dimension de la strate basique d'une variété de Kottwitz, sous des conditions simplificatrices. Dans le Chapitre 3 on établit des résultats partiels qui vont en direction d'une formule pour la dimension de la strate basique d'une variété de Kottwitz, dans des conditions beaucoup plus légères. Une *variété de Kottwitz* est une variété de Shimura de type PEL de type (A), et est associée à une algèbre de division avec une involution de seconde espèce. Ces variétés sont nettement plus simples que toute la classe des variétés de PEL de type (A), où l'endoscopie joue un rôle.

**Question (3).** Considérons à nouveau la strate basique des variétés de Kottwitz en des places de bonne réduction. Nous supposons maintenant que $p$ est complètement déployé dans le centre de l'algèbre à division $D$ qui vient avec la variété de Kottwitz. Au Chapitre 3, sous ces hypothèses, nous répondons à la question (4) par “oui” et on obtient, comme corollaire, la réponse “oui” à la question (3).

**Question (4).** Dans le chapitre 3, nous calculons l'objet $\sum_{i=0}^{\infty}(-1)^i \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_p}, \iota^*\mathcal{L})$ comme élément du groupe de Grothendieck de $\mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^p)) \times \overline{\mathbb{Q}}_\ell[\mathrm{Gal}(\overline{\mathbb{F}}_p/k)]$-modules [2]. Ici, $B$ est la strate basique d'une variété de Kottwitz (associé à une algèbre à division $D$) en un nombre premier $p$ tel que $D \otimes \mathbb{Q}_p$ est isomorphe à un produit d'algèbres de la forme $M_n(\mathbb{Q}_p)$. L'objet s'exprime en fonction de formes automorphes sur le groupe $G$ et certains polynômes de nature combinatoire (voir la réponse à la question (3)).

**Méthode.** Nous allons maintenant expliquer la nouvelle méthode que nous utilisons dans cette thèse. Nous commençons avec la formule de Kottwitz pour le nombre de points d'une variété de Shimura de type PEL-sur un corps fini (cf. [**57**, **59**]) :

(3.1)
$$\sum_{x' \in \mathrm{Fix}_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}(\overline{\mathbb{F}}_p)} \mathrm{Tr}(\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}, \mathcal{L}_x) = |\ker^1(\mathbb{Q}, G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\phi_\alpha) \, \mathrm{Tr}\, \xi_{\mathbb{C}}(\gamma_0).$$

Cette introduction n'est pas le lieu pour définir toutes les notations et définitions impliquées dans cette formule. Nous n'expliquerons ici que certains des éléments principaux. Il convient de mentionner d'abord que Kottwitz a uniquement prouvé cette formule pour les variétés $S$ de type PEL, lorsque le groupe est de type (A) ou (C). Pour les variétés de type PEL de type (D), Kottwitz ne prouve ni ne conjecture une telle formule [3].

- $f^{\infty p}$ est un opérateur de Hecke quelconque dans l'algèbre de Hecke $\mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^p))$ des fonctions localement constantes sur $G(\mathbb{A}_{\mathrm{f}}^p)$ (où $G$ est le groupe qui intervient dans la donnée de Shimura) ;
- $\Phi_{\mathfrak{p}}$ est l'élément de Frobenius géométrique dans le groupe de Galois $\mathrm{Gal}(\overline{\mathbb{F}}_p/k)$ ;
- $\alpha$ est un entier positif ;
- $\xi$ est une représentation complexe irréductible de $G_{\mathbb{C}}$, et $\mathcal{L}$ est le système local $\ell$-adique associé à la représentation $\xi$ ($\ell$ est un nombre premier fixé différent de $p$, et nous avons fixé, et supprimé, un isomorphisme entre $\mathbb{C}$ et $\overline{\mathbb{Q}}_\ell$) ;
- La somme du côté droit de l'Équation (3.1) porte sur les triplets de Kottwitz $(\gamma_0;\gamma,\delta)$. Ces triplets sont associés aux classes d'isogénie des variétés abéliennes virtuelles. L'élément $\gamma_0$ parcourt les classes de conjugaison stables $\mathbb{R}$-elliptiques de $G(\mathbb{Q})$.
- Pour la description des points $x'$ du point associé $x \in \mathrm{Sh}_K(\overline{\mathbb{F}}_p)$, voir l'Équation (2.3.3).

L'énoncé précis du résultat se trouve dans l'article de Kottwitz [**59**], voir en particulier §19 et l'introduction de cet article.

2. Il faut dire que l'on doit fixer un isomorphisme de $\mathbb{C}$ avec $\overline{\mathbb{Q}}_\ell$ pour avoir une action de l'algèbre de Hecke sur la cohomologie $\ell$-adique.

3. Dans l'article [**57**] il ne conjecture une formule que pour les groupes connexes ; dans l'article [**59**] il définit les variétés de type (D), mais, quand les preuves commencent, il exclut ce cas.

En regardant la formule de l'Équation (3.1) nous pouvons expliquer l'idée principale de notre méthode. La formule de Kottwitz concerne le nombre de points dans toute la variété de Shimura $\mathrm{Sh}_K$ modulo un nombre premier $\mathfrak{p}$ du corps réflex $E$. L'idée principale est de restreindre le côté droit de l'Équation (3.1) en comptant seulement les points dans une strate de Newton donnée. Ainsi, nous fixons un isocristal $b \in B(G_{\mathbb{Q}_p})$ avec des $G$-structures additionelles. Cet élément correspond à une classe de $\sigma$-conjugaison dans le groupe $G(L)$, où $L$ est la complétion de l'extension maximale non ramifiée de $\mathbb{Q}_p$, et $\sigma$ est l'élément de Frobenius de la complétion du corps réflex $E$ en la place $\mathfrak{p}$. Alors $b$ définit une strate $\mathrm{Sh}_{K,\mathfrak{p}}^b$ de $\mathrm{Sh}_K$ modulo $\mathfrak{p}$. Nous restreignons la somme dans l'Équation (3.1) sur les triplets de Kottwitz $(\gamma_0; \gamma, \delta)$ tels que $\delta$ définit l'isocristal $b$. Le côté gauche doit alors être limité aux points fixes de la correspondance $f^{p\infty} \times \Phi_{\mathfrak{p}}^{\alpha}$ agissant sur la $b$-ième strate de Newton $\mathrm{Sh}_{K,\mathfrak{p}}^b$ de $\mathrm{Sh}_{K,\mathfrak{p}}$. Les restrictions des deux côtés de l'Équation (3.1) sont égales, et nous obtenons une version $b$-restreinte de la formule de Kottwitz.

Dans son article de la conférence de Ann Arbor, Kottwitz montre comment (le côté droit de) l'Équation (3.1) se stabilise. Cet argument de stabilisation est également valable pour notre formule $b$-restreinte. Donc nous pouvons encore comparer la formule $b$-restreinte avec la formule des traces. Ce faisant, nous arrivons à une somme de traces sur des représentations automorphes des groupes endoscopiques de $G$. Notre méthode consiste à traduire une question donnée sur une strate de Newton, par la formule restreinte de Kottwitz, en une question sur les représentations automorphes, et de voir si nous pouvons répondre à cette question traduite. Nous montrons dans cette thèse que nous pouvons répondre à la question traduite dans certains cas. Par exemple, pour répondre à la question (1) ci-dessus, on doit montrer qu'une somme de traces de certains opérateurs de Hecke (transférés) agissant sur les représentations automorphes de groupes endoscopiques de $G$ est non nulle (Chapitre 4, voir ci-dessous).

Il se trouve que les questions traduites sont souvent des problèmes combinatoires. Essayons d'expliquer un de ces problèmes combinatoires, et comment nous le résolvons. À l'exception du Chapitre 4, nous avons restreint notre attention à la strate basique dans cette thèse. Dans cette section, nous limitons aussi notre attention à la strate basique. En outre, nous considérons une variété de Shimura "de Kottwitz". Nous restreignons l'Équation (3.1) à la strate basique. Par les arguments que nous avons esquissés ci-dessus, le côté droit de cette équation restreinte peut être comparé à une formule des traces. Une caractéristique des variétés de Kottwitz est que l'endoscopie ne joue pas de rôle. C'est pourquoi nous allons obtenir simplement une trace de la forme $\mathrm{Tr}((\chi_c^{G(\mathbb{Q}_p)} f_\alpha) f^p, \mathcal{A}(G))$. Ici $\mathcal{A}(G)$ est l'espace des formes automorphes sur le groupe $G$, $f^p$ est l'opérateur de Hecke en dehors de $p$, et en $p$ nous avons l'opérateur de Hecke $\chi_c^{G(\mathbb{Q}_p)} f_\alpha$. La fonction $f_\alpha$ est la *fonction de Kottwitz* [**54**]. Cette fonction est fondamentale, et Kottwitz a montré que l'on doit prendre cette fonction en $p$, si l'on veut que la trace $\mathrm{Tr}(f_\alpha f^p, \mathcal{A}(G))$ soit égale au côté gauche de l'Équation (3.1). Une fois que cela est établi, on peut faire appel aux théorèmes de Fujiwara et Grothendieck-Lefschetz afin de trouver l'identité :

$$\mathrm{Tr}(f_\alpha f^p, \mathcal{A}(G)) = \sum_{i=0}^{\infty} \mathrm{Tr}\left(f^{p\infty} \times \Phi_{\mathfrak{p}}^{\alpha}, \mathrm{H}^i_{\text{ét}}(\mathrm{Sh}_{K,\overline{\mathbb{F}}_p}, \overline{\mathbb{Q}}_\ell)\right).$$

C'est l'identité que Kottwitz utilise pour associer une représentation galoisienne à certaines formes automorphes pour le groupe $G$ dans l'article [**58**]. Nous avons restreint la formule à la strate basique, ce qui donne l'identité

$$\mathrm{Tr}((\chi_c^{G(\mathbb{Q}_p)} f_\alpha) f^p, \mathcal{A}(G)) = \sum_{i=0}^{\infty} \mathrm{Tr}\left(f^{p\infty} \times \Phi_{\mathfrak{p}}^{\alpha}, \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_p}, \overline{\mathbb{Q}}_\ell)\right),$$

où $B$ est la strate basique. Le problème combinatoire que nous avons mentionné ci-dessus est le calcul des traces compactes $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p)$ pour toute représentation automorphe $\pi$ contenue dans l'espace de formes automorphes $\mathcal{A}(G)$.

Dans son article sur le lemme fondamental [**22**], Clozel a donné une formule pour la trace compacte d'une fonction de Hecke sur les représentations irréductibles lisses des groupes réductifs $p$-adiques :

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi_p) = \sum_{P=MN} \varepsilon_P \,\mathrm{Tr}\left(\widehat{\chi}_N \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2})\right),$$

(la somme s'étend sur les sous-groupes paraboliques standard ; pour les autres notations, nous renvoyons le lecteur à la Proposition 2.1.2). Cependant, cette formule est une somme alternée impliquant tous les modules de Jacquet de la représentation. Il n'est pas facile d'évaluer la formule pour une représentation arbitraire d'une manière satisfaisante (du moins, l'auteur ne sait pas comment), pour deux raisons : (1) les modules de Jacquet sont très compliqués, (2) la somme est très redondante et beaucoup des termes s'annulent.

Avec seulement le formule de Clozel, nous ne pensons pas avoir assez d'information pour dire quelque chose d'intéressant. Dans cette thèse, nous travaillons souvent avec l'hypothèse supplémentaire que le centre de $F$ de l'algèbre à division $D$ se déploie en un compositum $F = \mathcal{K}F^+$, où $F^+$ est un corps de nombres totalement réel, et $\mathcal{K}$ est quadratique imaginaire. Nous supposons également que le nombre premier $p$ de réduction est déployé dans l'extension $\mathcal{K}/\mathbb{Q}$. Ces hypothèses nous permettent d'utiliser le changement de base quadratique. En appliquant le changement de base du groupe $G$ au groupe $G^+ = \mathrm{Res}_{\mathcal{K}/\mathbb{Q}} G_{\mathcal{K}}$, nous pouvons comparer les représentations automorphes $\pi \subset \mathcal{A}(G)$ avec des représentations automorphes du groupe général linéaire. Ces représentations automorphes sont discrètes, et Moeglin et Waldspurger ont classifié le spectre discret du groupe général linéaire. Cela nous donne une liste explicite de représentations possibles $\pi_p$ en $p$, et il suffit pour nos besoins de calculer les traces $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p)$ pour *ces* représentations $\pi_p$. Les représentations sont, à induction parabolique près, des représentations de Speh. Tadic a trouvé une expression explicite des représentations de Speh dans le groupe de Grothendieck des représentations lisses. Il prouve une formule explicite qui exprime toute représentation de Speh donnée en un somme alternée des représentations standard. Nous savons comment calculer les traces compactes sur les représentations standard. Ainsi, il ne reste plus qu'à calculer la somme alternée.

Malheureusement, il se trouve que la somme alternée restante n'est pas facile à calculer en général. Dans le Chapitre 2, nous avons travaillé avec des conditions choisies de sorte que la somme est facile (triviale) à calculer (donc nous évitons ce problème dans le Chapitre 2). Dans le Chapitre 3 nous travaillons sous l'hypothèse que $p$ est complètement déployée dans le corps $F^+$, la somme est alors aussi plus simple, mais non-triviale. Nous interprétons la somme comme une somme sur les polynômes associés à certains chemins dans $\mathbb{Q}^2$, et nous montrons, en utilisant le Lemme de Lindström-Gessel-Viennot bien connu en combinatoire, que la somme se réduit à une certaine somme sur des *chemins sans intersection*. Puis nous déterminons les représentations qui contribuent à la (somme alternée des espaces de) cohomologie de la strate basique.

## 4. Les résultats de cette thèse

Nous indiquons chapitre par chapitre les résultats principaux de cette thèse.

**Chapitre 1 : La courbe modulaire.** Ce chapitre d'introduction ne contient pas de nouveaux résultats. Le théorème principal que nous prouvons est classique et peut être déduit facilement des travaux de Deligne et Rapoport [**34**].

Nous avons écrit ce chapitre comme un exemple de la méthode que nous avons esquissé dans la section précédente. Nous démontrons le théorème suivant :

THÉORÈME (Deligne-Rapoport). *Soit $N$ un entier avec $N \geq 4$ et considérons la courbe modulaire $Y_1(N)$. Soit $p$ un nombre premier qui ne divise pas $N$. Nous écrivons $Y_1(N)_{\mathrm{ss}}$ pour le lieu supersingulier de $Y_1(N) \otimes \mathbb{F}_p$. Soit $X'(N)$ la compactification de la courbe correspondant au groupe $\Gamma_1(N) \cap \Gamma_0(p)$. Soit $\alpha$ un entier positif. Si $\alpha$ est pair, nous avons*

$$\#Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha}) = 1 + \mathrm{genre}(X'(N)) - 2 \cdot \mathrm{genre}(X_1(N)).$$

*Si $\alpha$ est impair, nous avons*

$$\#Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha}) = 1 + \sum_{\pi} \dim(\pi_{\mathrm{f}})^{K'} \cdot \varepsilon(\pi_p),$$

*où $\pi$ porte sur les représentations suivantes de $\mathrm{GL}_2(\mathbb{A})$. Nous écrivons $Z(\mathbb{R})^+$ pour l'ensemble des matrices diagonales dans $\mathrm{GL}_2(\mathbb{R})$ de la forme $\mathrm{diag}(x,x)$ avec $x \in \mathbb{R}^\times_{>0}$, et nous écrivons $L^2_0(\mathrm{GL}_2(\mathbb{Q})Z(\mathbb{R})^+\backslash\mathrm{GL}_2(\mathbb{A}_{\mathrm{f}}))$ pour l'espace des formes paraboliques muni de l'action de $\mathrm{GL}_2(\mathbb{A})$ par translations à droite. Alors $\pi$ porte sur les sous-espaces irréductibles de $L^2_0(\mathrm{GL}_2(\mathbb{Q})Z(\mathbb{R})^+\backslash\mathrm{GL}_2(\mathbb{A}_{\mathrm{f}}))$ avec*

- *$\pi_\infty$ est la série discrète holomorphe de poids 2 ;*
- *$\pi_p$ est un twist par un caractère non ramifié de la représentation de Steinberg de $\mathrm{GL}_2(\mathbb{Q}_p)$, $\varepsilon(\pi_p) = 1$ si $\pi_p \cong \mathrm{St}$ et $\varepsilon(\pi_p) = -1$ si $\pi_p \cong \mathrm{St} \otimes \varphi$ avec $\varphi$ le caractère quadratique non-ramifié.*

**Chapitre 2 : La strate basique de quelques variétés de Shimura simples.** Nous considérons une classe restreinte de certaines variétés de Shimura simples de type PEL, et nous considérons la strate de Newton en une place déployée de bonne réduction. Nous établissons une relation entre la cohomologie de la strate basique de la variété de Shimura et l'espace des formes automorphes sur le groupe $G$. Nous montrons que l'espace des formes automorphes décrit complètement la cohomologie de la strate basique comme module de Hecke, ainsi que l'action de l'élément de Frobenius.

Donnons maintenant l'énoncé précis. Soit $D$ une algèbre de division sur $\mathbb{Q}$ équipée d'un anti-involution $*$. On note $F$ le centre de l'algèbre $D$. Nous supposons que $F$ est un corps de multiplication complexe, que $*$ induit la conjugaison complexe sur le centre $F$ et que $D \neq F$. Nous supposons que $F$ est un compositum d'une extension quadratique imaginaire $\mathcal{K}$ de $\mathbb{Q}$ et du sous-corps totalement réel $F^+$ de $F$. Nous choisissons un morphisme $h_0$ de $\mathbb{R}$-algèbres de $\mathbb{C}$ dans $D_\mathbb{R}$ tel que $h_0(z)^* = h_0(\overline{z})$ pour tout nombre complexe $z$, et nous supposons que l'involution $x \mapsto h_0(i)^{-1}x^*h_0(i)$ sur $D_\mathbb{R}$ est positive (cf. Deligne [**31**, (2.1.1.2)]). Alors $(D, h)$ induit une donnée de Shimura $(G, X, h^{-1})$. Soit $K \subset G(\mathbb{A}_\mathrm{f})$ un sous-groupe compact ouvert de $G$ et $p$ un nombre premier tel que nous avons bonne réduction en $p$ (dans le sens de [**59**, §6]) et tel que le groupe $K$ se décompose en un produit $K_pK^p$ où $K_p \subset G(\mathbb{Q}_p)$ est hyperspécial et le groupe hors $p$, $K^p$, est suffisamment petit, pour qu'on puisse prendre $\mathrm{Sh}_K$ la variété de Shimura qui représente le problème des modules de variétés abéliennes de type PEL définie chez Kottwitz [**59**, §6]. Nous notons $\mathcal{A}(G)$ l'espace des formes automorphes sur $G$. Soit $\xi$ une représentation irréductible complexe algébrique de $G(\mathbb{C})$. Soit $f_\infty$ une fonction (quelconque) sur le groupe $G(\mathbb{R})$ ayant les intégrales orbitales stables prescrites par les identités dans [**54**]. Pour $f_\infty$ nous pouvons prendre une fonction d'Euler-Poincaré [**58**, Lemma 3.2] (modulo un certain scalaire explicite, cf. [*loc. cit.*]). Nous supposons que le nombre premier $p$ est déployé dans l'extension $\mathcal{K}/\mathbb{Q}$. Soit $B$ la strate basique de la réduction de la variété $\mathrm{Sh}_K$ modulo une place $\mathfrak{p}$ du corps réflex $E$ au-dessus de $p$, et soit $\mathbb{F}_q$ le corps résiduel de $E$ en $\mathfrak{p}$. Nous notons $\Phi_\mathfrak{p} \in \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_q)$ pour le Frobenius géométrique $x \mapsto x^{q^{-1}}$. Soit $\mathcal{L}$ la restriction en $B_{\overline{\mathbb{F}}_p,\text{ét}}$ du système local $\ell$-adique associé à $\xi$ sur $\mathrm{Sh}_{K,\overline{\mathbb{F}}_p,\text{ét}}$ [**59**, §6]. Soit $f^{\infty p}$ un opérateur de Hecke $K^p$-sphérique dans l'algèbre $\mathcal{H}(G(\mathbb{A}_\mathrm{f}^p))$, où $\mathbb{A}_\mathrm{f}^p$ est l'anneau des adèles finies dont la composante en $p$ est triviale. Enfin, nous supposons une condition simplificatrice sur l'isocristal basique $\mu$-admissible. Soit $b \in B(G_{\mathbb{Q}_p}, \mu)$ l'isocristal avec des $G$-structures additionelles correspondant à la strate basique. Le groupe $G(\mathbb{Q}_p)$ est égal à $\mathbb{Q}_p{}^\times \times \mathrm{GL}_n(F^+ \otimes \mathbb{Q}_p)$, et l'ensemble $B(G_{\mathbb{Q}_p})$ se décompose suivant les facteurs irréductibles de l'algèbre de $F^+ \otimes \mathbb{Q}_p$. Par conséquent, nous avons pour chaque $F^+$-place $\wp$ au-dessus de $p$ un isocristal $b_\wp \in B(\mathrm{GL}_n(F^+_\wp))$. La condition simplificatrice sur l'isocristal $b$ est, pour chaque $\wp$, la seule pente de $b_\wp$ avec multiplicité $> 1$ est la pente 0. Sous ces conditions, nous avons le théorème suivant :

THÉORÈME. *La trace de la correspondance $f^{p\infty} \times \Phi_{\mathfrak{p}}^{\alpha}$ agissant sur la somme alternée des espaces de cohomologie $\sum_{i=0}^{\infty}(-1)^i \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_p}, \iota^*\mathcal{L})$ est égale à*

$$|\operatorname{Ker}^1(G:\mathbb{Q})| P(q^\alpha) \left( \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1, \pi_p \text{ nr.}}} \zeta_\pi^\alpha \cdot \operatorname{Tr}(f^p, \pi^p) + \varepsilon \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p \text{ type St.}}} \zeta_\pi^\alpha \cdot \operatorname{Tr}(f^p, \pi^p) \right). \tag{4.1}$$

*pour tous les entiers positifs $\alpha$. La condition "$\pi_p$ de type Steinberg" dans l'Équation* (4.1) *signifie que, pour chaque $F^+$-place $\wp$ au-dessus de $p$ on a les conditions suivantes :*

(1) *si le composant en $\wp$ de l'isocristal basique n'est pas étale (i.e. a des pentes non nulles), alors $\pi_\wp$ est un twist par un caractère non ramifié de la représentation de Steinberg de $\mathrm{GL}_n(F^+_\wp)$ ;*

(2) *si le composant en $\wp$ est étale (toutes les pentes sont nulles), alors la représentation $\pi_\wp$ est non ramifiée et générique.*

*Le symbole $\varepsilon \in \{\pm 1\}$ dans l'Équation* (4.1) *est égal à $(-1)^{(n-1)\#\mathrm{Ram}_p^+}$ où $\mathrm{Ram}_p^+$ est l'ensemble des $F^+$-places $\wp$ divisant $p$ telles que l'isocristal $b_\wp$ n'est pas étale. Le nombre $\zeta_\pi$ est un certain $q$-nombre de Weil dont le poids dépend de $\xi$ (voir Lemme 2.3.11). Le symbole $P(q^\alpha)$ est une certaine fonction polynomiale, voir la Définition 2.3.12 et la discussion qui suit cette définition.*

Pour donner un idée de sa forme nous donnons dans cette introduction la fonction $P(q^\alpha)$ sous deux autres hypothèses simplificatrices (pour l'énoncé complet nous devons nous référer au Chapitre 2). Soit $n$ l'entier positif tel que $n^2$ est la dimension de l'algèbre de $D$ sur le corps $F$. Par la classification des groupes unitaires sur les nombres réels, le groupe $G(\mathbb{R})$ induit pour chaque $F^+$-place infinie $v$ un ensemble de nombres non-négatifs $\{p_v, q_v\}$ tels que $p_v + q_v = n$. Supposons dans cette introduction que $p_v = 0$ pour toute place $v$, sauf pour une unique $F^+$-place infinie $v_0$. Deuxièmement, nous supposons que $p$ est complètement déployé dans le corps $F^+$. Alors il existe un polynôme $\mathrm{Pol} \in \mathbb{C}[X]$ tel que $P(q^\alpha)$ est égal à $\mathrm{Pol}|_{X=q^\alpha}$. Notre condition sur l'isocristal basique correspond à la condition que le nombre $p_{v_0}$ soit premier avec $n$ (voir paragraphe §2.3.2). Nous noterons $s$ pour la signature $p_{v_0}$. Alors, le polynôme $P(q^\alpha)$ est égal à l'évaluation du polynôme

$$X \sum X_{i_1} X_{i_2} \cdots X_{i_s} \in \mathbb{C}[X, X_1, X_2, \ldots, X_n], \tag{4.2}$$

au point $X = q^{\alpha \frac{s(n-s)}{2}}, X_1 = q^{\alpha \frac{1-n}{2}}, X_2 = q^{\alpha \frac{3-n}{2}}, \ldots, X_n = q^{\alpha \frac{n-1}{2}}$. Dans la somme de l'Équation (4.2) les indices $i_1, i_2, \ldots, i_s$ portent sur l'ensemble $\{1, 2, \ldots, n\}$ et satisfont aux conditions

- $i_1 < i_2 < i_3 < \ldots < i_s$ ;
- $i_1 = 1$ ;
- Si $s > 1$ il y a une condition supplémentaire : Pour chaque sous-indice $j \in \{2, \ldots, s\}$ on a $i_j < 1 + \frac{n}{s}(j-1)$.

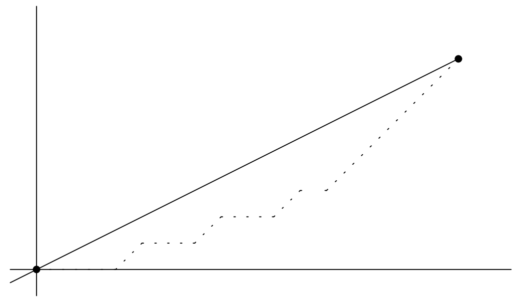

FIGURE 1. Calculer la trace compact de la fonction de Kottwitz $f_{n\alpha s}$ sur la représentation de Steinberg.

Dans le cas de Harris et Taylor [**45**] le polynôme $\mathrm{Pol}(q^\alpha)$ est égal à 1 (la strate basique est alors une variété *finie*).

La définition ci-dessus est courte, mais ne nous aide à comprendre ce qu'est ce polynôme. Dans la Figure 1, nous donnons une interprétation graphique pour $n = 16$ et $s = 8$. Nous traçons la ligne $\ell$ de pente $\frac{1}{2} = \frac{8}{16}$ passant par l'origine. Nous marquons l'origine $(0, 0)$ et le point $(16, 8)$. On considère certains chemins qui vont de l'origine au point $(16, 8)$. Ces chemins se composent en deux types d'étapes : celles qui vont vers l'est de la forme $(a, b) \to (a + 1, b)$ et celles qui vont vers le nord-est de la forme $(a, b) \to (a + 1, b + 1)$ (aucune autre étape n'est permise pour tracer les chemins). De plus les chemins doivent rester strictement sous de la ligne $\ell$. Soit $L$ un tel chemin, prenons le produit des puissances $p^{a\alpha}$ sur l'ensemble des étapes nord-est $(a, b) \to (a+1, b+1)$ qui font parti du chemin $L$. Ce produit est appellé *poids* de $L$ ; on le note $\mathrm{poids}(L)$. Le polynôme $P(q^\alpha)$ est égal à la somme des poids de tous les chemins qui vont de $(0, 0)$ vers le point $(16, 8)$.

Le lecteur remarquera que dans cet exemple nous avons mis de côte la condition selon laquelle $s$ est premier avec $n$. Dans le cas où $n$ et $s$ ont des diviseurs en commun, la formule ci-dessus donne toujours la trace compacte de la fonction de Kottwitz agissant sur la représentation de Steinberg (au signe près : on a $\mathrm{Tr}(\chi_c^{\mathrm{GL}_n(\mathbb{Q}_p)} f_{n\alpha s}, \mathrm{St}_{\mathrm{GL}_n(\mathbb{Q}_p)}) = (-1)^{n-1} P(p^\alpha))$. La formule pour la trace compacte sur la représentation triviale est presque la même, la seule chose qui change, c'est que, pour la représentation triviale, les chemins se trouvent aussi sous de la droit $\ell$, mais pas strictement : les chemins peuvent la toucher. Dans le cas où $n$ et $s$ sont premiers entre eux il n'y a pas de différence car il n'y a pas de point entier $(x, y)$ sur $\ell$ avec $0 < x < n$.

**Chapitre 3 : La strate basique et des exercices combinatoires.** Ce chapitre est la suite du Chapitre 2. Nous enlevons une hypothèse du théorème principal du chapitre précédent. Dans le dernier chapitre, nous avons (essentiellement) supposé que le polygone de Newton associé à la strate basique n'avait pas de point intégral autre que le point de début et le point final. Nous résolvons les problèmes combinatoires qui résultent de la supression de cette condition simplificatrice dans le cas où le nombre premier $p$ de réduction est complètement déployé dans le centre $F$ de l'algèbre à division $D$ qui définit la variété de Kottwitz.

Une conséquence de notre résultat final est une expression explicite de la fonction zêta de la strate basique. Les expressions sont en termes : (1) des formes automorphes sur le groupe $G$ de la donnée de Shimura, (2) du déterminant du facteur en $\mathfrak{p}$ de leur représentation galoisienne associée, et (3) des polynômes en $q^{\alpha}$, associés à certains chemins non-intersectant dans les treillis du plan $\mathbb{Q}^2$.

Avant que nous puissions donner l'énonce du résultat nous avons besoin d'introduire trois classes de représentations.

Considérons le groupe général linéaire $G_n = \mathrm{GL}_n(F)$ sur un corps local non-archimédien $F$.

Soient $x, y$ des entiers tels que $n = xy$. Nous définissons la représentation $\mathrm{Speh}(x, y)$ de $G_n$ : C'est l'unique quotient irréductible de la représentation $|\det|^{\frac{y-1}{2}}\mathrm{St}_{G_x} \times |\det|^{\frac{y-3}{2}}\mathrm{St}_{G_x} \times \cdots \times |\det|^{-\frac{y-1}{2}}\mathrm{St}_{G_x}$ où les produits "$\times$" signifient induction parabolique unitaire à partir du sous-groupe parabolique standard de $G_n$ avec chaque bloc de taille $x$. Une *représentation de Speh semi-stable* de $G_n$ est, par définition, une représentation isomorphe à $\mathrm{Speh}(x, y)$ pour des entiers positifs $x, y$ avec $n = xy$. Nous soulignons que nous n'avons pas défini toutes les représentations de Speh, nous avons seulement introduit celles qui sont semi-stables (ce qui est suffisant pour nos besoins ici).

Une représentation $\pi$ de $G_n$ est appelée *représentation rigide (semi-stable)* si elle est égale à un produit de la forme

$$\prod_{a=1}^{k} \mathrm{Speh}(x_a, y)(\varepsilon_a),$$

où $y$ est un diviseur de $n$ et $(x_a)$ est une partition de $\frac{n}{y}$, et les $\varepsilon_a$ sont des caractères unitaires non-ramifiés.

Une représentation $\pi$ du groupe $G(\mathbb{Q}_p) = \mathbb{Q}_p^\times \times \prod_{\wp|p} \mathrm{GL}_n(F_\wp^+)$ est appelé *représentation rigide (semi-stable)* si pour chaque $F^+$-place $\wp$ au-dessus de $p$, la composante $\pi_\wp$ est une représentation (semi-stable) rigide du groupe $\mathrm{GL}_n(F_\wp^+)$ dans la sens précédent :

$$\pi_\wp = \prod_{a=1}^{k} \mathrm{Speh}(x_{\wp,a}, y_\wp)(\varepsilon_{\wp,a}),$$

où deux conditions supplémentaires devraient être vraies : (1) $y_\wp = y_{\wp'}$ pour tout $\wp, \wp'|p$, et (2) le facteur de similitude $\mathbb{Q}_p^\times$ de $G(\mathbb{Q}_p)$ agit par un caractère non ramifié sur l'espace de $\pi$. Nous écrirons $y := y_\wp$ et on appelle l'ensemble des données $(x_{\wp,a}, \varepsilon_{\wp,a}, y)$ les *paramètres* de $\pi$.

Considérons une variété de Shimura de Kottwitz que nous avons introduit dans le paragraphe précédent. Cependant nous faisons deux changements :

– On oublie l'hypothèse sur les pentes de l'isocristal basique ;
– On ajoute la condition que le nombre premier $p$ est complètement déployé dans le centre de $F$ de $D$.

Nous avons alors :

THÉORÈME. *Soit $\alpha$ un entier positif. Alors*

$$\sum_{i=0}^{\infty}(-1)^i \operatorname{Tr}(f^{\infty p} \times \Phi_{\mathfrak{p}}^{\alpha}, \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_p}, \iota^*\mathcal{L})) = \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p \text{ est rigide}}} \operatorname{Tr}(\chi_c^G f_\alpha, \pi_p) \cdot \operatorname{Tr}(f^p, \pi^p). \tag{4.3}$$

On pourrait penser que le théorème ci-dessus est le résultat principal de ce chapitre, mais le travail n'est pas fini ici. Le but de ce chapitre est de calculer la trace compacte $\operatorname{Tr}(\chi_c^G f_\alpha, \pi_p)$ pour toute représentation rigide. Nous trouvons des expressions tout à fait explicites pour ces traces compactes en termes de chemins qui se ne coupent pas. Malheureusement, la définition de ces polynômes est trop technique pour être énoncée ici : on consultera le corps du chapitre pour les définitions. Nous nous contenterons d'un exemple d'un polynôme typique.

Considérons la représentation $\pi_p = \operatorname{Speh}(20, 4)$ de $\mathrm{GL}_{80}(\mathbb{Q}_p)$. Prenons deux copies du plan $\mathbb{Q}^2$ et traçons la ligne $\ell$ de pente $\frac{1}{2} = \frac{40}{80}$ passant par l'origine (voir la Figure 2). Dans la Figure 2, appelons $\ell_A$ la ligne sur le plan à gauche et $\ell_B$ la ligne sur le plan à droite. Sur la droite $\ell_A$ nous avons placé quatre points définis par :

$$\begin{aligned} \vec{x}_1 &:= (-8, -4) & \vec{y}_1 &:= (12, 6) \\ \vec{x}_3 &:= (-10, -5) & \vec{y}_3 &:= (10, 5) \end{aligned}$$

et sur $\ell_B$ quatre points définis par

$$\begin{aligned} \vec{x}_2 &:= (-9, -4\tfrac{1}{2}) & \vec{y}_2 &:= (11, 5\tfrac{1}{2}) \\ \vec{x}_4 &:= (-11, -5\tfrac{1}{2}) & \vec{y}_4 &:= (9, 4\tfrac{1}{2}). \end{aligned}$$

Ces points sont déterminés par des formules explicites à partir des segments de Zelevinsky de $\pi_p$. La pente des droites $\ell_A$ et $\ell_B$ est déterminée par le cocaractère de Shimura $\mu$. Les Figures 2A et 2B définiront chacune un polynôme ; voyons d'abord le définition du polynôme pour la Figure 2A (la définition du polynôme de la Figure 2B sera analogue). Comme le montre la figure, nous considérons des chemins qui relient le point $\vec{x}_3$ avec le point $\vec{y}_1$ et le point $\vec{x}_1$ avec le point $\vec{y}_3$. Ces chemins se composent de deux types d'étapes, les étapes vers l'est de la forme $(a, b) \to (a+1, b)$ et les étapes vers le nord-est de la forme $(a, b) \to (a+1, b+1)$ (aucune autre étape n'est permise dans les chemins). En outre, il y a deux conditions que les chemins doivent satisfaire : (C1) les chemins doivent rester strictement en-dessous de la ligne $\ell_A$ et, (C2) les chemins ne doivent pas se croiser. Nous appelons 2-*chemin* la donnée simultanée de deux chemins, l'un reliant les points $\vec{x}_3$ et $\vec{y}_1$, et l'autre reliant $\vec{x}_1$ et $\vec{y}_3$. Nous appelons *un* 2-*chemin de Dyck* un 2-chemin qui satisfait les conditions (C1) et (C2). A tout 2-chemin de Dyck $L$ on associe une certaine puissance de $p^\alpha$ ($\alpha$ est un entier positif fixé). Nous notons poids$(L)$ pour ce $p^\alpha$-puissance et nous l'appelons *poids de* $L$. Ce poids est défini comme suit. Pour $L$ donné, prenons le produit des $p^{a\alpha}$ sur l'ensemble des étapes nord-est $(a, b) \to (a + 1, b + 1)$ qui font partie du 2-chemin $L$. Le polynôme $P_A$ associée à la Figure 2A est alors la somme des poids de tous les 2-chemins de Dyck. Le polynôme associé à la Figure 2B est similaire ;

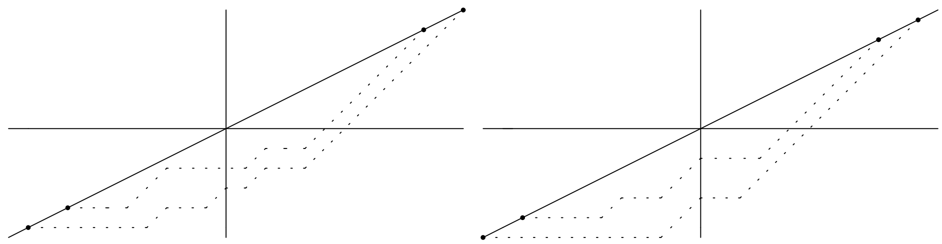

FIGURE 2. Exemple de chemins non-intersectants.

nous utilisons les points $\vec{x}_2, \vec{x}_4, \vec{y}_2, \vec{y}_4$. La trace compacte (de la fonction de Kottwitz $f_{n\alpha s}$) sur la représentation $\pi_p$ est alors le produit de $P_A$ avec $P_B$ (multiplié par un certain facteur de normalisation, que nous ignorons ici).

Un autre résultat de ce chapitre est le calcul de la dimension de la strate basique.

Avant d'énoncer notre résultat nous avons besoin d'introduire les nombres $s_v$. Plongeons le corps $F$ dans le corps $\mathbb{C}$. Considérons le sous-groupe $U$ formé des éléments $g \in G$ dont le facteur de similitude est égal à 1. Ce sous-groupe est obtenu par restriction à $\mathbb{Q}$ d'un groupe unitaire définie sur le corps $F^+$. Donc on a $U(\mathbb{R}) = \prod_{v \in \mathrm{Hom}(F^+, \mathbb{R})} U(s_v, n - s_v)$ pour des entiers $s_v \in \mathbb{Z}$ avec $0 \leq s_v \leq \frac{1}{2}n$.

THÉORÈME. *La dimension de $B$ est égal à :*

$$\sum_{v \in \mathrm{Hom}(F^+, \mathbb{C})} \left( \frac{s_v(1 - s_v)}{2} + \sum_{j=0}^{s_v - 1} \lceil j \frac{n}{s_v} \rceil \right).$$

**Chapitre 4 : Les strates de Newton sont non vides.** Considérons une variété de Shimura de type PEL et réduisons modulo un nombre premier $p$ de bonne réduction. La variété de Shimura paramétrise des variétés abéliennes en caractéristique $p$ avec certaines structures additionnelles de type PEL. À chaque variété abélienne nous pouvons associer son isocristal de Dieudonné. Les structures PEL sur la variété abélienne donne des structures PEL sur l'isocristal, et en tant que tels les isocristaux se situent dans la catégorie des "isocristaux avec structures additionnelles" (Kottwitz [**55**]). Nous regardons ces objets à isomorphisme près. Il n'est pas vrai que chaque $G$-isocristal résulte d'un point géométrique sur cette variété. En fait, il y a seulement un nombre fini d'isocristaux possibles ; depuis les travaux de Rapoport-Richarz et Kottwitz [**60**, **88**] nous savons qu'ils se trouvent tous dans un certain ensemble fini $B(G_{\mathbb{Q}_p}, \mu)$ d'isocristaux "admissibles", mais ils n'ont pas montré que $B(G_{\mathbb{Q}_p}, \mu)$ est *exactement* l'ensemble des possibilités : Il n'était pas clair que pour chaque élément $b \in B(G_{\mathbb{Q}_p}, \mu)$ il existe une variété abélienne en caractéristique $p$ avec structures additionnelles de type PEL dont ce module de Dieudonné rationnel est égal à $b$. Récemment Wedhorn et Viehmann [**104**] ont prouvé par des moyens géométriques que c'est effectivement le cas si le groupe de la donnée de Shimura est de type (A) ou (C). Dans ce chapitre, nous allons montrer que l'on peut également démontrer ce résultat en utilisant les formes automorphes et la formule de trace dans le cas où le groupe est de type (A). Au moment de la rédaction de ce chapitre, Sug Woo Shin, dans une conférence de BIRS, a annoncé une démonstration de ce résultat différent de celle de Viehmann-Wedhorn et de la nôtre.

En ce moment, nous sommes en train d'écrire la preuve pour le cas (C). Nous pensons que notre méthode donne également une preuve à certains variétés de Shimura de type Hodge, au moins dans les cas où le groupe est classique, si l'on peut démontrer pour ces variétes la formule de Kottwitz.

**Chapitre 5 : Équidistribution.** Nous démontrons un résultat d'équidistribution pour les opérateurs de Hecke agissant sur la strate basique des variétés de Kottwitz dans les cas où cette strate est une variété finie. Nous pouvons montrer que le taux de convergence est aussi bon que la borne qui provient de la conjecture de Ramanujan.

Considérons une variété de Kottwitz comme dans le Chapitre 2, mais faisons l'hypothèse supplémentaire que la strate basique est une variété finie. Nous supposons aussi que l'image de $K$ dans le cocentre de $G$ soit maximale.

Soit $A$ l'espace vectoriel complexe sur l'ensemble des points géométriques de la strate basique. Fixons une norme $|\cdot|$ sur l'espace vectoriel $A$. L'espace $A$ est un module sur l'algèbre de Hecke. Soit $T_{r,m}$ l'opérateur de Hecke dans l'algèbre $\mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^p))$ qui est obtenu par changement de base, de l'opérateur de Hecke habituel $T_{r,m}$ du groupe $G(\mathbb{A}_{\mathrm{f}}^p \otimes \mathcal{K})$ (qui est isomorphe à un produit de groupes linéaires genéraux). Le lecteur peut trouver la définition précise de cette suite d'opérateurs de Hecke dans la Section 5.2. Sur l'espace $A$ on définit l'endomorphisme "moyenne", Moy, qui à un vecteur $v$ associe sa moyenne sur les fibres de la flèche $\mathrm{Sh}_K(\overline{\mathbb{F}}_p) \to \pi_0(\mathrm{Sh}_K)(\overline{\mathbb{F}}_p)$.

Nous prouvons le résultat d'équidistribution ci-dessous :

Théorème. *Soit $v \in A$ un élément. Alors il existe une constante $C \in \mathbb{R}_{>0}$ ayant la proprieté suivante. Pour tout $\varepsilon > 0$, il existe un entier $M$, tel que pour tout entier $m > M$, sans facteur carré, et tout $r$ avec $1 \leq r \leq n-1$, nous avons*

$$\left| \frac{T_{r,m}(v)}{\deg(T_{r,m})} - \mathrm{Moy}(v) \right| \leq C m^{\varepsilon - [F:\mathbb{Q}] \frac{r(n-r)}{2}}.$$

Le théorème peut être prouvé aussi pour d'autres suites d'opérateurs de Hecke, mais — bien sûr — le taux de convergence dépend du choix de la suite.

Nous avons aussi un résultat partiel pour une large classe de variétés de Shimura de type PEL unitaires, mais toujours dans l'hypothèse où la strate basique est finie. Nous prévoyons d'être en mesure de prouver un résultat d'équidistribution, avec probablement un taux de convergence similaire, mais nous avons encore à estimer certains termes dans les expressions.

**Annexe A : Existence de représentations cuspidales.** Nous montrons que tout groupe réductif connexe $G$ sur un corps local non-archimédien a une représentation cuspidale complexe.

Nous n'avons pas utilisé ce résultat dans cette thèse, donc l'appendice est indépendant du reste de la thèse. Nous l'utilisons seulement pour le groupe général linéaire, pour lequel le résultat est bien connu. En fait, dans la littérature, il est souvent supposé que l'existence de représentations cuspidales est connue, mais nous n'avons pas trouvé de référence. Cette annexe pourrait combler cette lacune.

Nous avons besoin du résultat pour l'extension des résultats du chapitre 4 à certains variétés de Shimura de type Hodge. Actuellement, nous travaillons sur ce résultat, et cette annexe sera nécessaire dans cette preuve.

**Annexe B : Modules de Jacquet de représentations en échelle (avec Erez Lapid).** Nous calculons explicitement la semi-simplification des modules de Jacquet de représentations en échelle (anglais : "ladder representations").

Ce résultat est nécessaire (et presque suffisant) si on veut étendre les résultats des Chapitres 2 et 3 aux autres strates de Newton. Malheureusement, nous n'avons pas eu le temps de compléter ce travail. Nous avons donc choisi d'inclure le résultat sur les modules de Jacquet comme une annexe qui ne dépend pas du reste de la thèse.

L'énoncé précis du résultat n'est pas plus long que les premières pages de l'annexe B. Par conséquent, nous renvoyons le lecteur à l'annexe B pour le théorème.

CHAPTER 1

# The modular curve

We explain a new method to count points in the supersingular locus of the modular curves $Y_1(N)$. We will count the number of supersingular points in the set $Y_1(N)(\mathbb{F}_{p^\alpha})$, where $N$ is an integer greater or equal than 3, $\alpha$ is a positive integer, $p$ is a prime number which does not divide $N$, and $\mathbb{F}_{p^\alpha}$ is a finite field of order $p^\alpha$. The final result is Theorem 3.3.

Our computation of the number of supersingular points on $Y_1(N)$ is a variation on the classical method of Ihara-Langlands (refined by Kottwitz) [**24**, **46**, **47**, **62**, **69**, **70**, **76**]. This classical method computes the cardinality of the full set $Y_1(N)(\mathbb{F}_{p^\alpha})$ of elliptic curves over $\mathbb{F}_{p^\alpha}$ with $\Gamma_1(N)$-level structure. We alter the computation to calculate instead the number of supersingular elliptic curves with $\Gamma_1(N)$-level structure.

Our result is certainly not new: The final result of this chapter (Theorem 3.3) follows directly from the result of Deligne and Rapoport [**34**]. However, our argument is completely different from theirs, and in later chapters we show that our method also works for higher dimensional Shimura varieties. Thus, it is not really the end result of this chapter which is important, it is rather the method of proof.

This chapter is technically less demanding than the other chapters of this thesis. We try to avoid generality: We replace references to general arguments/theorems by short and simple calculations which are valid for $\mathrm{GL}_2$ but not necessarily for any other group. As a consequence, some of the statements that we prove in this chapter will be a special case of lemmas and propositions that we prove in later chapters.

Our aim in this introductory chapter is not to prove the most general result possible, even for $\mathrm{GL}_2$. We just want to explain the method for an easy example. In particular the reader will notice that we have included some simplifying conditions which are not really needed, but do make the text more readable.

**Notations:** The letter $G$ denotes the algebraic group $\mathrm{GL}_2$ over the integers $\mathbb{Z}$. The group $B$ is the standard Borel subgroup of $G$, $T \subset B$ is the standard maximal torus on the diagonal, and $Z \subset T$ is the center of $G$. We write $Z(\mathbb{R})^+$ for the topological neutral component of the group $Z(\mathbb{R})$. Let $I_p \subset G(\mathbb{Z}_p)$ the group consisting of those matrices $g \in G(\mathbb{Z}_p)$ such that $\overline{g} \in B(\mathbb{F}_p)$ (the standard Iwahori subgroup). The field $L$ is the completion of a maximal unramified extension $\mathbb{Q}_p^{\mathrm{nr}}$ of $\mathbb{Q}_p$. We write $\mathcal{O}_L$ for the ring of integers of $L$. Let $\alpha$ be a positive integer. Let $\mathbb{Q}_{p^\alpha} \subset L$ be the subfield of degree $\alpha$ over $\mathbb{Q}_p$, we let $\mathbb{Z}_{p^\alpha} \subset \mathbb{Q}_{p^\alpha}$ be the ring of integers of $\mathbb{Q}_{p^\alpha}$, and $\mathbb{F}_{p^\alpha}$ is the residue field of $\mathbb{Z}_{p^\alpha}$. Finally $\overline{\mathbb{Q}}_p$ is an algebraic closure of $\mathbb{Q}_p$

containing ${\mathbb{Q}_p}^{\mathrm{nr}}$, the subring $\overline{\mathbb{Z}}_p \subset \overline{\mathbb{Q}}_p$ is the ring of integers and $\overline{\mathbb{F}}_p$ is by definition the residue field of $\overline{\mathbb{Z}}_p$.

## 1. The modular curve

Consider the complex double half plane $\mathfrak{h}^{\pm} = \{z \in \mathbb{C} \mid \Im(z) \neq 0\}$ on which the group $G(\mathbb{R})$ acts by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+c}{bz+d}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G(\mathbb{R}), \quad z \in \mathfrak{h}^{\pm}.$$

We pick the point $i \in \mathfrak{h}^{\pm}$, and define $K_\infty \subset G(\mathbb{R})$ to be the stabilizer of $i$. We define the morphism $h \colon \mathbb{C}^\times \to G(\mathbb{R})$ by $(a+bi) \mapsto \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. The image of $h$ is the group $K_\infty$ and the orbit of $h$ under the conjugation action of $G(\mathbb{R})$ is equal to $\mathfrak{h}^{\pm}$. The couple $(G, \mathfrak{h}^{\pm})$ is a Shimura datum.

Let $N$ be an integer with $N \geq 4$. Let $K_1(N)$ be the subgroup of $G(\widehat{\mathbb{Z}})$ consisting of the matrices $g \in G(\widehat{\mathbb{Z}})$ such that $\overline{g} = \left(\begin{smallmatrix} 1 & * \\ 0 & * \end{smallmatrix}\right) \in G(\mathbb{Z}/N\mathbb{Z})$. We have the (complex points of the) Shimura variety $\mathrm{Sh}(G, K_1(N))$:

$$G(\mathbb{Q})\backslash G(\mathbb{A})/K_\infty K_1(N) = G(\mathbb{Q})\backslash \mathfrak{h}^{\pm} \times G(\mathbb{A}_{\mathrm{f}})/K_1(N). \tag{1.1}$$

The variety $\mathrm{Sh}(G, K_1(N))$ is equal to the modular curve $Y_1(N)$ over the reflex field $\mathbb{Q}$.

The curve $Y_1(N)$ has a natural model over the ring $\mathbb{Z}[1/N]$, for which we also write $Y_1(N)$. This model represents the following functor. For any scheme $S$ with $N \in \mathcal{O}_S(S)^\times$ the set $Y_1(N)(S)$ is equal to the set of equivalence classes of pairs $(E, P)$ consisting of an elliptic curves $E/S$ and $P \in E(S)$ a point of order $N$. Two pairs $(E_1, P_1)$, $(E_2, P_2)$ are *equivalent* if there is an $S$-isomorphism of elliptic curves $E_1 \xrightarrow{\sim} E_2$ sending the point $P_1$ to the point $P_2$.

## 2. The Ihara-Langlands method

We recall a classical theorem of Ihara, Langlands, Kottwitz and also Milne. This theorem expresses the number of points on the curve $Y_1(N)$ in terms of orbital integrals on the group $G(\mathbb{A})$.

THEOREM 2.1 (Ihara-Langlands-Kottwitz [**62**]). *Let $\alpha$ be a positive integer. Then we have that*

$$\# Y_1(N)(\mathbb{F}_{p^\alpha}) = \sum_\gamma v(\gamma) O_\gamma(f), \tag{2.1}$$

*where*

- *$\gamma$ ranges over a set of representatives for the set of the regular semi-simple $\mathbb{R}$-elliptic $G(\mathbb{Q})$-conjugacy classes;*

– $f_\infty$ *is a smooth function on* $G(\mathbb{R})$*, with compact support on* $G(\mathbb{R})/Z(\mathbb{R})$*, such that the following property holds. There exists a choice of Haar measures such that the orbital integral* $O_\gamma(f_\infty) = 0$ *for* $\gamma$ *regular semi-simple non-elliptic and* $O_\gamma(f_\infty) = \pm 1$ *for* $\gamma$ *regular elliptic. The sign of* $O_\gamma(f_\infty)$ *is* $-1$ *is* $\gamma$ *is central, and equal to* $1$ *otherwise;*
– $f_\alpha$ *is the function in the unramified Hecke algebra* $\mathcal{H}_0(G(\mathbb{Q}_p))$ *whose Satake transform is equal to* $p^{\alpha/2}(X^\alpha + Y^\alpha) \in \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]^{\mathfrak{S}_2} \cong \mathcal{H}_0(G(\mathbb{Q}_p))$*;*
– $f^{p,\infty}$ *is the characteristic function of the compact open subgroup* $K_1(N)^p \subset G(\mathbb{A}_{\mathrm{f}}^p)$*;*
– *we write* $f = f_\infty \otimes f_\alpha \otimes f^p$*;*
– $v(\gamma)$ *is the volume term* $\mathrm{Vol}(G_\gamma(\mathbb{Q})\backslash G_\gamma(\mathbb{A}_{\mathrm{f}}))$ *with respect to certain normalized Haar measures.*

The proof of the above theorem has been carried out in detail by Kottwitz in a course he gave in Orsay, see the notes [**62**] (cf. [**54**]) and see also the article [**76**]. Note that these results are also proved in the (published) articles [**58**,**59**], but in much greater generality than needed here. Clozel gives a summary of the argument for $\mathrm{GL}_2$ in his Bourbaki talk [**24**]. See also [**12**].

We will use a slightly stronger statement than Theorem 2.1. In fact, the proof of the above theorem gives more than the theorem states: Both the left hand side and the right hand side of Equation (2.1) decompose along isogeny classes, as follows. For any elliptic curve $E \in Y_1(N)(\mathbb{F}_{p^\alpha})$ we look at the subset $Y_1(N)(\mathbb{F}_{p^\alpha})(E) \subset Y_1(N)(\mathbb{F}_{p^\alpha})$ consisting of those $E' \in Y_1(N)(\mathbb{F}_{p^\alpha})$ such that $E$ is isogenous to $E'$. To $E$ we may associate, via Honda-Tate theory, an element $\gamma \in \mathrm{GL}_2(\mathbb{Q})$. We make the following complement to Theorem 2.1:

$$\#Y_1(N)(\mathbb{F}_{p^\alpha})(E) = v(\gamma) O_\gamma(f). \tag{2.2}$$

By taking the sum over all isogeny classes one will recover Theorem 2.1.

We use this last Equation to count supersingular elliptic curves.

## 3. The number of supersingular points on the modular curve

Define $Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha})$ to be the subset of $Y_1(N)(\mathbb{F}_{p^\alpha})$ consisting of the *supersingular* elliptic curves with $K_1(N)$-level structure. The goal of this section is to give an expression for the cardinal $\#Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha})$.

We restrict the sum in Theorem 2.1 to run only over those conjugacy classes $\gamma$ whose eigenvalues have the same $p$-adic valuation (cf. Equation (2.2)). The formula will then count the isogeny classes of supersingular elliptic curves. To achieve this, let $\chi$ be the characteristic function of the set $\Omega$ of elements $g \in G(\mathbb{Q}_p)$ whose eigenvalues all have the same $p$-adic absolute value. The set $\Omega \subset G(\mathbb{Q}_p)$ is invariant under $G(\mathbb{Q}_p)$-conjugation and open and closed. In particular the function $\chi f_\alpha$ lies in the space $C_c^\infty(G(\mathbb{Q}_p))$. One checks easily that the orbital integral $O_\gamma(\chi f_\alpha)$ is equal to the orbital integral $O_\gamma(f_\alpha)$ for elements $\gamma$ in $\Omega$ and that the orbital integral $O_\gamma(\chi f_\alpha)$ vanishes for any element $\gamma$ not lying in $\Omega$. Consequently the cardinal $\#Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha})$ is equal to the sum $\sum_\gamma v(\gamma) O_\gamma(\chi f)$.

The trace formula of Selberg[1] applied to the function $\chi f$ reads

$$\sum_\gamma v(\gamma) O_\gamma(\chi f) + \sum_\gamma \mathrm{Vol}(A(\mathbb{Q})\backslash A(\mathbb{A}_{\mathrm{f}})) O_\gamma(f^p \chi f_\alpha) \left( \frac{2}{|1 - t_1/t_2|_\infty} \right) \tag{3.1}$$
$$= \sum_\pi \mathrm{Tr}(\chi f, \pi),$$

where $\pi$ in the ranges over the irreducible subspaces of $L^2(Z(\mathbb{R})^+ G(\mathbb{Q})\backslash G(\mathbb{A}))$, and in the first sum $\gamma$ ranges over the semi-simple $G(\mathbb{Q})$ conjugacy classes which are $G(\mathbb{R})$-elliptic, and in the second sum $\gamma$ ranges over the rational elements $\gamma = \left(\begin{smallmatrix} t_1 & \\ & t_2 \end{smallmatrix}\right)$ of $A(\mathbb{Q})^+$ such that $|t_1|_\infty > |t_2|_\infty$.

The second large sum in Equation (3.1) is the *corrective term*; we claim that this corrective term vanishes (for our Hecke function). By the properties of the Satake transformation[2] the orbital integral $O_{\gamma_p}(f_\alpha)$ is non-zero only if $\gamma_p$ is elliptic or if one of its eigenvalues has non-zero $p$-adic valuation and the other one $p$-adic valuation equal to zero. Assume that the conjugacy class $\gamma_p$ contains an element the split torus $A(\mathbb{Q}_p)$ and assume that the integral $O_{\gamma_p}(f_\alpha \chi)$ is non-zero. The conjugacy class $\gamma_p$ lies in $A(\mathbb{Q}_p)$, and thus cannot be elliptic, and therefore the $p$-adic valuations of its eigenvalues are different (one is zero and the other one is $\alpha$). However, $\gamma$ is compact and therefore the valuations of its eigenvalues are equal. This is a contradiction, and thus the integral $O_{\gamma_p}(f_\alpha \chi)$ is zero for $\gamma_p \in A(\mathbb{Q}_p)$. This proves the claim.

The corrective term in Equation (3.1) vanishes and we obtain simply

$$\#Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha}) = \sum_\gamma v(\gamma) O_\gamma(\chi f) = \sum_\pi \mathrm{Tr}(\chi f, \pi). \tag{3.2}$$

In the following 3 subsections we will compute the traces $\mathrm{Tr}(\chi f, \pi)$ for all discrete automorphic representations $\pi$ of $G(\mathbb{A})$.

**3.1. A Local Computation at $p$.** Let us first focus on the trace at $p$ in Equation (3.2). To simplify notations, we write $G$ for the group $G(\mathbb{Q}_p)$, $T$ for $T(\mathbb{Q}_p)$, $B$ for $B(\mathbb{Q}_p)$, and $N$ for $N(\mathbb{Q}_p)$ in this subsection. The computation of the traces at $p$ is easy using Clozel's formula for compact traces (see [**22**, p. 259] or Proposition 2.1.2 of this thesis). The formula applied to $\mathrm{GL}_2$ states

$$\mathrm{Tr}(\chi f_\alpha, \pi) = \mathrm{Tr}(f_\alpha, \pi) - \mathrm{Tr}_T\left( \widehat{\chi}_N f_\alpha^{(B)}, \pi_N(\delta_B^{-1/2}) \right), \tag{3.3}$$

where we need to recall some definitions:

- The symbol $\pi$ is a smooth representation of $G$ of finite length.
- The $T$-representation $\pi_N$ is the *Jacquet module* of $\pi$, i.e. the $\mathbb{C}[T]$-module of $N$-coinvariants in $\pi$.

1. See Duflo & Labesse [**38**]. Note that this reference gives the trace formula for $\mathrm{PGL}_2$, not for $\mathrm{GL}_2$. The formula applies to the case at hand, because the automorphic representations $\pi$ for which $\mathrm{Tr}(\chi f, \pi)$ is non-zero have trivial central character ($\omega_\pi$ is trivial on $\det(K_1(N)^p) = \widehat{\mathbb{Z}}^{p,\times}$ and trivial on $\mathbb{Z}_p^\times$ by Proposition 3.1).

2. For a proof, see for example [**72**, thm 4.5.5], or the argument at Proposition 2.1.7.

– The character $\delta_B$ is the *modulus character* of $B$ with respect to a right[3] Haar measure. Explicitly, we have the formula $\delta(x) = |\det(x, \mathrm{Lie}(N))| = |ad^{-1}|$ if $x = \begin{pmatrix} a & \\ & d \end{pmatrix} \in T$.
– The function $\widehat{\chi}_N$ is the characteristic function of those matrices $\begin{pmatrix} a & \\ & d \end{pmatrix} \in T$ with $|ad^{-1}| < 1$. Note that the function $\widehat{\chi}_N$ is not defined in this manner in the reference [**22**, p. 259], but see Section 2.1.5 of this thesis for the proof that the function $\widehat{\chi}_N$ satisfies the above description in case the group is $\mathrm{GL}_2$.
– The function $f_\alpha^{(B)} \colon T \to \mathbb{C}$ is the *constant term* of $f$ at $B$; it is defined by

$$f_\alpha^{(B)}(t) = \delta_B^{-1/2}(t) \int_N f(tn)\mathrm{d}n,$$

for all elements $t$ of the torus $T$. Here the Haar measure on $N$ is normalized so that it is compatible with the Haar measure on $G$ via the Iwasawa decomposition $G = KAN$.

By definition, a representation $\pi$ of $G$ is *semi-stable* if it has invariant vectors for the standard Iwahori subgroup. A first consequence of Formula (3.3) is that $\mathrm{Tr}(\chi f_\alpha, \pi)$ is non-zero only if $\pi$ is semi-stable. Thus, there are no cuspidal representations which contribute. To see this: From Formula 3.3 follows that if the trace $\mathrm{Tr}(\chi f_\alpha, \pi)$ is nonzero, then the representation $\pi$ is unramified or the Jacquet module $\pi_N$ is nonzero. The result in Proposition 2.4 of [**16**] states that the vector space $(\pi_N)^{\mathbb{Z}_p^{\times 2}}$ is isomorphic to the vector space $\pi^{I_p}$. Thus, in both cases it follows that $\pi$ is *semi-stable*.

Assume from now on that the representation $\pi$ is semi-stable. These semi-stable representations are classified [**13**, thm 9.11] and divided into 3 groups:

(**1**) The *irreducible* representations of the form $\mathrm{Ind}_T^G(\chi)$, where $\chi \colon T \to \mathbb{C}^\times$ is an unramified character (the induction is unitary).

(**2**) The unramified, one-dimensional smooth representations.

(**3**) The semi-stable special representations. These are the twists of the Steinberg representation $\mathrm{St}_G$ by an unramified character of $G$.

We compute the compact traces on the representations in the above list:

PROPOSITION 3.1. *The following statements are true:*

(*i*) *Let $\chi$ be an unramified character of the torus $T$. Then:*

$$\mathrm{Tr}\left(\chi f_\alpha, \mathrm{Ind}_B^G(\chi)\right) = 0.$$

(*ii*) *Let $\phi$ be an unramified character of the group $G$. Then:*

$$\mathrm{Tr}\left(\chi f_\alpha, \mathbb{C}(\phi)\right) = \phi\begin{pmatrix} p & \\ & 1 \end{pmatrix}^{-\alpha}.$$

(*iii*) *Let $\phi$ be an unramified character of the group $G$. Then:*

$$\mathrm{Tr}\left(\chi f_\alpha, \mathrm{St}_G(\phi)\right) = -\phi\begin{pmatrix} p & \\ & 1 \end{pmatrix}^{-\alpha}.$$

3. We often refer to the book of Henniart and Bushnell [**13**] in this text. Note that in [*loc. cit*] the modulus character is normalized with respect to a *left* Haar measure, hence some sign differences appear.

PROOF. In each case we compute using Formula (3.3). Let us first make the function $\widehat{\chi}_N f_\alpha^{(B)}$ explicit (this function occurs in Formula (3.3)). We have

$$f_\alpha^{(B)} = p^{\alpha/2} \cdot \left( \mathbf{1}_{|t_1|=p^{-\alpha},|t_2|=1} + \mathbf{1}_{|t_1|=1,|t_2|=p^{-\alpha}} \right)$$

as function in the Hecke algebra $\mathcal{H}(T)$ of $T$. Hence

$$\widehat{\chi}_N f_\alpha^{(B)} = p^{\alpha/2} \cdot \mathbf{1}_{|t_1|=p^{-\alpha},|t_2|=1} \in \mathcal{H}(T).$$

We begin with case $(i)$. Let $w$ be the matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ in $G$. Let the characters $\chi_i \colon \mathbb{Q}_p^\times \to \mathbb{C}^\times$ for $i = 1, 2$ be such that $\chi\left(\begin{smallmatrix} a & b \\ & d \end{smallmatrix}\right) = \chi_1(a)\chi_2(d)$. Let $\chi^w$ be the character $w^{-1}\chi w$, i.e. the character on the torus $T$ given by $\chi^w\left(\begin{smallmatrix} d & \\ & a \end{smallmatrix}\right) = \chi_1(a)\chi_2(d)$. The Jacquet module $\pi_N(\delta_B^{-1/2})$ is[4] equal to $\mathbb{C}(\chi) \oplus \mathbb{C}(\chi^w)$. Hence

$$\mathrm{Tr}\left(\widehat{\chi}_N f_\alpha^{(B)}, \pi_N(\delta_B^{-1/2})\right) = p^{\alpha/2} \cdot \left(\chi_1(p^{-\alpha}) \cdot 1 + 1 \cdot \chi_2(p^{-\alpha})\right).$$

We have

$$\mathrm{Tr}(f_\alpha, \pi) = p^{\alpha/2} \cdot \left(\chi_1(p^{-\alpha}) + \chi_2(p^{-\alpha})\right).$$

Thus Clozel's Formula 3.3 implies that the trace $\mathrm{Tr}(\chi f_\alpha, \pi)$ vanishes.

We now do case $(ii)$. The representation $\pi$ is one-dimensional, isomorphic to $\mathbb{C}(\phi)$, where $\phi \colon G \to \mathbb{C}^\times$ is an unramified character. We have

$$\mathrm{Tr}(\widehat{\chi}_N f_\alpha^{(B)}, \pi_N) = \mathrm{Tr}(p^{\alpha/2} \cdot \mathbf{1}_{|t_1|=p^{-\alpha},|t_2|=1}, \mathbb{C}(\phi\delta_B^{-1/2})) = p^\alpha \phi\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-\alpha}.$$

We have

$$\mathrm{Tr}(f_\alpha, \pi) = p^{\alpha/2}(p^{\alpha/2} + p^{-\alpha/2})\phi\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-\alpha}.$$

Thus, by Formula 3.3:

$$\mathrm{Tr}(\chi f_\alpha, \pi) = \phi\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-\alpha}.$$

Assume for case $(iii)$ that $\pi \cong \mathrm{St}_G(\phi)$, where $\phi$ is an unramified character $G \to \mathbb{C}^\times$. We have an exact sequence $\mathbf{1}(\phi) \rightarrowtail \mathrm{Ind}_B^G(\mathbf{1})(\phi) \twoheadrightarrow \mathrm{St}_G(\phi)$. Therefore the trace $\mathrm{Tr}(f_\alpha\chi, \mathrm{Ind}_B^G(\mathbf{1})(\phi))$ is equal to the sum $\mathrm{Tr}(f_\alpha\chi, \mathrm{St}_G(\phi)) + \mathrm{Tr}(f_\alpha\chi, \mathbf{1}(\phi))$. The result now follows by combining $(i)$ and $(ii)$. This completes the proof. □

**3.2. The trace at infinity.** The trace at infinity $\mathrm{Tr}(f_\infty, \pi_\infty)$ is computed by Kottwitz for the determination of the Zeta function of the modular curve [**62**]. We recall the result in this subsection.

We define a certain discrete series representation $\pi_\infty^0$ of $G(\mathbb{R})$. Consider the induced representation $I(\chi) := \mathrm{Ind}_{B(\mathbb{R})}^{G(\mathbb{R})}(\chi)$ where $\chi\left(\begin{smallmatrix} t_1 & * \\ 0 & t_2 \end{smallmatrix}\right) := |t_1|^{\frac{1}{2}}|t_2|^{-\frac{1}{2}}$. The semi-simplification $I(\chi)^{\mathrm{ss}}$ is equal to the direct sum of the trivial representation of $G(\mathbb{R})$ and a discrete series representation $\pi_\infty^0$. This defines the representation $\pi_\infty^0$.

4. See for example the restriction-induction Lemma in [**13**, p. 63].

PROPOSITION 3.2. *Let $\pi_\infty$ be an irreducible admissible $G(\mathbb{R})$-representation. The trace of $f_\infty$ on $\pi$ vanishes unless the isomorphism class of the representation $\pi_\infty$ lies in the set $\{\mathbf{1}, \mathbf{1}(\mathrm{sign} \circ \det), \pi_\infty^0\}$. The trace of $f_\infty$ on $\mathbf{1}$ and $\mathbf{1}(\mathrm{sign}\circ\det)$ is equal to 1, and the trace of $f_\infty$ on $\pi_\infty^0$ is equal to $-1$.*

PROOF. For a proof, see the notes of Kottwitz [**62**]. □

**3.3. The number of supersingular points.** In this section we compute the number of supersingular points on the modular curve $Y_1(N)$ at a prime of good reduction.

We need to consider a certain finite cover of the curve $Y_1(N)$. We define:

$$K'(N) \stackrel{\text{def}}{=} \left\{ g \in G(\widehat{\mathbb{Z}}) \mid g \equiv \left(\begin{smallmatrix} 1 & * \\ & * \end{smallmatrix}\right) \mod N,\ g \equiv \left(\begin{smallmatrix} * & * \\ & * \end{smallmatrix}\right) \mod p \right\}.$$

We have $K'(N) = K_1(N)^p I_p$. Thus we have replaced the component $K_1(N)_p = G(\mathbb{Z}_p)$ at $p$ of the group $K_1(N)$ with the Iwahori group $I_p$. This way we get the compact open group $K' := K_1(N)^p I_p \subset G(\widehat{\mathbb{Z}})$. We let $Y'(Np)$ be the Shimura variety $\mathrm{Sh}(G, K')$, it is a smooth quasi-projective curve defined over $\mathbb{Q}$, and a finite cover of $Y_1(Np) \otimes \mathbb{Q}$. We write $X'(Np)$ for the compactification of $Y'(Np)$ (see [**51**, chap. 8]).

Let $\phi$ be the non-trivial unramified character of $G(\mathbb{Q}_p)$ whose square is 1. Define the constant $\varepsilon(\pi_p)$ for a smooth irreducible representation of $G(\mathbb{Q}_p)$ to be 1 if $\pi_p$ is isomorphic to $\mathrm{St}_G$, to be $-1$ if $\pi_p$ is isomorphic to $\mathrm{St}_G(\phi)$ and to be equal to 0 for all other representations.

THEOREM 3.3. *Let $\alpha$ be a positive integer. If $\alpha$ is even we have*

$$\# Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha}) = 1 + \mathrm{genus}(X'(Np)) - 2 \cdot \mathrm{genus}(X_1(N)).$$

*If $\alpha$ is odd we have*

$$\# Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha}) = 1 + \sum_{\pi} \dim(\pi_{\mathrm{f}})^{K'} \cdot \varepsilon(\pi_p),$$

*where $\pi$ ranges over those irreducible subspaces of $L^2_0(G(\mathbb{Q})Z(\mathbb{R})^+ \backslash G(\mathbb{A}_{\mathrm{f}}))$ with*
- *$\pi_\infty \cong \pi_\infty^0$;*
- *$\pi_p$ is an unramified twist of the Steinberg representation of $G(\mathbb{Q}_p)$.*

PROOF. By applying Proposition 3.1 and Equation (3.2) we see that the cardinal $\# Y_1(N)_{\mathrm{ss}}(\mathbb{F}_{p^\alpha})$ is equal to

$$\sum_{\pi} \mathrm{Tr}(\chi f, \pi) = \sum_{\pi, \pi_p \in (\mathbf{2})} \mathrm{Tr}(f_\infty f^p, \pi^p) \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-\alpha} + \sum_{\pi, \pi_p \in (\mathbf{3})} \mathrm{Tr}(f_\infty f^p, \pi^p) \left( -\phi_\pi \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-\alpha} \right), \tag{3.4}$$

where in each sum $\pi$ ranges over the irreducible $G(\mathbb{A})$-subspaces of $L^2(Z(\mathbb{R})^+ G(\mathbb{Q}) \backslash G(\mathbb{A}))$. The notation "$\pi_p \in (\mathbf{2})$" refers to the classification of semi-stable representations of $G(\mathbb{Q}_p)$ on page 21 and similarly for the notation "$\pi_p \in (\mathbf{3})$".

The discrete spectrum $L^2_{\mathrm{disc}}(Z(\mathbb{R})^+ G(\mathbb{Q}) \backslash G(\mathbb{A}))$ of $G$ decomposes as a direct sum of the cuspidal spectrum $L^2_0(Z(\mathbb{R})^+ G(\mathbb{Q}) \backslash G(\mathbb{A}))$ and the residual spectrum $L^2_{\mathrm{res}}(Z(\mathbb{R})^+ G(\mathbb{Q}) \backslash G(\mathbb{A}))$. The residual spectrum is the Hilbert direct sum of the spaces $\mathbb{C}(\varepsilon \circ \det)$, where $\varepsilon$ ranges over the characters of $\mathbb{R}_{>0}^{\times}\mathbb{Q}^{\times}\backslash\mathbb{A}^{\times}$. If $\pi$ is an irreducible cuspidal $G(\mathbb{A})$-representation then all its local factors are infinite dimensional[5]. Therefore the first sum on the right hand side of Equation (3.4) runs over the residual spectrum of $G$ and the second sum runs over the cuspidal spectrum.

Let $\pi = \mathbb{C}(\varepsilon \circ \det)$ be a residual automorphic representation of $G$ such that the trace $\mathrm{Tr}(f\chi, \pi)$ is nonzero. The character $\varepsilon_\infty$ is trivial on the set $\det(K_1(N)^p) = \widehat{\mathbb{Z}}^{p\times}$. By Proposition 3.1 the factor $\varepsilon_p$ must be unramified. Therefore the character $\varepsilon$ is unramified at all finite places and thus trivial. Applying Proposition 3.2 we obtain

$$\sum_{\pi, \pi_p \in (\mathbf{2})} \mathrm{Tr}(f_\infty f^p, \pi^p)\phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)^{-\alpha} = \mathrm{Tr}(f_\infty f^p f_\alpha, \mathbf{1}) = 1. \tag{3.5}$$

We will now evaluate the sum $\sum_\pi \mathrm{Tr}(f_\infty f^p, \pi^p)\left(-\phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)^{-\alpha}\right)$ where $\pi$ ranges over the cuspidal automorphic representations of $G$ (this is the last sum in Equation (3.4)). Thus assume that $\pi$ is cuspidal automorphic representation of $G$ such that the trace of the truncated function $\chi f$ on $\pi$ does not vanish. The central character $\omega_\pi$ of $\pi$ is trivial on the group $\det(K_1(N)^p)$ and also trivial on the group $\mathbb{Z}_p^\times$. Hence the character $\omega_\pi$ must be trivial, and therefore the square $\phi_\pi^2$ is trivial as well. Consequently, the value $\phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)$ is either 1 or $-1$. The representation at infinity $\pi_\infty$ is generic and thus infinite dimensional. By Lemma 3.2 we must have $\pi_\infty \cong \pi_\infty^0$, and therefore $\mathrm{Tr}(f_\infty, \pi_\infty) = -1$. The trace $\mathrm{Tr}(f^{p,\infty}, \pi^{p,\infty})$ is equal to $\dim((\pi^{p,\infty})^{K_1(N)^p})$. Thus the second sum in Equation (3.4) is equal to

$$\sum_{\pi, \pi_\infty \cong \pi_\infty^0, \pi_p \in (\mathbf{3})} \dim \pi_{\mathrm{f}}^{p, K_1(N)^p} \phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)^{-\alpha}. \tag{3.6}$$

Assume $\alpha$ is odd. Then the Theorem follows from the above Equation (3.6) and the remark that if $\pi_p$ is ramified at $p$, and semi-stable, then $\pi_p$ is either $\mathrm{St}_G$ or $\mathrm{St}_G(\phi)$.

Now assume $\alpha$ to be even so that the sign $\phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)^{-\alpha}$ is equal to 1. If the representation $\pi$ contributes to the sum in Equation (3.6), then $\pi_p$ is an unramified twist of the Steinberg representation and the dimension of the space $(\pi_p)^{I_p}$ is equal to 1. Hence we may write $\phi_\pi\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)^{-\alpha} = \dim(\pi_p)^{I_p}$ as both sides of this equation are equal to 1. The sum in Equation (3.6) simplifies to

$$\sum_{\pi, \pi_\infty \cong \pi_\infty^0, \pi_p \in (\mathbf{3})} \dim(\pi_{\mathrm{f}})^{K'}. \tag{3.7}$$

Drop for the moment the condition that $\pi_p \in (\mathbf{3})$. In the article [**30**] is proved that cuspidal automorphic representations of $\mathrm{GL}_2$ with factor $\pi_\infty^0$ at infinity correspond to cuspidal modular

5. The component at a place $v$ of a cuspidal automorphic representation of $\mathrm{GL}_n$ is *generic*, see [**93**] or [**48**]. In the *special case* of the group $\mathrm{GL}_2(\mathbb{Q}_p)$ a smooth irreducible representation is generic if and only if it is infinite dimensional, and similarly for $\mathrm{GL}_2(\mathbb{R})$.

forms:

$$\sum_{\pi, \pi_\infty \cong \pi_\infty^0} \dim(\pi_{\mathrm{f}})^{K'} = \dim S_2(\Gamma), \tag{3.8}$$

where $S_2(\Gamma)$ is the space of weight 2 modular forms for the congruence subgroup $\Gamma := K' \cap G(\mathbb{Z})$ of $G(\mathbb{Z})$. The value in Equation (3.7) is equal to the value in Equation (3.8) minus the following sum:

$$\sum_{\pi, \pi_\infty \cong \pi_\infty^0, \pi_p = \mathrm{unr}} \dim(\pi_{\mathrm{f}})^{K'}, \tag{3.9}$$

where with the abbreviation "$\pi_p = \mathrm{unr}$" we mean that the representation $\pi_p$ is unramified. For an unramified generic representation $\pi_p$ of $G(\mathbb{Q}_p)$ the dimension of the space $(\pi_p)^{I_p}$ is equal to 2. In particular, Equation (3.9) equals $2 \cdot \dim S_2(K_1(N))$ and the number of supersingular points on the modular curve $Y_1(N)$ is equal to $1 + \dim S_2(\Gamma) - 2 \cdot \dim S_2(K_1(N))$. This completes the proof. □

## 4. The Deligne and Rapoport model

We show that, for $\alpha$ even, Theorem 3.3 follows from the description of the reduction modulo $p$ of the curve $X_0(p)$ by Deligne and Rapoport [**34**].

Consider, on the category of eliptic curves over $\mathbb{Z}[1/Np]$, the moduli problem of elliptic curves with $K_1(N)$-structure. A priori this problem is only defined over $\mathbb{Z}[1/Np]$, but one extends its definition to the ring $\mathbb{Z}[1/N]$ [**51**, chap. 1] (one can even extend to $\mathbb{Z}$, see [*loc. cit.*]). In particular we have a model of the scheme $Y'(Np)$ over $\mathbb{Z}[1/N]$, and the compactification $X'(Np)$ is also defined over $\mathbb{Z}[1/N]$ [chap. 8, *loc. cit.*]. The curve $X'(Np)$ has semi-stable reduction at $p$ [**34**].

In Theorem 3.3 we established that the number of supersingular points on the modular curve $Y_1(N)$ is equal to

$$1 + \mathrm{genus}(X'(Np)) - 2 \cdot \mathrm{genus}(X_1(N)).$$

In this section we show that this formula agrees with the description of the supersingular points on $X_1(N)$ by [**34**, V.1.18].

Let $\eta$ be the generic point of $\mathrm{Spec}(\overline{\mathbb{Z}}_p)$, and let $s$ be the special point of $\mathrm{Spec}(\overline{\mathbb{Z}}_p)$. Then Deligne and Rapoport have proved that

$$X'(Np)_s = Z_1 \amalg_S Z_2,$$

where $Z_i := X_1(N)_s$ and $S \subset Y_1(N)_s \subset X_1(N)_s$ is the supersingular locus.

Let $i_1$ (resp. $i_2$) denote the inclusion of $Z_1$ (resp. $Z_2$) in $X'(Np)_s$. Consider the morphism

$$\mathcal{O}_{X'(Np)_s} \longrightarrow i_{1*}\mathcal{O}_{Z_1} \oplus i_{2*}\mathcal{O}_{Z_2}$$

of sheaves on $X'(Np)_s$. Its cokernel is a direct sum over all points $P \in Z_1 \cap Z_2$ of sky-scraper sheaves. At each supersingular point $P \in Z_1 \cap Z_2$ we may consider the induced mapping on the completed local rings

$$\widehat{\mathcal{O}_{X'(Np)_s,P}} \longrightarrow (i_{1*}\mathcal{O}_{Z_1} \oplus i_{2*}\mathcal{O}_{Z_2})^{\wedge}_{P}.$$

This mapping coincides with the reduction map

$$\frac{\overline{\mathbb{F}}_p[[x,y]]}{(x,y)} \longrightarrow \frac{\overline{\mathbb{F}}_p[[x]]}{(x)} \oplus \frac{\overline{\mathbb{F}}_p[[y]]}{(y)}$$

whose cokernel is identified with $\overline{\mathbb{F}}_p$ via the mapping

$$\frac{\overline{\mathbb{F}}_p[[x]]}{(x)} \oplus \frac{\overline{\mathbb{F}}_p[[y]]}{(y)} \longrightarrow \overline{\mathbb{F}}_p$$

defined by $(f,g) \mapsto f(0) - g(0)$. We have an exact sequence

$$0 \longrightarrow \mathcal{O}_{X'(Np)_s} \longrightarrow i_{1*}\mathcal{O}_{Z_1} \oplus i_{2*}\mathcal{O}_{Z_2} \longrightarrow \bigoplus_{P \in Z_1 \cap Z_2} \overline{\mathbb{F}}_p \longrightarrow 0.$$

The Euler-Poincaré characteristic is additive on exact sequences, and thus

$$\chi(X'(Np)_s, \mathcal{O}_s) + |Z_1 \cap Z_2| = \chi(Z_1, \mathcal{O}_{Z_1}) + \chi(Z_2, \mathcal{O}_{Z_2}).$$

We have $Z_1 \cong Z_2 \cong X_1(N)_{\overline{\mathbb{F}}_p}$, and we have

$$\chi(Z_1, \mathcal{O}_{Z_1}) = 1 - \text{genus}(X_1(N)).$$

The Euler characteristic $\chi(X'(Np)_s, \mathcal{O}_s)$ is equal to $1 - \text{genus}(X)$, and therefore

$$1 - \text{genus}(X) + \#Y_1(N)_{\text{ss}}(\overline{\mathbb{F}}_p) = 2 \cdot (1 - \text{genus}(X_1(N)),$$

which is equivalent to the formula we found in Theorem 3.3.

CHAPTER 2

# The cohomology of the basic stratum I

[*À paraître dans Mathematische Annalen [**63**].*]

We consider a restricted class of certain simple Shimura varieties called the Kottwitz varieties, and we study them modulo a split prime of good reduction. We assume (essentially) there are no integral points on the Newton polygon of the basic stratum (other than the begin and end point). In this setting we establish a relation between the cohomology of the *basic stratum* of the Shimura variety $S$ modulo $p$ and the space of automorphic forms on $G$. The space of automorphic forms completely describes the cohomology of the basic stratum as Hecke module, as well as the action of the Frobenius element. The main result of this chapter is Theorem 3.13.

Let us comment on the strategy of proof of the main theorem. The formula of Kottwitz for Shimura varieties of PEL-type [**59**] is an expression for the number of points over finite fields on these varieties at primes of good reduction. We truncate the formula of Kottwitz to only contain the conjugacy classes which are *compact at* $p$. Thus we count virtual Abelian varieties with additional PEL-type structure lying in the *basic stratum*. The stabilization argument of Kottwitz carried out in his Ann Arbor article [**57**] still applies because the notion of $p$-compactness is stable under stable conjugacy. After stabilizing we obtain a sum of stable orbital integrals on the group $G(\mathbb{A})$, which can be compared with the geometric side of the trace formula. Ignoring endoscopy and possible non-compactness of the variety, the geometric side is equal to the *compact* trace $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathcal{A}(G))$ as considered by Clozel in his article on the fundamental lemma [**22**]. Using base change and Jacquet-Langlands we compare this compact trace with the twisted trace of a certain truncated Hecke operator acting on automorphic representations of the general linear group. We arrive at a local combinatorial problem at $p$ to classify the contributing representations (rigid representations, Section 2), and the computation of the compact trace of the Kottwitz function on these representations (Section 1). The computation of these compact traces turns out to be easy because we assumed there is no integral point lying on the Newton polygon of the basic isocrystal.

The main theorem is established in Section 3. In Subsections 4.1 and 4.2 we deduce two applications, in the first we express the zeta function of the basic stratum in terms of automorphic data, in the second application we derive a dimension formula for the basic stratum. In the first Section §1 we carry out the necessary local computations at $p$. In Section §1 we also prove a vanishing result of the truncated constant terms of the Kottwitz function due to the imposed conditions on the basic isocrystal (Proposition 1.10). This result is the technical reason for the simplicity of the formula in Theorem 3.13: without the conditions on the basic isocrystal, the final theorem contains a more complicated conclusion and involves a larger class of representations at $p$ (see Chapter 3). In Section §2 we apply the Moeglin-Waldspurger classification to determine the smooth irreducible representations of the general linear group occurring as components of discrete automorphic representations at finite places of a number field. This result is important for the final argument in Section §3.

## 1. Local computations

In this section we compute the compact traces of the functions of Kottwitz against the representations of the general linear group that occur in the (alternating sum of the) cohomology of unitary Shimura varieties.

**1.1. Notations.** Let $p$ be a prime number and let $F$ be a non-Archimedean local field with residue characteristic equal to $p$. Let $\mathcal{O}_F$ be the ring of integers of $F$, let $\varpi_F \in \mathcal{O}_F$ be a prime element. We write $\mathbb{F}_q$ for the residue field of $\mathcal{O}_F$, and the number $q$ is by definition its cardinal. The symbol $G_n$ denotes the locally compact group $\mathrm{GL}_n(F)$. If confusion is not possible then we drop the index $n$ from the notation. We call a parabolic subgroup $P$ of $G$ *standard* if it is upper triangular, and we often write $P = MN$ for its standard Levi decomposition. We write $K$ for the hyperspecial subgroup $\mathrm{GL}_n(\mathcal{O}_F) \subset G$. Let $\mathcal{H}(G)$ be the Hecke algebra of locally constant compactly supported complex valued functions on $G$, where the product on this algebra is the one defined by the convolution integral with respect to the Haar measure giving the group $K$ measure 1. We write $\mathcal{H}_0(G)$ for the spherical Hecke algebra of $G$ with respect to $K$. Let $P_0$ be the standard Borel subgroup of $G$, let $T$ be the diagonal torus of $G$, and let $N_0$ be the group of upper triangular unipotent matrices in $G$.

We write $\mathbf{1}_{G_n}$ for the trivial representation and $\mathrm{St}_{G_n}$ for the Steinberg representation of $G_n$. If $P = MN \subset G$ is a standard parabolic subgroup, then $\delta_P$ is equal to $|\det(m, \mathfrak{n})|$, where $\mathfrak{n}$ is the Lie algebra of $N$. The induction $\mathrm{Ind}_P^G$ is *unitary* parabolic induction. The Jacquet module $\pi_N$ of a smooth representation is not normalized by convention, for us it is the space of coinvariants for the unipotent subgroup $N \subset G$. For the definition of the constant terms $f^{(P)}$ and the Satake transform we refer to the article of Kottwitz [**52**, §5]. The valuation $v$ on $F$ is normalized so that $p$ has valuation 1 and the absolute value is normalized so that $p$ has absolute value $q^{-1}$. Finally, let $x \in \mathbb{R}$ be a real number, then $\lfloor x \rfloor$ (*floor function*) (resp. $\lceil x \rceil$, *ceiling function*) denotes the unique integer in the real interval $(x-1, x]$ (resp. $[x, x+1)$).

Let $n \in \mathbb{Z}_{\geq 0}$ be a non-negative integer. A *composition* of $n$ is an element $(n_a) \in \mathbb{Z}_{\geq 1}^k$ for some $k \in \mathbb{Z}_{\geq 1}$ such that $n = \sum_{a=1}^k n_a$. We write $\ell(n_a)$ for $k$ and call it the *length* of the composition. The set of compositions $(n_a)$ of $n$ is in bijection with the set of standard parabolic subgroups of $G_n = \mathrm{GL}_n(F)$. Under this bijection a composition $(n_a)$ of $n$ corresponds to the standard parabolic subgroup

$$P(n_a) \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} g_1 & & * \\ & \ddots & \\ 0 & & g_k \end{pmatrix} \in G_n \,\middle|\, g_a \in G_{n_a} \right\} \subset G_n.$$

We also consider extended compositions. Let $k$ be a non-negative integer. An *extended composition of* $n$ of length $\ell(n_a) = k$ is an element $(n_a) \in \mathbb{Z}_{\geq 0}^k$ such that $n = \sum_{a=1}^k n_a$.

**1.2. Compact traces.** In this subsection we work in a slightly more general setting. We assume that $G$ is the set of $F$-points of a smooth reductive group $\underline{G}$ over $\mathcal{O}_F$. We pick a minimal parabolic subgroup $P_0$ of $G$ and we standardize the parabolic subgroups of $G$ with respect to $P_0$. A semisimple element $g$ of $G$ is called *compact* if for some (any) maximal torus $T$ in $G$ containing $g$ the absolute value $|\alpha(g)|$ is equal to 1 for all roots $\alpha$ of $T$ in $\mathfrak{g}$. We now wish to define compactness for the non semisimple elements $g \in G$. We first pass to the algebraic closure: An element $g \in \underline{G}(\overline{F})$ is *compact* if its semisimple part is compact. A rational element $g \in G$ is *compact* if it is compact when viewed as an element of $\underline{G}(\overline{F})$. Let $\chi_c^G$ be the characteristic function on $G$ of the set of compact elements $G_c \subset G$. The subset $G_c \subset G$ is open, closed and stable under stable conjugation. We wish to make the following remark: Let $M$ be a Levi subgroup of $G$ and let $g$ be an element of $M \subset G$. The condition "$g$ is compact for the group $M$" is not equivalent to "$g$ is compact for the group $G$". We need the two notions and therefore we put the group $G$ in the exponent $\chi_c^G$ to clearly distinguish between the two.

Let $f$ be a locally constant, compactly supported function on $G$. The *compact trace* of $f$ on the representation $\pi$ is defined by $\mathrm{Tr}(\chi_c^G f, \pi)$ where $\chi_c^G f$ is the *point-wise product*. We define $\overline{f}$ to be the conjugation average of $f$ under the maximal compact subgroup $K$ of $G$. More precisely, for all elements $g$ in $G$ the value $\overline{f}(g)$ is equal to the integral $\int_K f(kgk^{-1})\mathrm{d}k$ where the Haar measure is normalized so that $K$ has volume 1.

Let $P$ be a standard parabolic subgroup of $G$ and let $A_P$ be the split center of $P$, we write $\varepsilon_P = (-1)^{\dim(A_P/A_G)}$. Define $\mathfrak{a}_P$ to be $X_*(A_P)_{\mathbb{R}}$ and define $\mathfrak{a}_P^G$ to be the quotient of $\mathfrak{a}_P$ by $\mathfrak{a}_G$. To the parabolic subgroup $P$ we associate the subset $\Delta_P \subset \Delta$ consisting of those roots acting non trivially on $A_P$. We write $\mathfrak{a}_0 = \mathfrak{a}_{P_0}$ and $\mathfrak{a}_0^G = \mathfrak{a}_{P_0}^G$. For each root $\alpha$ in $\Delta$ we have a coroot $\alpha^\vee$ in $\mathfrak{a}_0^G$. For $\alpha \in \Delta_P \subset \Delta$ we send the coroot $\alpha^\vee \in \mathfrak{a}_0^G$ to the space $\mathfrak{a}_P^G$ via the canonical surjection $\mathfrak{a}_0^G \twoheadrightarrow \mathfrak{a}_P^G$. The set of these restricted coroots $\alpha^\vee|_{\mathfrak{a}_P^G}$ with $\alpha$ ranging over $\Delta_P$ form a basis of the vector space $\mathfrak{a}_P^G$. By definition the set of fundamental weights $\{\varpi_\alpha^G \in \mathfrak{a}_P^{G*} \mid \alpha \in \Delta_P\}$ is the basis of $\mathfrak{a}_P^{G*} = \mathrm{Hom}(\mathfrak{a}_P^G, \mathbb{R})$ dual to the basis $\{\alpha^\vee_{\mathfrak{a}_P^G}\}$ of coroots. We let $\tau_P^G$ be the characteristic function on the space $\mathfrak{a}_P^G$ of the *acute Weyl chamber*,

$$\mathfrak{a}_P^{G+} = \left\{ x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \alpha, x \rangle > 0 \right\}. \tag{1.1}$$

We let $\widehat{\tau}_P^G$ be the characteristic function on $\mathfrak{a}_P^G$ of the *obtuse Weyl chamber*,

$$^{+}\mathfrak{a}_P^G = \left\{ x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \varpi_\alpha^G, x \rangle > 0 \right\}. \tag{1.2}$$

Let $P = MN$ be a standard parabolic subgroup of $G$. Let $X(M)$ be the group of rational characters of $M$. The *Harish-Chandra mapping*[1] $H_M$ of $M$ is the unique map from $M$ to $\mathrm{Hom}_{\mathbb{Z}}(X(M), \mathbb{R}) = \mathfrak{a}_P$, such that the $q$-power $q^{-\langle \chi, H_M(m) \rangle}$ is equal to $|\chi(m)|_p$ for all elements $m$ of $M$ and rational characters $\chi$ in $X(M)$. We define the function $\chi_N$ to be the composition $\tau_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$, and we define the function $\widehat{\chi}_N$ to be the composition $\widehat{\tau}_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$. The functions $\chi_N$ and $\widehat{\chi}_N$ are locally constant and $K_M$-invariant, where $K_M = \underline{M}(\mathcal{O}_F)$.

Lemma 1.1. *Let $P = MN$ be a standard parabolic subgroup of $G$. Let $m$ be a semisimple element of $M$, then*

(1) *$\chi_N(m)$ is equal to 1 if and only if for all roots $\alpha$ in the set $\Delta_P$ we have $|\alpha(m)| < 1$;*

(2) *$\hat{\chi}_N(m)$ is equal to 1 if and only if for all roots $\alpha$ in the set $\Delta_P$ we have $|\varpi_\alpha(m)| < 1$.*

Proposition 1.2. *Let $\pi$ be an admissible $G$-representation of finite length, and let $f$ be an element of $\mathcal{H}(G)$. The trace $\mathrm{Tr}(f, \pi)$ of $f$ on the representation $\pi$ is equal to the sum $\sum_{P=MN} \mathrm{Tr}_{M,c} \left( \chi_N \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}) \right)$ where $P$ ranges over the standard parabolic subgroups of $G$.*

Proof. For the proof see [**21**, prop 2.1]. Another proof of this proposition is given in [**22**, p. 259–262]. □

Proposition 1.3. *Let $\pi$ be an admissible $G$-representation of finite length, and let $f$ be an element of $\mathcal{H}(G)$. The compact trace $\mathrm{Tr}(\chi_c^G f, \pi)$ of $f$ on the representation $\pi$ is equal to the sum $\sum_{P=MN} \varepsilon_P \mathrm{Tr}_M \left( \widehat{\chi}_N \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}) \right)$ where $P$ ranges over the standard parabolic subgroups of $G$.*

Proof. This is the Corollary to Proposition 1 in the article [**22**]. □

Remark. Proposition 1.2 and Proposition 1.3 are true for reductive groups over non-Archimedean local fields in general.

We record the following corollary. We have a parabolic subgroup $\underline{P_0} \subset \underline{G}$ such that $P_0 = \underline{P_0}(F)$. Let $I \subset \underline{G}(\mathcal{O}_F)$ be the group of elements $g \in \underline{G}(\mathcal{O}_F)$ that reduce to an element of the group $\underline{P_0}(\mathcal{O}_F/\varpi_F)$ modulo $\varpi_F$. The group $I$ is called *the standard Iwahori subgroup of* $G$. A smooth representation $\pi$ of $G$ is called *semi-stable* if it has a non-zero invariant vector under the subgroup $I$ of $G$.

Corollary 1.4. *Let $\pi$ be a smooth admissible representation of $G$ such that the trace $\mathrm{Tr}(\chi_c^G f, \pi)$ does not vanish for some spherical function $f \in \mathcal{H}_0(G)$. Then $\pi$ is semi-stable.*

1. In the definition of the Harish-Chandra map there are different sign conventions possible. For example [**1**] and [**44**] use the convention $q^{\langle \chi, H_M(m) \rangle} = |\chi(m)|_p$ instead. Our sign follows that of [**102**]. In the article [**22**] there is no definition of the Harish-Chandra map but we have checked that Clozel uses our normalization.

PROOF. (cf. [**43**, p. 1351–1352]). By Proposition 1.3 the trace $\mathrm{Tr}(\widehat{\chi}_N f^{(P)}, \pi_N(\delta_P^{-1/2}))$ is nonzero for some standard parabolic subgroup $P = MN$ of $G$. The function $\widehat{\chi}_N f^{(P)}$ is $K_M$-spherical, and therefore $\pi_N(\delta_P^{-1/2})$ is an unramified representation of $M$. In particular the representation $\pi_N$ has an invariant vector for the Iwahori subgroup $I$ of $M$. The Proposition 2.4 in [**16**] gives a linear bijection from the vector space $(\pi_N)^{\underline{M}(\mathcal{O}_F)}$ to the vector space $\pi^I$. Therefore the space $\pi^I$ cannot be 0. □

PROPOSITION 1.5. *Let $\Omega$ be an open and closed subset of $G$ invariant under conjugation by $G$. Let $P = MN$ be a standard parabolic subgroup of $G$. Let $\rho$ be an admissible representation of $M$ of finite length, and let $\pi$ be the induction $\mathrm{Ind}_P^G(\rho)$ of the representation $\rho$ to $G$. Then for all $f$ in $\mathcal{H}(G)$ the trace $\mathrm{Tr}(\chi_\Omega f, \pi)$ is equal to the trace $\mathrm{Tr}(\chi_\Omega(\overline{f}^{(P)}), \rho)$.*

PROOF. By the main theorem of [**37**] we have

$$\mathrm{Tr}_G(\chi_\Omega f, \pi) = \mathrm{Tr}_M((\chi_\Omega \overline{f})^{(P)}, \rho). \tag{1.3}$$

We prove that the functions $\chi_\Omega \cdot (\overline{f}^{(P)})$ and $(\chi_\Omega(\overline{f}^{(P)}))$ in $\mathcal{H}(M)$ have the same orbital integrals. Let $\gamma \in M$. Then the orbital integral $O_\gamma^M(\chi_\Omega \cdot (\overline{f}^{(P)}))$ equals $O_\gamma^M(\overline{f}^{(P)})$ if $\gamma \in \Omega$ and vanishes for $\gamma \notin \Omega$. By Lemma 9 in [**37**] we have

$$O_\gamma^M((\chi_\Omega(\overline{f})^{(P)})) = O_\gamma^M((\overline{\chi_\Omega f})^{(P)}) = D(\gamma) O_\gamma^G(\chi_\Omega f),$$

where $D(\gamma) = D_M(\gamma)^{-1/2} D_G(\gamma)^{1/2}$ is a certain Jacobian factor for which we do not need to know the definition; we refer to [*loc. cit*] for the definition. By applying Lemma 9 of [*loc. cit*] once more the orbital integral $O_\gamma^G(\chi_\Omega f)$ is equal to $O_\gamma^G(f) = D(\gamma) O_\gamma^M(f^{(P)})$ for $\gamma \in \Omega$ and the orbital integral is 0 for $\gamma \notin \Omega$. Therefore, the orbital integrals of the functions $\chi_\Omega \cdot (\overline{f}^{(P)})$ and $(\chi_\Omega \overline{f}^{(P)})$ agree.

Recall Weyl's integration formula for the group $M$: for any $h \in \mathcal{H}(M)$ we have

$$\mathrm{Tr}(h, \rho) = \sum_T \frac{1}{|W(M,T)|} \int_{T_{\mathrm{reg}}} \Delta_M(t)^2 \theta_\rho(t) O_t(h) \mathrm{d}t, \tag{1.4}$$

where $\theta_\rho$ is the Harish-Chandra character of $\rho$ and where $T$ runs over the Cartan subgroups of $M$ modulo $M$-conjugation, and $W(M,T)$ is the rational Weyl group of $T$ in $M$, see [**32**, p. 97] (cf. [**21**, p. 241]). The right hand side in Equation (1.4) depends only on the orbital integrals of the function $h$. Thus, two functions $h, h' \in \mathcal{H}(M)$ with the same orbital integrals have the same trace on all smooth $M$-representations of finite length. Therefore the $M$-trace $\mathrm{Tr}_M((\chi_\Omega f)^{(P)}, \rho)$ of the function $(\chi_\Omega f)^{(P)}$ against $\rho$ is equal to $\mathrm{Tr}_M(\chi_\Omega(f^{(P)}), \rho)$. By combining Equation (1.4) with Equation (1.3) we obtain the proposition. □

**1.3. The Kottwitz functions $f_{n\alpha s}$.** From this point onwards $G$ is the general linear group. Let $n$ and $\alpha$ be positive integers, and let $s$ be a non-negative integer with $s \leq n$. We call the number $s$ the *signature*, and we call the number $\alpha$ the *degree*. Let $\mu_s \in X_*(T) = \mathbb{Z}^n$

be the cocharacter defined by

$$(\underbrace{1,1,\dots,1}_{s},\underbrace{0,0,\dots,0}_{n-s}) \in \mathbb{Z}^n.$$

We write $A_n$ for the algebra $\mathbb{C}[X_1^{\pm 1},\dots,X_n^{\pm 1}]^{\mathfrak{S}_n}$. The function $f_{n\alpha s} \in \mathcal{H}_0(G)$ is the spherical function with

$$\mathcal{S}_G(f_{n\alpha s}) = q^{\alpha s(n-s)/2} \sum_{\nu \in \mathfrak{S}_n \cdot \mu_s} [\nu]^\alpha = q^{\alpha s(n-s)/2} \sum_{I \subset \{1,\dots,n\}, \#I = s} \prod_{i \in I} X_i^\alpha \in A_n$$

as Satake transform (cf. [**54**]). When $n, \alpha, s \in \mathbb{Z}_{\geq 0}$ are such that $n < s$, then we put $f_{n\alpha s} = 0$.

Definition 1.6. Let $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n} \in \mathbb{C}[X_1^{\pm 1},\dots,X_n^{\pm 1}]$ be a monomial. Then the *degree* of $X$ is $\sum_{i=1}^n e_i \in \mathbb{Z}$. We call an element of the algebra $\mathbb{C}[X_1^{\pm 1},\dots,X_n^{\pm 1}]$ homogeneous of degree $d$ if it is a linear combination of monomials of degree $d$. These notions extend to the algebras $\mathcal{H}_0(G)$ and $A_n$ via the isomorphism $\mathcal{H}_0(G) = A_n$ and the inclusion $A_n \subset \mathbb{C}[X_1^{\pm 1},\dots,X_n^{\pm 1}]$.

Lemma 1.7. *Let $f \in \mathcal{H}_0(G)$ be a homogeneous function of degree $d$. Then $f$ is supported on the set of elements $g \in G$ with $|\det g| = q^{-d}$.*

Proof. (cf. [**4**, p. 34 bottom]). The function $f^{(P_0)}$ is supported on the set of elements $t \in T$ with $|\det t| = q^{-d}$. Let $\chi$ be the characteristic function of the subset $\{g \in G \mid |\det g| = q^{-d}\} \subset G$. The Satake transform $(\chi f)^{(P_0)}$ is equal to $\chi|_T \cdot (f^{(P_0)})$. The function $\chi f$ is equal to $f$ by injectivity of the Satake transform. □

By taking $f = f_{n\alpha s}$ we obtain in particular:

Lemma 1.8. *The function $f_{n\alpha s}$ is supported on the set of elements $g \in G$ with $|\det g| = q^{-\alpha s}$.*

Proof. The Satake transform $\mathcal{S}_G(f_{n\alpha s})$ of the Kottwitz function $f_{n\alpha s}$ is homogeneous of degree $\alpha s$ in the algebra $A_n$. □

Lemma 1.9. *Let $P = MN$ be a standard parabolic subgroup of $G$ corresponding to the composition $(n_a)$ of $n$. Let $k$ be the length of this composition. The constant term of $f_{n\alpha s}$ at $P$ is equal to*

$$\sum_{(s_a)} q^{\alpha \cdot C(n_a, s_a)} \cdot (f_{n_1 \alpha s_1} \otimes f_{n_2 \alpha s_2} \otimes \cdots \otimes f_{n_k \alpha s_k}), \tag{1.5}$$

*where the sum ranges over all extended compositions $(s_a)$ of $s$ of length $k$. The constant $C(n_a, s_a)$ is equal to $\frac{s(n-s)}{2} - \sum_{a=1}^k \frac{s_a(n_a - s_a)}{2}$.*

Remark. In the above sum only the extended compositions $(s_a)$ of $s$ with $s_a \leq n_a$ participate: If $s_a > n_a$ for some $a$, then $f_{n_a \alpha s_a} = 0$ by our convention.

PROOF. (cf. [**81**, Prop. 4.2.1]). Let $I_a \subset \{1, 2, \dots, n\}$ be the blocks corresponding to the composition $(n_a)$. If $I$ is a subset of the index set $\{1, \dots, n\}$, then we write $X_I$ for the monomial $\prod_{i \in I} X_i \in \mathbb{C}[X_1, X_2, \dots, X_n]$ in this proof. Taking constant terms is transitive and the constant term of a spherical function is spherical. Therefore it suffices to prove that both $f_{n\alpha s}$ and the function in Equation (1.5) have the same Satake transform. We compute

$$\begin{aligned}
\sum_{(s_a)} q^{\alpha \cdot C(n_a, s_a)} \prod_{a=1}^{k} \mathcal{S}_G(f_{n_a \alpha s_a}) &= \sum_{(s_a)} q^{\alpha \cdot C(n_a, s_a)} \prod_{a=1}^{k} q^{\alpha \frac{s_a(n_a - s_a)}{2}} \sum_{I \subset I_a, \# I = s_a} X_I^{\alpha} \\
&= q^{\alpha \frac{s(n-s)}{2}} \sum_{(s_a)} \prod_{a=1}^{k} \sum_{I \subset I_a, \# I = s_a} X_I^{\alpha} \\
&= q^{\alpha \frac{s(n-s)}{2}} \sum_{(s_a)} \sum_{I} X_I^{\alpha} \qquad (I \subset \{1, \dots, n\},\ \forall a:\ |I \cap I_a| = s_a) \\
&= q^{\alpha \frac{s(n-s)}{2}} \sum_{I \subset \{1, \dots, n\}, |I| = s} X_I^{\alpha}.
\end{aligned}$$

This concludes the proof. □

**1.4. Truncation of the constant terms.** In this subsection we compute the truncated function $\chi_c^G(f_{n\alpha s}^{(P)})$. This result is crucial to determine the representations of $G$ contributing to the cohomology of the basic stratum of Shimura varieties associated to unitary groups.

PROPOSITION 1.10. *Let $P = MN$ be a standard parabolic subgroup of $G$, and let $(n_a)$ be the corresponding composition of $n$. Let $k$ be the length of the composition $(n_a)$ and let $d$ be the greatest common divisor of $n$ and $s$. The truncated constant term $\chi_c^G(f_{n\alpha s}^{(P)})$ is non-zero only if there exists a composition $(d_a)$ of $d$ such that for all indices $a$ the number $n_a$ is obtained from $d_a$ by multiplying with $\frac{n}{d}$. If such a composition $(d_a)$ exists, then the function $\chi_c^G(f_{n\alpha s}^{(P)})$ is equal to*

$$\chi_c^G(f_{n\alpha s}^{(P)}) = q^{\alpha \cdot C(n_a, s_a)} \cdot \left( \chi_c^{G_{n_1}} f_{n_1 \alpha s_1} \otimes \chi_c^{G_{n_2}} f_{n_2 \alpha s_2} \otimes \cdots \otimes \chi_c^{G_{n_k}} f_{n_k \alpha s_k} \right) \in \mathcal{H}_0(M), \tag{1.6}$$

*where $s_a = \frac{s}{d} \cdot d_a$ for all $a \in \{1, 2, \dots, k\}$, and the constant $C(n_a, s_a)$ equals $\frac{s(n-s)}{2} - \sum_{a=1}^{k} \frac{s_a(n_a - s_a)}{2}$.*

PROOF. By Lemma 1.9 the truncated constant term $\chi_c^G(f_{n\alpha s}^{(P)})$ is a sum of terms of the form $\chi_c^G(f_{n_1 \alpha s_1} \otimes \cdots \otimes f_{n_k \alpha s_k})$ where $(s_a)$ ranges over extended compositions of $s$. To prove the Proposition we describe precisely the extended compositions with non-zero contribution. Thus assume that one of those terms is non-zero; say the one corresponding to the extended composition $(s_a)$ of $s$. Let $m$ be a semisimple point in $M$ where this term does not vanish. Let $m_a \in G_{n_a}$ be the $a$-th block of $m$, and let $m_{a,1}, \dots, m_{a,n_1} \in \overline{F}$ be the set of eigenvalues of $m_a$. The element $m$ is compact not only in the group $M$, but also in the group $G$, and therefore the absolute value $|m_{a,i}|$ is equal to the absolute value $|m_{b,j}|$ for all indices $a, i, b$ and

$j$. In particular the value $|\det(m_a)|^{1/n_a}$ is equal to $|\det(m_b)|^{1/n_b}$. By Lemma 1.8 the absolute value of the determinant $\det(m_a)$ is equal to $q^{-\alpha s_a}$. Therefore the fraction $\frac{s_a}{n_a}$ is equal to the fraction $\frac{s_b}{n_b}$ for all indices $a$ and $b$. We claim that the fraction $\frac{s_a}{n_a}$ equals $\frac{s}{n}$. To see this, we have $n_b \frac{s_a}{n_a} = s_b$ for all indices $a, b$, and thus

$$n\frac{s_a}{n_a} = (n_1 + n_2 + \ldots + n_k)\frac{s_a}{n_a} = s_1 + s_2 + \ldots + s_k = s, \tag{1.7}$$

proving the claim. We have $\frac{d}{s} \cdot s_a = \frac{d}{s} \cdot n_a \cdot \frac{s}{n} = \frac{d}{n} \cdot n_a$. Because $n_a \cdot \frac{s}{n} = s_a$ is integral, the number $n_a \frac{d}{n} = s_a \frac{d}{s}$ is integral as well. This implies that the composition $(n_a)$ (resp. $(s_a)$) is obtained from the composition $(d_a) := (n_a \frac{d}{n})$ by multiplying with $\frac{n}{d}$ (resp. $\frac{s}{d}$). □

**1.5. The functions $\chi_N f^{(P)}_{n\alpha s}$ and $\widehat{\chi}_N f^{(P)}_{n\alpha s}$.** Let $P = MN$ be a standard parabolic subgroup of $G$. The functions $\chi_N f^{(P)}_{n\alpha s}$ and $\widehat{\chi}_N f^{(P)}_{n\alpha s}$ occur in the formulas for the compact traces on smooth representations of $G$ of finite length (see Proposition 1.2 and Proposition 1.3). For later computations it will be useful to have them determined explicitly.

PROPOSITION 1.11. *Let $P = MN$ be a standard parabolic subgroup of $G$, and let $(n_a)$ be the corresponding composition of $n$. Write $k$ for the length of the composition $(n_a)$. The following statements are true:*

(*i*) *The function $\chi_N f^{(P)}_{n\alpha s} \in \mathcal{H}_0(M)$ is equal to*

$$\sum_{(s_a)} q^{\alpha \cdot C(n_a, s_a)} \cdot (f_{n_1 s_1} \otimes f_{n_2 s_2} \otimes \cdots \otimes f_{n_k s_k}),$$

*where the sum ranges over all extended compositions $(s_a)$ of $s$ of length $k$ satisfying*

$$\frac{s_1}{n_1} > \frac{s_2}{n_2} > \ldots > \frac{s_k}{n_k}.$$

(*ii*) *The function $\widehat{\chi}_N f^{(P)}_{n\alpha s} \in \mathcal{H}_0(M)$ is equal to*

$$\sum_{(s_a)} q^{\alpha C(n_a, s_a)} \cdot (f_{n_1 \alpha s_1} \otimes f_{n_2 \alpha s_2} \otimes \cdots \otimes f_{n_k \alpha s_k}),$$

*where the sum ranges over all extended compositions $(s_a)$ of $s$ of length $k$ satisfying*

$$(s_1 + s_2 + \ldots + s_a) > \frac{s}{n}(n_1 + n_2 + \ldots + n_a),$$

*for all indices $a$ strictly smaller than $k$.*

PROOF. Let $H_i$ for $i = \{1, 2, \ldots, n\}$ denote the $i$-th vector of the canonical basis of the vector space $\mathfrak{a}_0 = \mathbb{R}^n$. The subset $\Delta_P$ of $\Delta$ is the subset consisting of the roots $\alpha_{n_1+n_2+\ldots+n_a}$ for $a \in \{1, 2, \ldots, k-1\}$. For any root $\alpha = \alpha_i|_{\mathfrak{a}_P}$ in $\Delta_P$ we have:

$$\varpi^G_\alpha = (H_1 + \cdots + H_i - \frac{i}{n}(H_1 + H_2 + \cdots + H_n))|_{\mathfrak{a}_P}. \tag{1.8}$$

Let $m$ be an element of the standard Levi subgroup $M$. By Lemma 1.1 the element $m$ lies in the obtuse Weyl chamber if and only if the absolute value $|\varpi^G_\alpha(m)|$ is smaller than 1 for all roots $\alpha$ in $\Delta_P$. By Equation (1.8) the evaluation at $m = (m_a)$ of the characteristic function $\widehat{\chi}_N(m)$ is equal to 1 if and only if

$$|\det(m_1)| \cdot |\det(m_2)| \cdots |\det(m_a)| < |\det(m)|^{\frac{n_1+n_2+\ldots+n_a}{n}} \tag{1.9}$$

for all indices $a \in \{1, \ldots, k-1\}$.

We determine the function $\widehat{\chi}_N f^{(P)}_{n\alpha s}$. Let $m = (m_a)$ be an element of $M$. Assume $m$ lies in the obtuse Weyl chamber (cf. Equation (1.9)). Let $(s_a)$ be an extended composition of $s$. Besides the condition $\widehat{\chi}_N(m) \neq 0$ we assume that $(f_{n_1\alpha s_1} \otimes f_{n_2\alpha s_2} \otimes \cdots \otimes f_{n_k\alpha s_k})(m) \neq 0$. By Lemma (1.8) the absolute value $|\det(m_a)|$ is equal to $q^{-s_a\alpha}$ for all indices $a$. By Equation (1.9) we thus have the equivalent condition

$$(s_1 + s_2 + \ldots + s_a) > \frac{s}{n}(n_1 + n_2 + \cdots + n_a) \tag{1.10}$$

for all indices $a \in \{1, \ldots, k-1\}$. We have proved that if the product of the obtuse function $\widehat{\chi}_N$ with the function $(f_{n_1\alpha s_1} \otimes f_{n_2\alpha s_2} \otimes \cdots \otimes f_{n_k\alpha s_k})$ is non-zero, then the extended composition $(s_a)$ satisfies Equation (1.10) for all indices $a < k$. Conversely, if the extended composition $(s_a)$ satisfies the conditions in Equation (1.10), then any element $m$ of $M$ with $|\det(m_a)| = q^{-s_a\alpha}$ satisfies $\widehat{\chi}_N(m) = 1$. This completes the proof of the proposition for the function $\widehat{\chi}_N f^{(P)}_{n\alpha s}$.

The proof for the function $\chi_N f^{(P)}_{n\alpha s}$ is the same: Instead of using Equation (1.9), one uses that $\chi_N(m)$ equals 1 if and only if $|\alpha(m)| < 1$ for all roots $\alpha \in \Delta_P$. Therefore the element $m$ lies in the acute Weyl chamber if and only if

$$|\det(m_1)|^{1/n_1} < |\det(m_2)|^{1/n_2} < \cdots < |\det(m_k)|^{1/n_k}. \tag{1.11}$$

This completes the proof. □

**1.6. Computation of some compact traces.** In this subsection we compute compact traces against the trivial representation and the Steinberg representation.

Definition 1.12. If $\pi$ is an unramified representation of some Levi subgroup $M$ of $G$ then we write $\varphi_{M,\pi} \in \widehat{M}$ for the *Hecke matrix* of this representation. We recall the definition of the Hecke matrix. For an unramified representation $\pi$ of $G$ there exists a smooth unramified character $\chi$ of the torus $T$ and a surjection $\mathrm{Ind}^G_{P_0}(\chi) \twoheadrightarrow \pi$. Fix such a character $\chi$ together with such a surjection. Let $\widehat{T}$ be the complex torus dual to $T$. We compose any rational cocharacter $F^\times \to T(F)$ with $\chi$, and then we evaluate this composition at the prime element $\varpi_F$. This yields an element of $\mathrm{Hom}(X_*(T), \mathbb{C}^\times)$. The set $\mathrm{Hom}(X_*(T), \mathbb{C}^\times)$ is equal to the set $X_*(\widehat{T}) \otimes \mathbb{C}^\times = \widehat{T}(\mathbb{C})$. Thus we have an element of $\widehat{T}(\mathbb{C})$ well-defined up to the action of the rational Weyl-group of $T$ in $M$. This element in $\widehat{T}(\mathbb{C})$ is the *Hecke matrix* $\varphi_{M,\pi} \in \widehat{M}$.

Proposition 1.13. *Let $f \in \mathcal{H}_0(G)$ be a spherical function on $G$. Let $\mathrm{St}_G$ be the Steinberg representation of $G$. The compact trace $\mathrm{Tr}(\chi^G_c f, \mathrm{St}_G)$ is equal to $\varepsilon_{P_0}\mathcal{S}_T(\widehat{\chi}_{N_0} f^{(P_0)})(\varphi_{T,\delta^{1/2}_{P_0}})$.*

PROOF. By Proposition 1.3 we have

$$\mathrm{Tr}(\chi_c^G f, \pi) = \sum_{P=MN} \varepsilon_P \,\mathrm{Tr}(\widehat{\chi}_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2})).$$

The normalized Jacquet module $(\mathrm{St}_G)_N(\delta_P^{-1/2})$ at a standard parabolic subgroup $P = MN$ is equal to an unramified twist of the Steinberg representation of $M$ (cf. [**5**, thm 1.7(2)]). Assume that the parabolic subgroup $P = MN \subset G$ is not the Borel subgroup. Then the representation $\mathrm{St}_M$ of $M$ is ramified while the function $\widehat{\chi}_N f^{(P)}$ is spherical. The contribution of $P$ thus vanishes and consequently only the term corresponding to $P_0$ remains in the above formula. The Jacquet module $(\mathrm{St}_G)_{N_0}$ is equal to $\mathbf{1}(\delta_{P_0})$. This completes the proof. □

LEMMA 1.14. *Let $P = MN$ be a standard parabolic subgroup of $G$ which is proper. Let $f \in \mathcal{H}_0(G)$ be a homogeneous spherical function of degree coprime to $n$. Then $\chi_c^G f^{(P)} = 0$.*

PROOF. Write $s$ for the degree of $f$. Let $(n_a)$ be the composition of $n$ corresponding to $P$. We may write $\chi_c^G = \chi_c^M \chi_M^G$ as functions on $M$, where $\chi_M^G \in C^\infty(M)$ is the characteristic function of the set of elements $m = (m_a) \in M = \prod_{a=1}^k G_{n_a}$ such that

$$|\det m_1|^{1/n_1} = |\det m_2|^{1/n_2} = \cdots = |\det m_k|^{1/n_k}. \tag{1.12}$$

We claim that $\chi_M^G f^{(P)} = 0$. Let $m = (m_a) \in M$ be an element such that $f^{(P)}(m) \neq 0$ and $\chi_M^G(m) \neq 0$. Thus Equation (1.12) is true for $(m_a)$. Let $s_a$ be the integer such that $|\det m_a| = q^{-s_a}$. From Equation (1.12) we obtain that $\frac{s_a}{n_a} = \frac{s_b}{n_b}$ for all indices $a$ and $b$. We have $s_1 + s_2 + \ldots + s_k = s$. Use the argument at Equation (1.7) to obtain $\frac{s_a}{n_a} = \frac{s}{n}$ for all indices $a$. We find in particular that $n_a \frac{s}{n}$ is an integer. Because $n$ and $s$ are coprime this implies that $n_a = n$, i.e. that $P = G$. This completes the proof. □

PROPOSITION 1.15. *Let $f \in \mathcal{H}_0(G)$ be a homogeneous function of degree $s$. Assume $s$ is prime to $n$. The compact trace* $\mathrm{Tr}(\chi_c^G f, \mathbf{1})$ *is equal to* $\varepsilon_{P_0} \mathrm{Tr}(\chi_c^G f, \mathrm{St}_G)$.

PROOF. For the trivial representation $\mathbf{1}$ of $G$ we have the character identity $\mathbf{1} = \sum_{P=MN} \varepsilon_P \varepsilon_{P_0} \mathrm{Ind}_P^G(\mathrm{St}_M(\delta_P^{-1/2}))$ holding in the Grothendieck group of $G$. By the Proposition 1.5 we have

$$\mathrm{Tr}(\chi_c^G f, \mathrm{Ind}_P^G(\mathrm{St}_M(\delta_P^{-1/2}))) = \mathrm{Tr}(\chi_c^G f^{(P)}, \mathrm{St}_M(\delta_P^{-1/2})).$$

By Lemma 1.14 we have $\chi_c^G f^{(P)} = 0$ if $P$ is proper. The statement follows. □

EXAMPLE. We claim that the polynomial $\mathcal{S}_T(\chi_{N_0} f_{n\alpha s}^{(P_0)})$ in the ring $\mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots X_n^{\pm 1}]$ is equal to the polynomial $q^{\alpha \frac{s(n-s)}{2}} \sum X_{i_1}^\alpha X_{i_2}^\alpha \cdots X_{i_s}^\alpha$ where the indices $i_1, i_2, \ldots, i_s$ in the sum range over the set $\{1, 2, \ldots, n\}$ and satisfy the conditions (1) $i_1 < i_2 < i_3 < \ldots < i_s$; (2) $i_1 = 1$; (3) If $s > 1$ then for each subindex $j \in \{2, \ldots, s\}$ we have $i_j < 1 + \frac{n}{s}(j-1)$. The verification is elementary from Equation (4.3) but let us give details anyway. Let $(s_i)$ be an extended composition of $s$ of length $n$ with $s_i \in \{0,1\}$ for all $i$ and assume that the monomial $M(s_i) := X_1^{\alpha s_1} X_2^{\alpha s_2} \cdots X_n^{\alpha s_n}$ occurs in $\mathcal{S}_T(\chi_{N_0} f_{n\alpha s}^{(P_0)})$ with a non-zero coefficient. We have

$$s_1 + s_2 + \ldots + s_i > \frac{s}{n} i \tag{1.13}$$

for all $i < n$. Define for each subindex $j \leq s$ the index $i_j$ to be equal to $\inf\{i : s_1 + s_2 + \ldots + s_i = j\}$. With this choice for $i_j$ we have $M(s_i) = X_{i_1}^{\alpha} X_{i_2}^{\alpha} \cdots X_{i_s}^{\alpha}$. Equation (1.13) forces $i_1 = 1$ and for all $j \leq s-1$ that $i_{j+1} - 1$ is equal to the supremum $\sup\{i : s_1 + s_2 + \ldots + s_i = j\}$. Consequently $j > \frac{s}{n}(i_{j+1} - 1)$ for all $j \leq s-1$. By replacing $j$ by $j-1$ in this last formula we obtain for all $j$ with $2 \leq j \leq s$ the inequality

$$i_j < 1 + (j-1)\frac{n}{s}.$$

In the inverse direction, starting from this inequality for all $j$ together with the condition "$i_1 = 1$" we may go back to the inequalities in Equation (1.13). This proves the claim.

Example. We have

$$\begin{aligned}
\mathrm{Tr}(\chi_c^G f_{n\alpha 1}, \mathbf{1}) &= 1 \\
\mathrm{Tr}(\chi_c^G f_{n\alpha 2}, \mathbf{1}) &= 1 + q^{\alpha} + q^{2\alpha} + \ldots + q^{\alpha(\lfloor \frac{n}{2} \rfloor - 1)}.
\end{aligned}$$

## 2. Discrete automorphic representations and compact traces

We introduce two classes of semi-stable representations, the Speh representations and the rigid representations which are certain products of Speh representations. Then we deduce from the Moeglin-Waldspurger classification the possible components at $p$ of discrete automorphic representations in the semi-stable case.

Let $x, y$ be integers such that $n = xy$. We define the representation $\mathrm{Speh}(x,y)$ of $G$ to be the unique irreducible quotient of the representation $|\det|^{\frac{y-1}{2}} \mathrm{St}_{G_x} \times |\det|^{\frac{y-3}{2}} \mathrm{St}_{G_x} \times \cdots \times |\det|^{-\frac{y-1}{2}} \mathrm{St}_{G_x}$ where the product means unitary parabolic induction from the standard parabolic subgroup of $G_n$ with $y$ blocks and each block of size $x$. A *semi-stable Speh representation* of $G$ is, by definition, a representation isomorphic to $\mathrm{Speh}(x,y)$ for some $x, y$ with $n = xy$. We emphasize that we did not introduce all Speh representations, we have introduced only the ones which are semi-stable.

A smooth representation $\pi_p$ of $G$ is called *semi-stable rigid* representation if it is isomorphic to a representation of the following form. Consider the following list of data

- $k \in \{1, 2, \ldots, n\}$;
- for each $a \in \{1, 2, \ldots, k\}$ an unitary unramified character $\varepsilon_a \colon G \to \mathbb{C}^{\times}$;
- for each $a \in \{1, 2, \ldots, k\}$ a real number $e_a$ in the open (real) interval $(-\frac{1}{2}, \frac{1}{2})$;
- positive integers $y, x_1, x_2, \ldots, x_k$ such that $\frac{n}{y} = \sum_{a=1}^{k} x_a$,

then we form the representation

$$\mathrm{Ind}_P^G \bigotimes_{a=1}^{k} \mathrm{Speh}(x_a, y)(\varepsilon_a|\cdot|^{e_a}),$$

where $P = MN \subset G$ is the parabolic subgroup corresponding to the composition $(yx_a)$ of $n$ and where the tensor product is taken along the blocks of $M = \prod_{a=1}^{k} G_{yx_a}$. We remark that these representations are irreducible.

THEOREM 2.1 (Moeglin-Waldspurger). *Let $F$ be a number field and let $v$ be a finite place of $F$. Let $\pi_v$ be the local factor at $v$ of a discrete (unitary) automorphic representation $\pi$ of $\mathrm{GL}_n(\mathbb{A}_F)$. Assume that $\pi_v$ is semi-stable. Then $\pi_v$ is a semi-stable rigid representation.*

REMARK. Let $\pi_v$ be a semi-stable component of a discrete automorphic representation, as considered in the Theorem 2.1. Then using the definition of rigid representation we associate, among other data, to $\pi_v$ the real numbers $e_a$ in the open interval $(-\frac{1}{2}, \frac{1}{2})$ (see above). The Ramanujan conjecture predicts that the numbers $e_a$ are 0. This conjecture is proved in the restricted setting of Section §3 where we work with automorphic representations occurring in the cohomology of certain Shimura varieties. Therefore the numbers $e_a$, which a priori could be there, will not play a role for us.

PROOF OF THEOREM 2.1. By the classification of the discrete spectrum of $\mathrm{GL}_n(\mathbb{A}_F)$ in [**80**] there exist

- a decomposition $n = xy$, $x, y \in \mathbb{Z}_{\geq 1}$;
- a cuspidal automorphic representation $\omega$ of $\mathrm{GL}_x(\mathbb{A}_F)$;
- a character $\varepsilon\colon \mathrm{GL}_n(\mathbb{A}_F) \to \mathbb{C}^\times$,

such that after twisting by $\varepsilon$, the representation $\pi$ is the irreducible quotient $J$ of the induced representation $I$ which is equal to $\mathrm{Ind}_{P_x(\mathbb{A}_F)}^{\mathrm{GL}_n(\mathbb{A}_F)} \left(\omega|\cdot|^{\frac{y-1}{2}}, \ldots, \omega|\cdot|^{\frac{1-y}{2}}\right)$. In this formula the induction is unitary and $P_x = M_x N_x \subset \mathrm{GL}_n$ is the standard parabolic subgroup of $\mathrm{GL}_n$ with $y$-blocks, each one of size $x \times x$. By applying the local component functor [**86**, prop 2.4.1] to the surjection $I \twoheadrightarrow J$ we obtain a surjection $I_v \twoheadrightarrow J_v$. The component at $v$ of $I_v$ is simply $\mathrm{Ind}_{P_x(F_v)}^{\mathrm{GL}_n(F_v)} \left(\omega_v|\cdot|^{\frac{y-1}{2}}, \ldots, \omega_v|\cdot|^{\frac{1-y}{2}}\right)$. The representation $\omega_v$ is a factor of a cuspidal automorphic representation of $\mathrm{GL}_x(\mathbb{A}_F)$ and therefore *generic*[2].

From this point onwards we work locally at $v$ only, so we drop the $\mathrm{GL}_n(F_v)$-notation and write simply $G_n$. By the Zelevinsky classification of $p$-adic representations [**105**] any generic representation is of the form $\sigma_1|\det|^{e_1} \times \sigma_2|\det|^{e_2} \times \cdots \times \sigma_k|\det|^{e_k}$ where the $\sigma_a$ are square integrable representations and the $e_a \in \mathbb{R}$ lie in the open interval $(-\frac{1}{2}, \frac{1}{2})$. The $\sigma_a$ are equal to the unique irreducible subquotient of a representation of the form $\rho \times \rho|\det|^2 \times \cdots \times \rho|\det|^{k-1}$ where $\rho$ is cuspidal and where the central character of $\rho|\det|^{\frac{k-1}{2}}$ is unitary. We assumed that

2. This follows from the results in [**93**], combined with the method in [**48**], see the discussion on the end of page 172 and beginning of page 173 in the introduction to [**93**].

$\pi_v$ is semi-stable. Therefore $\rho$ is semi-stable and cuspidal, and therefore a one-dimensional unramified character. This implies that $\sigma_a$ is equal to $\mathrm{St}_{G_{n_a}}(\varepsilon_a)$ for some $n_a \in \mathbb{Z}_{\geq 0}$ and some unramified unitary character $\varepsilon_a$ of $G_{n_a}$. Thus $\sigma_a$ is equal to $\mathrm{St}_{G_{n_1}}(\varepsilon_1)|\det|^{e_1} \times \cdots \times \mathrm{St}_{G_{n_k}}(\varepsilon_k)|\det|^{e_k}$. For the representation $I_v$ we obtain

$$\begin{aligned}
I_v &= \omega_v|\det|^{\frac{y-1}{2}} \times \cdots \times \omega_v|\det|^{\frac{1-y}{2}} \\
&= (\sigma_1|\det|^{e_1} \times \cdots \sigma_r|\det|^{e_r})\,|\det|^{\frac{y-1}{2}} \times \cdots \times (\sigma_1|\det|^{e_1} \times \cdots \sigma_r|\det|^{e_r})\,|\det|^{\frac{1-y}{2}} \\
&= \prod_{a=1}^{k} \left(\sigma_a|\det|^{e_a + \frac{y-1}{2}} \times \cdots \times \sigma_a|\det|^{e_a + \frac{1-y}{2}}\right).
\end{aligned}$$

For each $a$, the representation $\sigma_a|\det|^{e_a+\frac{y-1}{2}} \times \cdots \times \sigma_a|\det|^{e_a+\frac{1-y}{2}}$ has $\mathrm{Speh}(x_a, y)(\varepsilon_a|\det|^{e_a})$ as (unique) irreducible quotient. Thus we obtain a surjection $I_v \twoheadrightarrow \prod_{a=1}^k \mathrm{Speh}(x_a, y)(\varepsilon_a|\det|^{e_a})$. The representations $\mathrm{Speh}(x_a, y)(\varepsilon_a)$ are unitary and because $|e_a|$ is *strictly* smaller than $\frac{1}{2}$ it is impossible to have a couple of indices $(a, b)$ such that the representation $\mathrm{Speh}(x_a, y)(\varepsilon_a|\det|^{e_a})$ is a twist of $\mathrm{Speh}(x_b, y)(\varepsilon_b|\det|^{e_b})$ with $|\det|$. By the Zelevinsky segment classification it follows that the product $\prod_{a=1}^k \mathrm{Speh}(x_a, y)(\varepsilon_a|\det|^{e_a})$ is irreducible. By uniqueness of the Langlands quotient the representation $J_v$ is isomorphic to the product $\prod_{a=1}^k \mathrm{Speh}(x_a, y)(\varepsilon_a|\det|^{e_a})$, as required. □

PROPOSITION 2.2. *Let $\pi$ be a semi-stable rigid representation of $G = \mathrm{GL}_n(F)$ where $F$ is a finite extension of $\mathbb{Q}_p$. Let $f$ be a homogeneous function in $\mathcal{H}_0(G)$ of degree $s$ coprime to $n$, then the compact trace $\mathrm{Tr}(\chi_c^G f, \pi)$ vanishes unless $\pi$ is the trivial representation or the Steinberg representation.*

PROOF. Assume that $\mathrm{Tr}(\chi_c^G f, \pi)$ is non-zero. By Proposition 1.5 the compact trace of $\chi_c^G f^{(P)}$ against the representation $\bigotimes_{i=1}^k \mathrm{Speh}(x_i, y)(\varepsilon_a|\cdot|^{e_a})$ is non zero. The truncated constant term $\chi_c^G f^{(P)}$ vanishes if the parabolic subgroup $P \subset G$ is proper (Lemma 1.14). Therefore $\pi$ is a Speh-representation; say $x$ and $y$ are its parameters. The character formula of Tadic [**99**, p. 342] expresses $\pi$ as an alternating sum of induced representations:

$$|\det|^{\frac{x+y}{2}} u(\mathrm{St}_x, y) = \sum_{w \in \mathfrak{S}'_y} \varepsilon(w) \prod_{i=1}^{y} \delta[i, x + w(i) - 1] \in \mathcal{R}$$

(for notations see [*loc. cit*]). The compact trace on all these induced representations vanish unless they are induced from the parabolic subgroup $P = G$. This is true only if the representation $\delta[i, x + w(i) - 1]$ is the unit element in $\mathcal{R}$ for all indices except one, i.e. if $(x + w(i) - 1) - i + 1 = 0$. After simplifying we find that $w(i) = i - x$ for all indices $i$ except one. Make the assumption that $y > 1$. Then clearly, if $x > 1$, the number $i - x$ is non-positive for the indices $i = 1$ and $i = 2$. It then follows that $w(i)$ is non-positive for $i = 1$ or $i = 2$. However, that is impossible because $w$ is a permutation of the index set

$\{1, 2, \ldots, y\}$. The conclusion is that either $y = 1$ or $x = 1$. But then $\pi$ is the Steinberg or the trivial representation. □

## 3. The basic stratum of some Shimura varieties associated to division algebras

In this section we establish the main result of this chapter.

**3.1. Notations and assumptions.** As explained in the introduction, we place ourselves in a restricted version of the setting of Kottwitz in the article [**58**]. We start by copying some of the notations from that article. Let $D$ be a division algebra over $\mathbb{Q}$ equipped with an anti-involution $*$. Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ inside $\mathbb{C}$. Write $F$ for the center of $D$ and we embed $F$ into $\overline{\mathbb{Q}}$. We assume that $F$ is a CM field and we assume that $*$ induces the complex conjugation on $F$. We write $F^+$ for the totally real subfield of $F$ and we assume that $F$ decomposes into a compositum $\mathcal{K}F^+$ where $\mathcal{K}/\mathbb{Q}$ is quadratic imaginary. Let $n$ be the positive integer such that $n^2$ is the dimension of $D$ over $F$. Let $G$ be the $\mathbb{Q}$-group such that for each commutative $\mathbb{Q}$-algebra $R$ the set $G(R)$ is equal to the set of elements $x \in D \otimes_{\mathbb{Q}} R$ with $xx^* \in R^\times$. The mapping $c\colon G \to \mathbb{G}_{m,\mathbb{Q}}$ defined by $x \mapsto xx^*$ is called the *factor of similitudes*. Let $h_0$ be an algebra morphism $h_0\colon \mathbb{C} \to D_{\mathbb{R}}$ such that $h_0(z)^* = h_0(\overline{z})$ for all $z \in \mathbb{C}$. We assume that the involution $x \mapsto h_0(i)^{-1}x^*h_0(i)$ is positive. We restrict $h_0$ to $\mathbb{C}^\times$ to obtain a morphism $h$ from Deligne's torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{m,\mathbb{C}}$ to $G_{\mathbb{R}}$; we let $X$ be the $G(\mathbb{R})$ conjugacy class of[3] $h^{-1}$. Let $\mu \in X_*(G)$ be the restriction of $h \otimes \mathbb{C}\colon \mathbb{C}^\times \times \mathbb{C}^\times \to G(\mathbb{C})$ to the factor $\mathbb{C}^\times$ of $\mathbb{C}^\times \times \mathbb{C}^\times$ indexed by the identity isomorphism $\mathbb{C} \xrightarrow{\sim} \mathbb{C}$. We write $E \subset \overline{\mathbb{Q}}$ for the reflex field of this Shimura datum (see below for a description of $E$). We obtain varieties $\mathrm{Sh}_K$ defined over the field $E$ and these varieties represent corresponding moduli problems of Abelian varieties of PEL-type as defined in [**59**].

Let $p$ be a prime number where the group $G_{\mathbb{Q}_p}$ is unramified over $\mathbb{Q}_p$, and the conditions of [**59**, §5] are satisfied so that the moduli problem and the variety $\mathrm{Sh}_K$ extend to be defined over the ring $\mathcal{O}_E \otimes \mathbb{Z}_p$ [*loc. cit.*]. We assume that the prime $p$ *splits* in the field $\mathcal{K}$. Let $K \subset G(\mathbb{A}_{\mathrm{f}})$ be a compact open subgroup, of the form $K = K_pK^p$, with $K_p \subset G(\mathbb{Q}_p)$ hyperspecial (coming from the choice of a lattice and extra data, see [*loc. cit.*, §5]). Furthermore, we assume that $K^p \subset G(\mathbb{A}_{\mathrm{f}}^p)$ is small enough such that $\mathrm{Sh}_K/\mathcal{O}_E \otimes \mathbb{Z}_p$ is smooth [*loc. cit*, §5]. Fix an embedding $\nu_p\colon E \to \overline{\mathbb{Q}}_p$. The embedding $\nu_p$ induces an $E$-prime $\mathfrak{p}$ lying above $p$. We write $\mathbb{F}_q$ for the residue field of $E$ at the prime $\mathfrak{p}$.

Let $\xi$ be an irreducible algebraic representation over $\overline{\mathbb{Q}}$ of $G_{\overline{\mathbb{Q}}}$ and let $\mathcal{L}$ be the local system corresponding to $\xi \otimes \mathbb{C}$ on the variety $S_{K,\mathcal{O}_{E_\mathfrak{p}}}$. Let $\mathfrak{g}$ be the Lie algebra of $G(\mathbb{R})$ and let $K_\infty$ be the stabilizer subgroup in $G(\mathbb{R})$ of the morphism $h$. Let $f_\infty$ be a function at infinity whose stable orbital integrals are prescribed by the identities of Kottwitz in [**57**]; it can be taken

3. The reason for this sign is that the formula conjectured by Kottwitz in the article [**58**] turned out to be slightly mistaken. When Kottwitz proved his conjecture in [**59**] he found that a different sign should be used. However he did not change the sign in the conclusion of his theorem, rather he introduced it at the beginning by replacing $h$ by $h^{-1}$. We follow the conventions of Kottwitz because we refer to both articles constantly.

to be (essentially) an Euler-Poincaré function [**58**, Lemma 3.2] (cf. [**27**]). The function has the following property: Let $\pi_\infty$ be an $(\mathfrak{g}, K_\infty)$-module occurring as the component at infinity of an automorphic representation $\pi$ of $G$. Then the trace of $f_\infty$ against $\pi_\infty$ is equal to the Euler-Poincaré characteristic $\sum_{i=0}^\infty N_\infty(-1)^i \dim \mathrm{H}^i(\mathfrak{g}, K_\infty; \pi_\infty \otimes \xi)$, where $N_\infty$ is a certain explicit constant (cf. [**58**, p. 657, Lemma 3.2]). Let $\ell$ be an auxiliary prime number (different from $p$) and $\overline{\mathbb{Q}}_\ell$ an algebraic closure of $\mathbb{Q}_\ell$ together with an embedding $\overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_\ell$. We write $\mathcal{L}$ for the $\ell$-adic local system on $\mathrm{Sh}_{K, \mathcal{O}_{E_\mathfrak{p}}}$ associated to the representation $\xi \otimes \overline{\mathbb{Q}}_\ell$ of $G_{\overline{\mathbb{Q}}_\ell}$.

Because $p$ splits in the extension $\mathcal{K}/\mathbb{Q}$, the group $G_{\mathbb{Q}_p}$ splits into a direct product of general linear groups:

$$G_{\mathbb{Q}_p} \cong \mathbb{G}_{\mathrm{m}, \mathbb{Q}_p} \times \prod_{\wp | p} \mathrm{Res}_{F^+_\wp / \mathbb{Q}_p} \mathrm{GL}_{n, F^+_\wp}, \tag{3.1}$$

where the product ranges over the set of $F^+$-places above $\mathfrak{p}$. Observe that we wrote '$\cong$' and not '$=$'. The choice of an isomorphism amounts to the choice of, for each $F^+$-place $\wp$ of an $F$-place $\wp'$ above $\wp$. Recall that we have embedded $\mathcal{K}$ into $\mathbb{C}$ and that $F = \mathcal{K} \otimes F^+$. Therefore, we have in fact for each $\wp$ such an $\wp'$. We fix for the rest of this chapter in Equation 3.1 the isomorphism corresponding to this choice of $F$-primes above the $F^+$-primes above $p$. We write $T_{\mathbb{Q}_p} \subset G_{\mathbb{Q}_p}$ for the diagonal torus. Observe that the group $G_{\mathbb{Q}_p}$ has an obvious model over $\mathbb{Z}_p$; we will write $G_{\mathbb{Z}_p}$ for this model, and we assume $K_p = G_{\mathbb{Z}_p}(\mathbb{Z}_p)$.

The field $E$ is included in the field $F$. We copy Kottwitz's description of the reflex field $E$ (cf. [**58**, p. 655]). Consider the subgroup consisting of the elements $g \in G$ whose factor of similitudes is equal to 1. This subgroup is obtained by Weil restriction of scalars from an unitary group $U$ defined over the field $F^+$. Let $v \colon F^+ \to \mathbb{R}$ be an embedding and let $v_1, v_2$ be the two embeddings of $F$ into $\mathbb{C}$ that extend $v$. We associate a number $n_{v_1}$ to $v_1$ and a number $n_{v_2}$ to $v_2$ such that the group $U(\mathbb{R}, v)$ is isomorphic to the standard real unitary group $U(n_{v_1}, n_{v_2})$. The group $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ acts on the set of $\mathbb{Z}$-valued functions on $\mathrm{Hom}(F, \mathbb{C})$ by translations. The reflex field $E$ is the fixed field of the stabilizer subgroup in $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ of the function $v \mapsto n_v$.

We write $V(F^+) := \mathrm{Hom}(F^+, \overline{\mathbb{Q}})$. We identify $V(F^+)$ with $\mathrm{Hom}(F^+, \overline{\mathbb{Q}}_p)$ via the embedding $\nu_p$, and also with $\mathrm{Hom}(F^+, \mathbb{R})$ via the inclusion $F^+ \subset \mathbb{R}$. In particular $V(F^+)$ is a $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-set and a $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$-set. For every $F^+$-prime $\wp$ above $p$ we write $V(\wp)$ for the Galois orbit in $V(F^+)$ corresponding to $\wp$.

We have embedded the field $\mathcal{K}$ into $\mathbb{C}$, and thus each $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$-orbit in $V(F^+)$ contains a distinguished point, i.e. for each each embedding $v \colon F^+ \to \mathbb{C}$ we have a distinguished extension $v_1 \colon F \to \mathbb{C}$. We write $s_v$ for the number $n_{v_1}$. We define $s_\wp := \sum_{v \in V(\wp)} s_v$. We define $\mathrm{Unr}_p^+$ to be the set of $F^+$-places $\wp$ above $p$ such that $s_\wp = 0$, and $\mathrm{Ram}_p^+$ to be the set of $F^+$-places above $p$ such that $s_\wp > 0$. We work under one additional technical assumption: We assume that for every $\wp \in \mathrm{Ram}_p^+$ the number $s_\wp$ is coprime with $n$.

**3.2. Isocrystals and the basic stratum.** Write $\mathcal{A}_K$ for the universal Abelian variety over $\mathrm{Sh}_K$ and $\lambda, i, \overline{\eta}$ for its additional PEL type structures [**59**, §6]. Let $L$ be the completion of the maximal unramified extension of $\mathbb{Q}_p$ contained in $\overline{\mathbb{Q}}_p$. Then $E_{\mathfrak{p},\alpha}$ is a subfield of $L$. Let $\alpha > 0$ be a positive integer, write $E_{\mathfrak{p},\alpha} \subset \overline{\mathbb{Q}}_\ell$ for the unramified extension of degree $\alpha$ of $E_{\mathfrak{p}}$, and write $\mathbb{F}_{q^\alpha}$ for the residue field of $E_{\mathfrak{p},\alpha}$. We write $\sigma$ for the automorphism of $L$ acting by $x \mapsto x^p$ on the residue field of $L$. We write $V$ for the $D^{\mathrm{opp}}$-module with space $D$ where an element $d \in D^{\mathrm{opp}}$ acts on the left through multiplication on the right on the space $D$.

Let $x \in \mathrm{Sh}_K(\mathbb{F}_{q^\alpha})$ be a point. The rational Dieudonné module $\mathbb{D}(\mathcal{A}_{K,x})_\mathbb{Q}$ is an $(E_{\mathfrak{p},\alpha}/\mathbb{Q}_p)$-isocrystal. The couple $(\lambda, i)$ induces via the functor $\mathbb{D}(\square)_\mathbb{Q}$ additional structures on this isocrystal. There exists an isomorphism $\varphi \colon V \otimes E_{\mathfrak{p},\alpha} \xrightarrow{\sim} \mathbb{D}(\mathcal{A}_{K,x})_\mathbb{Q}$ of skew-Hermitian $B$-modules [**59**, p. 430], and via this isomorphism we can send the crystalline Frobenius on $\mathbb{D}(\mathcal{A}_{K,x})_\mathbb{Q}$ to a $\sigma$-linear operator on $V \otimes E_{\mathfrak{p},\alpha}$. This operator on $V \otimes E_{\mathfrak{p},\alpha}$ may be written in the form $\delta \cdot (\mathrm{id}_V \otimes \sigma)$ where $\delta \in G(E_{\mathfrak{p},\alpha})$ is independent of $\varphi$ up to $\sigma$-conjugacy. We also have the $L$-isocrystal $\mathbb{D}(\mathcal{A}_{K,x} \otimes \overline{\mathbb{F}}_q)_\mathbb{Q} = \mathbb{D}(\mathcal{A}_{K,x})_L$ inducing in the same manner an element of $G(L)$, well defined up to $\sigma$-conjugacy. Let $B(G_{\mathbb{Q}_p})$ be the set of all $\sigma$-conjugacy classes in $G(L)$ from [**55**]. This set classifies the $L$ isocrystals with additional $G_{\mathbb{Q}_p}$-structure up to isomorphism.

In the articles [**88**] and [**60**] there is introduced the subset $B(G_{\mathbb{Q}_p}, \mu_{\overline{\mathbb{Q}}_p}) \subset B(G_{\mathbb{Q}_p})$ of $\mu_{\overline{\mathbb{Q}}_p}$-admissible isocrystals. The point is that if an isocrystal arises from some element $x \in \mathrm{Sh}_K(\overline{\mathbb{F}}_q)$ then this isocrystal is always $\mu_{\overline{\mathbb{Q}}_p}$-admissible. The set $B(G_{\mathbb{Q}_p})$ can be described explicitly as follows. We have $G_{\mathbb{Q}_p} = \mathbb{G}_{\mathrm{m},\mathbb{Q}_p} \times \mathrm{Res}_{F^+_{\mathbb{Q}_p}/\mathbb{Q}_p} \mathrm{GL}_{n,F^+_{\mathbb{Q}_p}}$ inducing the decomposition

$$B(G_{\mathbb{Q}_p}) = B(\mathbb{G}_\mathrm{m}) \times \prod_{\wp | p} B(\mathrm{Res}_{F^+_\wp/\mathbb{Q}_p} \mathrm{GL}_{n,F^+_\wp}).$$

Write $\mu_\wp$ for the component at $\wp$ of the cocharacter $\mu$. Fix one $\wp | p$. There is the Shapiro bijection [**60**, Eq. 6.5.3]

$$B(\mathrm{Res}_{F^+_\wp/\mathbb{Q}_p} \mathrm{GL}_{n,F^+_\wp}, \mu_\wp) = B(\mathrm{GL}_{n,F^+_\wp}, \mu'_\wp)$$

where the right hand side is the set of $\sigma^{[F^+_\wp : \mathbb{Q}_p]}$-conjugacy classes in $\mathrm{GL}_n(L)$ and $\mu'_\wp$ is defined by

$$\mu'_\wp \stackrel{\mathrm{def}}{=} \sum_{v \in V(\wp)} (\underbrace{1, 1, \ldots, 1}_{s_v}, \underbrace{0, 0, \ldots, 0}_{n - s_v}) \in \mathbb{Z}^n.$$

There is an unique element $b \in B(G_{\mathbb{Q}_p})$ with the property that, for each $\wp$, the corresponding isocrystal $b'_\wp$ in $B(\mathrm{GL}_{n,F^+_\wp}, \mu'_\wp)$ has precisely one slope (i.e. $b$ is *basic*). This slope must then be $\frac{s_\wp}{n}$ because the end point of the Hodge polygon of $\mu'_\wp$ is $(n, s_\wp)$. The component of $b$ at the factor of similitude is the $\sigma$-conjugacy class equal to the set of elements $x \in L^\times$ whose valuation is equal to 1.

Lemma 3.1. *We have $(n, s_\wp) = 1$ if and only if the isocrystal $V_\wp$ is simple.*

PROOF. We have $V_\wp = V_\lambda^m$ where $V_\lambda$ is the simple object of slope $\lambda = \frac{s_\wp}{n}$. In case $s_\wp$ and $n$ are coprime this simple object is of height $n$; otherwise its height is strictly less than $n$, and it occurs with positive multiplicity. □

The isocrystal $b$ introduced above characterises the basic stratum $B \subset \mathrm{Sh}_{K,\mathbb{F}_q}$ as the reduced subscheme such that for all points $x \in B(\overline{\mathbb{F}}_q)$ the isocrystal associated to the Abelian variety $\mathcal{A}_{K,x}$ is equal to $b$. The variety $B$ is projective, but in general not smooth [4].

The Hecke correspondences on $\mathrm{Sh}_K$ may be restricted to the subvariety $\iota\colon B \hookrightarrow \mathrm{Sh}_{K,\mathbb{F}_q}$. The algebra $\mathcal{H}(G(\mathbb{A}_\mathrm{f}))$ and the Galois group $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$ act on the cohomology spaces $\mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \iota^*\mathcal{L})$, and these actions cummute with each other.

**3.3. The function of Kottwitz.** Let $\alpha$ be a positive integer. Let $E_{\mathfrak{p},\alpha}/E_\mathfrak{p}$ be an unramified extension of degree $\alpha$. We write $\phi_\alpha$ for the characteristic function of the double coset $G(\mathcal{O}_{E_{\mathfrak{p},\alpha}})\mu(p^{-1})G(\mathcal{O}_{E_{\mathfrak{p},\alpha}})$ in $G(E_{\mathfrak{p},\alpha})$. The function $f_\alpha$ is by definition obtained from $\phi_\alpha$ via base change from $G(E_{\mathfrak{p},\alpha})$ to $G(\mathbb{Q}_p)$. We call the functions $f_\alpha$ the *functions of Kottwitz*; these functions play a fundamental role in the point-counting formula of Kottwitz for the number of points of the variety $\mathrm{Sh}_K$ over finite fields. In this section we give an explicit description of these functions $f_\alpha$ of Kottwitz.

DEFINITION 3.2. Let $\wp$ be an $F^+$-place above $p$. We write $V_\alpha(\wp)$ for the set of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_{\mathfrak{p},\alpha})$-orbits in the set $V(\wp)$, and $V_\alpha(F^+)$ for the set of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_{\mathfrak{p},\alpha})$ orbits in the set $V(F^+)$. If $v \in V_\alpha(F^+)$ is such an orbit, then this orbit corresponds to a certain finite unramified extension $E_{\mathfrak{p},\alpha}[v]$ of $E_{\mathfrak{p},\alpha}$. Let $\alpha_v$ be the degree over $\mathbb{Q}_p$ of the field $E_{\mathfrak{p},\alpha}[v]$, we then have $E_{\mathfrak{p},\alpha}[v] = E_{\mathfrak{p},\alpha_v}$.

REMARK. Let $\overline{v}$ be an element of $V_\alpha(\wp)$, then the number $s_v$ is independent of the choice of representative $v \in \overline{v}$.

REMARK. Observe that if $F^+$ is Galois over $\mathbb{Q}$, then all the Galois orbits in $V(F^+)$ have the same length.

PROPOSITION 3.3. *The function $f_\alpha$ is given by*

$$f_\alpha = \mathbf{1}_{q^{-\alpha}\mathbb{Z}_p^\times} \otimes \bigotimes_{\wp|p} \prod_{v \in V_\alpha(\wp)} f_{n\alpha_v s_v}^{\mathrm{GL}_n(F_\wp^+)} \in \mathcal{H}_0(G(\mathbb{Q}_p)),$$

*where the product is the convolution product.*

PROOF. We have the $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-set $V(F^+) = \mathrm{Hom}(F^+, \overline{\mathbb{Q}}_p)$. This Galois set is unramified and we have the Frobenius $\sigma$ acts on $V(F^+)$. The Galois set $V(F^+)$ decomposes: $V(F^+) = \coprod_{\wp|p} V(\wp)$, where $V(\wp) := \mathrm{Hom}(F^+_\wp, \overline{\mathbb{Q}}_p)$. We have

$$F^+ \otimes E_{\mathfrak{p},\alpha} = \prod_{\wp|p} (E_{\mathfrak{p},\alpha_v})^{\# V_\alpha(\wp)}.$$

4. The only cases where we know it is smooth is when it is a finite variety.

Because $p$ splits in $\mathcal{K}$ we have $\mathcal{K} \subset \mathbb{Q}_p \subset E_{\mathfrak{p},\alpha}$ and therefore

$$G(E_{\mathfrak{p},\alpha}) = E_{\mathfrak{p},\alpha}^{\times} \times \prod_{\wp|p} \prod_{\overline{v} \in V_\alpha(\wp)} \mathrm{GL}_n(E_{\mathfrak{p},\alpha_v}). \tag{3.2}$$

Recall that, by the definition of the reflex field, if two elements $v, v' \in V(F^+)$ lie in the same $\sigma^{[E_{\mathfrak{p}}:\mathbb{Q}_p]}$-orbit, then $s_v = s_{v'}$. Thus, with respect to the decomposition in Equation (3.2) we may write

$$\phi_\alpha = \mathbf{1}_{p^{-1}\mathcal{O}_{E_{\mathfrak{p},\alpha}}} \otimes \bigotimes_{\wp|p} \bigotimes_{\overline{v} \in V_\alpha(\wp)} \mathbf{1}_{\mathrm{GL}_n(\mathcal{O}_{E_{\mathfrak{p},\alpha_v}})} \cdot \mu_{\overline{v}}(p^{-1}) \cdot \mathbf{1}_{\mathrm{GL}_n(\mathcal{O}_{E_{\mathfrak{p},\alpha_v}})} \in \mathcal{H}_0(G(E_{\mathfrak{p},\alpha})),$$

where $\mu_{\overline{v}}$ is the cocharacter $(\mu_{s_v})^{[E_{\mathfrak{p},\alpha_v}:E_{\mathfrak{p},\alpha}]} \in X_*(\mathrm{Res}_{E_{\mathfrak{p},\alpha_v}/E_{\mathfrak{p},\alpha}} \mathbb{G}_{\mathrm{m}}^n) = \mathbb{Z}^{n \cdot [E_{\mathfrak{p},\alpha_v}:E_{\mathfrak{p},\alpha}]}$.

The explicit description of $f_\alpha$ now follows by applying the base change morphism from the spherical Hecke algebra of the group $G(E_{\mathfrak{p},\alpha})$ to the spherical Hecke algebra of the group $G(\mathbb{Q}_p) = \mathbb{Q}_p^{\times} \times \prod_{\wp|p} \mathrm{GL}_n(F_\wp^+)$. This completes the proof. □

**3.4. An automorphic description of the basic stratum.** Let $\iota$ be the inclusion $B \hookrightarrow \mathrm{Sh}_{K,\mathbb{F}_q}$. For each positive integer $\alpha$ and each $f^{\infty p} \in \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^p))$ we define the constant

$$T_B(f^p, \alpha) \stackrel{\mathrm{def}}{=} \sum_{i=0}^{\infty} (-1)^i \operatorname{Tr}(f^{\infty p} \times \Phi_{\mathfrak{p}}^\alpha, \mathrm{H}_{\text{ét}}^i(B_{\overline{\mathbb{F}}_q}, \iota^* \mathcal{L})).$$

We write $f$ for the function $f^{\infty p} f_\alpha f_\infty$ in the Hecke algebra of $G$ and similarly for $\chi_c^{G(\mathbb{Q}_p)} f$ even though the truncation occurs only at $p$.

We first give an automorphic expression for the trace $T_B(f^p, \alpha)$ for all sufficiently large integers $\alpha$.

Proposition 3.4. *There exists an integer $\alpha_0$ depending on the function $f^p$ such that $T_B(f^p, \alpha)$ equals $\operatorname{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathcal{A}(G))$ for all $\alpha \geq \alpha_0$.*

Proof. The main theorem of the article [**59**] gives an equation of the form

$$|\operatorname{Ker}^1(\mathbb{Q}, G)| \cdot \sum_{x' \in \mathrm{Fix}_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}(\overline{\mathbb{F}}_q)} \operatorname{Tr}(\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}, \mathcal{L}_x) = \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\phi_\alpha) \operatorname{Tr} \xi_{\mathbb{C}}(\gamma_0), \tag{3.3}$$

the notations are from [*loc. cit*], see especially §19. (In the above formula the point $x$ associated to an $x' \in \mathrm{Fix}^B_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}$ is the image of $x'$ in $B$ via the canonical map $\mathrm{Fix}^B_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}} \to B$.) We restrict this formula to the basic stratum $B$ by considering on the right hand side only *basic* Kottwitz triples. In this context *basic* means that the stable conjugacy class $\gamma_0$ in $(\gamma_0;\gamma,\delta)$ is compact at $p$, or, equivalently that the isocrystal corresponding to $\delta$ is the basic isocrystal in $B(G_{\mathbb{Q}_p}, \mu)$. The elements $x' \in \mathrm{Fix}_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}(\overline{\mathbb{F}}_q)$ in the sum in the left hand side of the Equation then have to be restricted to range over the set of fix points $\mathrm{Fix}^B_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}$ of the correspondence $\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}$ acting on the variety $B$. Everything else remains unchanged. This follows from the arguments of Kottwitz given for the above equation (see [*loc. cit*, §19], cf. Scholze [**92**, Prop. 6.6]).

From Fujiwara's trace formula [**40**, thm 5.4.5] we obtain

$$T_B(f^p, \alpha) = \sum_{x' \in \mathrm{Fix}^B_{\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(\Phi_{\mathfrak{p}}^\alpha \times f^{\infty p}, \iota^* \mathcal{L}_x)$$

for $\alpha$ large enough; say that this formula is true for all $\alpha \geq \alpha_0$. Note that in Fujiwara's statement the integer $\alpha_0$ depends on the correspondence and the sheaf $\mathcal{L}$.

We recall the definition of the norm $\mathcal{N}$ of (certain) $\sigma$-conjugacy classes (cf. [**4**] [**53**, p. 799]). To any element $\delta \in G(F_\alpha)$ we associate the element $N(\delta) := \delta\sigma(\delta)\cdots\sigma^{\alpha-1}(\delta) \in G(F_\alpha)$. For any element $\delta \in G(F_\alpha)$, defined up to $\sigma$-conjugacy, with semi-simple norm $N(\delta)$ one proves (see [*loc. cit.*]) that $N(\delta)$ actually comes from a conjugacy class $\mathcal{N}(\delta)$ in the group $G(F)$.

The element $\delta \in G(E_{\mathfrak{p},\alpha})$ is called *$\sigma$-compact* if its norm $\mathcal{N}(g)$ is a compact conjugacy class in $G(\mathbb{Q}_p)$. Let $\chi_c^{G(\mathbb{Q}_p)}$ be the characteristic function on $G(\mathbb{Q}_p)$ of the subset of compact elements (cf. §1.2). We let $\chi_{\sigma c}^{G(E_{\mathfrak{p},\alpha})}$ be the characteristic function on $G(E_{\mathfrak{p},\alpha})$ of the set of $\sigma$-compact elements. Consequently $T_B(f^{\infty p}, \alpha)$ is equal to

$$\sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\chi_{\sigma c}^{G(E_{\mathfrak{p},\alpha})}\phi_\alpha)\,\mathrm{Tr}\,\xi_{\mathbb{C}}(\gamma_0)$$

where $(\gamma_0;\gamma,\delta)$ ranges over *all* Kottwitz triples. Kottwitz has pseudo-stabilized this formula:

$$\tau(G) \sum_{(\gamma_0;\gamma,\delta)} \sum_{\kappa \in \mathfrak{K}(I_0/\mathbb{Q})} \langle \alpha(\gamma_0;\gamma,\delta), s\rangle e(\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\chi_{\sigma c}^{G(E_{\mathfrak{p},\alpha})}\phi_\alpha)\,\mathrm{Tr}\,\xi_{\mathbb{C}}(\gamma_0)\cdot$$
$$\cdot \mathrm{Vol}(A_G(\mathbb{R})^0 \backslash I(\infty)(\mathbb{R}))^{-1}, \tag{3.4}$$

see [**57**, Eq. (7.5)]. By the base change fundamental Lemma (see [**22**] and [**56**]) the functions $\phi_\alpha$ and $f_\alpha$ have matching stable orbital integrals (the functions are *associated*). By construction of the function $\chi_{\sigma c}^{G(E_{\mathfrak{p},\alpha})}$ this is then also the case for the truncated functions $\chi_{\sigma c}^{G(E_{\mathfrak{p},\alpha})}\phi_\alpha$ and $\chi_c^{G(\mathbb{Q}_p)} f_\alpha$. The group $G$ arises from a division algebra and therefore the group $\mathfrak{K}(G_{\gamma_0}/\mathbb{Q})$ is trivial for any (semisimple) element $\gamma \in G(\mathbb{Q})$ [**58**, Lemma 2]. Let $\gamma_\infty$ be a semisimple element of $G(\mathbb{R})$. Then the stable orbital integral $\mathrm{SO}_{\gamma_\infty}(f_\infty)$ vanishes unless $\gamma_\infty$ is elliptic, in which case it is equal to $\mathrm{Vol}(A_G(\mathbb{R})^0 \backslash I(\mathbb{R}))^{-1} e(I)$, where $I$ denotes the inner form of the centralizer of $\gamma_\infty$ in $G$ that is anisotropic modulo the split center $A_G$ of $G$ [**58**, Lemma 3.1]. Consequently Equation (3.4) is equal to the stable formula $\tau(G)\sum_{\gamma_0} SO_{\gamma_0}(f^{\infty p}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha) f_\infty)$.

By the argument at [**58**, Lemma 4.1] the above stable formula is the geometric side of the trace formula for the group $G$ and the function $\chi_c^{G(\mathbb{Q}_p)} f$; therefore it is equal to the trace of $\chi_c^{G(\mathbb{Q}_p)} f$ on the space of automorphic forms $\mathcal{A}(G)$ on $G$. We have obtained that $T_B(f^p, \alpha)$ equals $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathcal{A}(G))$ for all $\alpha \geq \alpha_0$. □

DEFINITION 3.5. We call a smooth representation $\pi_p$ of $G(\mathbb{Q}_p)$ of *Steinberg type* if the following two conditions hold: (1) For all $F^+$-places $\wp$ above $p$ we have

$$\pi_\wp = \begin{cases} \mathrm{St}_{\mathrm{GL}_n(F_\wp^+)} \otimes \phi_\wp & \wp \in \mathrm{Ram}_p^+ \\ \text{Generic unramified} & \wp \in \mathrm{Unr}_p^+ \end{cases}$$

where $\phi_\wp$ is an unramified character. (2) The factor of similitudes ${\mathbb{Q}_p}^\times$ of $G(\mathbb{Q}_p)$ acts through an unramified character on the space of $\pi_p$.

LEMMA 3.6. *Let $\pi$ be an automorphic representation of $G$. Then $\pi$ is one-dimensional if the component $\pi_\wp$ is one-dimensional for some $F^+$-place $\wp$ above $p$.*

PROOF. Assume $\pi_\wp$ is one-dimensional. By twisting $\pi$ with a character we may assume that $\pi_\wp$ is the trivial representation. Let $H \subset G(\mathbb{A}_\mathrm{f})$ be a compact open subgroup such that $\pi^H \neq 0$. We embed $\pi$ in the space of automorphic forms on $G$. Then elements of $\pi$ are complex valued functions on $G(\mathbb{A})$. The group $U \subset G$ is the unitary group of elements whose factor of similitude is trivial, and this group $U$ arises by restriction of scalars from a unitary group $U'$ over $F^+$. Let $SU$ be the derived group of $U'$. Then $SU$ is a simply connected algebraic group over $F^+$. We may restrict the automorphic representation $\pi$ of $G$ to obtain a representation of the group $SU(\mathbb{A}_{F^+})$ (which is reducible in general). Let $h \in \pi$ be an element, then $h$ is a complex valued function on $G(\mathbb{A}_{F^+})$ invariant under the groups $SU(F^+), H$ and also under the group $SU(F^+_\wp)$ because $\pi_\wp$ is the trivial representation. By strong approximation for the group $SU$ we see that $SU(\mathbb{A}_{F^+})$ acts trivially on $h \in \pi^H$. Thus $SU(\mathbb{A}_{F^+})$ acts trivially on the space $\pi$. Therefore $\pi$ is an Abelian automorphic representation of $G$ and thus one-dimensional. □

PROPOSITION 3.7. *For all $\alpha \geq \alpha_0$ the trace $T_B(f^p, \alpha)$ is equal to*

$$\sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1, \pi_p = \mathrm{Unr}}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi) + \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p = \text{ St. type}}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi), \tag{3.5}$$

*where both sums range over the irreducible subspaces of $\mathcal{A}(G)$.*

PROOF. Fix throughout this proof an automorphic representation $\pi \subset \mathcal{A}(G)$ of $G$ such that $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi) \neq 0$. We base change $\pi$ to an automorphic representation $BC(\pi)$ of the algebraic group $\mathcal{K}^\times \times D^\times$. Here we are using that $D$ is a division algebra and therefore the second condition in Theorem A.3.1(b) of the Clozel-Labesse appendix in [**65**] is satisfied (cf. [**45**, §VI.2] and [**96**]). In turn we use the Jacquet-Langlands correspondence [**101**] (cf. [**45**, §VI.1] and [**6**]) to send $BC(\pi)$ to an automorphic representation $\Pi := JL(BC(\pi))$ of the $\mathbb{Q}$-group $G^+ = \mathrm{Res}_{\mathcal{K}/\mathbb{Q}}\mathbb{G}_\mathrm{m} \times \mathrm{Res}_{F,\mathbb{Q}}\mathrm{GL}_{n,F}$.

The transferred representation $\Pi$ is discrete and $\theta$-stable, meaning that $\Pi$ is isomorphic to the representation $\Pi^\theta$ obtained from $\Pi$ by precomposition $G^+(\mathbb{A}) \to G^+(\mathbb{A}) \to \mathrm{End}_\mathbb{C}(\Pi)^\times$ with $\theta$. Because $\Pi$ is a subspace of the space of automorphic forms $\mathcal{A}(G^+)$ it comes with a natural intertwining operator $A_\theta \colon \Pi \xrightarrow{\sim} \Pi^\theta$ induced from the action of $\theta$ on $\mathcal{A}(G^+)$ (here we are using that multiplicity one is true for the discrete spectrum of $G^+$). The group $G^+(\mathbb{Q}_p)$ is isomorphic to $G(\mathbb{Q}_p) \times G(\mathbb{Q}_p)$ and the representation $\Pi_p$ is isomorphic to $\pi_p \otimes \pi_p$. We have $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) \neq 0$. Therefore $\pi_p$ is semi-stable by Corollary 1.4. For each $F^+$-prime $\wp$ the component $\pi_\wp$ is equal to a component $\Pi_{\wp'}$ for some (any) $F$-place $\wp'$ above $\wp$. As the representation $\Pi$ is a *discrete* automorphic representation of the group $G^+(\mathbb{A})$ the component $\pi_\wp = \Pi_{\wp'}$ is a *semi-stable rigid* representation by the Moeglin-Waldspurger theorem (Theorem 2.1).

We prove a lemma before finishing the proof of Proposition 3.7.

LEMMA 3.8. *Assume that $\pi$ is infinite dimensional and that $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi) \neq 0$. Then the transferred representation $\Pi$ is cuspidal.*

PROOF. We use the divisibility conditions on $n$ and $s_\wp$ to see that $\Pi$ is cuspidal: Because of these conditions, the Proposition 2.2 implies that the component $\pi_\wp$ of $\pi$ at the prime $\wp$ is an unramified twist of either the trivial representation or of the Steinberg representation if $\wp$ lies in the set $\mathrm{Ram}_p^+$, i.e. if the basic isocrystal is not étale at $\wp$. The trivial representation is not possible by the Lemma 3.6 and the assumption that $\pi$ is infinite dimensional. There is at least one $\wp$ such that $b_\wp$ is not étale (thus $\mathrm{Ram}_p^+ \neq 0$), and therefore the discrete representation $\Pi$ is an unramified twist of the Steinberg representation at some finite $F^+$-place. By the Moeglin-Waldspurger classification of the discrete spectrum, the $G^+(\mathbb{A})$-representation $\Pi$ must be cuspidal. □

*Continuation of the proof of Proposition 3.7.* If the prime $\wp \in \mathrm{Unr}_p^+$ is such that the basic isocrystal at $\wp$ *is* étale at $\wp$ then the function $\chi_c^{\mathrm{GL}_n(F_\wp^+)} f_\wp$ is simply the unit of the spherical Hecke algebra, hence unramified, and therefore $\pi_\wp$ is an unramified representation; because $\pi_\wp$ occurs in a cuspidal automorphic representation of $G^+(\mathbb{A})$ the representation $\pi_\wp$ is furthermore generic by the result of Shalika [**93**]. By Lemmas 3.6 and 3.8 there are the following possibilities for $\pi$. Either $\pi$ is one-dimensional and the component $\pi_p$ is unramified, or $\pi$ is infinite dimensional, and the component $\pi_p$ is of Steinberg type. We have proved that $T_B(f^p, \alpha)$ is equal to

$$\sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi) + \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p = \text{ St. type}}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi), \tag{3.6}$$

where both sums range over the irreducible subspaces of $\mathcal{A}(G)$. □

The main theorem is now essentially established, we only need to expand the above sums slightly further using the calculations that we did in the first two sections.

We define a number $\zeta_{\pi_p} \in \mathbb{C}$ for the two types of representations at $p$ that occur in Equation (3.5): those of Steinberg type and the one-dimensional, unramified representations.

DEFINITION 3.9. Assume that $\pi_p = \mathbf{1}(\phi_p)$ is unramified and one-dimensional. We define

$$\zeta_{\pi_p} \overset{\text{def}}{=} \phi_c(q) \prod_{\wp \in \mathrm{Ram}_p^+} \phi_\wp(q^{s_\wp}) \in \mathbb{C}^\times, \tag{3.7}$$

where $\phi_c$ is the character by which the factor of similitude acts on the space of $\pi_p$. Assume that $\pi_p$ is of Steinberg type. Then for all $\wp \in \mathrm{Ram}_p^+$ we have $\pi_\wp \cong \mathrm{St}_{\mathrm{GL}_n(F_\wp^+)}(\phi_\wp)$ for some unramified character $\phi_\wp$ of $F_\wp^{+\times}$. Let $\phi_c$ be the character by which the factor of similitude of $G(\mathbb{Q}_p)$ acts on the space of $\pi_p$. We define $\zeta_{\pi_p}$ again by Equation (3.7).

DEFINITION 3.10. Let $\pi$ be a $\xi$-cohomological automorphic representation of $G$. The center $Z$ of $G$ contains the torus $\mathbb{G}_\mathrm{m}$. We may precompose the central character $\omega_\pi$ of $\pi$ with the inclusion $\mathbb{A}^\times \subset Z(\mathbb{A})$ to obtain a character $\mathbb{A}^\times \to \mathbb{C}^\times$. Let $w \in \mathbb{Z}$ be the unique integer such that the composition

$$\mathbb{A}^\times \longrightarrow \mathbb{C}^\times \xrightarrow{|\cdot|} \mathbb{R}^\times_{>0} \tag{3.8}$$

is the character $||\cdot||^{w/2}$.

LEMMA 3.11. *Let $\pi$ be a $\xi$-cohomological automorphic representation of $G$ which is either unramified and one-dimensional, or of Steinberg type at $p$. Then $\zeta_\pi$ is a Weil-$q$-number of weight $w/n$.*

PROOF. Let $\phi_\wp$ be the character of $\mathrm{GL}_n(F_\wp^+)$ as defined in Equation 3.9. Let $\omega_\pi$ be the central character of $\pi$. Then $\omega_{\pi,\wp} = \phi_\wp^n$ for all $\wp \in \mathrm{Ram}_p^+$, and at the factor of similitude of $G(\mathbb{Q}_p)$ we have $\omega_{\pi,c} = \phi_c^n$. Thus, the number $\zeta_{\pi_p}$ is an $n$-th root of the number

$$\eta_{\pi_p} := \omega_c(q) \prod_{\wp \in \mathrm{Ram}_p^+} \omega_{\pi,\wp}(q^{s_\wp}) \in \mathbb{C}^\times. \tag{3.9}$$

Thus, to prove that $\zeta_\pi$ is Weil-$q$-number, it suffices to prove that $\eta_{\pi_p}$ is a Weil-$q$-number. The central character $\omega_\pi$ is a Grössencharakter of the center $Z \subset G$. The center $Z$ of $G$ is the set of elements $z \in F^\times \subset D^\times$ such that the norm of $z$ down to $F^{+\times}$ lies in the subset $\mathbb{Q}^\times \subset F^{+\times}$. Because $\pi$ is $\xi_\mathbb{C}$ cohomological we have $\omega_{\pi,\infty} = \xi_\mathbb{C}^{-1}|_{Z(\mathbb{R})}$. Let $U_Z \subset Z$ be the subtorus consisting of elements in $F^\times$ whose norm to $F^{+\times}$ is equal to 1. We have an exact sequence $\mu_2 \rightarrowtail U_Z \times \mathbb{G}_{\mathrm{m},\mathbb{Q}} \twoheadrightarrow Z$ of algebraic groups over $\mathbb{Q}$, where the injection is the embedding on the diagonal and the surjection is the multiplication map $\varphi\colon (u,x) \mapsto ux$. We may restrict the character $\omega_\pi$ of $Z(\mathbb{A})$ to the group $U_Z(\mathbb{A}) \times \mathbb{A}^\times$ and we obtain in this manner a character $\omega_{\pi,1}$ of $U_Z(\mathbb{A})$ and a character $\omega_{\pi,2}$ of $\mathbb{A}^\times$.

The component at $p$ of the mapping $\varphi\colon U_Z(\mathbb{A}) \times \mathbb{A}^\times \to Z(\mathbb{A})$ is the identity mapping

$$U_Z(\mathbb{Q}_p) \times {\mathbb{Q}_p}^\times = F_{\mathbb{Q}_p}^{+\times} \times {\mathbb{Q}_p}^\times \longrightarrow F_{\mathbb{Q}_p}^{+\times} \times {\mathbb{Q}_p}^\times = Z(\mathbb{Q}_p).$$

Let $K_1 \times K_2 \subset U(\mathbb{A}) \times \mathbb{A}^\times$ be a compact open subgroup such that $\omega_{\pi,1}$ is $K_1$-spherical and such that $\omega_{\pi,2}$ is $K_2$-spherical. The group $U(\mathbb{Q})K_1 \backslash U(\mathbb{A})$ is *compact* and therefore the product

$$\prod_{\wp \in \mathrm{Ram}_p^+} \omega_{\pi,\wp}(q^{s_\wp}) \in \mathbb{C}^\times,$$

(cf. Equation (3.9)) is a Weil-$q$-number of weight 0. The group $\mathbb{Q}^\times K_2 \backslash \mathbb{A}^\times$ is non-compact and thus $\omega_{\pi,c}(q)$ is a Weil-$q$-number whose weight is $w$, where $w$ is defined in Definition 3.10. This completes the proof. $\square$

DEFINITION 3.12. We write $P(q^\alpha)$ for the trace $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \mathbf{1})$.

In general $P(q^\alpha)$ is not a polynomial in $q^\alpha$, it depends on $\alpha$ in the following manner. The explicit description of the function $f_\alpha$ from Proposition 3.3 shows

$$P(q^\alpha) = \prod_{\wp \in \mathrm{Ram}_p^+} \mathrm{Tr}\left( \chi_c^{\mathrm{GL}_n(F_\wp^+)} \prod_{v \in V(\wp)} f_{n\alpha_v s_v}^{\mathrm{GL}_n(F_\wp^+)}, \mathbf{1} \right). \tag{3.10}$$

The traces in the product in Equation (3.10) are computed in Subsection 1.6 (see Proposition 1.15).

REMARK. In general the function $P(q^\alpha)$ is not a polynomial in $q^\alpha$. The number $c_{\wp,\alpha}$ depends on the class of $\alpha$ in the group $\mathbb{Z}/M\mathbb{Z}$, where $M$ is large such that the algebra $F^+ \otimes E_{\mathfrak{p},M}$ is isomorphic to a product of copies of $E_{\mathfrak{p},M}$. For the $\alpha$ that range over the elements of a fixed class $\overline{c} \in \mathbb{Z}/M\mathbb{Z}$ there exists a polynomial $\mathrm{Pol}_{\overline{c}} \in \mathbb{C}[X]$ such that $P(q^\alpha) = \mathrm{Pol}_{\overline{c}}|_{X=q^\alpha}$.

THEOREM 3.13. *The trace of the correspondence $f^p \times \Phi_\mathfrak{p}^\alpha$ acting on the alternating sum of the cohomology spaces $\sum_{i=0}^\infty (-1)^i \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \iota^*\mathcal{L})$ is equal to*

$$|\,\mathrm{Ker}^1(\mathbb{Q}, G)| P(q^\alpha) \left( \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1, \pi_p = \mathrm{unr}}} \zeta_\pi^\alpha \cdot \mathrm{Tr}(f^p, \pi^p) + (-1)^{(n-1)\cdot \#\mathrm{Ram}_p^+} \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p = \text{ St. type}}} \zeta_\pi^\alpha \cdot \mathrm{Tr}(f^p, \pi^p) \right). \tag{3.11}$$

*for all positive integers $\alpha$.*

PROOF. Assume that $\alpha \geq \alpha_0$. In Proposition 3.7 we established that

$$T_B(f^p, \alpha) = \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1, \pi_p = \mathrm{Unr}}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi) + \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p = \text{ St. type}}} \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi).$$

Let $\pi$ be an automorphic representations contributing to one of the above two sums. We have

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \pi) = \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p)\, \mathrm{Tr}(f^p, \pi^p).$$

For $\pi_p$ there are two possibilities: (1) $\pi_p$ is one-dimensional, (2) $\pi_p$ is of Steinberg type. In the first case we have

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \zeta_\pi^\alpha \cdot P(q^\alpha).$$

In the second case we have

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \zeta_\pi^\alpha \cdot \prod_{\wp | p} \mathrm{Tr}(\chi_c^{\mathrm{GL}_n(F_\wp^+)} f_\wp, \mathrm{St}_{\mathrm{GL}_n(F_\wp^+)}).$$

By Proposition 1.15 we have $\mathrm{Tr}(\chi_c^{\mathrm{GL}_n(F_\wp^+)} f_\wp, \mathrm{St}_{\mathrm{GL}_n(F_\wp^+)}) = (-1)^{n-1} \mathrm{Tr}(\chi_c^{\mathrm{GL}_n(F_\wp^+)} f_\wp, \mathbf{1})$ and therefore

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \zeta_\pi^\alpha \cdot (-1)^{(n-1)\cdot \#\mathrm{Ram}_p^+} \cdot P(q^\alpha).$$

Thus Equation (3.11) is true for all $\alpha \geq \alpha_0$; observe that it must then be true for all $\alpha > 0$. This completes the proof. □

## 4. Applications

In this section we deduce two applications from our main theorem. We first deduce an expression for the zeta function of the basic stratum in terms of the cohomology of a complex Shimura variety. In the second application we deduce an explicit formula for the dimension of the variety $B/\mathbb{F}_q$.

**4.1. The number of points in $B$.** Let $I_p \subset K_p$ be the standard Iwahori subgroup at $p$. We use Theorem 3.13 to deduce a formula for the zeta function of $B$ in terms of the cohomology of the complex variety $\mathrm{Sh}_{K^p I_p}(\mathbb{C})$.

Corollary 4.1. *We have*

$$\#B(\mathbb{F}_{q^\alpha}) = |Ker^1(\mathbb{Q}, G)| N_\infty P(q^\alpha) \cdot \Bigg( \sum_{\mathbf{1}(\phi_p)} \sum_{i=0}^{\infty} (-1)^i \zeta_\pi^\alpha \dim \mathrm{H}^i(\mathrm{Sh}_{K^p I_p}(\mathbb{C}), \mathcal{L})[\mathbf{1}(\phi_p)]$$

$$+ (-1)^{(n-1)\#\mathrm{Ram}_p^+} \sum_{\pi_p \text{ St. type}} \sum_{i=0}^{\infty} (-1)^i \zeta_\pi^\alpha \dim \mathrm{H}^i(\mathrm{Sh}_{K^p I_p}(\mathbb{C}), \mathcal{L})[\pi_p^{I_p}] \Bigg) \tag{4.1}$$

*for all positive integers $\alpha$. The numbers $\zeta_\pi$ are roots of unity whose order is bounded by $n \cdot \#\left(Z(\mathbb{Q})\backslash Z(\mathbb{A}_\mathrm{f})/(K \cap Z(\mathbb{A}_\mathrm{f}))\right)$.*

Proof. Take $f^{\infty p} = \mathbf{1}_{K^p}$ and $\xi_\mathbb{C}$ the trivial representation of $G_\mathbb{C}$. By the Grothendieck-Lefschetz trace formula, the Theorem 3.13 provides an expression for the cardinal $\#B(\mathbb{F}_{q^\alpha})$ for all positive integers $\alpha$:

$$\#B(\mathbb{F}_{q^\alpha}) = P(q^\alpha) \Bigg( \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \dim(\pi)=1, \pi_\infty = \mathbf{1}}} \zeta_\phi^\alpha \cdot \dim\left(\pi^p\right)^{K^p} \Bigg) +$$

$$+ (-1)^{(n-1)\#\mathrm{Ram}_p^+} P(q^\alpha) \Bigg( \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p = \text{ St. type}}} \zeta_\pi^\alpha \cdot \mathrm{ep}(\pi_\infty) \dim\left(\pi_\mathrm{f}^p\right)^{K^p} \Bigg), \tag{4.2}$$

where $\mathrm{ep}(\pi_\infty)$ is the Euler-Poincaré characteristic $\sum_{i=0}^{\infty}(-1)^i \dim \mathrm{H}^i(\mathfrak{g}, K_\infty; \pi_\infty)$. The representation $\xi$ at infinity is trivial; therefore the component at infinity of the central character of any automorphic representation contributing to the sums in Equation (4.2) is trivial as well. Thus the numbers $\zeta_\pi^\alpha \in \mathbb{C}^\times$ are roots of unity. The first part of the statement now follows from the formula of Matsushima [**10**, Thm. VII.3.2]. The bound on the order of the roots of unity $\zeta_\pi$ follows from the proof of Lemma 3.11. □

Remark. Note that $\#B(\mathbb{F}_{q^\alpha})$ is (for sufficiently divisible $\alpha$) a sum of powers of $q^\alpha$. This suggests that $B$ may have a decomposition in affine cells as in the case of signatures $(n-1,1); (n,0), \dots (n,0)$.

**4.2. A dimension formula.** In this subsection we show that the dimension of the variety $B/\mathbb{F}_q$ can be deduced from Corollary 4.1. The strategy is to look for the highest order terms in the combinatorial polynomials that describe the compact traces on the representations that occur in the alternating sum of the cohomology of $B$.

Proposition 4.2. *The dimension of the variety $B/\mathbb{F}_q$ is equal to*

$$\sum_{\wp \in \mathrm{Ram}_p^+} \left( \sum_{v \in V(\wp)} \frac{s_v(1-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right).$$

Proof. The Galois group $\mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$ acts through a finite cyclic group on the set of geometric components of the variety $B/\mathbb{F}_q$. In particular the $\alpha$-th power of the Frobenius does not permute these components if $\alpha$ is sufficiently divisible, say divisible by $M \in \mathbb{Z}$ suffices. Assume from now on that $M$ divides $\alpha$. Then each irreducible component of the variety $B_{\mathbb{F}_{q^\alpha}}$ is a geometric component. Pick a component of maximal dimension and inside it a dense open affine subset. By Noether's normalization Lemma this affine subset is finite over an affine space $\mathbb{A}^d_{\mathbb{F}_{q^\alpha}}$ where $d$ is the dimension of $B$. Thus the number of $\mathbb{F}_{q^\alpha}$-points in $B$ is a certain constant times $q^{\alpha d}$ plus lower order terms. From Equation (4.1) we obtain a formula of the form $\#B(\mathbb{F}_{q^\alpha}) = P(q^\alpha) \cdot C$ where $C$ is a complicated constant equal to a difference of dimensions of cohomology spaces.

There are two ways to see that the constant $C$ is non-zero. First Fargues established in his thesis [**39**] that the basic stratum is non-empty, and thus the constant $C$ is non-zero. Second, we sketch an argument for non-emptiness of $B$ using Theorem 3.13. Use an existence theorem of automorphic representations (for example [**20**]) to find after shrinking the group $K$ at least one automorphic representation $\pi$ of $G$ contributing to the sums in Theorem 3.13. By base change and Jacquet-Langlands we can send any such automorphic representation to an automorphic representation of the general linear group (plus similitude factor). By strong multiplicity one for $\mathrm{GL}_n$ the contributing automorphic representations of $G$ are determined up to isomorphism by the set of local components outside any given finite set of places. Therefore we can find a Hecke operator $f^p$ acting by 1 on $\pi^p$ and by 0 on all other automorphic representations contributing to Equation (4.2) (which are *finite* in number). The trace of the correspondence $f^p \times \Phi_{\mathfrak{p}}^{r\alpha}$ acting on the cohomology of the variety $B$ is then certainly non-zero. In particular the variety has non-trivial cohomology and must be non-empty. Therefore the constant $C$ is non-zero.

For the determination of the dimension we forget about the constant $C$. By increasing $M$ (and thus $\alpha$) if necessary we may (and do) assume that the $E_{\mathfrak{p},\alpha}$-algebra $F^+ \otimes E_{\mathfrak{p},\alpha}$ is split. Then, by Proposition 3.3 we have

$$f_\alpha = \mathbf{1}_{q^{-\alpha}\mathbb{Z}_p^\times} \otimes \bigotimes_{\wp | p} \prod_{v \in V(\wp)} f_{n\alpha s_v}^{\mathrm{GL}_n(F_\wp^+)} \in \mathcal{H}_0(G(\mathbb{Q}_p)).$$

We make the formulas for the compact trace of $f_\alpha$ on the trivial representation and on the Steinberg representation explicit. Fix a $\wp$ and write $f_\wp$ for the component of the function $f_\alpha$ at the prime $\wp$. Write $z := \#V(\wp)$. The Satake transform $\mathcal{S}_T(\widehat{\chi}_{N_0} f_\wp^{(P_0)})$ is equal to the polynomial

$$q^{\alpha \sum_{v=1}^{z} \frac{s_v(n-s_v)}{2}} \sum_{(t_{1i}),(t_{2i}),\ldots,(t_{zi})} X_1^{\alpha(t_{11}+t_{21}+\ldots+t_{z1})} X_2^{\alpha(t_{12}+t_{22}+\ldots+t_{z2})} \cdots X_n^{\alpha(t_{1n}+t_{2n}+\ldots+t_{zn})} \tag{4.3}$$

in the ring $\mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots, X_n^{\pm 1}]$. In the above sum, for an index $v$ given, the symbol $(t_{vi})$ ranges over the extended compositions of the number $s_v$ of length $n$ with the following properties:

(C1) for each index $i$ we have $t_{vi} \in \{0, 1\}$;

(C1) define for each $i$ the number $t_i$ to be the sum $t_{1i} + t_{2i} + \ldots + t_{zi}$, then we have

$$t_1 + t_2 + \cdots + t_i > \frac{s_\wp}{n} i, \tag{4.4}$$

for every index $i \in \{1, 2, \ldots, n-1\}$.

The highest order term of the polynomial $P(q^{r\alpha})$ corresponds to extended composition $(t_i)$ of $s$ defined by the equalities

$$t_1 + t_2 + t_3 + \ldots + t_i = \left\lfloor i \frac{s_\wp}{n} \right\rfloor + 1$$

for all $i < n$. This extended composition gives the monomial

$$q^{\alpha \sum_{v \in V(\wp)} \frac{s_v(n-s_v)}{2}} X_1^\alpha X_{\lceil \frac{n}{s_\wp} \rceil}^\alpha X_{\lceil 2\frac{n}{s_\wp} \rceil}^\alpha \cdots X_{\lceil (s_\wp - 1)\frac{n}{s_\wp} \rceil}^\alpha$$

of the truncated Satake function $\mathcal{S}_T(\widehat{\chi}_{N_0} f_\wp^{(P_0)}) \in \mathbb{C}[X_*(T_\wp)]$. We evaluate this monomial at the Hecke matrix of the $T_\wp$-representation $\delta_{P_0}^{1/2}$ to obtain

$$q^{\alpha\left(\sum_{v \in V(\wp)} \frac{s_v(n-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \frac{2\lceil j\frac{n}{s_\wp} \rceil + 1 - n}{2}\right)} \in \mathbb{C}[q^\alpha].$$

By summing over all $\wp \in \mathrm{Ram}_p^+$ we see that the dimension of the variety $B$ is equal to

$$\sum_{\wp \in \mathrm{Ram}_p^+} \left( \sum_{v \in V(\wp)} \frac{s_v(1-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right).$$

This completes the proof. $\square$

CHAPTER 3

# The cohomology of the basic stratum II

We remove an hypothesis from the main Theorem of the previous chapter. In the previous chapter we proved a relation between the $\ell$-adic cohomology of the basic stratum of some simple Shimura varieties and the cohomology of the complex Shimura variety. These simple Shimura varieties are those of Kottwitz considered in his Inventiones article [**58**] on the construction of Galois representation. The varieties are associated to certain division algebras over $\mathbb{Q}$ with involution of the second kind; we call such varieties Kottwitz varieties. We proved the main theorem of the previous chapter assuming (essentially) that the Newton polygon associated to the basic stratum has no integral point other than the begin point and the end point. In this chapter we solve the resulting combinatorial problems when one removes this simplifying condition from the theorem in case the prime $p$ of reduction is split in the center of the division algebra defining the Kottwitz variety.

A consequence of our final result is an explicit expression for the zeta function of the basic stratum of Kottwitz's varieties at split primes of good reduction. The expressions are in terms of: (1) Automorphic forms on the group $G$ of the Shimura datum, (2) The determinant of the factor at $\mathfrak{p}$ of their associated Galois representations, and (3) Polynomials in $q^{\alpha}$ of combinatorial nature, associated to certain non-crossing lattice paths in the plane $\mathbb{Q}^2$.

As an application we deduce a formula for the dimension of the basic stratum. Our formula agrees with the conjecture from [**61**] (cf. [**87**, Conj. 7.5], [**17**]) for the dimension of the Newton strata, specialized to the cases considered in this chapter.

## 1. Notations

Let $p$ be a prime number and let $F$ be a non-Archimedean local field with residue characteristic equal to $p$. Let $\varpi_F \in \mathcal{O}_F$ be a prime element, and define $q := \#(\mathcal{O}_F/\varpi_F)$. We write $G_n$ for the topological group $\mathrm{GL}_n(F)$, and we write $\mathcal{H}(G)$ for the Hecke algebra of locally compact constantly supported functions on $G$. We often drop the index $n$ from the notation if confusion is not possible. We call a parabolic subgroup $P$ of $G$ *standard* if it is upper triangular, and we write $P = MN$ for its standard Levi decomposition. We write $K$ for the hyperspecial group $\mathrm{GL}_n(\mathcal{O}_F)$ and $\mathcal{H}_0(G)$ for the Hecke algebra of $G$ with respect to $K$. The group $P_0 \subset G$ is the standard Borel subgroup of $G$, $T$ is the diagonal torus of $G$, and $N_0$ is the group of upper triangular unipotent matrices in $G$.

We write $\widehat{G}, \widehat{T}, \widehat{M}, \ldots$ for the corresponding complex dual groups, $\widehat{G} = \mathrm{GL}_n(\mathbb{C})$, $\widehat{T} = (\mathbb{C}^\times)^n$, and so on. If $\pi$ is an unramified representation of some Levi subgroup $M$ of $G$ then we write $\varphi_{M,\pi} \in \widehat{M}$ for the Hecke matrix of this representation.

Let $n$ be a positive integer. A *partition* of $n$ is a finite, *non-ordered* list of non-negative numbers whose sum is equal to $n$. A *composition* of $n$ is a finite, *ordered* list of positive numbers whose sum is equal to $n$. Recall that the compositions of $n$ correspond to the standard parabolic subgroups of $G$.

We write $A$ for the ring $\mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots, X_n^{\pm 1}]^{\mathfrak{S}_n}$. The Satake transform $\mathcal{S}$ provides an isomorphism from $\mathcal{H}_0(G)$ onto the ring $A$.

Let $n$ and $\alpha$ be positive integers, and let $s$ be a non-negative integer with $s \leq n$. We call the number $s$ the *signature*, and we call the number $\alpha$ the *degree*. The function $f_{n\alpha s} \in \mathcal{H}_0(G)$ is the spherical function whose Satake transform is

$$q^{\alpha s(n-s)/2} \sum_{\nu \in \mathfrak{S}_n \cdot \mu_s} [\nu]^\alpha = q^{\alpha s(n-s)/2} \sum_{I \subset \{1, \ldots, n\}, \# I = s} \prod_{i \in I} X_i^\alpha \in A. \tag{1.1}$$

We put $f_{n\alpha s} = 0$ when $n, \alpha, s \in \mathbb{Z}_{\geq 0}$ are such that $n < s$. We will call $f_{n\alpha s}$ a *simple Kottwitz function*. The *composite Kottwitz functions* $f_{n\alpha\sigma}$ are obtained from partitions $\sigma$ of $s$ as follows. Let $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_r)$ be a partition of $s$. Then we write $f_{n\alpha\sigma} \in \mathcal{H}_0(G)$ for the convolution product $f_{n\alpha\sigma_1} * f_{n\alpha\sigma_2} * \cdots * f_{n\alpha\sigma_r} \in \mathcal{H}_0(G)$.

We write $\chi_c^G$ for the characteristic function on $G$ of the subset of compact elements. Let $\pi$ be a smooth $G$-representation of finite length and $f$ a locally constant, compactly supported function on $G$. Then we write $\mathrm{Tr}(\chi_c^G f, \pi)$ for the *compact trace* [**22**] of $f$ against $\pi$.

Let $m, m' \in \mathbb{Z}_{\geq 1}$. If $\pi$ (resp. $\pi'$) is a smooth admissible representation of $G_m$ (resp. $G_{m'}$), then we write $\pi \times \pi'$ for the $G_{m+m'}$-representation parabolically induced (unitary induction) from the representation $\pi \otimes \pi'$ of the standard Levi subgroup consisting of two blocks, one of size $m$, and the other one of size $m'$. The tensor product $\pi \otimes \pi'$ in the above formula is taken along the blocks of this Levi subgroup. We write $\mathcal{R}$ for the direct sum $\bigoplus_{n \in \mathbb{Z}_{\geq 0}} \mathrm{Groth}(G_n)$ with the convention that $G_0$ is the trivial group. The group $G_0$ has one unique irreducible representation $\sigma_0$ (the space $\mathbb{C}$, with trivial action). The operation "direct sum of representations" together with the product "$\times$" turns the vector space $\mathcal{R}$ into a commutative $\mathbb{C}$-algebra with $\sigma_0$ as unit element. We call it the *ring of Zelevinsky*.

The ring of Zelevinsky has an involution $\iota$, called the *Zelevinsky involution*. Aubert [**5**] gave a refined definition of this involution, making sense for all reductive groups. The involution is defined by $X^\iota := \sum_{P=MN} \varepsilon_P \mathrm{Ind}_P^G(X_N(\delta_P^{-1/2}))$ for all $X \in \mathcal{R}$. With 'involution' we mean that $\iota$ is an automorphism of the complex algebra $\mathcal{R}$ and it is of order two: $\iota^2 = \mathrm{Id}_{\mathcal{R}}$.

We write $\nu$ for the absolute value morphism from $\mathrm{GL}_1(F) = F^\times$ to $\mathbb{C}^\times$. By a *segment* $S = \langle x, y \rangle$ we mean a set of numbers $\{x, x+1, \ldots, y\}$ where $x, y \in \mathbb{Q}$ and where we need to explain the conventions in case $y \leq x$. In case $y$ is strictly smaller than $x - 1$, then $\langle x, y \rangle = \emptyset$; in case $x$ is equal to $y$, then the segment $\langle x, y \rangle = \{x\}$ has one element. We have one unusual convention: For $y = x - 1$ we define the segment $\langle x, y\rangle$ to be the set $\{\star\}$ of one element containing a distinguishing symbol "$\star$". The *length* $\ell(S)$ of a segment $S = \langle x, y\rangle$ is defined to be $y - x + 1$. Thus the segment $\{\star\}$ has length 0, the segment $\{x\}$ has length 1, the segment $\{x, x+1\}$ has length 2, etc. We put $\ell\langle x, y\rangle = -1$ in case $y < x - 1$.

For any segment $\langle x, y\rangle$ with $y \geq x$ we write $\Delta\langle x, y\rangle$ for the unique irreducible quotient of the induced representation $\nu^x \times \nu^{x+1} \times \cdots \times \nu^y$. We define $\Delta\{\star\}$ to be $\sigma_0$ (the one-dimensional representation of the trivial group $\mathrm{GL}_0(F)$), and we define $\Delta\langle x, y\rangle$ to be 0 in case $y < x - 1$. For any segment $S$ of non-negative length the object $\Delta S$ is a representation of the group $\mathrm{GL}_n(F)$, where $n$ is the length of $S$.

For the standard properties of segments we refer to Zelevinsky's work [**105**] (cf. [**90**]), but note that our conventions are slightly different, because we allow rational numbers in the segments and we have the segment $\{\star\}$. We mention that this difference is there only for notational purposes, and that it does not change the mathematics.

For any finite ordered list of segments $S_1, S_2, \ldots, S_t$ we have the product representation $\pi := (\Delta S_1) \times (\Delta S_2) \times \cdots \times (\Delta S_t)$. Observe that, due to our conventions, in case $S_a = \{\star\}$ for some $a$, then $\Delta S_a$ is the unit in $\mathcal{R}$, and

$$\pi = (\Delta S_1) \times (\Delta S_2) \times \cdots \times \widehat{(\Delta S_a)} \times \cdots \times (\Delta S_t) \in \mathcal{R}, \tag{1.2}$$

where the hat means that we leave the corresponding factor out of the product. In case $S_b = \emptyset$ for some index $b$, then we have $\pi = 0$ in $\mathcal{R}$.

In the combinatorial part of this chapter the representations of interest are the Speh representations. We recall their definition here. Let $t, h$ be positive integers such that $n = th$. We define $\mathrm{Speh}(h, t)$ to be the (unique) irreducible quotient of the representation $\mathrm{St}_{G_h}\nu^{\frac{t-1}{2}} \times \cdots \times \mathrm{St}_{G_h}\nu^{\frac{1-t}{2}}$. This representation has $t$ segments, $S_a = \langle x_a, y_a\rangle$, $a = 1, \ldots, t$, where

$$x_a = \frac{t-h}{2} - (a-1) \quad \text{and} \quad y_a = \frac{t+h}{2} - a.$$

Observe that, for each index $a$, we have $\ell S_a = h$. Furthermore, for each index $a < t$, we have $x_{a+1} = x_a - 1$ and $y_{a+1} = y_a - 1$.

If $P = MN \subset G$ is a standard parabolic subgroup of $G$, then we have the spherical functions $\chi_N, \widehat{\chi}_N$ in $\mathcal{H}_0(M)$ associated to the acute and obtuse Weyl chambers. We refer to Equations (2.1.1) and (2.1.2) for the precise definition and explicit description of these functions.

## 2. Computation of some compact traces

In this section we compute the compact traces $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$ of the simple Kottwitz functions $f$ on a certain class of representations $\pi$. This class will be sufficiently large to contain all smooth representations that occur in the cohomology of (basic) strata of unitary Shimura varieties at primes of good reduction.

We will follow the following strategy to compute $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$. A semistable representation $\pi$ of $G$ is called *standard* if it is isomorphic to a product of essentially square integrable representations. The computation of the compact trace $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$ on a square-integrable representation is easy, and using van Dijk's formula adapted for compact traces Proposition 2.1.5, we easily deduce formulas for compact traces on the standard representations. Any semistable irreducible representation $\pi$ may be [1] (uniquely) written as a sum $\pi = \sum_I c_I \cdot I \in \mathcal{R}$ where $I$ ranges over the standard representations, and the coefficients $c_I \in \mathbb{C}$ are 0 for nearly all $I$. We have

$$\mathrm{Tr}(\chi_c^G f, \pi) = \sum_I c_I \,\mathrm{Tr}(\chi_c^G f, I).$$

Thus, there are two steps to compute $\mathrm{Tr}(\chi_c^G f, \pi)$: (Prob1) Know the coefficients $c_I$ and (Prob2) Make the sum $\sum_I c_I \,\mathrm{Tr}(\chi_c^G f, I)$. The first problem (Prob1) is related to the Kazhdan-Lustzig conjecture [2]. The "Kazhdan-Lustzig Theorem" of Beilinson-Bernstein [**8**] (and [**50**]) interprets the multiplicity of any given irreducible representation $\pi$ in the representation $I$. The Kazhdan-Lustzig Theorem interprets this multiplicity as the dimension of certain intersection cohomology spaces, and also as the value at $q = 1$ of certain Kazhdan-Lustzig polynomials.

For the irreducible representations $\pi$ contributing to the cohomology of Newton strata of unitary Shimura varieties we will not have to deal with problem (Prob1). The Theorem of Moeglin-Waldspurger [**80**] (cf. (Theorem 2.2.1)) for the discrete spectrum of the general linear group implies that these representations must be of a very particular kind (*rigid representations*). Any rigid representation is a product of unramified twists of Speh representations in $\mathcal{R}$, and therefore we restrict our attention to these Speh representations only. Tadic has solved the first problem (Prob1) for the Speh representations. The coefficients $c_I$ turn out to be $-1, 0$ or $1$ for these representations (precise statement in Theorem 2.1). Therefore, we are mostly concerned with the second problem (Prob2).

**2.1. Tadic's determinantal formula.** We recall an important character formula of Tadic for the Speh representations. This formula is a crucial ingredient for our computations.

Let $S_1 = \langle x_1, y_1\rangle, S_2 = \langle x_2, y_2\rangle, \ldots, S_t = \langle x_t, y_t\rangle$ be an ordered list of segments defining a representation of the group $G = \mathrm{GL}_n(F)$. Let $\mathfrak{S}_t$ be the symmetric group on $\{1, 2, \ldots, t\}$. For any $w \in \mathfrak{S}_t$ we define the number $n_a^w$ to be $y_a - x_{w(a)} + 1$. We have

$$\sum_{a=1}^k n_a^w = \left(\sum_{a=1}^k y_a\right) - \left(\sum_{a=1}^k x_{w(a)}\right) + k = \left(\sum_{a=1}^k y_a\right) - \left(\sum_{a=1}^k x_a\right) + k = \sum_{a=1}^k n_a = n. \tag{2.1}$$

The numbers $n_a^w$ need not be positive. We define $\mathfrak{S}_t' \subset \mathfrak{S}_t$ to be subset consisting of those permutations $w \in \mathfrak{S}_t$ such that the numbers $n_a^w$ are positive or 0. If the permutation $w$ lies

1. Zelevinsky proved in [**105**] that the standard representations form a basis of $\mathcal{R}$ as complex vector space.
2. This conjecture is a Theorem, see [**19**, Thm. 8.6.23]

in the subset $\mathfrak{S}'_t \subset \mathfrak{S}_t$, then $(n_a^w)$ is a composition of $n$. Assuming that $w \in \mathfrak{S}'_t$ we will write $P_w = M_w N_w$ for the parabolic subgroup of $G$ corresponding to the composition $(n_a^w)$.

Let $w \in \mathfrak{S}'_t$. We define the segments $S_1^w := \langle x_{w(1)}, y_1 \rangle, S_2^w := \langle x_{w(2)}, y_2 \rangle, \dots, S_t^w := \langle x_{w(t)}, y_t \rangle$. We have $\ell(S_a^w) = n_a^w$. We let $\Delta_w$ be the representation of $M_w$ defined by $(\Delta S_1^w) \otimes \cdots \otimes (\Delta S_t^w)$, where the tensor product is taken along the blocks of $M_w$. The representation $I_w$ is defined to be the product $\Delta S_1^w \times \Delta S_2^w \times \cdots \Delta S_t^w$, i.e. it is the (unitary) parabolic induction $\mathrm{Ind}_{P_w}^G \Delta_w$ of $\Delta_w$ to $G$. In case $w \in \mathfrak{S}_t \backslash \mathfrak{S}'_t$ we define both $\Delta_w$ and $I_w$ to be 0.

Remark. It is possible that $S_a^w = \{\star\}$ for some permutation $w$. In that case the representation $\Delta S_a^w$ is the unit element $\sigma_0$ of $\mathcal{R}$, and thus can be left out of the product that defined $I_w$ (cf. Equation (1.2)).

In these notations we have the following theorem:

Theorem 2.1 (Tadic). *Let $\pi$ be a Speh representation of $G$ and let $S_1 = \langle x_1, y_1 \rangle, S_2 = \langle x_2, y_2 \rangle, \dots S_t = \langle x_t, y_t \rangle$ be its segments. The representation $\pi$ satisfies Tadic's determinantal formula*

$$\pi = \sum_{w \in \mathfrak{S}_t} \mathrm{sign}(w) I_w.$$

Proof. This Theorem was frst proved by Tadic in [**99**] for Speh representations with a difficult argument. Chenevier and Renard simplified the proof and observed that the above expression is a determinant of a matrix with coefficients in Zelevinsky's ring $\mathcal{R}$. Also Badulescu gave a simpler proof of Theorem 2.1 in the note [**7**] using the Moeglin-Waldspurger algorithm [**79**]. Recently Lapid and Minguez [**71**, Thm. 1] extended the formula to the larger class of ladder representations. □

Remark. Our formulation of Theorem 2.1 is weaker than the theorem proved by the above authors, because we consider only *semistable* Speh representations. (They have a similar statement also for the non semistable Speh/ladder representations.)

By the definition of the subset $\mathfrak{S}'_t \subset \mathfrak{S}_t$ we have for all $w \in \mathfrak{S}_t$ that $I_w \neq 0$ if and only if $w \in \mathfrak{S}'_t$, and thus we may as well index over the elements $w \in \mathfrak{S}'_t$ in the sum in the above Theorem. In the cases where the inclusion $\mathfrak{S}'_t \subset \mathfrak{S}_t$ is strict, the subset $\mathfrak{S}'_t$ is practically never a subgroup of $\mathfrak{S}_t$, it will neither be closed under composition nor contain inverses of elements.

**2.2. Lattice paths and the Steinberg representation.** In this section we will express the compact trace of the functions $f_{n\alpha s}$ on the Steinberg representation in terms of certain lattice paths in $\mathbb{Q}^2$.

We fix throughout this section a positive integer $\alpha$, called the *degree*. This integer will play only a minor role in the computations of this section as it affects only the weights of the paths. The degree will become more important later.

Let $A^+$ be the polynomial ring $\mathbb{C}[q^a|a \in \mathbb{Q}]$ of rational, formal powers of the variable $q$. Equivalently, $A^+$ is the complex group ring $\mathbb{C}[\mathbb{Q}^+]$ of the additive group $\mathbb{Q}^+$ underlying $\mathbb{Q}$. A *path* $L$ in $\mathbb{Q}^2$ is a sequence of points $\vec{v}_0, \vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ such that $\vec{v}_{i+1} - \vec{v}_i = (1,0)$ (*east*), or $\vec{v}_{i+1} - \vec{v}_i = (1,1)$ (*north-east*). The starting point of $L$ is $\vec{v}_0$ and the end point is $\vec{v}_r$; the number $r$ is the *length*. An eastward step $(1,0)$ has weight 1 and a north-eastward step $(a,b) \to (a+1, b+1)$ has weight $q^{-\alpha \cdot a} \in A^+$. The *weight of the path* $L$ is defined to be the product in $A^+$ of the weights of its steps.

Remark. We allow paths of length zero; such a path consists of one point $\vec{v}_0$ and no steps. The weight of a path of length 0 is equal to 1. The paths of length 0 correspond to compact traces on the special segments $\{\star\}$ introduced earlier.

Let $L$ be a path in $\mathbb{Q}^2$. Connect the starting point $\vec{v}_0$ of $L$ with its end point $\vec{v}_r$ via a straight line $\ell$. Then $L$ is called a *Dyck path* if all of its points $\vec{v}_a$ lie on or below the line $\ell$ in the plane $\mathbb{Q}^2$. The Dyck path is called *strict* if none of its points $\vec{v}_a$ other than the initial and end point, lies on the line $\ell$.

Let $\vec{x}, \vec{y}$ be two points in $\mathbb{Q}^2$. Then we write $\mathrm{Dyck_s}(\vec{x}, \vec{y}) \in A^+$ for the sum of the weights of all the strict Dyck paths that go from the point $\vec{x}$ to the point $\vec{y}$. We call the polynomial $\mathrm{Dyck_s}(\vec{x}, \vec{y})$ the *strict Dyck polynomial.* There are also non-strict Dyck polynomials $\mathrm{Dyck}(\overline{x}, \overline{y})$ but we are not concerned with those in this subsection; they are important for the computation of compact traces on the trivial representation.

Let $f \in \mathcal{H}_0(G)$ be a function. We abuse notation and write $\widehat{\chi}_N \mathcal{S}(f)$ for the $T$-Satake transform of the function $\widehat{\chi}_N f^{(P_0)}$. This truncation $\widehat{\chi}_N f$ of an element $f \in A$ is best understood graphically.

We first extend the notion of a path slightly to the concept of a graph. A *graph* in $\mathbb{Q}^2$ is a sequence of points $\vec{v}_0, \vec{v}_1, \dots, \vec{v}_r$ with $\vec{v}_{i+1} - \vec{v}_i = (1, e)$, where $e$ is an integer. Thus the paths are those graphs with $e \in \{0, 1\}$ for each of its steps. We define the *weight* of a step $(a, b) \to (a + 1, b + e)$ to be $q^{-\alpha \cdot e \cdot a} \in A^+$, and the *weight of a graph* is the product of the weights of its steps.

To a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n} \in \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \dots, X_n^{\pm 1}]$, with $e_i \in \mathbb{Z}$ and $\sum_{i=1}^n e_i = s$ we associate the graph $\mathcal{G}_X$ with points

$$\vec{v}_0 := \ell(\tfrac{1-n}{2}), \quad \vec{v}_i := \vec{v}_0 + (i, e_n + e_{n-1} + \ldots + e_{n+1-i}) \in \mathbb{Q}^2, \tag{2.2}$$

for $i = 1, \dots, n$. If $x \in \mathbb{Q}$, then we write $\ell(x)$ for the point $(x, \frac{s}{n}x)$ on the line $\ell$. Because the sum $\sum_{i=1}^n e_i$ is equal to $s$, the end point of the graph is

$$\ell(\tfrac{1-n}{2}) + (n, s) = \ell(\tfrac{n-1}{2} + 1) \in \mathbb{Q}^2.$$

We have $\widehat{\chi}_{N_0} X = X$ if and only if[3]

$$e_1 + e_2 + \ldots + e_i > \tfrac{s}{n} i, \tag{2.3}$$

for all indices $i < n$ (and if $\widehat{\chi}_{N_0} X \neq X$, then $\widehat{\chi}_{N_0} M = 0$). The condition in Equation (2.3) is true if and only if the graph defined in Equation (2.2) lies strictly below the straight line $\ell \subset \mathbb{Q}^2$ of slope $\frac{s}{n}$ going through the origin. Furthermore, the evaluation of $X$ at the point

$$\left(q^{\frac{1-n}{2}}, q^{\frac{3-n}{2}}, \ldots, q^{\frac{n-1}{2}}\right) \in \widehat{T}, \tag{2.4}$$

equals the weight of the graph $\mathcal{G}_X$.

Remark. The reader might find it strange that in Equation (2.2) we let the graphs go in the inverse direction. Why did we make this convention? We made this convention because we want graphs that stay below the line $\ell$. If we draw the graphs using the 'natural' formula, then we get a graph above the line $\ell$ going from right to left. So why not consider only graphs that stay above $\ell$? Of course this is equivalent, but later, when we compute the graph for the trivial representation, we get get graphs whose 'natural' formula stays below $\ell$ and goes from left to right. Thus, either way, we have to invert directions.

Lemma 2.2. *Consider the representation $\pi = \mathbf{1}_T(\delta_{P_0}^{1/2})$ of the group $T$. Let $f$ be a function in the spherical Hecke algebra of $T$. Then the trace of $f$ against $\pi$ is equal to the evaluation of $\mathcal{S}(f) \in A$ at the point*

$$\left(q^{\frac{1-n}{2}}, q^{\frac{3-n}{2}}, \ldots, q^{\frac{n-1}{2}}\right) \in \widehat{T},$$

Proof. The character $\delta_{P_0}^{1/2}$ on $T$ is equal to

$$T \ni (t_1, t_2, \ldots, t_n) \longmapsto |t_1|^{\frac{n-1}{2}} |t_2|^{\frac{n-3}{2}} \cdots |t_n|^{\frac{n-1}{2}} \in \mathbb{C}^\times.$$

To any (rational) cocharacter $\nu \in X_*(T)$ we may associate the composition $(\delta_{P_0}^{1/2} \circ \nu) \colon F^\times \to T \to \mathbb{C}^\times$. We evaluate this composition at the prime element $\varpi_F \in F^\times$. Thus we have an element of the set

$$\mathrm{Hom}(X_*(T), \mathbb{C}^\times) = \mathrm{Hom}(X^*(\widehat{T}), \mathbb{C}^\times) = X_*(\widehat{T}) \otimes_{\mathbb{Z}} \mathbb{C}^\times = \widehat{T}(\mathbb{C}), \tag{2.5}$$

where the last isomorphism is given by

$$X_*(\widehat{T}) \otimes_{\mathbb{Z}} \mathbb{C}^\times \ni \nu \otimes z \longmapsto \nu(z) \in \widehat{T}(\mathbb{C}).$$

We have $T = (F^\times)^n$ and thus we have the standard basis $e_i$ on $X_*(T)$. This corresponds to the standard basis $e_i$ on $X_*(\widehat{T})$ via the first two equalities in Equation (2.5). If we take $\nu = e_i$

3. This is true because the fundamental weights $\varpi_{\alpha_i}^G$ of the general linear group are of the form $H_1 + \cdots + H_i - \frac{i}{n}(H_1 + H_2 + \ldots + H_n)$ on $\mathfrak{a}_0$. The statement follows also directly from the conclusion made at Equation (2.1.10).

in $(\delta_{P_0}^{1/2} \circ \nu)(\varpi_F)$ then we get

$$(\delta_{P_0}^{1/2} \circ e_i)(\varpi_F) = |\varpi_F|^{\frac{n-1}{2} - i + 1} = q^{\frac{1-n}{2} + i - 1}.$$

This completes the verification. □

Let $\ell \subset \mathbb{Q}^2$ be the line of slope $\frac{s}{n}$ through $0 \in \mathbb{Q}^2$ that we introduced earlier. We write $\ell(x)$ for the point $(x, \frac{s}{n}x)$ on $\ell$ if $x \in \mathbb{Q}$.

LEMMA 2.3. *The compact trace* $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \mathrm{St}_G)$ *on the Steinberg representation is equal to the polynomial* $(-1)^{n-1} q^{s(n-s)/2} \cdot \mathrm{Dyck_s}(\ell(\frac{1-n}{2}), \ell(\frac{n-1}{2} + 1)) \in A^+$.

PROOF. The proof is a translation of a result that we obtained in Chapter 2:

$$\mathrm{Tr}(\chi_c^G f, \mathrm{St}_G) = (-1)^{n-1} \, \mathrm{Tr}\left(\widehat{\chi}_{N_0} f_{n\alpha s}, \mathbf{1}_T(\delta_{P_0}^{1/2})\right), \tag{2.6}$$

(see Proposition 2.1.13). NB: We wrote $\delta_{P_0}^{1/2}$ and not $\delta_{P_0}^{-1/2}$; the additional sign is there because the Jacquet module at $P_0$ of the Steinberg representation $\mathrm{St}_G$ is $\mathbf{1}_T(\delta_{P_0})$.

In case $f = f_{n\alpha s}$ then every monomial $X$ occurring in $\mathcal{S}(f)$ is multiplicity free[4], and therefore the graph $\mathcal{G}_X$ is in fact a *path*. The above construction $X \mapsto \mathcal{G}_X$ provides a bijection between the monomials that occur in $\mathcal{S}(f)$ and the possible paths that go from the point $\vec{v}_0$ to the point $\vec{v}_r$. Finally $\widehat{\chi}_{N_0} X \neq 0$ if and only if the corresponding path is a Dyck path (see Equation (2.3)). This completes the proof. □

Compact traces are compatible with twists:

LEMMA 2.4. *Let* $\chi$ *be an unramified character of* $F^\times$, $\pi$ *a smooth irreducible* $G$ *representation, and* $f_{n\alpha\sigma} \in \mathcal{H}_0(G)$ *a function of Kottwitz. Then* $\mathrm{Tr}(\chi_c^G f_{n\alpha\sigma}, \pi \otimes \chi) = \chi(\varpi_F^{\alpha s}) \cdot \mathrm{Tr}(\chi_c^G f, \pi)$.

PROOF. Lemma 2.1.8. □

LEMMA 2.5. *Assume that* $\pi$ *is an essentially square integrable representation of the form* $\Delta S$, *where* $S = \langle x, y \rangle$ *is a segment of length* $n$. *Then*

$$\mathrm{Tr}(\chi_c^G f, \Delta\langle x, y\rangle) = (-1)^{n-1} \cdot q^{\frac{s(n-s)}{2}} \cdot \mathrm{Dyck_s}(\ell(x), \ell(y+1)).$$

PROOF. The representation $(\Delta S) \otimes \nu^{-x + \frac{1-n}{2}}$ is the Steinberg representation, and so Lemma 2.3 applies to it. The result then follows from Lemma 2.4. □

4. Multiplicity free in the sense that no variable $X_i$ occurs with exponent $e_i > 1$ in $X$.

**2.3. Lattice $t$-paths and standard representations.** We describe the compact traces on the standard representations of $G$ using "$t$-paths".

Let $t$ be a positive integer. Let $\vec{x} = (\vec{x}_a)$ and $\vec{y} = (\vec{y}_a)$ be two ordered lists of points in $\mathbb{Q}^2$, both of length $t$. A *$t$-path from $\vec{x}$ to $\vec{y}$* is the datum consisting of, for each index $a \in \{1, 2, \ldots, t\}$, a path $L_a$ from the point $\vec{x}_a$ to the point $\vec{y}_a$. A $t$-path $(L_a)$ is called a *Dyck $t$-path* if all the paths $L_a$ are Dyck paths. The Dyck path $(L_a)$ is called *strict* if, for each index $a$, no point $\vec{v}_i$ of $L_a$ other than $\vec{v}_0$ and $\vec{v}_r$ lies on the line $\ell$. The *weight* $\mathrm{weight}(L_a)$ of a $t$-path $(L_a)$ is the product of the weights of the paths $L_a$, where $a$ ranges over the set $\{1, 2, \ldots, t\}$. We extend the definition of the strict Dyck polynomial $\mathrm{Dyck_s}(\vec{x}, \vec{y}) \in A^+$ also to $t$-paths: The polynomial $\mathrm{Dyck_s}(\vec{x}, \vec{y}) \in A^+$ is by definition the sum of the weights of the strict Dyck $t$-paths from the points $(\vec{x}_a)$ to the points $(\vec{y}_a)$. We have

$$\mathrm{Dyck_s}(\vec{x}, \vec{y}) = \prod_{a=1}^{t} \mathrm{Dyck_s}(\vec{x}_a, \vec{y}_a) \in A^+. \tag{2.7}$$

LEMMA 2.6. *Let $S_1 = \langle x_1, y_1 \rangle, S_2 = \langle x_2, y_2 \rangle, \ldots, S_t = \langle x_t, y_t \rangle$ be a list of segments and let $I$ be the representation $(\Delta S_1) \times (\Delta S_2) \times \cdots \times (\Delta S_t)$. Then the compact trace $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, I)$ is equal to $(-1)^{n-t}\mathrm{Dyck_s}(\vec{x}, \vec{y})$, where for the indices $a = 1, \ldots, t$ we have $\vec{x}_a := \ell(x_a)$ and $\vec{y}_a := \ell(y_a + 1)$.*

REMARK. The sign $(-1)^{n-t}$ is equal to $\varepsilon_{M \cap P_0}$, where $M$ is the standard Levi subgroup of $G$ corresponding to the composition $\sum_{a=1}^{t} \ell(n_a)$ of $n$.

PROOF. Let $P$ be the parabolic subgroup of $G$ corresponding to the composition $n = \sum_{a=1}^{t} \ell(S_a)$ of $n$. Let $\chi_M^G$ be the characteristic function on $M$ of the subset of elements $m \in M$ such that $\langle \varpi_\alpha^G, H_M(m) \rangle = 0$ for all $\alpha \in \Delta_P$. By the integration formula of van Dijk for compact traces Proposition 2.1.5 we have

$$\begin{aligned} \mathrm{Tr}(\chi_c^G f, I) &= \mathrm{Tr}\left(\chi_c^G f_{n\alpha s}^{(P)}, (\Delta S_1) \times (\Delta S_2) \times \cdots \times (\Delta S_t)\right) \\ &= \mathrm{Tr}\left(\chi_c^M \chi_M^G f_{n\alpha s}^{(P)}, (\Delta S_1) \times (\Delta S_2) \times \cdots \times (\Delta S_t)\right). \end{aligned} \tag{2.8}$$

We proved in Proposition 2.1.10 that the function $\chi_c^G f_{n\alpha s}^{(P)}$ is equal to

$$q^{\alpha C(n_a, s_a)} f_{n\alpha s_1} \otimes f_{n\alpha s_2} \otimes \cdots \otimes f_{n\alpha s_t}, \tag{2.9}$$

where $s_a := \frac{n_a}{n} s$, and

$$C(n_a, s_a) := \frac{s(n-s)}{2} - \sum_{a=1}^{t} \frac{s_a(n_a - s_a)}{2}.$$

The constant term in Equation (2.9) vanishes in case one of the numbers $s_a$ is non-integral. We have $(\chi_c^G f_{n\alpha s}^{(P)})^{(P_0 \cap M)} = \chi_M^G f_{n\alpha s}^{(P_0)}$. Consequently, one may rewrite the trace in Equation (2.8) to the product

$$q^{\alpha C(n_a, s)} \prod_{a=1}^{t} \mathrm{Tr}(\chi_c^{G_{n_a}} f_{n_a \alpha s_a}, \Delta S_a),$$

By Lemma 2.5 we obtain

$$q^{\alpha C(n_a,s)} \prod_{a=1}^{t} (-1)^{n_a-1} q^{\frac{n_a(n_a-s_a)}{2}\alpha} \mathrm{Dyck}_{\mathrm{s}}(\ell(x_a), \ell(y_a+1)).$$

Note that the condition that $s_a$ is integral precisely corresponds to the condition that the vertical distance between the point $\vec{y}_a$ and the point $\vec{x}_a$ has to be integral before paths can exist. Therefore the expression in this last Equation simplifies to the one stated in the Lemma and the proof is complete. □

**2.4. Non-crossing paths.** We express the compact traces on Speh representations in terms of non-crossing lattice paths.

We call a $t$-path $(L_a)$ *crossing* if there exists a couple of indices $a, b$ with $a \neq b$ such that the path $L_a$ has a point $\vec{v} \in \mathbb{Q}^2$ in common with the path $L_b$. There is an important condition:

- The point $\vec{v}$ of crossing must appear in the list of points $\vec{v}_{a,i}$ that define $L_a$ and it must also occur in the list of points $\vec{v}_{b,i}$ that define $L_b$.

(Because we work with rational coordinates, the point of intersection could be a point lying halfway a step of a path (for example). We are ruling out such possibilities.)

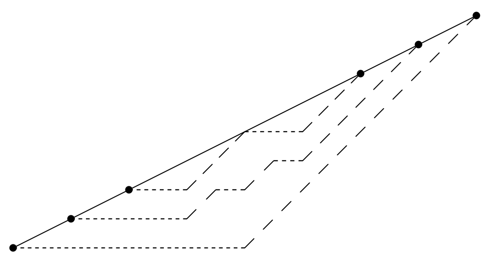

FIGURE 1. An example of a 3-path corresponding to the representation $\pi$ of $\mathrm{GL}_{54}(F)$ defined by the segments $\langle 3, 20\rangle$, $\langle 2, 19\rangle$ and $\langle 1, 18\rangle$. We take $s = 27$ and we take the permutation $w = (13) \in \mathfrak{S}'_3$. The 3 dots on the lower left hand corner are the points $\vec{x}_1, \vec{x}_2$ and $\vec{x}_3$ in $\mathbb{Q}^2$ respectively; the points $\vec{y}_1, \vec{y}_2$ and $\vec{y}_3$ are in the upper right corner. Observe that this 3-path is non-strict.

We write $\mathrm{Dyck}^+_{\mathrm{s}}(\vec{x}, \vec{y})$ for the sum of the weights of the *non-crossing* strict Dyck $t$-paths. Let $\pi$ be the Speh representation of $G$ associated to the Zelevinsky segments $\langle x_1, y_1\rangle, \langle x_2, y_2\rangle, \ldots, \langle x_t, y_t\rangle$ with $x_1 > x_2 > \ldots > x_t$ and $y_1 > y_2 > \ldots > y_t$. We define the points $\vec{x}_a := \ell(x_a) \in \mathbb{Q}^2$ and $\vec{y}_a := \ell(y_a + 1) \in \mathbb{Q}^2$, for $a = 1, 2, \ldots, t$. The group $\mathfrak{S}_t$ acts on the free $\mathbb{Q}^2$-module $\mathbb{Q}^{2t} = (\mathbb{Q}^2)^t$ by sending the $a$-th standard basis vector $e_a \in (\mathbb{Q}^2)^t$ to the basis vector $e_{w(a)} \in (\mathbb{Q}^2)^t$. Thus if we have the vector $\vec{x} \in \mathbb{Q}^{2t}$, then we get the new vector $\vec{x}^w$ whose $a$-th coordinate $\vec{x}^w_a \in \mathbb{Q}^{2t}$ is equal to $w(a)$-th coordinate of the vector $\vec{x}$.

REMARK. The difference $\frac{s}{n} \cdot (y_a + 1) - \frac{s}{n} \cdot x_a$ need not be integral. In that case there do not exist paths from the point $\vec{x}^w_a \in \mathbb{Q}^2$ to $\vec{y}_a \in \mathbb{Q}^2$.

Let $\pi$ be a Speh representation of type $(h,t)$. The points $\vec{x}_a \in \mathbb{Q}^2$ and $\vec{y}_a \in \mathbb{Q}^2$ lie on the line $\ell \subset \mathbb{Q}^2$, and the point $\vec{x}_a$ lies on the left of the point $\vec{y}_a$ with horizontal distance $y_a + 1 - x_a = \ell(S_a) = h$. The two lists of points may overlap: There could exist couples of indices $(a,b)$ such that $\vec{x}_a = \vec{y}_b$. All points $\vec{x}_a$ and $\vec{y}_b$ are *distinct* if we have $h \geq t$ (cf. Figure 1).

Assume $h \geq t$. Then, because all the points $\vec{x}_a$, $\vec{y}_b$ are distinct, there is no permutation $w \in \mathfrak{S}_t$ such that one of the segments $S_a^w = \langle x_{w(a)}, y_a \rangle$ is empty or equal to $\{\star\}$ for some index $a$. In particular we have $\mathfrak{S}'_t = \mathfrak{S}_t$.

Definition 2.7. To any point $\vec{v} \in \mathbb{Q}^2$ we associate the *invariant* $\rho(\vec{v}) := p_2(\vec{v}) \in \mathbb{Q}/\mathbb{Z}$ where $p_2 \colon \mathbb{Q}^2 \to \mathbb{Q}$ is projection on the second coordinate.

Remark. The horizontal distance between the point $\vec{x}_b$ and the point $\vec{y}_a$ is integral for all indices. Therefore the invariant of the first coordinate is not of interest. However, the vertical distance is the number $s_a^w = \frac{s}{n} n_a^w \in \mathbb{Q}$, which certainly need not be integral.

Using this invariant we define a particular permutation $w_0 \in \mathfrak{S}_t$:

Definition 2.8. Assume $h \geq t$ and *assume* that for each invariant $\rho \in \mathbb{Q}/\mathbb{Z}$ the number of indices $a$ such that the point $\vec{x}_a$ has invariant $\rho$ is equal to the number of indices $a$ such that the point $\vec{y}_a$ has invariant $\rho$. The element $w_0 \in \mathfrak{S}_t$ is the unique permutation such that for all indices $a, b$ we have

(2.10)
$$\Big(a < b \quad \text{and} \quad \rho(\vec{x}_a) = \rho(\vec{x}_b)\Big) \Longrightarrow \Big(w_0^{-1}(a) > w_0^{-1}(b) \quad \text{and} \quad \rho(\vec{y}_a) = \rho(\vec{y}_b) = \rho(\vec{x}_a)\Big).$$

Remark. Observe that the permutation $w_0$ depends on the integer $s$ because the heights of the points $\vec{x}_a$, $\vec{y}_a$, and therefore also their invariants depend on $s$.

Remark. If our assumption on the invariants $\rho(\vec{x}_a)$ and $\rho(\vec{y}_a)$ in Definition 2.8 is *not* satisfied, then the permutation $w_0$ cannot exist because it has to induce bijections between sets of different cardinality.

One could also define the permutation $w_0 \in \mathfrak{S}_t$ inductively: First the index $w_0^{-1}(t) \in \{1, 2, 3, \ldots, t\}$ is the minimal index $b$ such that the points $\vec{x}_t$ and $\vec{y}_b$ have the same invariant. Next, the index $w_0^{-1}(t-1) \in \{1, 2, 3, \ldots, t\}$ is the minimal index $b$, different from $w_0^{-1}(t)$, such that $\vec{x}_a$ and $\vec{y}_b$ have the same invariant. And so on: $w_0^{-1}(t-i) \in \{1, 2, 3, \ldots, t\}$ is the minimal index $b$ different from the previously chosen indices $w_0^{-1}(t)$, $w_0^{-1}(t-1)$, $\ldots$, $w_0^{-1}(t-i+1)$, such that the points $\vec{y}_b$ and $\vec{x}_{t-i}$ have the same invariant.

Lemma 2.9. *Let $\pi$ be a Speh representation with parameters $h, t$ with $h \geq t$. Let $d$ be the greatest common divisor of $n$ and $s$ and write $m$ for the quotient $\frac{n}{d}$. Define the points $\vec{x}_a := \ell(x_a)$ and $\vec{y}_a := \ell(y_a + 1)$. Let $d$ be the greatest common divisor of $n$ and $s$, and write $m$ for the number $\frac{n}{d} \in \mathbb{Z}$. The following two statements are equivalent:*

(*i*) *for each invariant* $\rho \in \mathbb{Q}/\mathbb{Z}$ *the number of indices* $a$ *such that the point* $\vec{x}_a$ *has invariant* $\rho$ *is equal to the number of indices* $a$ *such that the point* $\vec{y}_a$ *has invariant* $\rho$*;*

(*ii*) $m$ *divides* $t$ *or* $m$ *divides* $h$*.*

REMARK. The number $m$ is the order of the element $\frac{s}{n}$ in the torsion group $\mathbb{Q}/\mathbb{Z}$.

PROOF. We first claim that "$m|t \Rightarrow (i)$". We have

$$\rho(\vec{x}_{a+1}) = \rho(\vec{x}_a) - \tfrac{s}{n} \in \mathbb{Q}/\mathbb{Z} \tag{2.11}$$

and the same relation for the points $\vec{y}_a$. Therefore, if $m$ divides $t$, then the possible classes of the points $\vec{x}_a$ are equally distributed over the subset $\frac{s}{n}\mathbb{Z}/\mathbb{Z} \subset \mathbb{Q}/\mathbb{Z}$, and every invariant occurs precisely $\frac{t}{m}$ times. The same statement also holds for the points $\vec{y}_a$, and in particular $(i)$ is true. This proves the claim.

We now claim that "$m|h \Rightarrow (i)$". Assume $m|h$. Then the invariants of $\vec{x}_a$ and $\vec{y}_a$ are the same for all indices $a$. Thus $(i)$ is true.

We prove that "$(m \nmid t$ and $m \nmid h) \Rightarrow ((i)$ is false)". Assume $m \nmid t$ and $m \nmid h$. We first reduce to the case where $t < m$. Assume $t \geq m$. Consider the elements

$$\rho(\vec{x}_1), \rho(\vec{x}_2), \dots, \rho(\vec{x}_m), \quad \text{and} \quad \rho(\vec{y}_1), \rho(\vec{y}_2), \dots, \rho(\vec{y}_m) \in \mathbb{Q}/\mathbb{Z}. \tag{2.12}$$

By Equation (2.11) every possible class in $\frac{s}{n}\mathbb{Z}/\mathbb{Z}$ occurs precisely once in both lists. Thus, the truth value of $(i)$ is not affected if we remove the elements $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ and $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m$ from the respective lists. Renumber the indices and repeat this argument until $t < m$. Because we assumed that $t$ did not divide $m$ there remains a positive number of elements in the list $\vec{x}_a$ and $\vec{y}_a$. We renumber so that the indices range from 1 to $t$. Then we have reduced to the case where $1 \leq t < m$.

Now look at the two lists $\rho(\vec{x}_1), \rho(\vec{x}_2), \dots, \rho(\vec{x}_t)$ and $\rho(\vec{y}_1), \rho(\vec{y}_2), \dots, \rho(\vec{y}_t)$. In both lists every class in $\mathbb{Q}/\mathbb{Z}$ occurs *at most* once. We assumed that $m$ does not divide $h$, and therefore $\rho(\vec{x}_1) \neq \rho(\vec{y}_1)$. If there does not exist an index $b$ such that $\rho(\vec{x}_1) = \rho(\vec{y}_b)$, then (i) is false and we are done. Thus assume $\rho(\vec{x}_1) = \rho(\vec{y}_b)$ for some $1 < b \leq t$. By Equation (2.11) we then have $\rho(\vec{y}_{b-1}) = \rho(\vec{y}_b) + \frac{s}{n} = \rho(\vec{x}_1) + \frac{s}{n}$. The invariant $\xi := \rho(\vec{x}_1) + \frac{s}{n} \in \mathbb{Q}/\mathbb{Z}$ does not occur in the list $\rho(\vec{x}_1), \rho(\vec{x}_2), \dots, \rho(\vec{x}_t)$. Thus, we have found an invariant, namely $\xi$, occurring once in the list of invariants of the elements $\vec{y}_a$ and does not occur in the list of invariants of the elements $\vec{x}_a$. This contradicts $(i)$ and completes the proof. □

THEOREM 2.10. *Let* $\pi$ *be a Speh representation with parameters* $h, t$ *with* $h \geq t$*. Let* $d$ *be the greatest common divisor of* $n$ *and* $s$ *and write* $m$ *for the quotient* $\frac{n}{d}$*. Define the points* $\vec{x}_a := \ell(x_a)$ *and* $\vec{y}_a := \ell(y_a + 1)$*. The compact trace* $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$ *on* $\pi$ *is non-zero if and only if* $m$ *divides* $t$ *or* $m$ *divides* $h$*, and if the compact trace is non-zero, then it is equal to* $(-1)^{n-t}\mathrm{sign}(w_0) q^{\frac{s(n-s)}{2}\alpha} \mathrm{Dyck}_{\mathrm{s}}^{+}(\vec{x}^{w_0}, \vec{y})$*, where the permutation* $w_0 \in \mathfrak{S}_t$ *depends on* $s$ *and is defined in Definition 2.8.*

Remark. Perhaps one could extend this Theorem to obtain formulas for compact traces on Ladder representations as considered by Minguez and Lapid in [**71**].

Proof of Theorem 2.10. For a technical reason we assume that $0 < s < n$. In case $s = 0$ we have $f_{n\alpha s} = \mathbf{1}_{\mathrm{GL}_n(\mathcal{O}_F)}$. All the elements in $\mathrm{GL}_n(\mathcal{O}_F)$ are compact and therefore $\chi_c^G f_{n\alpha 0} = f_{n\alpha 0}$. The compact trace becomes the usual trace and the theorem is easy. A similar argument applies in case $s = n$. Thus we may indeed assume $0 < s < n$.

By Theorem 2.1 the compact trace $\mathrm{Tr}(\chi_c^G f, \pi)$ is equal to the combinatorial sum $\sum_{w \in \mathfrak{S}_t} \mathrm{sign}(w)\,\mathrm{Tr}(\chi_c^G f, I_w)$ for any Hecke operator $f \in \mathcal{H}(G)$. We apply it to the Kottwitz functions $f = f_{n\alpha s}$. We have $\mathfrak{S}'_t = \mathfrak{S}_t$ because $h \geq t$. Let $w \in \mathfrak{S}_t$. Recall that van Dijk's formula is also true for truncated traces Proposition 2.1.5, and thus for any $w \in \mathfrak{S}'_t$ the trace $\mathrm{Tr}(\chi_c^G f, I_w)$ equals $\mathrm{Tr}(\chi_c^G f^{(P_w)}, \Delta_w)$. Thus we have the formula

$$\mathrm{Tr}(\chi_c^G f, \pi) = \sum_{w \in \mathfrak{S}_t} \mathrm{sign}(w) \cdot \mathrm{Tr}(\chi_c^G f^{(P_w)}, \Delta_w). \tag{2.13}$$

By Lemma 2.6 we get for $f = f_{n\alpha s}$,

$$\mathrm{Tr}(\chi_c^G f, \pi) = q^{\frac{s(n-s)}{2}\alpha} \sum_{w \in \mathfrak{S}_t} \mathrm{sign}(w) \cdot \varepsilon_{P_0 \cap M_w} \cdot \mathrm{Dyck_s}(\vec{x}^w, \vec{y}). \tag{2.14}$$

We apply a standard combinatorial argument [5]. Put the *lexicographical order* $<$ on $\mathbb{Q}^2$:

$$\forall \vec{u}, \vec{v} \in \mathbb{Q}^2: \quad (\vec{u} < \vec{v}) \Longleftrightarrow (\vec{u}_1 < \vec{v}_1 \text{ or } (\vec{u}_1 = \vec{v}_1 \text{ and } \vec{u}_2 < \vec{v}_2)).$$

Let $(L_a)$ be a strict Dyck $t$-path from the points $\vec{x}^w$ to the points $\vec{y}$, and assume that $(L_a)$ has at least one point of crossing. Let $\vec{v} \in \mathbb{Q}^2$ be the point chosen among the points of crossing which is minimal for the lexicographical order on $\mathbb{Q}^2$. Let $(a, b)$ a couple of different indices, minimal for the lexicographical order on the set of all such couples, such that $\vec{v}$ lies on the path $L_a$ and also on the path $L_b$. We define a new path $L'_a$, defined by following the steps of $L_b$ until the point $\vec{v}$ and then following the steps of the path $L_a$. We define $L'_b$ by following $L_a$ until the point $\vec{v}$ and then continuing the path $L_b$. For the indices $c$ with $c \neq a, b$ we define $L'_c := L_c$. Observe that $(L'_a)$ is a $t$-path from the points $\vec{x}^{(ab)w}$ to the points $\vec{y}$. Furthermore, it is a Dyck path (with respect to this *new* configuration of points), and we have $\mathrm{weight}(L_a) = \mathrm{weight}(L'_a)$ because the weight is the product of the weights of the steps, and only the order of the steps has changed in the construction $(L_a) \mapsto (L'_a)$. The construction is self-inverse: If we apply the construction to the path $(L'_a)$ then we re-obtain $(L_a)$. Both paths $(L_a)$ and $(L'_a)$ occur in the sum of Equation (2.14). The sign $\varepsilon_{P_0 \cap M_w}$ is equal to $(-1)^{n-1}(-1)^{t-1}(-1)^{\#\{c \in \{1,2,\ldots,t\}\,|\,\vec{x}_{w(c)} = \vec{y}_c\}}$. By the assumption that $h \geq t$, the points in the list $\vec{x}$ are all different to the points in the list $\vec{y}$, and therefore the sign $\varepsilon_{P_0 \cap M_w}$ equals

5. The *Lindström-Gessel-Viennot Lemma.* The argument appears in many (almost) equivalent forms in the literature. We learned and essentially copied it from Stanley's book [**97**, Thm 2.7.1]. Note however that, strictly speaking, the Theorem 2.7.1 there does not apply as stated at this point in our argument. In the paragraph that follows we show that Stanley's argument may be adapted so that it does apply to our situation.

$(-1)^{n-t}$ (and does not depend on the permutation $w$). The sign of the permutation $w$ is opposite to the sign of $(ab)w$. Consequently, the contributions of the paths $(L_a)$ and $(L'_a)$ to Equation (2.14) cancel, and only the non-crossing paths remain in the sum. We find

$$\mathrm{Tr}(\chi_c^G f, \pi) = (-1)^{n-t} q^{\frac{s(n-s)}{2}} \sum_{w \in \mathfrak{S}_t} \mathrm{sign}(w) \cdot \mathrm{Dyck}_s^+(\vec{x}^w, \vec{y}). \tag{2.15}$$

We need a second notion of crossing paths, called *topological intersection*. Here we mean that, when the $t$-path $L$ is drawn in the plane $\mathbb{Q}^2$ there is a point $\vec{x} \in \mathbb{Q}^2$ lying on two paths $L_a$, $L_b$ occurring in $L$. Because we allow rational coordinates, topological intersection is not the same as intersection: It is easy to give an example of a 2-path, which, when drawn in the plane $\mathbb{Q}^2$ has one topological intersection point $\vec{x} \in \mathbb{Q}^2$ but the point $\vec{x}$ does not occur in the lists of points $\vec{v}_{1,0}, \vec{v}_{1,1}, \dots, \vec{v}_{1,r_1}$, $\vec{v}_{2,0}, \vec{v}_{2,1}, \dots, \vec{v}_{2,r_2}$ defining the 2-path. Such paths are considered non-crossing under our definition, even though they may have topological intersection points[6].

We claim that there is at most one permutation $w \in \mathfrak{S}_t$ such that the polynomial $\mathrm{Dyck}^+(\vec{x}^w, \vec{y})$ is non-zero, and that this permutation is the one we defined in Definition 2.8. Let $\mathfrak{S}''_t$ be the set of all permutations such that $\mathrm{Dyck}^+(\vec{x}^w, \vec{y}) \neq 0$, and assume that $\mathfrak{S}''_t$ contains an element $w \in \mathfrak{S}''_t$. We first make the following observation:

(Obs) To any point $\vec{v} \in \mathbb{Q}^2$ we associated the *invariant* $\rho(\vec{v}) := p_2(\vec{v}) \in \mathbb{Q}/\mathbb{Z}$. The horizontal distance between the point $\vec{x}_{w(a)}$ and the point $\vec{y}_a$ is the number $n_a^w$. The vertical distance is the number $s_a^w = \frac{s}{n} n_a^w \in \mathbb{Q}$. Because $w \in \mathfrak{S}''_t$ there exists a path from the point $\vec{x}_{w(a)}$ to the point $\vec{y}_a$. Consequently $s_a^w$ is integral. This implies that $\rho(\vec{x}_{w(a)}) = \rho(\vec{y}_a)$ for all indices $a$ and in particular the invariant of the point $\vec{x}_{w(a)}$ is independent of $w \in \mathfrak{S}''_t$.

We show inductively that $w$ is uniquely determined. We start with showing that the index $w^{-1}(t) \in \{1, 2, \dots, t\}$ is determined. We claim that $w^{-1}(t) \in \{1, 2, \dots, t\}$ is the minimal index such that the point $\vec{y}_{w^{-1}(t)}$ has the same invariant as $\vec{x}_t$. To see that this claim is true, suppose for a contradiction that it is false, i.e. assume the index $w^{-1}(t)$ is *not* minimal. Then there is an index $b$ strictly smaller than $w^{-1}(t)$ such that $\vec{y}_b$ has the same invariant as $\vec{x}_t$. By the observation (Obs) there exists an index $a \neq t$ such that $\vec{x}_a$ has the same invariant as $\vec{x}_t$ and such that $\vec{x}_a$ is connected to $\vec{y}_b$. Draw a picture (see Figure 2) to see that the paths $L_a$ and $L_t$ must intersect topologically. But, by construction, the invariants of $\vec{x}_a$ and $\vec{x}_t$ are the same. Therefore, any topological intersection point of the paths $L_a$ and $L_t$ is a point of crossing. Thus, the paths $L_a$ and $L_b$ are crossing. This is a contradiction, and therefore the claim is true. Thus the value $w^{-1}(t)$ is determined.

We now look at the index $t-1$. The point $\vec{x}_{t-1}$ is connected to the point $\vec{y}_{w^{-1}(t-1)}$. We claim that $w^{-1}(t-1) \in \{1, 2, \dots, t\}$ is the minimal index, different from $w^{-1}(t)$, such that

6. If one uses the wrong, topological notion of intersection, then the proof breaks at 8 lines below Equation (2.14): The constructed 'path' $(L'_c)$ is not a path.

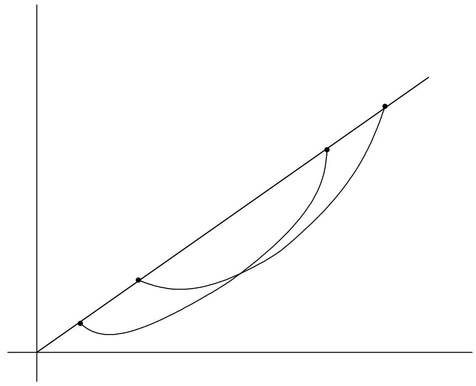

FIGURE 2. The leftmost point $\vec{x}_t$ is connected to the third point $\vec{y}_{w^{-1}(t)}$, and the second point $\vec{x}_a$ is connected to the last point $\vec{y}_b$. Any 2-path staying below the line $\ell$ must self-intersect topologically.

$\vec{y}_{w^{-1}(t-1)}$ has the same invariant as $\vec{x}_{t-1}$. The proof of this claim is the same as the one we explained for the index $t$. We may repeat the same argument for the remaining indices $t-2$, $t-3$, etc. Consequently $w$ is uniquely determined by its properties, and equal to the permutation $w_0$ defined in Definition 2.8.

We proved that if the set $\mathfrak{S}_t''$ is non-empty, then it contains precisely one element, and this element is equal to $w_0$. Therefore, if the compact trace does not vanish, then $m$ must divide $t$ or $m$ divides $h$ by Lemma 2.9. Inversely, assume that $m$ divides $t$ or $m$ divides $h$. The permutation $w_0 \in \mathfrak{S}_t$ exists by Lemma 2.9. We claim that $\mathrm{Dyck}_{\mathrm{s}}^{+}(\vec{x}^{w_0}, \vec{y}) \neq 0$, so that $w_0 \in \mathfrak{S}_t''$. To prove this, it suffices to construct one non-crossing $t$-path from the points $\vec{x}^{w_0}$ to the points $\vec{y}$. This is easy (see Figure 3): Let $a$ be an index, and write $n_a^{w_0}$ for the horizontal distance between $\vec{x}_a^{w_0}$ and $\vec{y}_a$ and $s_a^{w_0}$ for the vertical distance. The path $L_a$ from $\vec{x}_a^{w_0}$ to $\vec{y}_a$ is defined to be the path taking $n_a^{w_0} - s_a^{w_0}$ horizontal eastward steps, and then $s_a^{w_0}$ diagonal northeastward steps. Then $(L_a)$ is a strict non-crossing $t$-path and therefore $\mathrm{Dyck}_{\mathrm{s}}^{+}(\vec{x}^{w_0}, \vec{y})$ is non-zero. This completes the proof. $\square$

## 3. A dual formula

The argument for Theorem 2.10 extends to the case where $h \leq t$. This computation more complicated, because the permutation $w \in \mathfrak{S}_t$ that contributes to Equation (2.15) is no longer unique and the signs $\varepsilon_{P_0 \cap M_w}$ in Equation (2.14) depend on the contributing permutations $w$ (these signs are independent of $w$ only in case $h \geq t$). We don't reproduce the computation here, because there is a more elegant approach using the duality of Zelevinsky.

The Zelevinsky dual of a Speh representation with parameters $(h, t)$ is a Speh representation with the role of the parameters inversed, thus of type $(t, h)$. Furthermore, taking the Zelevinsky dual of the formula of Tadic yields a new character formula, now in terms of duals

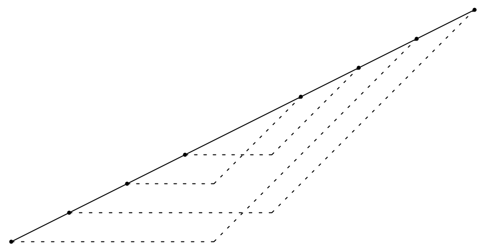

FIGURE 3. An example of a non-crossing 4-path $(L_a)$ in case $\frac{s}{n} = \frac{1}{2} \in \mathbb{Q}/\mathbb{Z}$ and $t = 4$. For each $a$, the path $L_a$ first takes $n_a^{w_0} - s_a^{w_0}$ horizontal steps and then $s_a^{w_0}$ vertical steps. Note that paths with the same invariant do not intersect.

of standard representations. Of course, the Zelevinsky dual of a standard representation is not standard, rather it is an unramified twist of products in $\mathcal{R}$ of one dimensional representations. Therefore, we compute first the compact trace on the one dimensional representations, then use van Dijk's theorem, Proposition 2.1.5, to obtain formulas for products in $\mathcal{R}$ of one dimensional representations, and finally use the dual of Tadic's formula to compute the compact traces on Speh representations with $h \le t$ (opposite inequality to Theorem 2.10). We will then have computed the formula for all Speh representations. This approach seems longer but that is not true: The individual steps we take also appear in an equivalent form in our original computation.

**3.1. The trivial representation.** We compute the compact traces of spherical Hecke operators acting on the trivial representation of $G$. We recall some definitions on roots and convexes from [**73**, §1] and [**68**, Chap. 1].

Let $P$ be a standard parabolic subgroup of $G$. Let $A_P$ be the center of $P$. We write $\varepsilon_P = (-1)^{\dim(A_P/A_G)}$. We define $\mathfrak{a}_P := X_*(A_P) \otimes \mathbb{R}$. If $P \subset P'$ then we have $A_{P'} \subset A_P$ and thus an induced map $\mathfrak{a}_{P'} \to \mathfrak{a}_P$. We write $T = A_{P_0}$. We define $\mathfrak{a}_P^{P'}$ to be the quotient of $\mathfrak{a}_P$ by $\mathfrak{a}_{P'}$. We write $\mathfrak{a}_0 = \mathfrak{a}_{P_0}$ and $\mathfrak{a}_0^G = \mathfrak{a}_{P_0}^G$.

We write $\Delta$ for the set of simple roots of $T$ occurring in the Lie algebra of $N_0$. For each root $\alpha$ in $\Delta$ we have a coroot $\alpha^\vee$ in $\mathfrak{a}_0^G$. We write $\Delta_P \subset \Delta$ for the subset of $\alpha \in \Delta$ acting non-trivially on $A_P$. For $\alpha \in \Delta_P \subset \Delta$ we send the coroot $\alpha^\vee \in \mathfrak{a}_0^G$ to the space $\mathfrak{a}_P^G$ via the canonical surjection $\mathfrak{a}_0^G \twoheadrightarrow \mathfrak{a}_P^G$. The set of these restricted coroots $\alpha^\vee|_{\mathfrak{a}_P^G}$ with $\alpha$ ranging over $\Delta_P$ form a basis of the vector space $\mathfrak{a}_P^G$. By definition the set of fundamental weights $\{\varpi_\alpha^G \in \mathfrak{a}_P^{G*} \mid \alpha \in \Delta_P\}$ is the basis of $\mathfrak{a}_P^{G*} = \mathrm{Hom}(\mathfrak{a}_P^G, \mathbb{R})$ dual to the basis $\{\alpha^\vee_{\mathfrak{a}_P^G}\}$ of coroots. Recall that we have the *acute* and *obtuse* Weyl chambers of $G$. The *acute chamber* $\mathfrak{a}_P^{G+}$ is the set of $x \in \mathfrak{a}_P^G$ such that $\langle \alpha, x\rangle > 0$ for all roots $\alpha \in \Delta_P$. The *obtuse chamber* ${}^+\mathfrak{a}_P^G$ is the set of $x \in \mathfrak{a}_P^G$ such that we have the inequality $\langle \varpi_\alpha^G, x\rangle > 0$ for all fundamental weights $\varpi_\alpha^G$, associated to $\alpha \in \Delta_P$. We need another chamber, defined by ${}^{\le}\mathfrak{a}_P^G = \left\{x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \varpi_\alpha^G, x\rangle \le 0\right\}$. We call this chamber the *closed opposite obtuse Weyl chamber*. Let ${}^{\le}\widehat{\tau}_P^G$ be the characteristic function on

$\mathfrak{a}_P$ of this chamber. Let $H_M\colon M \to \mathfrak{a}_P$ be the *Harish-Chandra* mapping, normalized such that $|\chi(m)|_p = q^{-\langle \chi, H_M(m)\rangle}$ for all rational characters $\chi$ of $M$. We define the function $\xi_c^G$ on $M_0 = T$ to be the composition ${}^{\leq}\widehat{\tau}_{P_0}^G \circ (\mathfrak{a}_{P_0} \twoheadrightarrow \mathfrak{a}_{P_0}^G) \circ H_{M_0}$.

If $f \in \mathcal{H}_0(G)$ is a function whose Satake transform is the function $h \in A$, then we often abuse notation, and write $\xi_c^G h$ for the Satake transform of the function $\xi_c^G f^{(P_0)}$, and similarly for the functions $\chi_N f$ and $\widehat{\chi}_N f$ if $f \in \mathcal{H}_0(M)$.

The following Proposition and proof are valid for any split reductive group $G$ over a non-Archimedean local field.

Proposition 3.1. *Let $f$ be a function in the Hecke algebra $\mathcal{H}_0(G)$. The compact trace $\mathrm{Tr}(\chi_c^G f, \mathbf{1}_G)$ is equal to $\mathrm{Tr}\left(\xi_c^G f^{(P_0)}, \mathbf{1}_T(\delta_{P_0}^{-1/2})\right)$.*

Proof. For comfort we prove the proposition under the additional assumption that $G$ is its own derived group. We have

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathbf{1}) = \sum_{P=MN} \varepsilon_P \,\mathrm{Tr}(\widehat{\chi}_N f^{(P)}, \mathbf{1}(\delta_P^{-1/2})).$$

Recall that we have the notation $\varphi_{M,\rho} \in \widehat{M}$ for the Hecke matrix of a representation $\rho$ of $M$. The Hecke matrix $\varphi_{M,\delta_P^{-1/2}}$ is conjugate in $\widehat{M}$ to the Hecke matrix $\varphi_{T,\delta_P^{-1/2}\delta_{P_0\cap M}^{-1/2}} = \delta_{T,\delta_{P_0}^{-1/2}} \in \widehat{T} \subset \widehat{M}$. Recall that the Satake transform is defined by the composition of the morphism $f \mapsto f^{(P_0)}$ with the obvious isomorphism $\mathcal{H}_0(T) \cong \mathbb{C}[X_*(T)]$ (the Satake transformation for $T$). Therefore

$$\mathrm{Tr}(\widehat{\chi}_N f^{(P)}, \mathbf{1}(\delta_P^{-1/2})) = \mathcal{S}(\widehat{\chi}_N f^{(P_0)})(\varphi_{T,\delta_{P_0}^{-1/2}}).$$

Using linearity of the Satake transform we obtain

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathbf{1}) = \mathcal{S}\left(\sum_{P=MN} \varepsilon_P \widehat{\chi}_N f^{(P_0)}\right)(\varphi_{T,\delta_{P_0}^{-1/2}}).$$

Thus we have to compute the function $\sum_{P=MN} \varepsilon_P \widehat{\chi}_N$ on the group $T$. By definition we have

$$\widehat{\chi}_N = \widehat{\tau}_P^G \circ H_M.$$

Let $W_M$ be the rational Weyl group of $T$ in $M$. Let $t \in T$. Then

$$H_M(t) = \frac{1}{\# W_M} \sum_{w \in W_M} w H_T(t).$$

Thus $\widehat{\chi}_N(t) = 1$ if and only if

$$\forall \alpha \in \Delta_P : \quad \sum_{w \in W_M} \langle \varpi_\alpha^G, w H_T(t)\rangle > 0.$$

We have for all $\alpha \in \Delta_P$ the inequality $\langle \varpi_\alpha^G, H_T(t)\rangle > 0$ if and only if we have $\langle \varpi_\alpha^G, w H_T(t)\rangle > 0$ for all $w \in W_M$. Therefore, we have on the group $T$

$$\widehat{\chi}_N = \widehat{\tau}_P^G \circ H_T.$$

Thus

$$\sum_{P=MN} \varepsilon_P \widehat{\chi}_N = \left( \sum_{P=MN} \varepsilon_P \widehat{\tau}_P^G \right) \circ H_T.$$

By inclusion-exclusion we have

$$\sum_{P=MN} \varepsilon_P \widehat{\tau}_P^G = {}^{\leq}\widehat{\tau}_P^G.$$

This proves the proposition in case $G = G_{\rm der}$. It is easy to deduce the statement from the case $G = G_{\rm der}$. □

REMARK. Consider the space $I$ of locally constant functions from $G/P_0$ to $\mathbb{C}$, and equip $I$ with the $G$-action through right translations. Then, with an argument similar to the one above, one may compute the compact traces on the irreducible subquotients $V$ of $C$. Recall from Borel and Wallach [**10**] that these representations are all mutually non-isomorphic and occur with multiplicity one in $I$. Borel and Wallach describe the representations $V$ precisely; they are indexed by the standard parabolic subgroups of $G$.

**3.2. The dual formula.** In this subsection we prove the dual version of Theorem 2.10.

LEMMA 3.2. *Let $T_1 = \langle u_1, v_1 \rangle$, $T_2 = \langle u_2, v_2 \rangle$, $\dots$, $T_h = \langle u_h, v_h \rangle$ be a list of segments and consider the representation $J := (\Delta T_1)^\iota \times (\Delta T_2)^\iota \times \cdots \times (\Delta T_h)^\iota$. Then $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$ is equal to $q^{s(n-s)/2}\mathrm{Dyck}(\vec{u}, \vec{v})$, where $\vec{u}_a = \ell(u_a)$ and $\vec{v}_a = \ell(v_a + 1)$ for $a = 1, 2, \dots, t$.*

REMARK. Recall that for the compact trace on the Steinberg representation, $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \mathrm{St}_G)$ we had the sign $\varepsilon_{P_0}$ multiplied with a *strict* Dyck polynomial. In case $n$ and $s$ are coprime, then any Dyck polynomial from the point $\ell(\frac{1-n}{2})$ to the point $\ell(\frac{n-1}{2} + 1)$ is strict; consequently the trace on Steinberg and trivial representation differ only by the sign $\varepsilon_{P_0}$.

PROOF. The proof is the same as the proof for Lemma 2.3, replacing the result in Equation (2.6) with the result from Proposition 3.1. However, we repeat the argument for verification purposes (one has to be careful with the signs).

Assume first that $h = 1$ and that $\pi$ is the trivial representation of $G$. In the previous subsection we proved that

$$\mathrm{Tr}\left(\chi_c^G f_{n\alpha s}, \pi\right) = \mathrm{Tr}\left(\xi_c^G f_{n\alpha s}^{(P_0)}, \mathbf{1}_T(\delta_{P_0}^{-1/2})\right).$$

To a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n} \in \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \dots X_n^{\pm 1}]$ with $e_i \in \mathbb{Z}$ and $\sum_{i=1}^n e_i = s$ we associate the graph $\mathcal{G}_X$ with points

$$\vec{v}_0 := \ell(\tfrac{1-n}{2}), \quad \vec{v}_i := \vec{v}_0 + (i, e_1 + e_2 + \ldots + e_i) \in \mathbb{Q}^2, \tag{3.1}$$

for $i = 1, 2, \dots, n$. We have $\xi_c^G X = X$ if and only if

$$e_1 + e_2 + \cdots + e_i \leq \tfrac{s}{n} i, \tag{3.2}$$

for all indices $i < n$, and $\xi_c^G X = 0$ otherwise. The evaluation of $X$ at the point

$$\left(q^{\frac{n-1}{2}}, q^{\frac{n-3}{2}}, \ldots, q^{\frac{1-n}{2}}\right) \tag{3.3}$$

equals the weight[7] of the graph $\mathcal{G}_X$.

The trace of $f_{n\alpha s}$ against the representation $\mathbf{1}_T(\delta_{P_0}^{-1/2})$ is equal to the evaluation of $f_{n\alpha s}$ at the point in Equation (3.3) (use Lemma 2.2 but notice that the signs are different). The monomials $X$ occurring $\mathcal{S}(f_{n\alpha s})$ yield paths from the point $\ell(\frac{1-n}{2}) \in \mathbb{Q}^2$ to the point $\ell(\frac{n-1}{2}+1)$. The condition in Equation (3.2) is true if and only if the graph $\mathcal{G}_X$ lies (non-strictly) below the line $\ell$. Therefore we have

$$\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \mathbf{1}_G) = q^{\frac{s(n-s)}{2}} \mathrm{Dyck}(\ell(\tfrac{1-n}{2}), \ell(\tfrac{n-1}{2}+1)).$$

By twisting with the character $\nu^{-x+\frac{1-n}{2}}$ as we did in Lemma 2.5 we find

$$\mathrm{Tr}\left(\chi_c^G f, (\Delta\langle u, v\rangle)^\iota\right) = q^{\frac{s(n-s)}{2}} \mathrm{Dyck}(\ell(x), \ell(y+1)),$$

for all segments $\langle u, v\rangle$. Finally the argument in Lemma 2.6 may be repeated to find the compact traces on duals of standard representations as stated in the Lemma. □

THEOREM 3.3. *Let $\pi$ be a Speh representation with parameters $h, t$ with $h \le t$. Let $d$ be the greatest common divisor of $n$ and $s$ and write $m$ for the quotient $\frac{n}{d}$. Let $T_a = \langle u_a, v_a\rangle$ be the segments of $\pi^\iota$. Define the points $\vec{u}_a := \ell(u_a)$ and $\vec{v}_a := \ell(v_a + 1)$. The compact trace $\mathrm{Tr}(\chi_c^G f_{n\alpha s}, \pi)$ is non-zero if and only if $m$ divides $h$ or $m$ divides $t$. Assume that the compact trace is non-zero, then it is equal to $\mathrm{sign}(w_0) q^{\frac{s(n-s)}{2}\alpha} \mathrm{Dyck}^+(\vec{u}^{w_0}, \vec{v})$, where the permutation $w_0 \in \mathfrak{S}_h$ is defined in Definition 2.8.*

PROOF. Let $\pi^\iota$ be the representation dual to the representation $\pi$. After dualizing the formula of Tadic for $\pi^\iota$ we obtain an expression of the form

$$\pi = \sum_{w \in \mathfrak{S}_h} \mathrm{sign}(w) I_w^\iota. \tag{3.4}$$

The involution $\iota$ on $\mathcal{R}$ commutes with products. Therefore, if $T_1, \ldots, T_h$ are the Zelevinsky segments of the dual representation $\pi^\iota$, then $I_w^\iota$ is equal to $(\Delta T_1)^\iota \times (\Delta T_2)^\iota \times \cdots (\Delta T_k)^\iota$. By Lemma 3.2 we obtain

$$\mathrm{Tr}(\chi_c^G f_{n\alpha s}, I_w^\iota) = q^{s(n-s)/2} \mathrm{Dyck}(\vec{u}^w, \vec{y}).$$

A crucial remark is that the points $\vec{u}$ and $\vec{v}$ are all different because we assume that $h \le t$. Therefore one may repeat the argument in the proof of Theorem 2.10 using the dual formula in Equation (3.4); one only has to interchange $t$ with $h$ and every occurrence of the word "strict Dyck $t$-path" with "Dyck $h$-path", as the paths that describe the compact traces on (products in $\mathcal{R}$ of) trivial representations are not necessary strict. □

7. Equation (3.3) differs from Equation (2.4) by a sign in the exponents. However, observe also that the graph in Equation (3.1) is traced in the direction opposite to the graph in Equation (2.2).

## 4. Return to Shimura varieties

In Chapter 2 we proved a formula for the basic stratum of certain Shimura varieties associated to unitary groups, subject to a technical condition on the Newton polygon of the basic stratum (that it has no non-trivial integral points). In the previous sections we have completely resolved the combinatorial issues that arise if you remove this condition in case $p$ is totally split in the center of the division algebra. We may now essentially repeat the argument from Chapter 2 to obtain the description of the cohomology if there is no condition on the Newton polygon of the basic stratum. A large part of the argument remains the same, that part will only be sketched and we refer to Chapter 2 for the details.

**4.1. Notations and assumptions.** Let $\mathrm{Sh}_K/\mathcal{O}_E\otimes\mathbb{Z}_{(p)}$ be a Kottwitz variety [**58**]. Here we have fixed the following long list of notations and assumptions:

(1) Let $D$ be a division algebra over $\mathbb{Q}$;

(2) $F$ is the center of $D$, assume $F$ is a CM field of the form $F = \mathcal{K}F^+ \subset \overline{\mathbb{Q}}$, where $F^+$ is totally real, and $\mathcal{K}/\mathbb{Q}$ is quadratic imaginary;

(3) $*$ is an anti-involution on $D$ inducing complex conjugation on $F$;

(4) $n \in \mathbb{Z}_{\geq 0}$ is such that $\dim_F(D) = n^2$;

(5) $G$ is the $\mathbb{Q}$-group with $G(R) = \left\{x \in D_R^\times | g^*g \in R^\times\right\}$ for every commutative $\mathbb{Q}$-algebra $R$;

(6) $h$ is an algebra morphism $h\colon \mathbb{C} \to D_\mathbb{R}$ such that $h(z)^* = h(\overline{z})$ for all $z \in \mathbb{C}$;

(7) the involution $x \mapsto h(i)^{-1}x^*h(i)$ on $D_\mathbb{R}$ is positive;

(8) $X$ is the $G(\mathbb{R})$ conjugacy class of the restriction of $h$ to $\mathbb{C}^\times \subset \mathbb{C}$;

(9) $\mu \in X_*(G)$ is the restriction of $h\otimes\mathbb{C}\colon \mathbb{C}^\times\times\mathbb{C}^\times \to G(\mathbb{C})$ to the factor $\mathbb{C}^\times$ of $\mathbb{C}^\times\times\mathbb{C}^\times$ indexed by the identity isomorphism $\mathbb{C} \xrightarrow{\sim} \mathbb{C}$;

(10) $E \subset \overline{\mathbb{Q}}$ is the reflex field of this Shimura datum $(G, X, h^{-1})$;

(11) $\xi$ is an (any) irreducible algebraic representation over $\overline{\mathbb{Q}}$ of $G_{\overline{\mathbb{Q}}}$;

(12) Let $f_\infty$ be a function at infinity having its stable orbital integrals prescribed by the identities of Kottwitz in [**57**]; it can be taken to be (essentially) an Euler-poincaré function [**58**, Lemma 3.2] (cf. [**27**]). The function has the following property: Let $\pi_\infty$ be an $(\mathfrak{g}, K_\infty)$-module occurring as the component at infinity of an automorphic representation $\pi$ of $G$. Then the trace of $f_\infty$ against $\pi_\infty$ is equal to the Euler-Poincaré characteristic $\sum_{i=0}^\infty N_\infty(-1)^i \dim \mathrm{H}^i(\mathfrak{g}, K_\infty; \pi_\infty\otimes\xi)$, where $N_\infty$ is a certain explicit constant (cf. [**58**, p. 657, Lemma 3.2]).

(13) $p$ is a prime number where $\mathrm{Sh}_K$ has *good reduction* [**59**, § 5], and we assume that $p$ is *split* in $\mathcal{K}/\mathbb{Q}$;

(14) $K \subset G(\mathbb{A}_{\mathrm{f}})$ is a compact open subgroup, small enough that $\mathrm{Sh}_K/\mathcal{O}_E \otimes \mathbb{Z}_p$ is smooth and such that $K$ decomposes as $K^pK_p$ where $K^p$ is a compact open subgroup of $G(\mathbb{A}_{\mathrm{f}}^p)$ and $K_p$ is a hyperspecial compact open subgroup of $G(\mathbb{Q}_p)$.

(15) $\nu_p \colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ is a fixed embedding, $\nu_\infty \colon \overline{\mathbb{Q}} \to \mathbb{C}$ is another fixed embedding, the fields $F, F^+, E, \mathcal{K}$ are all embedded into $\mathbb{C}$;

(16) $\mathfrak{p}$ is the $E$-prime induced by $\nu_p$;

(17) $\mathbb{F}_q$ is the residue field of $E$ at the prime $\mathfrak{p}$ and $\overline{\mathbb{F}}_q$ is the residue field of $\overline{\mathbb{Q}}$ at $\nu_p$; for every positive integer $\alpha$, $E_{\mathfrak{p},\alpha} \subset \overline{\mathbb{Q}}_p$ is the unramified extension of $E_{\mathfrak{p}}$ of degree $\alpha$; $\mathbb{F}_{q^\alpha}$ is the residue field of $E_{\mathfrak{p},\alpha}$;

(18) $\iota \colon B \hookrightarrow \mathrm{Sh}_{K,\mathbb{F}_q}$ is the basic stratum [**87**] (cf. [**39, 59, 60, 88**]);

(19) $\chi_c^G$ is the characteristic function on $G(\mathbb{Q}_p)$ of the subset of compact elements (cf. [**22**]);

(20) $\ell$ is a prime number and $\overline{\mathbb{Q}}_\ell$ an algebraic closure of $\mathbb{Q}_\ell$ together with an embedding $\overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_\ell$;

(21) $\mathcal{L}$ is the $\ell$-adic local system on $\mathrm{Sh}_K/\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$ associated to the representation $\xi \otimes \overline{\mathbb{Q}}_\ell$ of $G_{\overline{\mathbb{Q}}_\ell}$ [**59**, p. 393];

(22) $U \subset G$ is the subgroup of elements with trivial factor of similitudes;

(23) for each infinite $F^+$-place $v$, the number $s_v$ is the unique integer $0 \leq s_v \leq \frac{1}{2}n$ such that $U(\mathbb{R}) \cong \prod_v U(s_v, n - s_v)$;

(24) the embedding $\overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_p$ induces an action of the group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ on the set of infinite $F^+$-places. For each $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-orbit $\wp$ we define the number $s_\wp := \sum_{v \in \wp} s_v$, and we write $\sigma_\wp$ for the partition $(s_v)_{v \in \wp}$ of the number $s_\wp$;

(25) the function $f_\alpha$ is the function of Kottwitz [**57**] associated to $\mu$ (cf. Proposition 2.3.3).

Remark. The second condition (2) is particular for our arguments, and does not occur in [**58**].

**4.2. The main argument.** We compute the factors $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p)$ occurring in Theorem 4.3 below. We need to introduce two classes of representations:

Definition 4.1. Consider the general linear group $G_n$ over a non-Archimedean local field. Then a representation $\pi$ of $G_n$ is called a (semistable) *rigid representation* if it is equal to a product of the form

$$\prod_{a=1}^{k} \mathrm{Speh}(x_a, y)(\varepsilon_a) \in \mathcal{R},$$

where $y$ is a divisor of $n$ and $(x_a)$ is a composition of $\frac{n}{y}$, and $\varepsilon_a$ are unramified unitary characters.

DEFINITION 4.2. A representation $\pi$ of the group $G(\mathbb{Q}_p) = \mathbb{Q}_p{}^\times \times \prod_{\wp|p} \mathrm{GL}_n(F_\wp^+)$ is called a *rigid* representation if for each $F^+$-place $\wp$ above $p$ the component $\pi_\wp$ is a (semistable) rigid representation of $\mathrm{GL}_n(F_\wp^+)$ in the previous sense:

$$\pi_\wp = \prod_{a=1}^{k} \mathrm{Speh}(x_{\wp,a}, y_\wp)(\varepsilon_{\wp,a}) \in \mathcal{R},$$

where two additional conditions hold: (1) $y_\wp = y_{\wp'}$ for all $\wp, \wp'|p$, and (2) the factor of similitudes $\mathbb{Q}_p{}^\times$ of $G(\mathbb{Q}_p)$ acts through an unramified character on the space of $\pi$. We write $y := y_\wp$ and call the set of data $(x_{\wp,a}, \varepsilon_{\wp,a}, y)$ the *parameters* of $\pi$.

REMARK. Recall that we work in the semistable setting, both notions of rigid representations that we introduced above in the semistable setting also have a natural variant in the non-semistable case.

THEOREM 4.3. *Let $\alpha$ be a positive integer. Assume the conditions (1)-(25) from §5.1. Then*

$$\sum_{i=0}^{\infty} (-1)^i \,\mathrm{Tr}(f^{\infty p} \times \Phi_{\mathfrak{p}}^\alpha, \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \iota^*\mathcal{L})) = \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p \text{ rigid}}} \mathrm{Tr}(\chi_c^G f_\alpha, \pi_p) \cdot \mathrm{Tr}(f^p, \pi^p). \tag{4.1}$$

REMARK. Using recent results obtained with Lapid (see Appendix B) it is possible to extend the above theorem to the other Newton strata. However the result will be combinatorially complicated. We hope to include this result soon.

PROOF OF THEOREM 4.3. Write $T(f^p, \alpha)$ for the left hand side of Equation (4.1). By Proposition 3.4 of Chapter 2 we have

$$T(f^p, \alpha) = \mathrm{Tr}(\chi_c^G f_\infty f_\alpha f^p, \mathcal{A}(G)), \tag{4.2}$$

for all sufficiently large integers $\alpha$. To simplify notations, we write $f := f_\infty f_\alpha f^p$.

Let $\pi \subset \mathcal{A}(G)$ be an automorphic representation of $G$ contributing to the trace $\mathrm{Tr}(\chi_c^G f, \mathcal{A}(G))$. During the proof of Proposition 2.3.4 of the previous chapter we explained that $\pi$ may be base changed to an automorphic representation $BC(\pi)$ of the algebraic group $\mathcal{K}^\times \times D^\times$, and that, in turn, $BC(\pi)$ may be send to an automorphic representation $\Pi := JL(BC(\pi))$ of the $\mathbb{Q}$-group $G^+ = \mathcal{K}^\times \times \mathrm{GL}_n(F)$.

The representation $\Pi$ is a *discrete* automorphic representation of the group $G^+(\mathbb{A})$, and $\Pi$ is semistable at $p$. The classification of Moeglin-Waldspurger implies that $\pi_\wp$ is the irreducible quotient of the induced representation

$$\mathrm{Ind}_{P(\mathbb{A}_F)}^{\mathrm{GL}_n(\mathbb{A}_F)} \left( \omega|\cdot|^{\frac{y-1}{2}}, \dots, \omega|\cdot|^{\frac{1-y}{2}} \right),$$

where $P \subset \mathrm{GL}_n$ is the homogeneous standard parabolic subgroup having $y$ blocks, and each block is of size $n/y$; the inducing representation $\omega$ is a cuspidal automorphic representation of $\mathrm{GL}_{n/y}(\mathbb{A}_F)$.

The representation $\Pi$ comes from an automorphic representation of the group $G$ via Jacquet Langlands and base change. Therefore, $\Pi$ is cohomological and conjugate self dual. These properties descend, up to twist by a character, to the representation $\omega$. The Ramanujan conjecture is proved to be true for the representation $\omega$ by the articles [**14**, **25**, **95**]. Thus the components $\omega_v$ of $\omega$ are *tempered* representations. Note that, of course, the components $\Pi_v$ are *not* tempered if $\Pi$ is not cuspidal.

An easy computation shows that $\pi_\wp$ is a rigid representation for all $F^+$-places $\wp$ dividing $p$ (Theorem 2.2.1). This means that there exists a positive divisor $y$ of $n$, a composition $\frac{n}{y} = \sum_{a=1}^k x_a$, and unramified unitary characters $\varepsilon_a$ such that

$$\pi_\wp \cong \mathrm{Ind}_{P(F_\wp^+)}^{\mathrm{GL}_n(F_\wp^+)} \bigotimes_{a=1}^r \mathrm{Speh}(x_a, y)(\varepsilon_a), \tag{4.3}$$

where $P \subset \mathrm{GL}_n$ is the standard parabolic subgroup corresponding to the composition $(x_a y)$ of $n$, and the tensor product is along the blocks of the standard Levi factor $M$ of $P$. In Equation (4.3) the number $y$ is of *global* nature and does not depend on $\wp$. The other data, $k$, $(x_a)$ and $\varepsilon_a$ do depend on the place $\wp$. □

We work under the condition that $p$ is split in the center $F$ of the algebra $D$. Because the prime $p$ is completely split in the extension $F/\mathbb{Q}$ we have by Proposition 2.3.3 that

$$f_\alpha = \mathbf{1}_{q^{-\alpha}} \otimes \bigotimes_{v \in \mathrm{Hom}(F^+, \mathbb{R})} f_{n\alpha s_v}^{\mathrm{GL}_n(\mathbb{Q}_p)} \in \mathcal{H}_0(G(\mathbb{Q}_p)),$$

where the numbers $s_v$ are the signatures of the unitary group (cf. subsection 1). We compute

$$\begin{aligned}
&\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \\
&= \prod_{v \in \mathrm{Hom}(F^+, \mathbb{R})} \mathrm{Tr}\left( \chi_c^{\mathrm{GL}_n(\mathbb{Q}_p)} f_{n\alpha s_v}, \mathrm{Ind}_{P(\mathbb{Q}_p)}^{\mathrm{GL}_n(\mathbb{Q}_p)} \bigotimes_{a=1}^r \mathrm{Speh}(x_v, y)(\varepsilon_{v,a}) \right) \\
&= \left( \prod_{v \in \mathrm{Hom}(F^+, \mathbb{R})} \prod_{a=1}^r \varepsilon_{v,a}(q^{-s_v \frac{y \cdot x_a}{n} \alpha}) \right) \cdot \\
&\quad \cdot \prod_{v \in \mathrm{Hom}(F^+, \mathbb{R})} \mathrm{Tr}\left( \chi_c^{\mathrm{GL}_n(\mathbb{Q}_p)} f_{n\alpha s_v}, \mathrm{Ind}_{P(\mathbb{Q}_p)}^{\mathrm{GL}_n(\mathbb{Q}_p)} \bigotimes_{a=1}^r \mathrm{Speh}(x_{v,a}, y) \right).
\end{aligned}$$

Write $\zeta_\pi^\alpha \in \mathbb{C}$ for the product $\prod_v \prod_a \varepsilon_a(q^{-s_v \frac{y \cdot x_a}{n} \alpha})$. The polynomial

$$\mathrm{Tr}\left( \chi_c^{\mathrm{GL}_n(F_\wp^+)} f_{n\alpha\sigma_\wp}, \mathrm{Ind}_{P(F_\wp^+)}^{\mathrm{GL}_n(F_\wp^+)} \bigotimes_{a=1}^r \mathrm{Speh}(x_a, y) \right) \in \mathbb{C}[q^\alpha], \tag{4.4}$$

is computed in Theorems 2.10 and 3.3 to be a polynomial defined by the weights of certain non-intersecting lattice paths. In particular the trace in Equation (4.4) vanishes unless the number

$$m_{v,a} := \frac{y \cdot x_{\wp,a}}{\gcd\left(y \cdot x_{\wp,a}, \frac{y \cdot x_{\wp,a}}{n} s_\wp\right)} = \frac{y \cdot x_{\wp,a}}{\gcd(n, s_\wp)} \tag{4.5}$$

is an integer, and divides either $x_{\wp,a}$ or $y$. We make the *assumption* that the compact trace $\mathrm{Tr}(\chi_c^G f_\alpha, \pi_p)$ is non-zero and therefore these divisibility relations are satisfied.

The number $\zeta_\pi^\alpha \in \mathbb{C}$ is determined by the central character $\omega_\pi \colon Z(\mathbb{A}) \to \mathbb{C}^\times$ of the automorphic representation $\pi$ via the Equation:

$$\omega_\pi(x)^{\alpha s/n} = \varepsilon_s(q^\alpha) \cdot \prod_{\wp|p} \prod_{a=1}^{r} \varepsilon_{\wp,a}\left(q^{-s_\wp \frac{y \cdot x_a}{n} \alpha}\right) = \zeta_\pi^\alpha, \tag{4.6}$$

where $\varepsilon_s$ is the contribution of the factor of similitudes, and $x \in Z(\mathbb{A})$ is the following element of the center $Z$ of $G$:

$$x := (1) \times [q, (q_\wp)_\wp] \in Z(\mathbb{A}^p) \times \left[\mathbb{Q}_p{}^\times \times F_{\mathbb{Q}_p}^{+\times}\right] = Z(\mathbb{A}).$$

The divisibility relations in Equation (4.5) assure that taking the rational power $s/n$ of $\omega_\pi(x)$ on the left hand side makes sense.

REMARK. The number $\zeta_\pi$ is a Weil-$q$-number of weight determined by the local system $\mathcal{L}$, cf. Equation (2.3.10).

DEFINITION 4.4. We call a rigid representation $\pi_p$ of $G(\mathbb{Q}_p)$ of *$B$-type* if for all $\wp$, $\gcd(n, s_\wp)$ divides the product $y \cdot x_{\wp,a}$. Furthermore, for each $F^+$-prime $\wp$ and each index $a$, the number $m_{\wp,a}$ divides either $y$ or $x_{\wp,a}$.

We have proved that only the $B$-type representations contribute to the (alternating sum of the cohomology spaces) of $B$. Let $\pi_p$ be a $B$-type representation of $G(\mathbb{Q}_p)$. Then we write

$$\mathrm{Pol}(\pi) \stackrel{\text{def}}{=} \mathrm{Tr}\left(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p\right) \in \mathbb{C}[q^\alpha].$$

We computed this polynomial in the first 4 sections of this chapter. Explicitely, it is the product over all $\wp$, over all indices $a$ of the polynomial

$$\varepsilon \cdot q^{\frac{s_\wp(n-s_\wp)}{2}\alpha} \cdot \mathrm{Dyck}^+(\vec{x}^{w_0}, \vec{y}), \tag{4.7}$$

where the lists of points $\vec{x}, \vec{y} \in \mathbb{Q}^2$ are defined by:

(1) If $x_a \leq y$, then $\vec{x}$, $\vec{y}$ are of length $x_a$, and for each $b$ we have

$$\vec{x}_b := \ell\left(\frac{x_a - y}{2}\right), \quad \text{and} \quad \vec{y}_b := \ell\left(\frac{x_a + y}{2}\right),$$

(2) if $x_a \geq y$, then $\vec{x}$, $\vec{y}$ are of length $y$, and for each $b$ we have

$$\vec{x}_b := \ell\left(\frac{y - x_a}{2}\right), \quad \text{and} \quad \vec{y}_b := \ell\left(\frac{x_a + y}{2}\right),$$

where $\ell \subset \mathbb{Q}^2$ is the line of slope $\frac{s_\wp}{n}$ going through the origin. The notation does not show, but the points $\vec{x}$, $\vec{y}$ and the permutation $w_0$ depend on $\wp$. The symbol $w_0$ is a permutation in the group $\mathfrak{S}_{\min(x_a, y)}$ and is determined by Definition 2.8. The symbol $\varepsilon$ in Equation (4.7) is a sign and is equal to

$$\nu \cdot \mathrm{sign}(w_0), \tag{4.8}$$

where the sign $\nu$ is equal to $(-1)^{n-x_a}$ if $x_a \leq n$ and it is equal to 1 otherwise.

**4.3. Application: A dimension formula.** In Chapter 2 we explained that Formula (4.1) gives a formula for the number of points in $B$ if one takes $f^p = \mathbf{1}_{K^p}$ and $\mathcal{L}$ equal to the trivial local system. Using this simplified formula we proved in Proposition.2.4.2 a dimension formula for the basic stratum. We now extend this result to the Shimura varieties satisfying conditions (1)-(25) from the first subsection, with $p$ completely split in $F^+$.

Take $f^p = \mathbf{1}_{K^p}$ and $\mathcal{L}$ in Theorem 4.3 so that the right hand side of Equation (4.1) counts the number of points in $B$ over finite fields. We computed the class of representations at $p$ contributing to this formula. Each representation $\pi_p$ at $p$ contributes with a certain function $P(q^\alpha)$ to the zeta function of $B$. We call the *order* of $\pi_p$ the order of the function $P(q^\alpha)$ (as function in $q^\alpha$).

PROPOSITION 4.5. *The trivial representation $\pi_p = \mathbf{1}$ contributes with the largest order to the right hand side of Equation* (4.1).

REMARK. In the statement of the proposition, we mean 'largest order' in the non-strict sense. In general there are multiple representations contributing to the formula with the same order.

The order of the trivial representation is easily computed, it is equal to:

$$\sum_{\wp | p} \left( \sum_{v \in \wp} \frac{s_v(1-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right), \tag{4.9}$$

(cf. Equation (2.4.4))

PROOF OF PROPOSITION 4.5. Let $\pi_p$ be a unitary rigid representation. Pick one $\wp | p$. The component $\pi_\wp$ is a rigid representation of $G_\wp := \mathrm{GL}_n(F_\wp^+)$.

Assume first that $\pi_\wp$ is a Speh representation. We assume that $h \leq t$, so we will work in the dual setting. Treatment of the non-dual case is essentially the same (see Eq. (4.25) at the end of this argument below). Let $T_1 = \langle u_1, v_1 \rangle$, $T_2 = \langle u_2, v_2 \rangle$, $\ldots$, $T_h = \langle u_h, v_h \rangle$ be the segments of the Zelevinsky dual $\pi_\wp^\iota$ of $\pi_\wp$. By Tadic's formula the compact trace $\mathrm{Tr}(\chi_c^{G_\wp} f_\wp, \pi_\wp)$ is an alternating sum of compact traces $\mathrm{Tr}(\chi_c^{G_\wp} f_\wp, I_w^\iota)$ on Zelevinsky duals of certain standard representations $I_w$. The traces $\mathrm{Tr}(\chi_c^{G_\wp} f_\wp, I_w^\iota)$ can be described using graphs as we explained in the first section. The intuition is that the closer the graph is to the line $\ell$, the larger its weight is, and we claim that the largest weight is attained by trivial representation. More precisely, we claim that for all $f \in \mathcal{H}_0(G_\wp)$ and for all permutations $w \in \mathfrak{S}_h$ we have

$$\operatorname{Ord}(\operatorname{Tr}(\chi_c^{G_\wp} f, I_w^\iota)) \leq \operatorname{Ord}(\operatorname{Tr}(\chi_c^{G_\wp} f, \mathbf{1}_{G_\wp})), \tag{4.10}$$

where with $\operatorname{Ord}(h) \in \mathbb{Q}$ of an element $h \in A^+$ we mean the largest element $x \in \mathbb{Q}$ such that $q^x$ occurs as a monomial in the expression of $h$ with non-zero coefficient. By Proposition 3.1 and Dijk's integration formula for compact traces we have

$$\operatorname{Tr}(\chi_c^{G_\wp} f, I_w^\iota) = q^{\frac{s(n-s)}{2}} \cdot \sum_{X,\ \xi_c^{G_\wp} \chi_{M_w}^{G_\wp} X \neq 0} c_X \cdot \mathcal{G}_X(\vec{u}^w, \vec{v}) \in A^+,$$

where $X$ ranges over the monomials $X \in A$ of the Satake transform $\mathcal{S}(f)$ of $f$, $c_X \in \mathbb{C}$ is their coefficient and where we should explain the notation $\mathcal{G}_X(\vec{u}^w, \vec{v})$. The symbol $\vec{u}$ denotes the list of points $\vec{u}_a := \ell(u_a) \in \mathbb{Q}^2$ for $a = 1, \ldots h$ and the list of points $\vec{v}$ is defined by $\vec{v}_a := \ell(v_a + 1) \in \mathbb{Q}^2$. The symbol $\mathcal{G}_X$ is the graph of the monomial $X$ as defined in the first section. Recall however that $\mathcal{G}_X$ is only well-defined up to the definition of its starting point. The representation $I_w^\iota$ is obtained by induction from a one dimensional representation of a standard Levi subgroup $M_w$ of $G_\wp$. Let $(n_a^w)$ be the corresponding composition of $n$, and let $k_w$ be the length of this composition. We cut the graph $\mathcal{G}_X$ into $k_w$ pieces, the first piece contains the first $n_1^w$ steps of $\mathcal{G}_X$, the second piece contains the next block of $n_2^w$ steps of $\mathcal{G}_X$ and so on. Thus instead of one graph $\mathcal{G}_X$ we now have $k_w$ graphs, $\mathcal{G}_{X,1}^w, \mathcal{G}_{X,2}^w, \ldots, \mathcal{G}_{X,k_w}^w$, all well defined up to their starting points. We let the starting point of the graph $\mathcal{G}_{X,1}^w$ be $\vec{u}_1^w$, the starting point of the graph $\mathcal{G}_{X,2}^w$ is by definition $\vec{u}_2^w$, and so on. Then $\mathcal{G}_{X,1}^w$, $\mathcal{G}_{X,2}^w$, ... are well defined graphs in $\mathbb{Q}^2$, and due to our definition of starting points, we have

$$\prod_{a=1}^{k_w} \operatorname{weight}(\mathcal{G}_{X,a}^w) = \operatorname{Tr}\left(X, (I_w^\iota)_{N_0}(\delta_{P_0}^{-1/2})\right) \in A^+. \tag{4.11}$$

The condition $\xi_c^{G_\wp} \chi_{M_w}^{G_\wp} X \neq 0$ on $X$ means precisely that the graphs $\mathcal{G}_{X,a}^w$ have endpoint equal to $\vec{v}_a$ and that these graphs do not cross, but may touch, the line $\ell$.

Starting from the monomial $X$ we can also defined a second graph $\mathcal{H}_X$, such that

$$\operatorname{weight}(\mathcal{H}_X) = \operatorname{Tr}\left(X, \mathbf{1}(\delta_{P_0}^{-1/2})\right) \in A^+.$$

This graph has starting point $\vec{x} = \ell(\frac{1-n}{2})$ and end point $\vec{y} = \ell(\frac{n-1}{2} + 1)$; the steps of $\mathcal{H}_X$ are defined by Formula (2.2).

We now claim that

$$\prod_{a=1}^{h} \operatorname{weight}(\mathcal{G}_{X,a}^w) \leq \operatorname{weight}(\mathcal{H}_X) \in A^+, \tag{4.12}$$

for the obvious meaning of '$\leq$'.

Before we prove the claim, let us first show a simple fact of graphs. Let $\mathcal{G}$ be any graph in $\mathbb{Q}^2$. Then we have, for any point $(a,b) \in \mathbb{Q}^2$ that

$$\mathrm{Ord}(\mathcal{G} + (x,y)) = \mathrm{Ord}(\mathcal{G}) - x \cdot \mathrm{Height}(\mathcal{G}), \tag{4.13}$$

where the height of $\mathcal{G}$, $\mathrm{Height}(\mathcal{G})$, is the vertical distance between the initial point of $\mathcal{G}$ and its end point. This formula is easily seen to be true: The order $\mathrm{Ord}(\mathcal{G})$ is equal to the sum of $-a \cdot e$ over all diagonal steps $(a,b) \to (a+1, b+e)$ occurring in the graph $\mathcal{G}$. Adding the point $(x,y)$ to $\mathcal{G}$ amounts to changing $-a \cdot e$ to $-(a+x)e$ in the definition of the order of $\mathcal{G}$. Thus the order of $\mathcal{G}$ is shifted by the sum, over all diagonal steps $(a,b) \to (a+1, b+e)$, of the value $-xe$. This gives the formula in Equation (4.13).

We now return to the graphs $\mathcal{G}_X$ and $\mathcal{H}_X$ introduced earlier. We cut $\mathcal{H}_X$ into $h$ consecutive graphs. The first graph $\mathcal{H}_{X,1}$ consists of the first $n_1^w$ steps of $\mathcal{H}_X$, the second graph $\mathcal{H}_{X,2}$ consists of the second block of $n_2^w$ steps of $\mathcal{H}_X$, and so on. The graphs $\mathcal{G}_{X,a}$ have the same shape as the graphs $\mathcal{H}_{X,a}$, but they are shifted (the graphs are constructed starting from the same monomial $X$). Therefore we have the relations:

$$(\forall a): \quad \mathcal{H}_{X,a} = \mathcal{G}_{X,a} - \ell(u_{w(a)}) + \ell(\tfrac{1-n}{2} + n_1^w + \ldots n_{a-1}^w), \tag{4.14}$$

(we subtract the initial point of $\mathcal{G}_{X,a}$, and then add the initial point of $\mathcal{H}_{X,a}$); in the above formula we have the convention that

$$n_1^w + n_2^w + \ldots + n_{a-1}^w = 0,$$

in case $a = 1$. Note also that

$$\mathrm{Ord}(\mathcal{H}_X) = \sum_{a=1}^{h} \mathrm{Ord}(\mathcal{H}_{X,a}),$$

and similarly for $\mathcal{G}_X$. By Equations (4.13) and (4.14) we have

$$\mathrm{Ord}(\mathcal{H}_{X,a}) = \mathrm{Ord}(\mathcal{G}_{X,a}) - u_{w(a)} \cdot s_a^w + (\tfrac{1-n}{2} + n_1^w + \ldots + n_{a-1}^w) \cdot s_a^w,$$

where $s_a^w := n_a^w \cdot \frac{s}{n} = \mathrm{Height}(\mathcal{G}_{X,a}) = \mathrm{Height}(\mathcal{H}_{X,a})$. Thus we have to compute the following expression

$$C(w) = \frac{s}{n} \sum_{a=1}^{h} \left( \frac{1-n}{2} + n_1^w + n_2^w + \ldots + n_{a-1}^w - u_{w(a)} \right) n_a^w. \tag{4.15}$$

To show that Equation (4.12) is true, we show that $C(w) \leq 0$ for all permutations $w$.

To prove that $C(w) \leq 0$, we may ignore the factor $\frac{s}{n}$ in the above expression. We prove in two steps that $C(w) \leq 0$ for all $w$. We first determine the permutation $w$ such that the value $C(w)$ is maximal (Step 1). Then we compute for this particular permutation the value $C(w)$, and observe that it is non-positive (Step 2).

We begin with Step 1. We want to determine $w$ such that $C(w)$ is maximal. Let us first simplify the expression somewhat. The expression $C(w)$ is maximal for $w$ if and only if

$$\sum_{a=1}^{h}\left(n_1^w + n_2^w + \ldots + n_{a-1}^w - u_{w(a)}\right) n_a^w, \tag{4.16}$$

is maximal. To derive (4.16) we used [8], that the sum $\sum_{a=1}^{t} \frac{n-1}{2} n_a^w$ equals $n\frac{1-n}{2}$ and therefore this sum does not depend on $w$. (Similar arguments will appear also below.) We have

$$n_a^w = \left(\frac{t+h}{2} - a\right) - \left(\frac{h-t}{2} - (w(a)-1)\right) + 1 = t - a + w(a). \tag{4.17}$$

and

$$u_{w(a)} = \frac{t-h}{2} - (w(a)-1). \tag{4.18}$$

We plug Equations (4.17) and (4.18) into Equation (4.16) to get

$$\sum_{a=1}^{h}\left((t-1+w(1)) + \ldots + (t-(a-1)+w(a-1)) - \frac{t-h}{2} + (w(a)-1)\right) n_a^w$$

As before, this expression is maximal for $w$, if and only if the expression

$$\sum_{a=1}^{h}\left(w(1)-1+w(2)-2+\ldots+w(a-1)-(a-1)+w(a)\right)(t-a+w(a)) \tag{4.19}$$

is maximal. Equation (4.19) is maximal for $w$ if and only if the expression

$$\sum_{a=1}^{h}\left(w(1)-1+w(2)-2+\ldots+w(a-1)-(a-1)+w(a)\right)(w(a)-a) \tag{4.20}$$

is maximal. We may rewrite (4.20) to

$$\sum_{a=1}^{h}(w(1)+w(2)+\ldots+w(a))a - \sum_{a=1}^{h}(1+2+\ldots+(a-1))w(a) \tag{4.21}$$

We rearrange the first sum as follows. Count for each index $a$ the coefficient of $w(a)$ to get

$$\sum_{a=1}^{h}(w(1)+w(2)+\ldots+w(a))a = \sum_{a=1}^{h}\rho(h+1-a)w(a),$$

where

$$\rho(a) := 1+2+3+\ldots+a = \tfrac{1}{2}a(a+1).$$

Thus (4.21) equals

$$\sum_{a=1}^{h}(\rho(t+1-a) - \rho(a-1))w(a)$$

8. See Equation (2.1), but note that by duality the roles of $h$ and $t$ are switched.

The function $\nu(a)$ defined by $\nu(a) = \rho(h+1-a) - \rho(a-1)$, is strictly decreasing in $a$ because

$$\nu(a+1) - \nu(a) = -(h+1).$$

We are looking for $w$ such that

$$\sum_{a=1}^{h} \nu(a) \cdot w(a)$$

is maximal, with $\nu(a)$ a strictly decreasing function for $a \in \{1, 2, \dots, h\}$. This maximum is attained by the permutation $w$ defined by $a \mapsto h + 1 - a$. This completes Step 1.

We now do Step 2. Thus we have $w(a) = h + 1 - a$ for all indices $a \in \{1, 2, \dots, h\}$. We compute the sum

$$C(w) = \sum_{a=1}^{h} \left( \frac{1-n}{2} + n_1^w + n_2^w + \ldots + n_{a-1}^w - u_{w(a)} \right) n_a^w. \tag{4.22}$$

We have

$$\begin{aligned} n_a^w &= y_a - u_{w(a)} + 1 = \left( \frac{h+t}{2} - a \right) - \left( \frac{h-t}{2} - (w(a) - 1) \right) + 1 \\ &= t - a + w(a) = t - a + (h+1) - a = t + h + 1 - 2a, \end{aligned} \tag{4.23}$$

and we have

$$u_{w(a)} = \frac{t-h}{2} - (w(a) - 1) = \frac{t-h}{2} - (h - a). \tag{4.24}$$

Note also that,

$$n = \sum_{a=1}^{h} n_a^w = \sum_{a=1}^{h} t + h + 1 - 2a.$$

(cf. Equation (2.1)). Thus, Equation (4.22) becomes

$$\sum_{a=1}^{h} \left( \frac{1-n}{2} + \left( \sum_{b=1}^{a-1} t + h + 1 - 2b \right) - \left( \frac{t-h}{2} - (h-a) \right) \right) \cdot (t + h + 1 - 2a)$$

An easy (but somewhat lengthy) computation shows that this last formula simplifies to $n(h - t)$. By assumption we have $h \le t$. We conclude that the value in Equation (4.22) is non-positive, which is what we wanted to show. We have now established the claim in Equation (4.12).

With the same proof, but using the non-dual instead, one can show that

$$\mathrm{Ord}(\mathrm{Tr}(\chi_c^{G_\wp} f, I_w)) \le \mathrm{Ord}(\mathrm{Tr}(\chi_c^{G_\wp} f, \mathrm{St}_{G_\wp})), \tag{4.25}$$

is true for all representation $I_w$ occurring in Tadic's formula for Speh representations $\pi$ with $h \ge t$. Because

$$\mathrm{Ord}(\mathrm{Tr}(\chi_c^{G_\wp} f, \mathrm{St}_{G_\wp})) \le \mathrm{Ord}(\mathrm{Tr}(\chi_c^{G_\wp} f, \mathbf{1}_{G_\wp})),$$

the inequality of Equation (4.10) is true for all Speh representations. We leave it to the reader to deduce that Equation (4.10) also holds for products of Speh representations, and also for the rigid representations of $G_\wp$ (with the characters $\varepsilon_a$ trivial).

We return to the group $G(\mathbb{Q}_p)$ and the full representation $\pi_p$. The compact trace $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p)$ is the product of the traces on the components,

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \mathrm{Tr}_{\mathbb{Q}_p^\times}(f_s, \pi_s) \cdot \prod_{\wp | p} \mathrm{Tr}(\chi_c^{G_\wp} f_\wp, \pi_\wp),$$

(the first term in the product is the contribution of the factor of similitudes). We proved that all the terms of this product are bounded by the trace on the trivial representation. This is then also true for the entire product. □

We now deduce a formula for the dimension.

THEOREM 4.6. *The dimension of the basic stratum $B$ is equal to*

$$\sum_{\wp | p} \left( \sum_{v \in \wp} \frac{s_v(1-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right).$$

PROOF. Apply Proposition 4.5 and Theorem 4.3 to find

$$\dim(B) \leq \sum_{\wp | p} \left( \sum_{v \in \wp} \frac{s_v(1-s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right).$$

We now prove the opposite inequality. We return to the final formula we found in Theorem 4.3:

$$\sum_{i=0}^{\infty} (-1)^i \,\mathrm{Tr}(f^{\infty p} \times \Phi_{\mathfrak{p}}^\alpha, \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \iota^* \mathcal{L})) = \sum_{\substack{\pi \subset \mathcal{A}(G) \\ \pi_p \text{ rigid}}} \mathrm{Tr}(\chi_c^G f_\alpha, \pi_p) \cdot \mathrm{Tr}(f^p, \pi^p). \tag{4.26}$$

We take in this formula $f^p$ and $\mathcal{L}$ of the following form. Let $p_1$ be a prime number with

- $p_1$ is different from $\ell, p$;
- the group $G$ splits over $\mathbb{Q}_{p_1}$;
- the group $K$ splits into a product $K_{p_1} K^{p_1}$ of a hyperspecial group at $p_1$ and a compact open subgroup $K^{p_1} \subset G(\mathbb{A}_{\mathrm{f}}^{p_1})$ outside $p_1$.

We take

- $f^{pp_1} = \mathbf{1}_{K^{pp_1}}$;
- $f_{p_1}$ is an arbitrary $K_{p_1}$-spherical function;
- $\mathcal{L} = \overline{\mathbb{Q}}_\ell$ (the trivial local system).

There exist only a finite number of representations $\pi_{p_1}$ contributing to Equation (4.26), and one of these representations is the trivial representation. Thus we may find a spherical Hecke operator $f_{p_1} \in \mathcal{H}(G(\mathbb{Q}_{p_1}))$ such that

$$\mathrm{Tr}(f_{p_1}, \pi_{p_1}) = \begin{cases} 1 & \pi_{p_1} \cong \mathbf{1}_{G(\mathbb{Q}_{p_1})} \\ 0 & \text{otherwise,} \end{cases}$$

for all representations $\pi_{p_1}$ occurring in Equation (4.26). We consider the Hecke operator $f^p := \mathbf{1}_{K^{pp_1}} \otimes f_{p_1}$ in Equation (4.26). By construction, any automorphic representation $\pi \subset \mathcal{A}(G)$ contributing to Equation (4.26) has $\pi_{p_1} \cong \mathbf{1}_{G(\mathbb{Q}_{p_1})}$. By a strong approximation argument, the representation $\pi$ is one dimensional [9], and in particular Abelian. Consequently, at the prime $p \neq p_1$, the representation $\pi_p$ is a twist of $\mathbf{1}_{G(\mathbb{Q}_p)}$ by an unramified character $\chi_p$. Because the representation $\xi$ at infinity is trivial, the character $\chi_p$ is of finite order. Therefore there exists an integer $r > 0$ such that, whenever $r$ divides $\alpha$, we have

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \pi_p) = \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \mathbf{1}),$$

for all representations $\pi_p$ contributing to Equation (4.26). From now on we consider only $\alpha$ such that $r|\alpha$. The right hand side of Equation (4.26) simplifies to

$$C \cdot \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \mathbf{1}),$$

where $C$ is some non-zero constant. Thus for our choice of $f^{\infty p}$ the trace

$$\sum_{i=0}^{\infty} (-1)^i \, \mathrm{Tr}(f^{\infty p} \times \Phi_{\mathfrak{p}}^{\alpha}, \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \overline{\mathbb{Q}}_\ell)) \tag{4.27}$$

grows with the order of the trivial representation. View $\sum_{i=0}^{\infty} (-1)^i \mathrm{H}^i_{\text{ét}}(B_{\overline{\mathbb{F}}_q}, \overline{\mathbb{Q}}_\ell)$ as a virtual representation of the group $(\Phi_{\mathfrak{p}}^r)^{\mathbb{Z}}$, and write it as a linear combination of the characters of this group. The character of highest order occurring in this expression determines the dimension of the variety $B$. By the conclusion in Equation (4.27) there occurs a character whose order is at least $\mathrm{Ord}(\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_\alpha, \mathbf{1}))$. This means that

$$\dim(B) \geq \sum_{\wp | p} \left( \sum_{v \in \wp} \frac{s_v(1 - s_v)}{2} + \sum_{j=0}^{s_\wp - 1} \lceil j \frac{n}{s_\wp} \rceil \right).$$

This completes the proof of the Theorem. □

Remark. The above formula confirms the conjecture for the dimension of the basic stratum specialized to the cases we consider. See for example [**61**].

**4.4. Application: Vanishing of the cohomology.** In Chapter 2 we assumed that the signatures $s_\wp$ are coprime to the number $n$. Under these conditions the cohomology of the basic stratum is very simple: Locally at the prime $p$, only the trivial representation and (essentially) the Steinberg representation contribute to Expression (4.1). In fact this is true in a larger class of cases:

Corollary 4.7. *Assume there is one $F^+$-place $\wp$ above $p$ such that $s_\wp$ is coprime to $n$. Then only the Steinberg representation and the trivial representation contribute to the formula in Equation* (4.1).

9. See for example Lemma 3.6 in Chapter 2, although this result is of course well known.

PROOF. This follows directly form the definition of rigid representation of the group $G(\mathbb{Q}_p)$. □

REMARK. In Chapter 2 we assumed that, for all $\wp$, the number $s_\wp$ is coprime to $n$ or $s_\wp$ is equal to 0 or $n$. Only under this larger assumption the compact trace on the Steinberg representation coincides with the compact trace on the trivial representation (up to sign), just as in Chapter 2. In the above Corollary this need not be the case.

**4.5. Application: Euler-Poincaré characteristics.** Finally we have a remark on the Euler-Poincaré characteristic of the variety $B$. The evaluation at $q = 1$ of our formula gives the expression of the Euler-Poincaré characteristic. Thus to compute the Euler-Poincaré characteristic we get the combinatorial problem to compute, apart from dimensions of spaces of automorphic forms, the *number* of non-intersecting Dyck paths. This problem has been considered in an equivalent forms in the literature; a good starting point are the books of Stanley [**97**] and the references therein.

## 5. Examples

We end this chapter with some examples. Let us first explain why we need the condition that $p$ splits completely in the center of $D$.

**5.1. Products of simple Kottwitz functions.** To study the reduction modulo $p$ of unitary Shimura varieties, the simple Kottwitz functions $f_{n\alpha s}$ as we defined them in Equation (1.1) are *not* enough. These functions count only points of unitary Shimura varieties if the group $G$ of the Shimura datum is of the following kind. Consider a unitary Shimura variety associated to a division algebra $D$ as in the previous section. Let $U$ in $G$ be the subgroup of elements whose factor of similitudes is equal to one. Then $U$ is a unitary group and $U(\mathbb{R})$ is isomorphic to a product of standard unitary groups $U(p_\tau, q_\tau)$ with $\tau$ ranging over the infinite places of the maximal totally real subfield $F^+$ of the center $F$ of $D$. The function $f_{n\alpha s}$ counts points on the reduction of $\mathrm{Sh}_K$ modulo $p$ if we have $p_\tau = 0$ or $q_\tau = 0$ for all $F^+$-places $\tau$, but with one $F^+$-place excluded. For the excluded $F^+$-place $\tau_0$ we must have $p_{\tau_0} = s$ or $p_{\tau_0} = n - s$. For unitary Shimura varieties with several non-zero signatures at infinity, one will need to consider products of the functions $f_{n\alpha s}$ for several different values of $s$.

REMARK. Compact traces do *not* commute with products of Hecke operators.

EXAMPLE. Let us assume that there are two infinite $F^+$-places $\tau_0, \tau_1$ with $p_{\tau_0} = p_{\tau_1} = 1$ and that $p_\tau = 0$ for all other $\tau$. Choose embeddings $\mathbb{C} \supset \overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_p$, so the group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ acts on the set of infinite places of $F^+$. Assume the places $\tau_0$ and $\tau_1$ lie in the same $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-orbit and assume $\alpha$ is sufficiently divisible such that the $E_{\mathfrak{p},\alpha}$-algebra $F^+ \otimes E_{\mathfrak{p},\alpha}$ is split. Then the function counting points in the set $\#\mathrm{Sh}_K(\mathbb{F}_{q^\alpha})$ is (essentially) the convolution product $f = f_{n\alpha 1} * f_{n\alpha 1} \in \mathcal{H}_0(\mathrm{GL}_n(F))$, where $F$ is some finite extension of $\mathbb{Q}_p$. An easy computation shows that $f = 2q^{\alpha} f_{n\alpha 2} + f_{n(2\alpha)1}$, and therefore $\mathrm{Tr}(\chi_c^G f, \mathbf{1}_G) = 2(q + q^1 + \ldots + q^{\alpha\lfloor \frac{n}{2} \rfloor}) + 1$. Consequently the number of points in the basic stratum over the field $\mathbb{F}_{q^\alpha}$ is the product of the above polynomial times a cohomological expression depending only on the class of the degree $\alpha$ in the group $\mathbb{Z}/h\mathbb{Z}$, where $h$ is related to the class number of the cocenter of $G$ (Corollary 2.4.1). In particular the variety is of dimension $\lfloor \frac{n}{2} \rfloor$ in this case. If we assume instead that $\tau_0$ and $\tau_1$ lied in a *different* $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-orbit, then the basic stratum of $\mathrm{Sh}_K$ is a *finite* variety. Whether or not $\tau_0$ and $\tau_1$ lie in the same $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-orbit is a condition on how the prime $p$ decomposes as a product of prime ideals in the ring of integers $\mathcal{O}_{F^+}$ of $F^+$. Thus, roughly speaking, the form of the function $\alpha \mapsto \mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f, \mathbf{1}_{G(\mathbb{Q}_p)})$ depends only on two pieces of information: (1) The signatures of the unitary group at infinity, and (2) how the prime $p$ decomposes in $F^+$.

**5.2. Two different prime factors.** Assume $F^+$ is of degree 2 over $\mathbb{Q}$ and $n$ is a product of two primes $x, y$ with $x < y$. Let $U \subset G$ be the subgroup of elements whose factor of similitudes is trivial. We assume $U(\mathbb{R})$ is isomorphic to $U(x, n-x)(\mathbb{R}) \times U(y, n-y)(\mathbb{R})$. The reflex field $E$ of the Shimura datum coincides with the field $F$.

There are two cases to consider, either the prime $p$ where we reduce $\mathrm{Sh}_K$ splits in $F^+$ or $p$ is inert (but unramified). Assume that $p$ splits, then $G(\mathbb{Q}_p) = \mathbb{Q}_p^\times \times \mathrm{GL}_n(\mathbb{Q}_p) \times \mathrm{GL}_n(\mathbb{Q}_p)$. Recall that we picked an embedding $\nu_p \colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$. Therefore the factors of the product $\mathrm{GL}_n(\mathbb{Q}_p) \times \mathrm{GL}_n(\mathbb{Q}_p)$ are ordered: the embbeding $\nu_p$ identifies the two $F^+$-places $\tau_1, \tau_2$ at infinity with the two $F^+$-places $\wp_1, \wp_2$ above $p$. Via the isomorphism $U(\mathbb{R}) \cong U(x, n-x)(\mathbb{R}) \times U(y, n-y)(\mathbb{R})$ we associate to $\tau_1, \tau_2$ (and thus to $\wp_1, \wp_2$) a signature equal to $x$ or $y$. Assume that $\wp_1$ (and $\tau_1$) correspond to $x$ and $\wp_2$ (and $\tau_2$) correspond to $y$. Similarly, the first factor of the group $\mathrm{GL}_n(\mathbb{Q}_p) \times \mathrm{GL}_n(\mathbb{Q}_p)$ corresponds to $\wp_1$ and the second factor corresponds to $\wp_2$.

The $B$-type representations of $G(\mathbb{Q}_p)$ are the representations contributing to the cohomology of the basic stratum. Ignoring the factor of similitudes, the $B$-type representations of $G(\mathbb{Q}_p)/\mathbb{Q}_p^\times = \mathrm{GL}_n(\mathbb{Q}_p) \times \mathrm{GL}_n(\mathbb{Q}_p)$ are:

$$\mathrm{Speh}(x, y)(\varepsilon) \otimes \prod_{a=1}^{k} \mathrm{Speh}(x_a, y)(\varepsilon_a), \tag{5.1}$$

$$\prod_{a=1}^{k} \mathrm{Speh}(y_a, x)(\varepsilon_a) \otimes \mathrm{Speh}(x, y)(\varepsilon), \tag{5.2}$$

$$\mathrm{St}_G(\varepsilon) \otimes \mathrm{St}_G(\varepsilon'), \tag{5.3}$$

$$\mathbf{1}_G(\varepsilon) \otimes \mathbf{1}_G(\varepsilon'), \tag{5.4}$$

where, in these equations the number $k$ can, a priori, be any positive number. In Equation (5.1), the symbol $(x_a)$ ranges over the compositions of the prime $x$ and in Equation (5.2), the symbol $(y_a)$ ranges over the compositions of the prime $y$. The symbols $\varepsilon, \varepsilon', \varepsilon_a$ denote arbitrary, unrelated, unramified unitary characters.

**5.3. Some explicit polynomials.** We specialize our first example further, and assume that $x = 2$ and $y = 3$, so $U(\mathbb{R}) \cong U(2,4)(\mathbb{R}) \times U(3,3)(\mathbb{R})$. We write down the polynomials $\mathrm{Tr}(\chi_c^G f_\alpha, \pi_p) \in A^+$ for the representations that occur. The unramified characters $\varepsilon, \varepsilon_a, \varepsilon'$ and the factor of similitudes have no influence on the form of the polynomials, so we leave them out.

The computation of the compact traces on the representations $\pi_p = \mathrm{Speh}(3,2)$ and $\pi_p = \mathrm{Speh}(2,3)$ is done in the Figures 4 and 5. Recall that the computuation on $\mathrm{Speh}(3,2)$ is done via the segments of its dual representation $\mathrm{Speh}(2,3)$. The Zelevinsky segments of the representation $\mathrm{Speh}(2,3)$ are $\left\{-\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right\}$ and $\left\{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}\right\}$. To compute the compact traces we consider the line $\ell$ in $\mathbb{Q}^2$ of slope $\frac{s}{n}$ and consider the weights of non-crossing lattice paths. In our case there are two possible slopes, slope $\frac{1}{2}$ and slope $\frac{1}{3}$; these yield several different polynomials.

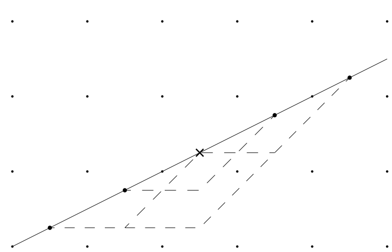

FIGURE 4. The compact trace on the representation $\mathrm{Speh}(3,2)$ with respect to the function $f_{6\alpha 3}$. We have $\frac{s}{n} = \frac{1}{2}$, and $\vec{x}_1 = \ell(-\frac{1}{2})$, $\vec{x}_2 = \ell(-\frac{3}{2})$, $\vec{y}_1 = \ell(\frac{3}{2}+1)$ and $\vec{y}_2 = \ell(\frac{1}{2}+1)$. The permutation $w_0$ is equal to $(12)$. We see that there are two Dyck 2-paths going from the points $\vec{x}^{w_0}$ to the points $\vec{y}$, and one of those paths is non-strict because it touches the line $\ell$. Therefore $\mathrm{Dyck}_{\mathrm{s}}^+(\vec{x}^{w_0}, \vec{y}) = q^{-1/2\alpha - 1/2\alpha - 3/2\alpha} = q^{-5/2\alpha}$ and $\mathrm{Dyck}^+(\vec{x}^{w_0}, \vec{y}) = q^{-5/2\alpha} + q^{-3/2\alpha}$. We conclude: $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 3}, \mathrm{Speh}(3,2)) = (-1)^{n-t} \mathrm{sign}(w_0) q^{\frac{s(n-s)}{2}\alpha} q^{-5/2\alpha} = -q^{2\alpha}$.

In the illustrations we found that

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 3}, \mathrm{Speh}(3,2)) = -q^{-2\alpha}$$
$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 2}, \mathrm{Speh}(3,2)) = -q^{3\alpha}.$$

Using the duality and the computation in the figures, we find that

$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 3}, \mathrm{Speh}(2,3)) = q^{2\alpha} + q^{3\alpha}$$
$$\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 2}, \mathrm{Speh}(2,3)) = q^{2\alpha}.$$

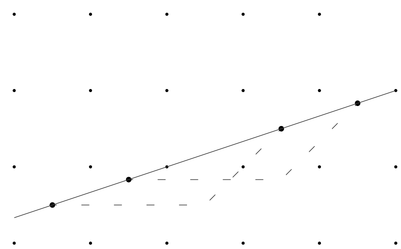

FIGURE 5. The compact trace on the representation $\pi_p = \mathrm{Speh}(3,2)$ with respect to the function $f_{6\alpha 2}$. We have $\frac{s}{n} = \frac{1}{3}$, and $\vec{x}_1 = \ell(-\frac{1}{2})$, $\vec{x}_2 = \ell(-\frac{3}{2})$, $\vec{y}_1 = \ell(\frac{3}{2}+1)$ and $\vec{y}_2 = \ell(\frac{1}{2}+1)$. The permutation $w_0$ is the trivial permutation. There is one Dyck 2-path going from the points $\vec{x}^{w_0}$ to the points $\vec{y}$ and this 2-path is strict. Therefore $\mathrm{Dyck}^+_{\mathrm{s}}(\vec{x}^{w_0}, \vec{y}) = \mathrm{Dyck}^+(\vec{x}^{w_0}, \vec{y}) = q^{-1/2\alpha - 3/2\alpha} = q^{-\alpha}$. We conclude: $\mathrm{Tr}(\chi_c^{G(\mathbb{Q}_p)} f_{6\alpha 2}, \mathrm{Speh}(3,2)) = (-1)^{n-t} \mathrm{sign}(w_0) q^{\frac{s(n-s)}{2}} = -q^{3\alpha}$.

By drawing the picture, we see in a similar manner to the illustrations that

$$\mathrm{Tr}(\chi_c^{G_4} f_{4\alpha 2}, \mathrm{Speh}(2,2)) = q^{3\alpha}. \tag{5.5}$$

and

$$\begin{aligned}
\mathrm{Tr}(\chi_c^{G_6} f_{6\alpha 2}, \mathbf{1}_{G_6}) &= 1 + q^{\alpha} + q^{2\alpha} \\
\mathrm{Tr}(\chi_c^{G_6} f_{6\alpha 2}, \mathrm{St}_{G_6}) &= -(q^{\alpha} + q^{2\alpha}) \\
\mathrm{Tr}(\chi_c^{G_6} f_{6\alpha 3}, \mathbf{1}_{G_6}) &= 1 + q^{\alpha} + 2q^{2\alpha} + q^{3\alpha} \\
\mathrm{Tr}(\chi_c^{G_6} f_{6\alpha 3}, \mathrm{St}_{G_6}) &= -(1 + q^{\alpha}). \qquad (5.6)
\end{aligned}$$

The representations at $p$ occurring in the alternating sum of the cohomology of the basic stratum are (up to twists):

$$\mathrm{Speh}(2,3) \otimes \mathrm{Speh}(2,3), \quad \mathrm{Speh}(2,3) \otimes (\mathbf{1}_{G_3} \times \mathbf{1}_{G_3});$$

$\mathrm{Speh}(3,2) \otimes \mathrm{Speh}(3,2)$, $\quad (\mathrm{Speh}(2,2) \times \mathbf{1}_{G_2}) \otimes \mathrm{Speh}(3,2) \quad (\mathbf{1}_{G_2} \times \mathrm{Speh}(2,2)) \otimes \mathrm{Speh}(3,2)$, $(\mathbf{1}_{G_2} \times \mathbf{1}_{G_2} \times \mathbf{1}_{G_2}) \otimes \mathrm{Speh}(3,2)$;

$$\mathrm{St}_{G_6} \otimes \mathrm{St}_{G_6};$$

$$\mathbf{1}_{G_6} \otimes \mathbf{1}_{G_6}.$$

Let us ignore the factor of similitudes of the group $G(\mathbb{Q}_p)$. On the group $\mathrm{GL}_2(\mathbb{Q}_p) \times \mathrm{GL}_2(\mathbb{Q}_p)$ the function of Kottwitz is equal to $f_{6\alpha 2} \otimes f_{6\alpha 3} \in \mathcal{H}_0(\mathrm{GL}_n(\mathbb{Q}_p)) \otimes \mathcal{H}_0(\mathrm{GL}_n(\mathbb{Q}_p))$. With the formulas we gave above the compact traces on the representations in this list are now all explicit.

CHAPTER 4

# Non-emptiness of the Newton strata

Recently Wedhorn and Viehmann [**104**] have proved through geometric means that, for a Shimura variety of PEL type of type (A) or (C), the Newton strata at a prime of good reduction are non-empty. We reprove this result using automorphic forms and the trace formula in case the group is of type (A). At the time of writing this chapter we learned that Sug Woo Shin also found a proof of this theorem with yet another method.

Let us explain our method of proof. The formula of Kottwitz for the number of points on a Shimura variety modulo $p$ can be restricted to count the number of points in any given Newton stratum. Thus, it suffices that this restriction be non-zero. Kottwitz rewrites the formula in terms of stable orbital integrals on certain endoscopic groups of $G$. This stable expression coincides with the geometric side of the stable trace formula. The geometric side equals the spectral side, so we get a sum over the endoscopic groups of $G$ of certain truncated, transferred Hecke operators acting on automorphic representations of these endoscopic groups. (The truncation is defined by the element of $B(G_{\mathbb{Q}_p}, \mu)$.) A general objective is to try and work out this expression; one will then get a description (of the alternating sum) of the cohomology of the Newton strata. Here we have aimed at a simpler goal: We do not describe the cohomology of the Newton stratum defined by $b \in B(G_{\mathbb{Q}_p}, \mu)$, we only show that the cohomology does not vanish, so that the corresponding Newton stratum must be non-empty.

We pick one very particular Hecke operator $f^p$ and carry out the computation sketched above only for this particular Hecke operator. We choose our Hecke operator with care, so that all the proper endoscopy vanishes and that in the end, after applying a simple version of the trace formula, we arrive at a sum of certain $b$-truncated traces on cuspidal automorphic representations of the quasi-split inner form $G^*$ of the group $G$ (Equation (7.11)):

$$\sum_{\Pi} m(\Pi) \cdot \mathrm{Tr}((f^p)^{G^*(\mathbb{A}_{\mathrm{f}}^p)}, \Pi^p) \, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p). \tag{0.7}$$

We choose the function $f^p$ so that, based on general conjectures, we *expect* that there is precisely *one* automorphic representation $\Pi_0$ contributing to this sum (for $\alpha$ sufficiently divisible). Therefore no cancellations occur and the sum is non-zero. We do not prove these general conjectures. However, we show that there is *at least one* contributing representation $\Pi_0$, and that for any other hypothetical $\Pi$ contributing to Equation (0.7), the quotient

$$\frac{m(\Pi) \cdot \mathrm{Tr}((f^p)^{G^*(\mathbb{A}_{\mathrm{f}}^p)}, \Pi^p) \, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)}{m(\Pi_0) \cdot \mathrm{Tr}((f^p)^{G^*(\mathbb{A}_{\mathrm{f}}^p)}, \Pi_0^p) \, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})}, \tag{0.8}$$

is a positive real number (here $\alpha$ is sufficiently divisible). Then, the sum in Equation (0.7) is non-zero. Thus the formula of Kottwitz does not vanish as well, and this means that the corresponding Newton stratum is nonempty.

An important step in the argument is showing that the representation $\Pi_0$ exists. In particular we have to find a local representation $\Pi_{0,p}$ at $p$ such that we have $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p}) \neq 0$. In the first section we find a set of such representations $\Pi_0$ with positive Plancherel measure. General theory of automorphic forms then assures the existence of a global automorphic representation $\Pi_0$ lying in our Plancherel set.

## 1. Isocrystals

We start this preliminary section with some notations. Let $p$ be a prime number and let $F$ be a finite extension of $\mathbb{Q}_p$. Let $\mathcal{O}_F$ be the ring of integers of $F$, let $\varpi_F \in \mathcal{O}_F$ be a prime element. We write $\mathbb{F}_q$ for the residue field of $\mathcal{O}_F$, and the number $q$ is by definition its cardinality. We fix an algebraic closure $\overline{\mathbb{Q}}_p$ of $F$, and we let $F_\alpha$ be the unramified extension of $F$ of degree $\alpha$ in $\overline{\mathbb{Q}}_p$. Let $G$ be a smooth reductive group over $\mathcal{O}_F$ (then $G_F$ is an unramified group [**100**]). We fix a minimal parabolic subgroup $P_0$ of $G$, and we standardize the parabolic subgroups of $G$ with respect to $P_0$. We write $T \subset P_0$ for the Levi component of $P_0$ and $N_0$ for the unipotent part, so that we have $P_0 = TN_0$. We call a parabolic subgroup $P$ of $G$ *standard* if it contains $P_0$, and we write $P = MN$ for its standard Levi decomposition. We write $K$ for the hyperspecial subgroup $G(\mathcal{O}_F) \subset G(F)$. Let $\mathcal{H}(G)$ be the Hecke algebra of locally constant compactly supported complex valued functions on $G(F)$, where the product on this algebra is the one defined by the convolution integral with respect to the Haar measure giving the group $K$ measure 1. We write $\mathcal{H}_0(G)$ for the spherical Hecke algebra of $G$ with respect to $K$. We write $\rho$ for the half sum of the positive roots of $G$.

We write $Z \subset G$ for the center of $G$, and we write $A \subset Z$ for the split center. Similarly $Z_M$ (resp. $Z_P$) is the center of the Levi-subgroup $M$ (resp. parabolic subgroup $P$); and we write $A_M$ (resp. $A_P$) for the split center of $M$. We write $A_0$ for $A_{P_0} \subset T$. We write $\mathfrak{a}_0 := X_*(A_0) \otimes \mathbb{R}$, and $C_0$ for the closed, positive chamber in $\mathfrak{a}_0$:

$$C_0 := \{x \in \mathfrak{a}_0 \mid \text{for all roots } \alpha \text{ in } \Delta(A_0, \mathrm{Lie}(N_0))\colon\ \langle x, \alpha \rangle \geq 0\}\,.$$

Let $B(G)$ be the set of $\sigma$-conjugacy classes in $G(L)$, where $L$ is the completion of the maximal unramified extension of $F$ and $\sigma$ is the arithmetic Frobenius of $L$ over $F$. Let $\mu \in X_*(T)$ be a $G$-dominant minuscule cocharacter. Recall that Kottwitz has defined the subset $B(G,\mu) \subset B(G)$ of $\mu$-admissible isocrystals [**60**, **88**].

Let $\mathbb{D}$ be the protorus over $F$ with character group given by $X_*(\mathbb{D}) = \mathbb{Q}$ and trivial Galois action. For any $b \in G(L)$ we have an unique morphism $\nu_b\colon \mathbb{D}_L \to G_L$ characterized by the following property: For every algebraic representation $(\rho, V)$ of $G$ on a finite dimensional vector space $V$ the composition $\rho \circ \nu_b$ determines the slope filtration on $(V \otimes L, \rho(b)(1 \otimes \sigma_L))$ [**55**, §4]. Replacing $b$ by a $\sigma$-conjugate amounts to conjugating $\nu_b$ with some $G(L)$-conjugate.

Moreover, one can replace $b$ so that $\nu_b$ has image inside the torus $A_{0,L}$, so that $\nu_b$ defines an element of $\mathfrak{a}_0$ [**60**, p. 267] [**88**, 1.7]. Write $\overline{\nu}_b$ for the unique element of $C_0$ whose orbit under the Weyl group meets $\nu_b$. The morphism $\overline{\nu}_b$ is called the *slope morphism* and the mapping $B(G) \to C_0, b \mapsto \overline{\nu}_b$ is called the *Newton map*. Note that the mapping $b \mapsto \overline{\nu}_b$ is not injective in general (it is injective in case $G = \mathrm{GL}_n(F)$).

Recall that we fixed an embedding $F \subset \overline{\mathbb{Q}}_p$. For each finite subextension $F' \subset \overline{\mathbb{Q}}_p$ of $F$ we have the unique mapping $H_T\colon T(F') \to X_*(T)_{\mathbb{R}}$ such that $q_{F'}^{-\langle \chi, H(t)\rangle} = |\chi(t)|$ for all $t \in T(F')$, where $q_{F'}$ is the cardinal of the residue field of $F'$, and the norm is normalized so that $|p|$ equals $q_{F'}^{-e}$ where $e$ is the ramification index of $F'/F$. By taking the union over all $F'$ we get a mapping $H_T\colon T(\overline{\mathbb{Q}}_p) \to X_*(T)_{\mathbb{R}}$. Consider the composition $H_A$ defined by $T(\overline{\mathbb{Q}}_p) \to X_*(T)_{\mathbb{R}} \to X_*(A)_{\mathbb{R}} = \mathfrak{a}_0$. Let $G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}} \subset G(\overline{\mathbb{Q}}_p)$ be the subset of semisimple elements. If $g \in G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}}$, then we may conjugate $g$ to an element $g'$ of $T(\overline{\mathbb{Q}}_p)$ and then consider $H_A(g') \in \mathfrak{a}_0$. This element of $\mathfrak{a}_0$ is only defined up to conjugacy, but we can take a representative in the, closed positive Weyl chamber $H(g) \in C_0^+$ which is well-defined. Thus we have a map $\Phi\colon G(\overline{\mathbb{Q}}_p)_{\mathrm{ss}} \to C_0$ defined on the semisimple elements. We extend the definition of $\Phi$ to $G(\overline{\mathbb{Q}}_p)$ by defining $\Phi(g) := \Phi(g_{\mathrm{ss}})$, where $g_{\mathrm{ss}}$ is the semisimple part of the element $g \in G(\overline{\mathbb{Q}}_p)$. We restrict to $G(F) \subset G(\overline{\mathbb{Q}}_p)$ to obtain the mapping $\Phi\colon G \to C_0$. In Proposition 1.1 we establish a relation between the map $\Phi$ and the Newton polygon mapping of isocrystals.

We recall the definition of the norm $\mathcal{N}$ of (certain) $\sigma$-conjugacy classes (cf. [**4**] [**53**, p. 799]). To any element $\delta \in G(F_\alpha)$ we associate the element $N(\delta) := \delta\sigma(\delta)\cdots\sigma^{\alpha-1}(\delta) \in G(F_\alpha)$. For any element $\delta \in G(F_\alpha)$, defined up to $\sigma$-conjugacy, with semi-simple norm $N(\delta)$ one proves (see [*loc. cit.*]) that $N(\delta)$ actually comes from a conjugacy class $\mathcal{N}(\delta)$ in the group $G(F)$.

PROPOSITION 1.1. *Let $\alpha$ be a positive integer and let $\delta \in G(F_\alpha)$ be an element of semi-simple norm, defined up to $\sigma$-conjugacy. Let $\gamma \in G(F)$ be an element in the conjugacy class $\mathcal{N}(\delta)$, and let $b$ be the isocrystal with additional $G$-structure defined by $\delta$. Then $\overline{\nu}_b = \alpha \cdot \Phi(\gamma) \in C_0$.*

PROOF. We first prove the case where $G$ is the general linear group. If $G = \mathrm{GL}_{n,F}$, then an isocrystal "with additional $G$-structure" is simply an isocrystal, i.e. a pair $(V, \Phi)$ where $V$ is an $n$-dimensional $L$ vector space and $\Phi$ is a $\sigma$-linear bijection from $V$ onto $V$. Because $b$ is induced by some $\delta \in G(F_\alpha)$, we may find a $F_\alpha$-vector space $V'$ together with a $\sigma$-linear bijection $\Phi'\colon V' \to V'$ such that $(V, \Phi)$ is obtained from $(V, \Phi)$ by extending the scalars $V = V' \otimes_{F_\alpha} L$ and $\Phi(v' \otimes l) := \Phi'(v') \otimes \sigma(l)$. Then $(V', \Phi')$ is an $F_\alpha$-space in the terminology of Demazure [**35**], and a theorem of Manin gives the relation $\overline{\nu}_b = \alpha \cdot \Phi(\gamma)$ (cf. [**35**, p. 90]).

Drop the assumption that $G = \mathrm{GL}_n$. Pick a representation $\rho\colon G \to \mathrm{GL}_V$ of $G$ in some finite dimensional $\mathbb{Q}_p$-vector space $V$. Then, by the statement for $\mathrm{GL}_n$, we see that $\alpha \cdot \Phi_{\mathrm{GL}_n}(\rho(\gamma))$ determines the slope filtration on the space $(V \otimes L, \rho(b)(1 \otimes \sigma_L))$. Thus $\rho \circ \overline{\nu}_b = \alpha \cdot \Phi_{\mathrm{GL}_n}(\rho(\gamma))$ for all $\rho$, and then the equality is also true for the group $G$. □

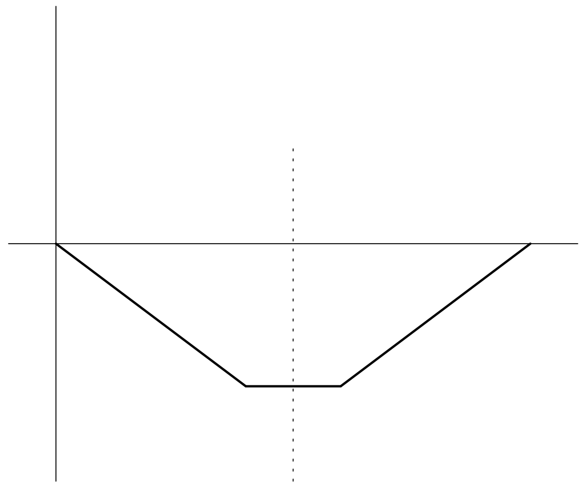

FIGURE 1. The dark line is an example of the Newton polygon of an isocrystal $b$ with additional $U_{10}^*$-structure. The horizontal line from $(0,0)$ to $(10,0)$ is the Newton polygon of the basic isocrystal. The vertical dotted line indicates the mirror symmetry of the Newton polygons of the $G$-isocrystals.

We now study the set $B(G)$ where $G$ is an unramified unitary group over $F$ splitting over the extension $F_2/F$. The absolute root system of $G$ is isomorphic to the usual root system in $\mathbb{R}^n$ of type A (cf. Bourbaki [**11**, chap. 6]), and the non-trivial element of the group $\mathrm{Gal}(F_2/F)$ acts on $\mathbb{R}^n$ via the operator $\theta$ defined by $(x_1, x_2, \ldots, x_n) \longmapsto (-x_n, -x_{n-1}, \ldots, -x_1)$. The space $\mathfrak{a}_0$ is the subspace of $\theta$ invariant elements in $\mathbb{R}^n$, thus it is equal to the set of $(x_i) \in \mathbb{R}^n$ with $x_i = -x_{n+1-i}$ for all indices $i$. The dimension of this space is equal to $\lfloor n/2 \rfloor$.

Whenever $b \in B(G)$ is an isocrystal with $G$-structure, we have its slope morphism $\overline{\nu}_b \in C_0$. We may view the slope morphism $\overline{\nu}_b$ as an $\theta$-invariant element of $\mathbb{R}^n$. This way we get the slopes $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $b$. These slopes are just the coordinates of the vector $\overline{\nu}_b \in \mathbb{R}^n$. We order them so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. These slopes satisfy the property $\lambda_i = -\lambda_{n+1-i}$. We associate to these slopes the Newton polygon $\mathcal{G}_b$ of $b$. The Newton polygon is by definition the continuous piecewise linear function from the real interval $[0, n]$ to $\mathbb{R}$ with the property that the only points where it is possibly not differentiable are the integral points $[0, n] \cap \mathbb{Z}$; the value of $\mathcal{G}_b$ at these points is defined by: $\mathcal{G}_b(0) := 0$ and $\mathcal{G}_b(i) := \lambda_1 + \lambda_2 + \ldots + \lambda_i$. Observe that, due to the $\theta$-invariance, we have $\mathcal{G}_b(n) = \lambda_1 + \ldots + \lambda_n = 0$. Furthermore the graph (or polygon) $\mathcal{G}_b$ is symmetric around the vertical line that goes through the point $(\frac{n}{2}, 0)$. In Figure 1 we show a typical unitary Newton polygon. In particular negative slopes may occur, which does not happen for the group $\mathrm{GL}_n(F)$ nor for the group $\mathrm{Gsp}_{2g}(F)$.

Let us now determine what the Hodge polygons looks like. The minuscule cocharacter $\mu$ is defined over $\overline{F}$, and is given by

$$\mu = (\underbrace{0, 0, \dots, 0}_{n-s}, \underbrace{1, 1, \dots, 1}_{s}) \in \mathbb{Z}^n \subset \mathbb{R}^n,$$

for some integer $s$ with $0 \leq s \leq n$. To define the set $B(G, \mu)$ Kottwitz [**60**, §6] takes the average of $\mu$ under the Galois action to get

$$\overline{\mu} := \tfrac{1}{2}(\mu + \theta(\mu)) = (\underbrace{-\tfrac{1}{2}, -\tfrac{1}{2}, \dots, -\tfrac{1}{2}}_{s'}, \underbrace{0, 0, \dots, 0}_{n-2s'}, \underbrace{\tfrac{1}{2}, \tfrac{1}{2}, \dots, \tfrac{1}{2}}_{s'}) \in \mathfrak{a}_0 \subset \mathbb{R}^n,$$

where $s' := \min(s, n-s)$. To this element $\overline{\mu} \in \mathbb{R}^n$ we may associate in the same manner a graph $\mathcal{G}_\mu$ as in Figure 1. Then $b \in B(G)$ lies in $B(G, \mu)$ if and only if the end point of $\mathcal{G}_b$ is $(n, 0)$ and if $\mathcal{G}_b$ lies above [1] the graph $\mathcal{G}_\mu$.

## 2. PEL datum

Let $G/\mathbb{Q}$ be a unitary group of similitudes arising from a PEL type Shimura datum [**59**, §5]. We recall briefly the definition of $G$ from [*loc. cit.*]. Let $B/\mathbb{Q}$ be a finite dimensional simple algebra and write $F$ for its center, and assume that $F$ is a CM field. Let $*$ be a positive involution on $B$ over $\mathbb{Q}$ inducing on $F$ the complex conjugation. Write $F^+ \subset F$ for the fixed field of $*$ on $F$. Let $V$ be a nonzero finitely generated left $B$-module. Let $(\cdot,\cdot)$ be a non-degenerate $\mathbb{Q}$-valued alternating form on $V$ such that $(bv, w) = (v, b^* w)$ for all $v, w \in V$ and all $b \in B$. Then $G/\mathbb{Q}$ is the algebraic group with for all commutative $\mathbb{Q}$-algebras $R$:

$$G(R) = \left\{ g \in \mathrm{End}_B(V)^\times | \exists c(g) \in R^\times : (g\cdot, g\cdot) = c(g)(\cdot,\cdot) \text{ on V} \right\}. \tag{2.1}$$

Let $G_1 \subset G$ be the kernel of the similitudes ratio. Then $G_1$ is obtained by restriction of scalars of a unitary group $G_0$ defined over the totally real field $F^+$ (following the notations of [*loc. cit.*]. The group $G_{1,\mathbb{Q}_p}$ is isomorphic to a product of groups

$$G_{1,\mathbb{Q}_p} \cong \prod_{\wp | p} G_{1,\wp}, \tag{2.2}$$

where $\wp$ ranges over the $F^+$-places above $p$, and where the group $G_{1,\wp}$ is either the restriction of scalars to $\mathbb{Q}_p$ of $\mathrm{GL}_{n,F_\wp^+}$ or of an unramified unitary group over $F_\wp^+$. We will study the group $G_{1,\mathbb{Q}_p}$ factor by factor. Thus, in this chapter we will need to work not only with unramified unitary groups, but with the slightly more general class of groups of the form $\mathrm{Res}_{F'_\wp/F_\wp} U$, where $F'_\wp/F_\wp$ is some unramified extension and $U$ is an unramified unitary group over $F'_\wp$. The study of isocrystals over these groups reduces quickly to the study of isocrystals over the group $U$ (which we did above), by the Shapiro bijection (cf. [**60**, 6.5.3]):

$$B(\mathrm{Res}_{F'_\wp/F_\wp} U) = B_{F'_\wp}(U), \tag{2.3}$$

1. Lies above in the non-strict sense, the two graphs may touch, or even be the same (the ordinary case).

where we have added the subscript "$F'_\wp$" in the right hand side to indicate that there we work with $\sigma'$-conjugacy classes, where $\sigma'$ is the arithmetic Frobenius of $\overline{\mathbb{Q}}_p$ over $F'_\wp$. Under the Shapiro bijection the subset $B(\mathrm{Res}_{F'_\wp/F_\wp}U, \mu_\wp)$ corresponds to the subset $B_{F'_\wp}(U, \mu'_\wp)$ of $B_{F'_\wp}(U)$, where $\mu'_\wp$ is defined by

$$\mu'_\wp \stackrel{\text{def}}{=} \sum_{v \in V(\wp)} (\underbrace{1, 1, \dots, 1}_{s_v}, \underbrace{0, 0, \dots, 0}_{n-s_v}) \in \mathbb{Z}^n.$$

Thus, the combinatorics for isocrystals with $\mathrm{Res}_{F'/F}U$-structure is almost the same as the combinatorics for the case $F' = F$; only the Hodge polygons are slightly more complicated.

We recall briefly how the functions of Kottwitz $\phi_\alpha$ and $f_\alpha$ are constructed [**59**, §5] [**57**, p. 173]. Let $E$ be the reflex field, let $p$ be a prime number where the Shimura variety has good reduction in the sense of [**59**, §6]. In particular the field $E$ is unramified at $p$; let $\mathfrak{p}$ be an $E$-prime above $p$. Write $E_\mathfrak{p}$ for the completion of $E$ at $\mathfrak{p}$, fix an embedding $E_\mathfrak{p} \subset \overline{\mathbb{Q}}_p$ and let for each positive integer $\alpha$, the field $E_{\mathfrak{p},\alpha} \subset \overline{\mathbb{Q}}_p$ be the unramified extension of $E_\mathfrak{p}$ of degree $\alpha$. In the PEL datum there is fixed a $*$-morphism $h\colon \mathbb{C} \to \mathrm{End}(B)^{\mathrm{opp}}_\mathbb{R}$. This morphism induces a morphism of algebraic groups from Deligne's torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_\mathrm{m}$ to the group $G_\mathbb{R}$. Tensor this morphism with $\mathbb{C}$ to get a morphism from $\mathbb{G}_\mathrm{m} \times \mathbb{G}_\mathrm{m}$ to $G_\mathbb{C}$ and then restrict to the factor $\mathbb{G}_\mathrm{m}$ of the product $\mathbb{G}_\mathrm{m} \times \mathbb{G}_\mathrm{m}$ corresponding to the identity $\mathbb{R}$-isomorphism $\mathbb{C} \to \mathbb{C}$. This way we obtain a cocharacter $\mu \in X_*(G)$. We quote from Kottwitz's article at Ann Arbor, p. 173: The $G(\mathbb{C})$ conjugacy class of $\mu$ gives a $G(\overline{\mathbb{Q}}_p)$ conjugacy class of morphisms fixed by the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_{\mathfrak{p},\alpha})$. Let $S_\alpha$ be a maximal $E_{\mathfrak{p},\alpha}$-split torus in the group $G$ over the ring of integers $\mathcal{O}_{E_{\mathfrak{p},\alpha}}$. Using Lemma (1.1.3) of [**54**] we choose $\mu$ so that it factors through $S_\alpha$. Then $\phi_\alpha = \phi_{G,\mu,\alpha}$ is the characteristic function of the double coset $G(\mathcal{O}_{\mathfrak{p},\alpha})\mu(p^{-1})G(\mathcal{O}_{\mathfrak{p},\alpha})$. The function $f_\alpha = f_{G,\mu,\alpha}$ is by definition the base change [**4**, **77**] of $\phi_\alpha$ from the group $G(E_{\mathfrak{p},\alpha})$ to the group $G(\mathbb{Q}_p)$.

## 3. Truncated traces

We revert to the general notations of the beginning of the first section, thus $G$ is a connected unramified reductive group over a local field. In this section we introduce the concept of truncated traces of smooth representations with respect to elements of the set $B(G)$, i.e. the isocrystals with additional $G$-structure. We will then compute these truncated traces on the Steinberg representation and on the trivial representation.

Using the mapping $\Phi$ from the previous section we define the truncated traces with respect to an arbitrary element $b \in B(G)$:

Definition 3.1. Let $\nu \in C_0$. We define:

$$\Omega^G_\nu \stackrel{\text{def}}{=} \{g \in G \mid \exists \lambda \in \mathbb{R}_{>0} :\ \Phi(g) = \lambda \cdot \nu \in C_0\}. \tag{3.1}$$

We let $\chi^G_\nu$ be the characteristic function on of the subset $\Omega^G_\nu$ of $G$. Let $b \in B(G)$ be an isocrystal with additional $G$-structure. Then we will write $\chi^G_b := \chi^G_{\overline{\nu}_b}$ and $\Omega^G_b := \Omega^G_{\overline{\nu}_b}$.

REMARK. The Newton mapping $B(G) \ni b \mapsto \overline{\nu}_b \in C_0$ is *injective* for a simply connected, connected quasi-split reductive group over a non-Archimedean local field [**60**, §6].

Let $P = MN$ be a standard parabolic subgroup of $G$ and let $A_P$ be the split center of $P$, we write $\varepsilon_P = (-1)^{\dim(A_P/A_G)}$. To the parabolic subgroup $P$ we associate the subset $\Delta_P \subset \Delta$ consisting of those roots acting non trivially on $A_P$. Define $\mathfrak{a}_P$ to be $X_*(A_P)_\mathbb{R}$ and define $\mathfrak{a}_P^G$ to be the quotient of $\mathfrak{a}_P$ by $\mathfrak{a}_G$, and define $\mathfrak{a}_P^+$ by

$$\mathfrak{a}_P^+ := \{x \in \mathfrak{a}_P \mid \text{for all roots } \alpha \text{ in } \Delta_P\text{: } \langle x, \alpha\rangle > 0\}.$$

We recall the definition of the obtuse and acute Weyl-chambers [**68**,**102**]. Let $P$ be a standard parabolic subgroup of $G$. We write $\mathfrak{a}_0 = \mathfrak{a}_{P_0}$ and $\mathfrak{a}_0^G = \mathfrak{a}_{P_0}^G$. For each root $\alpha$ in $\Delta$ we have a coroot $\alpha^\vee$ in $\mathfrak{a}_0^G$. For $\alpha \in \Delta_P \subset \Delta$ we send the coroot $\alpha^\vee \in \mathfrak{a}_0^G$ to the space $\mathfrak{a}_P^G$ via the canonical surjection $\mathfrak{a}_0^G \twoheadrightarrow \mathfrak{a}_P^G$. The set of these restricted coroots $\alpha^\vee|_{\mathfrak{a}_P^G}$ with $\alpha$ ranging over $\Delta_P$ form a basis of the vector space $\mathfrak{a}_P^G$. By definition the set of fundamental weights $\{\varpi_\alpha \in \mathfrak{a}_P^{G*} \mid \alpha \in \Delta_P\}$ is the basis of $\mathfrak{a}_P^{G*} = \mathrm{Hom}(\mathfrak{a}_P^G, \mathbb{R})$ dual to the basis $\{\alpha^\vee_{\mathfrak{a}_P^G}\}$ of coroots. We let $\tau_P^G$ be the characteristic function on the space $\mathfrak{a}_P^G$ of the *acute Weyl chamber*,

$$\mathfrak{a}_P^{G+} = \left\{x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \alpha, x\rangle > 0\right\}. \tag{3.2}$$

We let $\widehat{\tau}_P^G$ be the characteristic function on $\mathfrak{a}_P^G$ of the *obtuse Weyl chamber*,

$${}^+\mathfrak{a}_P^G = \left\{x \in \mathfrak{a}_P^G \mid \forall \alpha \in \Delta_P \ \langle \varpi_\alpha^G, x\rangle > 0\right\}. \tag{3.3}$$

We define the function $\chi_N$ to be the composition $\tau_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$, and we define the function $\widehat{\chi}_N$ to be the composition $\widehat{\tau}_P^G \circ (\mathfrak{a}_P \twoheadrightarrow \mathfrak{a}_P^G) \circ H_M$. The functions $\chi_N$ and $\widehat{\chi}_N$ are locally constant and $K_M$-invariant, where $K_M = M(\mathcal{O}_F)$.

Let $b \in B(G)$ be an isocrystal with additional $G$ structure and let $\overline{\nu}_b \in C_0$ be its slope morphism. For any standard parabolic subgroup $P \subset G$ we have the subset $\mathfrak{a}_P^+ \subset C_0$. Let $P_b$ be the standard parabolic subgroup of $G$ such that $\overline{\nu}_b \in \mathfrak{a}_{P_b}^+$. We call the group $P_b$ the subgroup of $G$ *contracted* by the isocrystal $b \in B(G)$. These groups are precisely the parabolic subgroups appearing in the Kottwitz decomposition of the set $B(G)$ (see [**60**, 5.1.1]). We write $P_b = M_b N_b$ for the standard decomposition of $P_b$.

We write $\pi_{0P}$ for the projection from the space $\mathfrak{a}_0$ onto $\mathfrak{a}_P$, it sends an element $X \in \mathfrak{a}_0$ to its average under the action of the Weyl group.

We introduce a certain characteristic function on $G$ associated to the isocrystal $b \in B(G)$:

DEFINITION 3.2. Let $P_b = M_b N_b$ be the standard parabolic subgroup of $G$ contracted by $b$. We define $\eta_b$ to be the characteristic function on $G$ of the set of elements $g \in G$ such that there exists a $\lambda \in \mathbb{R}_{>0}^\times$ such that $\pi_{0P}(\Phi(g)) = \lambda \cdot \overline{\nu}_b \in \mathfrak{a}_P^+$.

REMARK. If the isocrystal $b$ is basic, then we have $P = G$, and the element $\overline{\nu}_b \in C_0$ is central. Therefore the function $\eta_b$ is *spherical*.

In case the isocrystal $b \in b(G)$ is basic then $\chi_b^G$ coincides with $\eta_b \chi_c^G$:

LEMMA 3.3. *Let $b \in B(G)$ be a basic isocrystal. Then we have $\chi_b^G = \eta_b \chi_c^G$.*

PROOF. Let $g \in G$, and consider $\Phi(g) \in C_0$. Then $g$ is compact if and only if it contracts $G$ as parabolic subgroup (which means that $\Phi(g)$ lies in $\mathfrak{a}_G \subset C_0^+$). Assume $g$ is compact. Then $\chi_b^G(g) = 1$ if and only if the slope morphism $\overline{\nu}_b$ of $b$ lies in $\mathfrak{a}_G$, i.e. if and only if the centralizer of the slope morphism of $b$ is equal to $G$. But that means that $b$ is basic. Conversely, assume $b$ is basic. Then its slope morphism is central, thus $\chi_c^G(g) = 1$ if and only if $g$ contracts $G$, i.e. $g$ is compact. Furthermore we have $\eta_b(g) = 1$ because $\Phi(g)$ equals $\overline{\nu}_b$ up to a positive scalar. This completes the proof. □

We call the collection of subsets $\Omega_b^G$ for $b \in B(G)$ the *Newton polygon stratification* of the group $G$. For our proofs we will also need to study another stratification, called the *Casselman stratification* of $G$:

DEFINITION 3.4. Let $Q$ be a standard parabolic subgroup of $G$. We define $\Omega_Q^G \subset G$ to be the subset of elements $g \in G$ contracting [**22**, §1] a parabolic subgroup conjugate to $Q$. Write $\chi_Q^G$ for the characteristic function on $G$ of the subset $\Omega_Q^G \subset G$. These sets $\Omega_Q^G$ form the Casselman stratification of $G$.

For truncated traces with respect to the Casselman stratification we have:

PROPOSITION 3.5. *Let $Q = LU$ be a standard parabolic subgroup of $G$. Let $f \in \mathcal{H}(G)$ be a locally constant function with compact support. Then we have* $\mathrm{Tr}(\chi_Q^G f, \pi) = \mathrm{Tr}(\chi_U \chi_c^L \overline{f}^{(Q)}, \pi_U(\delta_Q^{-1/2}))$.

PROOF. By the Proposition [**22**, prop 1.1] on compact traces, for all functions $f$ on $G$, the full trace $\mathrm{Tr}(f, \pi)$ is equal to the sum of compact traces $\sum \mathrm{Tr}_M(\chi_c^M \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}))$, where the sum ranges over the standard parabolic subgroups $P = MN$ of $G$. Consider only those functions of the form $\chi_Q^G f \in \mathcal{H}(G)$. Then we obtain that the trace $\mathrm{Tr}(\chi_Q^G f, \pi)$ is equal to the sum $\sum \mathrm{Tr}_M(\chi_c^M \chi_Q^G \overline{f}^{(P)}, \pi_N(\delta_P^{-1/2}))$ where $P = MN$ ranges over the standard parabolic subgroups of $G$. Observe that $\chi_c^M \chi_Q^G = 0$ if $P \neq Q$. Therefore only the term corresponding to $P = Q$ remains in the sum. This completes the proof. □

Let us now explain the relation between the Casselman stratification and the Newton stratification. The following Proposition gives the relation between the Casselman stratification of $G$ and the Newton stratification:

PROPOSITION 3.6. *For all $b \in B(G)$ we have $\Omega_b^G \subset \Omega_{P_b}^G$.*

PROOF. Assume that $g \in \Omega_b^G$. Then $\Phi(g) = \lambda \overline{\nu}_b \in \mathfrak{a}_0$. Let $P$ be the standard parabolic subgroup of $G$ conjugate to the parabolic subgroup of $G$ contracted by $g$. Then $\overline{\nu}_b = \lambda \Phi(g) \in \mathfrak{a}_P^+$. Then, by definition, $P$ is the parabolic subgroup contracted by $b$. This completes the proof. □

Example. The inclusion $\Omega_b^G \subset \Omega_{P_b}^G$ is strict in general. Consider for example the case $G = \mathrm{GL}_{n,\mathbb{Q}_p}$ to see that it is non-strict only in particular cases, such as when $n = 2$. In the particular case of the Shimura varieties of Harris-Taylor [**45**], the Casselman stratification also separates the isocrystals.

We will now compute the truncated trace on the Steinberg representation.

Definition 3.7. Let $\xi_b^{St}$ be the characteristic function on $T$ defined by $\xi_b^{St} := \widehat{\chi}_{N_0 \cap M_b} \chi_{N_b} \eta_b$, where with the notation $\widehat{\chi}_{N_0 \cap M_b}$ we mean the characteristic function on the Levi subgroup $M_b \subset G$, corresponding to the obtuse chamber relative to the minimal parabolic subgroup of $M_b$.

Proposition 3.8. *Let $f \in \mathcal{H}_0(G)$ be a spherical Hecke operator. Then we have*

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \varepsilon_{P_0 \cap M_b} \mathrm{Tr}_T(\xi_b^{St} f^{(P_0)}, \mathbf{1}(\delta_{P_b}^{-1/2} \delta_{P_0 \cap M_b}^{1/2})). \tag{3.4}$$

Proof. Write $P = MN$ for the parabolic subgroup contracted by the isocrystal $b$. We compute:

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \mathrm{Tr}_M(\chi_b^G \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2})), \tag{3.5}$$

(Proposition 3.5). Let $b_M \in B(M)$ be a $G$-regular basic element such that its image in $B(G)$ is equal to $b$ [**55**, prop. 6.3]. By [*loc. cit.*] the set of all such $b_M$ are $G$-conjugate. As functions on $M$ we have $\chi_b^G \chi_N = \chi_{b_M}^M \chi_N$. Therefore we may simplify Equation (3.5) to $\mathrm{Tr}_M(\chi_{b_M}^M \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2}))$. By Lemma 3.3 the latter trace equals $\mathrm{Tr}_M(\chi_c^M \eta_b \chi_N f^{(P)}, (\mathrm{St}_G)_N(\delta_P^{-1/2}))$. In Chapter 2 we computed the compact traces on the Steinberg representation for all spherical Hecke operators. By Proposition 2 1.13 we get

$$\mathrm{Tr}(\chi_b^G f, \mathrm{St}_G) = \varepsilon_{P_0 \cap M} \mathrm{Tr}(\widehat{\chi}_{N_0 \cap M} \eta_b \chi_N f^{(P_0)}, \mathbf{1}(\delta_P^{-1/2} \delta_{P_0 \cap M}^{1/2})). \tag{3.6}$$

This completes the proof. □

In the same way one may compute the truncated traces on the trivial representation. We have to introduce two more notations. Let $\widehat{\chi}_{N_0 \cap M_b}^{\leq}$ be the characteristic function on $M_b$ corresponding to the negative closed obtuse chamber in $\mathfrak{a}_P$. Then we define:

Definition 3.9. Let $b \in B(G)$ be an isocrystal. We define $\xi_b^{\mathbf{1}} := \widehat{\chi}_{N_0 \cap M_b}^{\leq} \chi_{N_b} \eta_b$.

Proposition 3.10. *We have* $\mathrm{Tr}(\chi_b^G f, \mathbf{1}) = \mathrm{Tr}_T(\xi_b^{\mathbf{1}} f^{(P_0)}, \mathbf{1}(\delta_{P_0}^{-1/2}))$.

Proof. The proof of Proposition 3.8 may be repeated without change up to Equation (3.6). Replace the result in that last Equation with the result from Proposition 3.1 from [**64**]. This Proposition gives the compact trace on the trivial representation for any Hecke operator (and any unramified group). □

Remark. With a method similar to the above one may compute the truncated traces on the irreducible subquotients of the $G$-representation on the space $C^\infty(G/P_0)$ of locally constant functions on $G/P_0$.

## 4. The class of $\mathfrak{R}(b)$-representations

For the global applications to Shimura varieties we find a class representations $\mathfrak{R}_1(b)$ of positive Plancherel density on which the truncated trace of the Kottwitz functions are non-zero. In fact we take for *most* of the isocrystals $b \in B(G, \mu)$ simply the Steinberg representation at $p$, but there are some exceptions where the truncated trace on the Steinberg representation vanishes; in those cases we take a different representation.

Let $G$ be a connected, reductive unramified group over $\mathbb{Q}_p$, let $P_0$ be a Borel subgroup of $G$. Let $T$ be the Levi-component of $P_0$. Then $T$ is a maximal torus in $G$, and let $W$ be the absolute Weyl group of $T$ in $G$. Let $\mu \in X_*(T)$ be a minuscule cocharacter.

We write in this section $E$ for an arbitrary, finite unramified extension of $\mathbb{Q}_p$. In later sections, the field $E$ that we consider here will be the completion of the reflex field at a prime of good reduction. We fix an embedding of $E$ into $\overline{\mathbb{Q}}_p$, and for each positive integer $\alpha$ we write $E_\alpha \subset \overline{\mathbb{Q}}_p$ for the unramified extension of degree $\alpha$ of $E$.

Definition 4.1. (cf. [**54**]). Let $\alpha$ be a positive integer, and $E_\alpha$ the unramified extension of $E$ of degree $\alpha$ contained in $\overline{\mathbb{Q}}_p$. We write $W_\alpha$ for the subgroup $W(G(E_\alpha), T(E_\alpha))$ of $W$. Write $S_\alpha$ for a maximal $E_\alpha$-split subtorus of $G_{E_\alpha}$. We define $\phi_{G,\mu,\alpha} \in \mathcal{H}_0(G(E_\alpha))$ to be the spherical function whose Satake transform is equal to

$$p^{-\alpha\langle\rho_G,\mu\rangle} \sum_{w \in W_\alpha/\mathrm{stab}_{W_\alpha}(\mu)} [w(\mu)] \in \mathbb{C}[X_*(S_\alpha)]^{W_\alpha}, \tag{4.1}$$

where $\mathrm{stab}_{W_\alpha}(\mu) \subset W_\alpha$ is the stabilizer of $\mu$ in the group $W_\alpha$. We define $f_{G,\mu,\alpha}$ to be the function obtained from $\phi_{G,\mu,\alpha}$ via base change from the group $G(E_\alpha)$ to the group $G(F^+)$. We call $f_{G,\mu,\alpha}$ the *function of Kottwitz*.

Remark. Kottwitz proves in [**54**] that the definition of the Kottwitz functions $f_{G,\mu,\alpha}$ and $\phi_{\mathbb{G},\mu,\alpha}$ coincide with the definition that we gave at the end of section 2.

Remark. We note that the notation for the functions $f_{G,\mu,\alpha}$ and $\phi_{G,\mu,\alpha}$ is slightly abusive, as they also depend on the field $E$. Because confusion will not be possible, we have decided to drop the field $E$ from the notations.

Proposition 4.2. *Let $P = MN$ be a standard parabolic subgroup of $G$. We have*

$$f^{(P)}_{G,\mu,\alpha} = q^{-\alpha\langle\rho_G - \rho_M,\mu\rangle} \sum_{w \in W_\alpha/\mathrm{stab}_{W_\alpha}(\mu)W_{M,\alpha}} f_{M,w(\mu),\alpha} \in \mathcal{H}_0(M),$$

*where* $\mathrm{stab}_{W_\alpha}(\mu)W_{M,\alpha} \subset W_\alpha$ *is the subgroup of* $W_\alpha$ *generated by the group* $W_{M,\alpha}$ *of the Weyl group of* $T(E_\alpha)$ *in* $M(E_\alpha)$ *and the stabilizer subgroup of* $\mu$ *in* $W_\alpha$.

Proof. Compute the Satake transform of both sides to see that they are equal. □

The integer $\alpha$ will later be the degree of the finite field over which we will count points in the Newton stratum. In this chapter we only want to show that the Newton-strata are non-empty. Therefore, we will take $\alpha$ large so that the combinatorial problems simplify (large in the divisible sense).

We make the function of Kottwitz explicit in case $G$ is either the restriction of scalars of a general linear group over $F^+$ or the restriction of scalars of an unramified unitary group over $F^+$. From this point onwards we assume that we are in one of the following two cases:

$$G = \begin{cases} \mathrm{Res}_{F^+/\mathbb{Q}_p}(\mathrm{GL}_{n,F^+}) & \text{(linear type)} \\ \mathrm{Res}_{F^+/\mathbb{Q}_p}(U) & \text{(unitary type)} \end{cases} \tag{4.2}$$

where $F^+/\mathbb{Q}_p$ is a finite unramified extension, and where $U/F^+$ is an unramified unitary group, outer form of $\mathrm{GL}_{n,F^+}$. These groups $G$ occur as the components in the product decomposition in Equation (2.2). We assume that the cocharacter $\mu \in X_*(T)$ arises from a PEL-type datum, as we have explained in the discussion below Equation (2.1).

We begin with the linear case. We have a cocharacter $\mu \in X_*(T)$ (see below Equation (2.2)). Thus, for each $\mathbb{Q}_p$-embedding $v$ of $F^+$ into $\overline{\mathbb{Q}}_p$ we get a cocharacter $\mu_v$ of the form

$$(\underbrace{1, 1, \ldots, 1}_{s_v}, \underbrace{0, 0, \ldots, 0}_{n-s_v}) \in \mathbb{Z}^n.$$

To each such integer $s_v$ we associate the spherical function $f_{n\alpha s_v}$ on $\mathrm{GL}_n(F^+)$ whose Satake transform is defined by

$$\mathcal{S}(f_{n\alpha s_v}) = q^{\frac{s(n-s)}{2}\alpha} \sum_{i_1 < i_2 < \ldots < i_{s_v}} X_{i_1}^\alpha X_{i_2}^\alpha \cdots X_{i_{s_v}}^\alpha \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]. \tag{4.3}$$

We write $V_\alpha$ for the set of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/E_\alpha)$-orbits in the set $\mathrm{Hom}(F^+, \overline{\mathbb{Q}}_p)$. If $v \in V_\alpha$ is such an orbit, then this orbit corresponds to a certain finite unramified extension $E_\alpha[v]$ of $E_\alpha$. Let $\alpha_v$ be the degree over $\mathbb{Q}_p$ of the field $E_\alpha[v]$, we then have $E_\alpha[v] = E_{\alpha_v}$. The function $f_\alpha$ is given by

$$f_\alpha = \prod_{v \in V_\alpha} f_{n\alpha_v s_v}^{\mathrm{GL}_n(F^+)} \in \mathcal{H}_0(G(\mathbb{Q}_p)), \tag{4.4}$$

where the product is the convolution product (cf. Proposition 2.3.3).

Let us now assume that we are in the unitary case (cf. Equation (4.2)). We will make the function $f_{G,\mu,\alpha}$ explicit only in case $\alpha$ is even. To obtain the function of Kottwitz on $G$, we have to apply base change from $G(E_\alpha)$ to $G(\mathbb{Q}_p)$. Assume that $\alpha$ is even. Let $\mathbb{Q}_{p^2}$ be the quadratic unramified extension of $\mathbb{Q}_p$ contained in $\overline{\mathbb{Q}}_p$. The base change factors over the composition of base changes $G(E_\alpha) \rightsquigarrow G(\mathbb{Q}_{p^2}) \rightsquigarrow G(\mathbb{Q}_p)$. The base change of $\phi_\alpha$ to $G(\mathbb{Q}_{p^2})$ is a function of the form $f_{G^+,\mu,(\alpha/2)}$ on the group $G^+ = \mathrm{Res}_{\mathbb{Q}_{p^2}/\mathbb{Q}_p}(G_{\mathbb{Q}_{p^2}})$. Explicitly, the last quadratic base change $G(\mathbb{Q}_{p^2}) \rightsquigarrow G(\mathbb{Q}_p)$ is given by:

$$\Psi\colon \mathbb{C}[X_1^{\pm 1}, \dots, X_n^{\pm 1}]^{\mathfrak{S}_n} \longrightarrow \mathbb{C}[X_1^{\pm 1}, \dots, X_n^{\pm 1}]^{\mathfrak{S}_m \rtimes (\mathbb{Z}/2\mathbb{Z})^m},$$

$$X_i \longmapsto \begin{cases} X_i & 1 \le i \le \lfloor n/2 \rfloor, \\ 1 & i = \lfloor \frac{n}{2} + 1 \rfloor, \text{ and } n \text{ is odd}, \\ X_{n+1-i}^{-1} & n+1-\lfloor n/2 \rfloor \le i \le n, \end{cases} \tag{4.5}$$

where $m := \lfloor \frac{n}{2} \rfloor$ (cf. [**77**]). Thus we get $f_{G,\mu,\alpha} = \Psi f_{G^+,\mu,(\alpha/2)}$.

Lemma 4.3. *Let $G$ be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2). Let $\pi$ be a generic unramified representation of $G$, and $f = f_{G,\mu,\alpha}$ a function of Kottwitz, and $b \in B(G)$ an isocrystal. Let $\alpha \in \mathbb{Z}_{>0}$ be an integer, sufficiently divisible such that $W_\alpha$ is the absolute the Weyl group of $T$ in $G$. Then, the truncated trace $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \pi)$ is non-zero if and only if there exists some $w \in W_\alpha$ and some $\lambda \in \mathbb{R}_{>0}^\times$ such that $w(\mu) = \lambda \overline{\nu}_b \in \mathfrak{a}_0^G$.*

Remark. In case $G$ is the general linear group, then there exists a pair $w \in W, \lambda \in \mathbb{R}_{>0}^\times$ such that $w(\mu) = \lambda \overline{\nu}_b$ if and only if the slopes $\lambda_i$ of $b$ all lie in the set $\{0, 1\}$.

Proof. We have $\pi = \mathrm{Ind}_T^G(\rho)$, where $\rho$ is some smooth character of the torus $T$. By van Dijk's formula for truncated traces (Proposition 2.1.1) we have $\mathrm{Tr}(\chi_b^G f, \pi) = \mathrm{Tr}(\chi_b^G f^{(P_0)}, \rho)$. The truncation operation $h \mapsto \chi_b^G h$ on $\mathcal{H}_0(T)$, corresponds via the Satake transform to an operation on $\mathbb{C}[X_*(T)]$ sending certain monomials $[M] \in \mathbb{C}[X_*(T)]$ associated to elements $M \in X_*(T)$ to zero, and leaves certain other monomials invariant. Thus to compute the trace $\mathrm{Tr}(\chi_b^G f^{(P_0)}, \rho)$ one takes the set of monomials $[w(\mu)]$, $w \in W$ occurring in $f^{(P_0)}$, and removes some of them (maybe all), and then evaluate those that are left at the Hecke matrix of $\rho$. The lemma now follows from the observation that $\chi_b^G \mathcal{S}_T^{-1}[w(\mu)] \neq 0$ if and only if $w(\mu) = \lambda \overline{\nu}_b$ for some positive scalar $\lambda \in \mathbb{R}_{>0}^\times$. This completes the proof. □

We have to distinguish further between (essentially) two cases at $p$. The case the group is the general linear group, and the case where the group is the unramified unitary group. We begin with the general linear group.

Proposition 4.4. *Let $G$ be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2), and assume it is of linear type, so $G(\mathbb{Q}_p) = \mathrm{GL}_n(F^+)$. Let $b \in B(G, \mu)$ be a $\mu$-admissible isocrystal having the property that the number of slopes equal to $0$ is at most $1$, and the number of slopes equal to $1$ is also at most $1$. Let $\chi$ be an unramified character of $\mathrm{GL}_n(F^+)$. Then, for $\alpha$ sufficiently divisible, we have $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \mathrm{St}_G(\chi)) \neq 0$.*

Remark. In the proof of the Proposition we use the divisibility of $\alpha$ at two places. First, it simplifies the function of Kottwitz (cf. Equation (4.4)). Second, we want $\alpha$ sufficiently divisible so that the Weyl group $W(T(E_\alpha), G(E_\alpha))$ relative to the field $E_\alpha$ is the full Weyl group.

REMARK. In case the isocrystal $b$ has two or more slopes with value 0 (or 1), then the truncated trace of the Kottwitz function on the Steinberg representation vanishes.

PROOF. By Proposition 3.8 we have to show that the function $\xi_b^{\mathrm{St}} f_{G,\mu,\alpha}^{(P_0)}$ does not vanish. Recall that the function $f_{G,\mu,\alpha}$ is obtained from a function $\phi_\alpha$ through base change from the group $\mathrm{GL}_n(F^+ \otimes E_\alpha)$. Let us first assume that the $E_\alpha$-algebra $F^+ \otimes E_\alpha$ is a field. In that case we have that $f_{G,\mu,\alpha} = f_{n\alpha s}$ in the notations from Chapter 2, i.e. $\mathcal{S}(f_{G,\mu,\alpha})$ is (up to scalar) an elementary symmetric function in the Satake algebra,

$$\mathcal{S}(f_{G,\mu,\alpha}) = q^{\frac{s_v(n-s_v)}{2}\alpha} \sum_{i_1 < i_2 < \ldots < i_s} X_{i_1}^{d\alpha} X_{i_2}^{d\alpha} \cdots X_{i_s}^{\alpha} \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]. \tag{4.6}$$

We have to show that under the truncation operation $h \mapsto \xi_b^{\mathrm{St}} h$ on $\mathcal{H}(T)$ at least one of the monomials remains in Equation (4.6). Observe that the scalars in front of the monomials in Equation (4.6) all have the same sign, and that to get the truncated trace on the Steinberg representation we evaluate these monomials at a certain, nonzero point. Thus, the only problem is to see that there is at least one monomial $X$ occurring in $\mathcal{S}(f_{G,\mu,\alpha})$ and surviving the truncation $X \mapsto \xi_b^{\mathrm{St}} X$. At this point it will be useful to give a graphical interpretation of this truncation process.

A remark on the notation: With $\xi_b^{\mathrm{St}} X$ for $X$ a monomial in the Satake algebra of $T$, we mean the element $\mathcal{S}_T(\xi_b^{\mathrm{St}} \mathcal{S}_T^{-1}(X))$ of the Satake algebra of $T$. Below we will use similar conventions for the truncations $\chi_N X$, $\widehat{\chi}_{N_0 \cap M_b} X$ and $\eta_b X$.

A *graph* in $\mathbb{Z}^2$ is a sequence of points $\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_r$ with $\vec{v}_{i+1} - \vec{v}_i = (1, e)$, where $e$ is an integer. To a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n} \in \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots, X_n^{\pm 1}]$, with $e_i \in \mathbb{Z}$ and $\sum_{i=1}^n e_i = s$ we associate the graph $\mathcal{G}_X$ with points

$$\vec{v}_0 := (0,0), \quad \vec{v}_i := \vec{v}_0 + (i, e_n + e_{n-1} + \ldots + e_{n+1-i}) \in \mathbb{Z}^2, \tag{4.7}$$

for $i = 1, \ldots, n$. Because the sum $\sum_{i=1}^n e_i$ is equal to $s$, we see that the end point of the graph is $(n, s)$. The function $f_{G\alpha\mu}$ is (up to scalar) the elementary symmetric function of degree $s$ in $n$ variables, thus its monomials correspond precisely to the set of graphs that start at the point $(0,0)$, have end point $(n, s)$ and satisfy $\vec{v}_{i+1} - \vec{v}_i \in \{(1,0), (1,1)\}$ for all $i$.

To the slopes $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ of the isocrystal $b$ we associate the graph $\mathcal{G}_b$ with points

$$\vec{v}_0 := (0,0), \quad \vec{v}_i := \vec{v}_0 + (i, \lambda_1 + \lambda_2 + \ldots + \lambda_i) \in \mathbb{Z}^2, \tag{4.8}$$

for $i = 1, \ldots, n$. (Remark: To obtain the usual convex picture of the Newton polygon we had to invert the order of the vector $e_1, \ldots, e_n$ in Equation (4.7). Without the inversion we would be considering concave polygons.)

We may now explain the truncation $X \mapsto \xi_b^{St} X$ in terms of graphs. We have $\xi_b^{\mathrm{St}} X = X$ or $\xi_b^{\mathrm{St}} X = 0$. We claim that we have $\xi_b^{\mathrm{St}} X = X$ if the following conditions hold:

(**C1**) We have $\mathcal{G}_b(n) = \lambda \mathcal{G}_X(n)$ for some positive scalar $\lambda \in \mathbb{R}_{>0}$;

(**C2**) For every break point $x \in \mathbb{Z}^2$ of $\mathcal{G}_b$ the point $x$ also lies on the graph $\lambda \mathcal{G}_X$;

(**C3**) Outside the set of breakpoints of $\mathcal{G}_b$, the graph $\lambda\mathcal{G}_X$ lies strictly below the graph $\mathcal{G}_b$.

Thus, in short: $\mathcal{G}_X$ lies below $\mathcal{G}_b$ and the set of contact points between the two graphs is precisely the begin point, end point and the set of break points of $\mathcal{G}_b$. See also Chapter 3 this construction in an analogous context.

Remark. In the claim above we say "if" and not "if and only if". The conditions (**C1**), (**C2**) and (**C3**) are stronger than the condition $\xi_b^{\mathrm{St}} X = X$. In Lemmas 4.5, 4.6 and 4.7 below we give conditions (**C1'**), (**C2'**) and (**C3'**) which, when taken together, are equivalent to "$\xi_b^{\mathrm{St}} X = X$". However (**C1**, **C2**, **C3**) is not equivalent to (**C1'**, **C2'**, **C3'**). If one would one replace condition (**C3'**) with the stronger condition

$$(\mathbf{C3''}) \quad \text{We have } \overline{\mathcal{G}}_x = \lambda\mathcal{G}_b \text{ for some } \lambda \in \mathbb{R}_{>0},$$

then we have $(\mathbf{C1}, \mathbf{C2}, \mathbf{C3}) \iff (\mathbf{C1'}, \mathbf{C2'}, \mathbf{C3''})$.

Because the above fact is crucial for the argument, let us prove the claim with complete details. Let $X = (e_1, e_2, \ldots, e_n) \in \mathbb{Z}^n = X_*(A_0)$. We want to express the condition $\xi_b^{\mathrm{St}} X = X$ in terms of $\mathcal{G}_X$. The Satake transform for the maximal torus $T = (\mathrm{Res}_{F^+/\mathbb{Q}_p}\mathbb{G}_{\mathrm{m}})^n$ is simply

$$\begin{aligned} \mathcal{H}_0(T) &\xrightarrow{\sim} \mathbb{C}[X_1^{\pm 1}, X_2^{\pm 1}, \ldots, X_n^{\pm 1}], \\ \mathbf{1}_{(p^{-e_1}\mathcal{O}_{F^+}^\times)\times(p^{-e_2}\mathcal{O}_{F^+}^\times)\times\cdots\times(p^{-e_n}\mathcal{O}_{F^+}^\times)} &\longmapsto X^{e_1}X^{e_2}\cdots X^{e_n}. \end{aligned} \tag{4.9}$$

We have $\xi_b^{\mathrm{St}} = \widehat{\chi}_{N_0\cap M_b}\chi_{N_b}\eta_b$. Let $(n_a)$ be the composition of $n$ corresponding to the standard parabolic subgroup $P_b$ of $G$. Let $g = (g_1, \ldots, g_n) \in T$ such that $\chi_{N_b}(g) = 1$. Explicitly, this means that

$$|g_1 g_2 \cdots g_{n_1}|^{1/n_1} < |g_{n_1+1}g_{n_1+2}\cdots g_{n_1+n_2}|^{1/n_2} < \ldots < |g_{n_{k-1}}g_{n_{k-1}+1}\cdots g_n|^{1/n_k}, \tag{4.10}$$

(cf. Equation (2.1.11)). In terms of the graph $\mathcal{G}_X$ of $X$ this means the following. We have $X \in \mathfrak{a}_0$ and we have the projection $\pi_{0,P_b}(X)$ of $X$ in $\mathfrak{a}_{P_b}$ (obtained by taking the average under the action of the Weyl group of $M_b$). We write $\overline{\mathcal{G}}_X$ for the graph of $\pi_{0,P_b}(X) \in \mathfrak{a}_{P_b} \subset \mathfrak{a}_0$. This graph $\overline{\mathcal{G}}_X$ is obtained from the graph $\mathcal{G}_X$ as follows. Consider the list of points

$$p_0 := (0,0), \quad p_1 := (n_1, \mathcal{G}_X(n_1)), \quad p_2 := (n_1+n_2, \mathcal{G}_X(n_1+n_2)), \quad \ldots \quad p_k := (n, \mathcal{G}_X(n)). \tag{4.11}$$

Connect, using a straight line, the point $p_0$ with $p_1$, and with another straight line, the point $p_1$ with $p_2$, etc, to obtain the graph $\overline{\mathcal{G}}_X$ from $\mathcal{G}_X$. From Equations (4.9) and (4.10) we get:

Lemma 4.5. *For a monomial $X$ we have $\chi_{N_b}X = X$ if and only if the following condition is true:*

(**C1'**) The graph $\overline{\mathcal{G}}_X$ is convex.

(Remark: We have $\chi_{N_b}X = 0$ if condition (**C1'**) is not satisfied. This remark also applies to Lemmas 4.6 and 4.7.)

Before discussing the function $\widehat{\chi}_{N_0 \cap M_b}$, let us first discuss in detail the maximal case, i.e. the function $\widehat{\chi}_{N_0}$ for the group $G$ (cf. Proposition 2.1.11). We have $\mathfrak{a}_0 = \mathbb{R}^n$, write $H_1, \ldots, H_n$ for the basis of $\mathfrak{a}_0^*$ dual to the standard basis of $\mathbb{R}^n$. Write $\alpha_i$ for root $H_i - H_{i+1}$ in $\mathfrak{a}_0^*$. We have

$$\varpi_{\alpha_i}^G = \left( H_1 + H_2 + \ldots + H_i - \frac{i}{n} \left( H_1 + H_2 + \ldots + H_n \right) \right) \in \mathfrak{a}_0^{G*}. \tag{4.12}$$

Thus, for a monomial $X = X_1^{e_1} X_2^{e_2} \cdots X_n^{e_n}$ the condition $\langle \varpi_{\alpha_i}^G, X \rangle > 0$ corresponds to

$$e_1 + e_2 + \ldots + e_i > \frac{i}{n} \left( e_1 + e_2 + \ldots + e_n \right) \tag{4.13}$$

Thus we obtain

$$\mathcal{G}_X(n + 1 - i) > \frac{i}{n} s, \tag{4.14}$$

where $s$ is the degree of $X$, i.e. $s = \sum_{i=1}^n e_i$. Demanding that $\langle \varpi_\alpha^G, X \rangle$ is positive for all roots $\alpha$ of $G$, is demanding that the graph $\mathcal{G}_X$ lies strictly below the straight line connecting the point $(0, 0)$ with the point $(n, \mathcal{G}_X(n))$. (We get 'below' and not 'above' due to the inversion "$e_i \mapsto e_{n+1-i}$" in Equation (4.7).)

We now turn to the function $\widehat{\chi}_{N_0 \cap M_b}$. The group $M_b$ decomposes into a product of general linear groups, say it corresponds to the composition $(n_a)$ of the integer $n$. Thus, the condition

$$\forall \alpha \in \Delta^{M_b} : \langle \varpi_\alpha^{M_b}, X \rangle > 0,$$

is the condition in Equation (4.13) but, then for each of the blocks of $M_b$ individually. The conclusion is :

LEMMA 4.6. *For any monomial $X$ we have $\widehat{\chi}_{N_0 \cap X_b} \cdot X = X$ if and only if the following condition is true:*

**(C2')** The graph $\mathcal{G}_X$ lies below $\overline{\mathcal{G}}_X$ and the two graphs touch precisely at the points $p_i$.

The condition "$\eta_b X = X$" means $\pi_{0,P_b}(\Phi(g))$ equals $\lambda \overline{\nu}_b$ for all $g$ lying in the support of the function $\mathcal{S}_T^{-1}(X)$ on the group $T$. By the explicit formula for the Satake transform (Equation (4.9)), the condition is equivalent to the existence of a permutation $w \in \mathfrak{S}_n$ such that the vector

$$\Bigg( \underbrace{e_{w(1)} + e_{w(2)} + \cdots}_{n_1}, \underbrace{e_{w(n_1+1)} + e_{w(n_1+2)} + \cdots}_{n_2}, \ldots, \underbrace{e_{w(n_1+n_2+\ldots+n_{k-1}+1)} + \cdots}_{n_k} \Bigg) \in \mathfrak{a}_{P_b},$$

is a positive scalar multiple of the vector $\overline{\nu}_b$. Using earlier notations we get:

LEMMA 4.7. *For any monomial $X$ we have $\eta_b X = X$ if and only if the following condition is true:*

**(C3')** There exists an element $w \in \mathfrak{S}_n$ such that $\overline{\mathcal{G}}_{w(X)} = \lambda \mathcal{G}_b$ for some $\lambda \in \mathbb{R}_{>0}$.

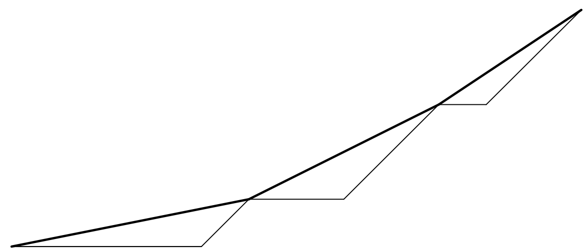

FIGURE 2. The dark line is an example of the Newton polygon of an isocrystal $b$ with additional $\mathrm{GL}_{12}(F^+)$-structure whose slope morphism is $(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3})$. The thin line is a $\xi_b^{\mathrm{St}}$-admissible path. For this Newton polygon there exist precisely two admissible paths. In general one takes the 'ordinary' path starting with horizontal steps within the blocks where the Newton polygon is of constant slope, and ending with diagonal steps.

To prove the claim we show that the group of conditions (**C1**), (**C2**) and (**C3**) implies the group of conditions (**C1'**), (**C2'**) and (**C3'**).

Thus, assume the conditions (**C1**), (**C2**) and (**C3**) are true for the monomial $X$. The parabolic subgroup $P_b$ is contracted by the isocrystal $b$. Thus the set of breakpoints of the polygon $\mathcal{G}_b$ is equal to the set

$$q_0 = (0,0), \quad q_1 = (n_1, \mathcal{G}_b(n_1)), \quad q_2 = (n_1+n_2, \mathcal{G}_b(n_1+n_2)) \quad \ldots \quad q_k = (n, \mathcal{G}_b(n)).$$

By condition (**C1**) there is a $\lambda \in \mathbb{R}_{>0}$ such that $\mathcal{G}_b(n) = \lambda \mathcal{G}_X(n)$. By conditions (**C2**) and (**C3**) the set $\{q_0, \ldots, q_n\}$ is then precisely the set of points where the graph $\lambda \mathcal{G}_X$ touches the graph $\mathcal{G}_b$. Taking averages, we get the relation $\overline{\mathcal{G}}_b = \lambda \overline{\mathcal{G}}_X$. We have $\mathcal{G}_b = \overline{\mathcal{G}}_b$ (because $P_b$ is associated to $b$), and therefore $\mathcal{G}_b = \lambda \overline{\mathcal{G}}_X$. Thus condition (**C3'**) is true for $w = \mathrm{Id} \in \mathfrak{S}_n$. The condition (**C2'**) is now implied by (**C2**) and (**C3**). Finally we prove condition (**C1'**). We have $\lambda \overline{\mathcal{G}}_X = \mathcal{G}_b$, and the graph $\mathcal{G}_b$ is convex. Thus $\overline{\mathcal{G}}_X$ is convex. The three conditions (**C1'**), (**C2'**) and (**C3'**) are now verified, and therefore the claim is true.

The monomials $M$ occurring in $\mathcal{S}(f_{G,\mu,\alpha})$ corresponds to the set of graphs from $(0,0)$ to $(n,s)$ whose steps consist of diagonal, north-eastward steps, or horizontal, eastward steps. Thus, it suffices to show that there exists a graph satisfying the conditions (**C1**), (**C2**) and (**C3**) above. This is indeed possible under the condition on the slopes of $\lambda_i$ of $b$ (see Figure 2 for the explanation). This completes the proof in case $F^+ \otimes E_\alpha$ is a field.

We now drop the assumption that the algebra $F^+ \otimes E_\alpha$ is a field. By Proposition 2.3.3 there exists a sufficiently large integer $M \geq 1$ such that for all degrees $\alpha$ divisible by $M$, the function $f_{G\mu\alpha}$ is (up to a scalar) a convolution product of the form $\prod_{i=1}^r f_{n\alpha s_i}$, where $r = [F^+ : \mathbb{Q}_p]$ and $(s_i)$ is a certain given composition of an integer $s$ of length $r$. Any monomial occurring in $\mathcal{S}(f_{n\alpha s})$ also occurs in the product $\prod_{i=1}^r \mathcal{S}(f_{n\alpha s_i})$ with a positive coefficient. Thus we may write $\prod_{i=1}^r f_{n\alpha s_i} = f_{n\alpha s} + R \in \mathcal{H}(G)$ for some function $R \in \mathcal{H}(G)$, whose Satake transform is a linear combination of monomials, with all coefficients positive. Consequently, to check that the truncated trace of $\prod_{i=1}^r f_{n\alpha s_i}$ on the Steinberg representation is non-zero, it suffices to check that the truncated trace of $f_{n\alpha s}$ on Steinberg is non-zero. This completes the proof. □

PROPOSITION 4.8. *Let $G$ be an algebraic group over $\mathbb{Q}_p$ defined as in Equation* (4.2), *and assume it is of linear type. Let $b \in B(G,\mu)$ be a $\mu$-admissible isocrystal. Let $m_0$ be the number of indices $i$ such that $\lambda_i = 0$, and let $m_1$ be the number of indices $i$ such that $\lambda_i = 1$. Write $m := n - m_0 - m_1$. Let $\pi_{m_0}$ (resp. $\pi_{m_1}$) be any generic unramified representation of $\mathrm{GL}_{m_0}(F^+)$ (resp. $\mathrm{GL}_{m_1}(F^+)$), and $\chi$ an unramified character of $\mathrm{GL}_m(F^+)$. Let $P$ be the standard parabolic subgroup of $G$ with 3 blocks, the first of size $m_1$, the second of size $m$ and the last one of size $m_3$. Then for $\alpha$ sufficiently divisible we have*

$$\mathrm{Tr}\left(\chi_b^G f_{G,\mu,\alpha}, \mathrm{Ind}_P^G\left(\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)}(\chi) \otimes \pi_{m_0}\right)\right) \neq 0.$$

REMARK. We have abused language slightly saying that $P$ has 3 blocks. We could have $m$, $m_0$ or $m_1$ equal to 0, in which case $P$ has less than 3 blocks. If one of the numbers $m, m_0$ or $m_1$ is 0, then one simply removes the corresponding factor from tensor product $\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)}(\chi) \otimes \pi_{m_0}$, and one induces from a parabolic subgroup with two blocks (or one block).

PROOF OF PROPOSITION 4.8. By van Dijk's formula for truncated traces (Proposition 2.1.5), we get a trace on $M$:

$$\mathrm{Tr}\left(\chi_b^G f_{G,\mu,\alpha}^{(P)}, \pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0}\right). \tag{4.15}$$

By Proposition 4.2 we have

$$f_{G,\mu,\alpha}^{(P)} = q^{-\alpha\langle\rho_G - \rho_M, \mu\rangle} \sum_{w \in W_{\alpha})/\mathrm{stab}_{W_{\alpha})}(\mu) W_{M,\alpha}} f_{M,w(\mu),\alpha} \in \mathcal{H}_0(M). \tag{4.16}$$

The intersection $\Omega_{\overline{\nu}_b}^G \cap M$ is equal to a union $\bigcup \Omega_{w(\overline{\nu}_b)}^M$ with $w$ ranging over the permutations $w \in W$ such that $w(\overline{\nu}_b)$ is $M$-positive. Consequently, if we plug Equation (4.16) into Equation (4.15), then we get a large sum, call it $(\star)$, of traces of functions $f_{M,w(\mu),\alpha}$ against a representation of the form $\pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0}$. All the signs are the same in this large sum $(\star)$, therefore it suffices that there is at least one non-zero term. Take $b_M \in B(M)$ the isocrystal whose slope morphism is $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ in the $M$-positive chamber of $\mathfrak{a}_0$. Then $b_M$ has only slopes 0 on the first block of $M$ and only slopes 1 on the third block, and all its slopes $\neq 0, 1$ are in the second block. The trace $\mathrm{Tr}(\chi_{b_M}^M f_{M,\mu,\alpha}, \pi_{m_1} \otimes \mathrm{St}_{\mathrm{GL}_m(F^+)} \otimes \pi_{m_0})$ occurs as a term in the expression $(\star)$. By Lemma 4.3 and Proposition 4.4 this term is non-zero. This completes the proof. □

We now establish the cases where the group is an unramified unitary group over $F^+$ (unitary type, cf. Equation (4.2)).

LEMMA 4.9. *Let $G$ be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2), and assume it is of unitary type. Let $b \in B(G, \mu)$ be an $\mu$-admissible isocrystal whose slope morphism $\overline{\nu}_b \in \mathfrak{a}_0$ has no coordinate equal to 0 and no coordinate equal to 1. Then, for $\alpha$ sufficiently divisible, the trace $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \mathrm{St}_{G(\mathbb{Q}_p)})$ is non-zero.*

PROOF. We use the explicit description $f_{G,\mu,\alpha} = \Psi f_{G^+,\mu,(\alpha/2)}$ of the Kottwitz function that we gave in Equation (4.5). Assume the algebra $F^+ \otimes E_\alpha$ is a field; then the base change mapping from $G(F_\alpha^+) \to G(F_2)$ is given by $X_i \mapsto X_i^{\alpha/2}$ on the Satake algebras. Over $F_\alpha^+$, the Weyl group $W_\alpha$ is equal to $\mathfrak{S}_n$ with its natural action on $\mathbb{R}^n$. The formula for the base change mapping $\Psi$ from Equation (4.5) also makes sense over the Satake algebras of the maximal split tori, i.e. we have a map $\Psi$ from the algebra $\mathbb{C}[X_1^{\pm 1}, \dots, X_n^{\pm 1}]$ to $\mathbb{C}[X_1^{\pm 1}, \dots, X_n^{\pm 1}]$. The monomials occurring in $f_{G,\mu,\alpha}$ are those monomials of the form $\Psi[w(\mu)]$ where $w$ is some element of $\mathfrak{S}_n$. The Weyl group translates $[w(\mu)]$ of $[\mu]$ correspond to all paths from $(0,0)$ to $(n,s)$, and the monomials of the form $\Psi[w(\mu)] = [w(\mu)] + [\theta(w\mu)]$ correspond to all paths from $(0,0)$ to $(n,0)$ staying below the horizontal line with equation $y = s$, and above the horizontal line with equation $y = -s$. The truncation $\chi_b^{G(\mathbb{Q}_p)} \Psi[w(\mu)]$ is non-zero if the path $\mathcal{G}$ of $\Psi[w(\mu)]$ lies below $\mathcal{G}_b$ and the set of contact points between the two graphs is precisely the initial point, end point and the set of break points of $\mathcal{G}_b$. This is the same condition as had for the general linear group (see above Equation (4.9)) because the root systems are the same. Such graphs exist in case $b$ has no slopes equal to $-1, 0$ or 1 (draw a picture). Consequently $\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha} \neq 0$, and then also $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \mathrm{St}_G) \neq 0$ by Proposition 3.8.

Forget the assumption that $F^+ \otimes E_\alpha$ is a field. We proceed just as we did for the general linear group (cf. Lemma 4.4), we write $f_{G,\mu,\alpha} = A + R$, where $R$ is a function whose Satake transform is a linear combination of monomials in the Satake algebra with all coefficients positive, and $A$ is a function for which we already know that its truncated trace on the Steinberg representation does not vanish. This completes the proof. □

PROPOSITION 4.10. *Let $G$ be an algebraic group over $\mathbb{Q}_p$ defined as in Equation (4.2), and assume it is of unitary type. Let $b \in B(G, \mu)$ be an isocrystal with slopes $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ (cf. the discussion below Proposition 1.1). Let $n = m_1 + m_2 + m_3$ be the composition of $n$ such that the first block of $m_1$ slopes $\lambda_i$ satisfy $\lambda_i = -1$, the second block of slopes $\lambda_i$ satisfy $-1 < \lambda_i < 1$ and is of size $m_2$, the third block of slopes $\lambda_i$ satisfy $\lambda_i = 1$ and is of size $m_3$. We have $m_1 = m_3$. Let $P = MN$ be the standard parabolic subgroup of $G$ corresponding to this composition of $n$, thus $M$ is a product of two groups, $M = M_1 \times M_2$, where $M_1 = \mathrm{GL}_{m_1}(F^+)$ is a general linear group and $M_2$ is an unramified unitary group. For $\alpha$ sufficiently divisible the trace $\mathrm{Tr}(\chi_b^{G(F^+)} f_{G,\mu,\alpha}, \bullet)$ against the representation $\mathrm{Ind}_{P(F^+)}^{G(F^+)}(\pi_{m_1} \otimes \mathrm{St}_{m_2}(\chi))$ is non-zero if $\pi_{m_1}$ is an unramified generic representation and $\chi$ an unramified character of $\mathrm{GL}_{m_2}(F^+)$.*

REMARK. The group $M_1$ could be trivial. This happens in case $-1 < \lambda < 1$ for all indices $i$. When $M_1$ is trivial, the considered representation is simply an unramified twist of the Steinberg representation.

PROOF. The proof is the same as the proof in case of the general linear group (cf. Proposition 4.8): one easily reduces the statement to Lemma 4.9. □

Let now $G/\mathbb{Q}$ be an unitary group of similitudes arising from a Shimura datum of PEL-type (cf. Equation (2.1)), and let $G_1 \subset G$ be the kernel of the factor of similitudes. The group $G_1$ is defined over a totally real field $F^+$, and defined with respect to a quadratic extension $F$ of $F^+$, which is a CM field. Let $A_0 \subset G$ be a maximally split torus, then we may write $A_0 = \mathbb{G}_{\mathrm{m}} \times A_0'$ (*not* a direct product), where $A_0' \subset G_1$ be the maximally split torus of $G_1$ defined by $G_1 \cap A_0$. At $p$ we have a decomposition of $F^+ \otimes \mathbb{Q}_p$ into a product of fields $F^+_\wp$, where $\wp$ ranges over the primes above $p$. Let $p$ be a prime number where $G$ is unramified. The group $G_{1,\mathbb{Q}_p}$ is of the form $G_{1,\mathbb{Q}_p} \cong \prod_\wp \mathrm{Res}_{F^+_\wp/\mathbb{Q}_p} G_{1,\wp}$, where the group $G_{1,\wp}$ is either an unramified unitary group over $F^+_\wp$, or the general linear group. In the first case we call the $F^+$-prime $\wp$ *unitary* and in the second case we call the prime *linear*.

Consider an isocrystal $b \in B(G)$. To $b$ we may associate its slope morphism $\overline{\nu}_b \in \mathfrak{a}_0$. Let $A'_{0,\wp} \subset G_{1,\wp}$ be the $\wp$-th component of $A_0'$; it is a split maximal torus in $G_{1,\wp}$, and write $\mathfrak{a}_0(\wp) := X_*(A'_{0,\wp})$. The space $\mathfrak{a}_0$ decomposes along the split center and the $F^+$-primes $\wp$ above $p$: $\mathfrak{a}_0 = \mathbb{R} \times \prod_\wp \mathfrak{a}_0(\wp)$. Thus we can speak for each $\wp$ of the $\wp$-component $\overline{\nu}_{b,\wp}$ of $\overline{\nu}_b$. In case $\wp$ is linear, the Proposition 4.8 gives us a class of representations $\pi'_\wp$ of $G_{1,\wp}(\mathbb{Q}_p)$ such that the $b_\wp$-truncated trace on $\pi'_\wp$ does not vanish. In case $\wp$ is unitary, we get such a class $\pi'_\wp$ from Proposition 4.10. Let $\pi'$ be the representation of $G_1(\mathbb{Q}_p)$ obtained from the factors $\pi_\wp$ by taking the tensor product.

DEFINITION 4.11. We write $\mathfrak{R}_1(b)$ for the just constructed class of $G_1(\mathbb{Q}_p)$-representations $\pi'$.

REMARK. The set of representations $\mathfrak{R}_1(b)$ has positive Plancherel measure in the set of $G_1(\mathbb{Q}_p)$ representations, and the $b$-truncated trace of the Kottwitz function on these representations does not vanish by construction.

We now extend the class $\mathfrak{R}_1(b)$ to a class of $G(\mathbb{Q}_p)$-representations, as follows:

DEFINITION 4.12. Let $\pi \in \mathfrak{R}_1(b)$. Then $\pi$ is an $H(\mathbb{Q}_p)$-representation; let $\omega_\pi$ be its central character, thus $\omega_\pi$ is a character of $Z_1(\mathbb{Q}_p)$. Assume $\chi$ is a character of $Z(\mathbb{Q}_p)$ extending $\omega_\pi$. Then we may extend the representation $\pi$ to a representation $\pi\chi$ of the group $H(\mathbb{Q}_p)Z(\mathbb{Q}_p)$. We define $\mathfrak{R}_1(b)'$ to be the set of $H(\mathbb{Q}_p)Z(\mathbb{Q}_p)$-representations of the form $\pi\chi$. Not all the inductions $\mathrm{Ind}_{H(\mathbb{Q}_p)}^{G(\mathbb{Q}_p)}(\pi\chi)$ have to be irreducible, we ignore the reducible ones. We define $\mathfrak{R}(b)$ to be the set of representations $\Pi$ isomorphic to an irreducible induction $\mathrm{Ind}_{H(\mathbb{Q}_p)Z(\mathbb{Q}_p)}^{G(\mathbb{Q}_p)}(\pi\chi)$ with $\pi\chi \in \mathfrak{R}_1(b)'$.

The required non-vanishing property of the representations in $\mathfrak{R}_b$ will be shown in the next section.

## 5. Local extension

We need to extend from $G_1(\mathbb{Q}_p)$ to the group $G(\mathbb{Q}_p)$. Let $Z$ be the center of the group $G$. Consider the morphism of algebraic groups $\psi\colon G_{1,\mathbb{Q}_p} \times Z_{\mathbb{Q}_p} \twoheadrightarrow G_{\mathbb{Q}_p}$; the group $\mathrm{Ker}(\psi)$ is the center $Z_1$ of the group $G_1$, so

$$\mathrm{Ker}(\psi) = \prod_{\wp} \begin{cases} \mathbb{G}_{\mathrm{m}} & \wp \text{ is linear} \\ U_1^* & \wp \text{ is unitary,} \end{cases} \tag{5.1}$$

where $U_1^*$ is the unramified non-split form of $\mathbb{G}_{\mathrm{m}}$ over $F_\wp^+$. Over $\mathbb{Q}$, $Z$ is defined by $Z(\mathbb{Q}) = \{x \in F^\times | N_{F/F^+}(x) \in \mathbb{Q}^\times\}$. Using Equation (5.1), the long exact sequence for Galois cohomology and Shapiro's lemma, the group $G(\mathbb{Q}_p)/G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ maps injectively into the group $(\mathbb{Z}/2\mathbb{Z})^t$, where $t$ is the number of unitary places of $F^+$ above $p$.

Write $\mu' \in X_*(T)$ for the cocharacter of the maximal torus $(T \cap G_1) \cap Z$ of $G_1 \times Z$ obtained from $\mu$ via restriction. Let $f_{G_1 \times Z, \mu', \alpha}$ be the corresponding function of Kottwitz on the group $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$. Furthermore we write $\chi_b^{G_1 \times Z}$ for the characteristic function on $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$ of elements $(g, z)$ such that we have $\chi_b^{G(\mathbb{Q}_p)}(gz) = 1$. We prove the following statement:

Proposition 5.1. *Fix a representation $\pi_0$ of $G_1(\mathbb{Q}_p)$. Let $\Pi$ be a smooth irreducible representation of $G(\mathbb{Q}_p)$ containing the representation $\pi_0$ of $G_1(\mathbb{Q}_p)$ upon restriction to $G_1(\mathbb{Q}_p) \times Z(\mathbb{Q}_p)$. Assume the central character of $\Pi$ is of finite order. Then, for all sufficiently divisible $\alpha$, we have*

$$\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi) = t(\Pi)\, \mathrm{Tr}(\chi_b^{G_1 \times Z} f_{G_1 \times Z, \mu, \alpha}, \pi_0)$$

*where $t(\Pi)$ is a positive real number.*

Before proving Proposition 5.1 we first establish some technical results. We fix smooth models of $G, G_1, Z$, etc. over $\mathbb{Z}_p$ (and use the same letter for them). We have the exact sequence $Z_1 \rightarrowtail Z \times G_1 \twoheadrightarrow G$, so the cokernel of $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)$ in $G(\mathbb{Q}_p)$ is a subgroup of $H^1(\mathbb{Q}_p, Z_1) \cong (\mathbb{Z}/2\mathbb{Z})^t$, where $t$ is the number of unitary places.

Lemma 5.2. *The mapping $G_1(\mathbb{Z}_p) \times Z(\mathbb{Z}_p) \to G(\mathbb{Z}_p)$ is surjective.*

Proof. We have an exact sequence $Z_1 \rightarrowtail G_1 \times Z \twoheadrightarrow G$ of algebraic groups over $\mathrm{Spec}(\mathbb{Z}_p)$. Thus we get $Z_1(\mathbb{F}_p) \rightarrowtail G_1(\mathbb{F}_p) \times Z(\mathbb{F}_p) \to G(\mathbb{F}_p) \to H^1(\mathbb{F}_p, Z_1)$. The group $Z_1$ is a torus and therefore connected. By Lang's theorem we obtain $H^1(\mathbb{F}_p, Z_1) = 1$. Thus the mapping $G_1(\mathbb{F}_p) \times Z(\mathbb{F}_p) \to G(\mathbb{F}_p)$ is surjective. By Hensel's lemma the mapping $G_1(\mathbb{Z}_p) \times Z(\mathbb{Z}_p) \to G(\mathbb{Z}_p)$ is then also surjective. □

Lemma 5.3. *The function of Kottwitz $f_{G,\mu,\alpha}$ has support on the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \subset G(\mathbb{Q}_p)$.*

Proof. Define $\chi$ on $G(\mathbb{Q}_p)$ to be the characteristic function of the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \subset G(\mathbb{Q}_p)$. The mapping $Z \times G_1 \to G$ is surjective on $\mathbb{Z}_p$-points, and therefore $\chi$ is spherical. The functions $\chi f^G_{G,\mu,\alpha}$ and $f_{G,\mu,\alpha}$ are then both spherical functions and to show that they are equal it suffice to show that their Satake transforms agree (the Satake transform is injective). We have $\mathcal{S}(\chi f_{G,\mu,\alpha}) = \chi|_{A(\mathbb{Q}_p)}\mathcal{S}(f_{G,\mu,\alpha})$, where $A$ is a maximal split torus of $G$, $\chi|_{A(\mathbb{Q}_p)}$ is the characteristic function of the subset $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \cap A(\mathbb{Q}_p) \subset A(\mathbb{Q}_p)$. Observe that, in fact, $Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p) \cap A(\mathbb{Q}_p) = A(\mathbb{Q}_p)$. This implies $\chi|_{A(\mathbb{Q}_p)}\mathcal{S}(f_{G,\mu,\alpha}) = \mathcal{S}(f_{G,\mu,\alpha})$, and shows that $\chi f_{G,\mu,\alpha}$ and $f_{G,\mu,\alpha}$ have the same Satake transform. This completes the proof of the lemma. □

We now turn to the proof of Proposition 5.1.

Proof of Proposition 5.1. By Clifford theory [**103**, thm 2.40] the representation $\Pi$ restricted to $G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ is a finite direct sum of irreducible representations $\pi_i$, where $\pi_i$ satisfies $\pi_i(g) = \pi_0(x_i g x_i^{-1})$ for some $x_i$ not depending on $g$. We clarify that in this finite direct sum multiplicities may occur. As characters on $G_1(\mathbb{Q}_p)Z(\mathbb{Q}_p)$ we may write $\theta_\Pi = \sum_{i=1}^t \theta_{\pi_i}\omega_i$, where $\theta_{\pi_i}$ is the Harish-Chandra character of $\pi_i$, viewed as a $G_1(\mathbb{Q}_p)$-representation, and $\omega_i$ is the central character of $\pi_i$. Using Lemma 5.3 we may now compute:

$$\begin{aligned}
\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi) &= \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}\theta_\Pi \mathrm{d}g \\
&= \sum_{i=1}^t \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}\theta_{\pi_i}\omega_i \mathrm{d}g \\
&= \sum_{i=1}^t \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f^{x_i^{-1}}_{G,\mu,\alpha}\theta_{\pi_0}\omega_0 \mathrm{d}g, \qquad (5.2)
\end{aligned}$$

where $f^{x_i^{-1}}_{G,\mu,\alpha}$ is the conjugate of $f_{G,\mu,\alpha}$ by $x_i^{-1}$. Note, however, that the function of Kottwitz is stable under the action of the Weyl group of $G$. Therefore $f^{x_i^{-1}}_{G,\mu,\alpha} = f_{G,\mu,\alpha}$. We get the expression:

$$t \int_{Z(\mathbb{Q}_p)G_1(\mathbb{Q}_p)} \chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}\theta_{\pi_0}\omega_0 \mathrm{d}g.$$

On the other hand we have

$$0 \neq \mathrm{Tr}(\chi_b^{Z\times G_1} f_{Z\times G_1,\mu',\alpha}, \pi_0) = \int_{Z(\mathbb{Q}_p)\times G_1(\mathbb{Q}_p)} \chi_b^{Z\times G_1} f_{Z\times G_1,\mu',\alpha}[\theta_{\pi_0}\times\omega_0]\mathrm{d}g.$$

We compute the right hand side:

$$\begin{aligned}
&\int_{\frac{Z(\mathbb{Q}_p)\times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}} \int_{Z_1(\mathbb{Q}_p)} (\chi_b^{Z\times G_1} f_{Z\times G_1,\mu',\alpha}[\theta_{\pi_0}\times\omega_0])(zz_1, hz_1)\mathrm{d}z_1 \frac{\mathrm{d}(z,h)}{\mathrm{d}z_1} \\
&= \int_{\frac{Z(\mathbb{Q}_p)\times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}} \chi_b^{Z\times G_1} \int_{Z_1(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1, hz_1)\mathrm{d}z_1 (\theta_{\pi_0}\omega_0)(z,h) \frac{\mathrm{d}(z,h)}{\mathrm{d}z_1}. \qquad (5.3)
\end{aligned}$$

We claim that

$$\int_{Z_1(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1, hz_1)\mathrm{d}z_1 = f_{G,\mu,\alpha}(z,h). \tag{5.4}$$

The map $Z \times G_1 \to G$ is surjective on $\mathbb{Z}_p$-points, and therefore the function

$$\int_{Z_1} f_{Z\times G_1,\mu',\alpha}(zz_1, hz_1)\mathrm{d}z_1$$

is $G(\mathbb{Z}_p)$-spherical. Therefore, to show that Equation (5.4) is true, it suffices to show that the Satake transforms of these functions agree.

We compute the Satake transform of the left hand side:

$$\begin{aligned}
\delta_{P_0}^{-1}\int_{N_0(\mathbb{Q}_p)}\int_{Z_1(\mathbb{Q}_p)} & f_{Z\times G_1,\mu',\alpha}(zz_1n_0, hz_1n_0)\mathrm{d}z_1\mathrm{d}n_0 \\
&= \delta_{P_0}^{-1}\int_{Z_1(\mathbb{Q}_p)}\int_{N_0(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1n_0, hz_1n_0)\mathrm{d}n_0\mathrm{d}z_1 \\
&= \int_{Z_1(\mathbb{Q}_p)} \delta_{P_0}^{-1}\int_{N_0(\mathbb{Q}_p)} f_{Z\times G_1,\mu',\alpha}(zz_1n_0, hz_1n_0)\mathrm{d}n_0\mathrm{d}z_1 \\
&= \int_{Z_1(\mathbb{Q}_p)} \left(f_{Z\times G_1,\mu',\alpha}\right)^{(P_0)}(zz_1, hz_1)\mathrm{d}z_1
\end{aligned}$$

By Definition 4.1 the last expression is equal to $f^{(P_0)}_{G,\mu,\alpha}(z,h)$. This proves Equation (5.4). We may continue with Equation (5.3) to obtain

$$\int_{\frac{Z(\mathbb{Q}_p)\times G_1(\mathbb{Q}_p)}{Z_1(\mathbb{Q}_p)}} \chi_b^{Z\times G_1} f_{G,\mu,\alpha}\theta_{\pi_0}\omega_0\frac{\mathrm{d}(z,h)}{\mathrm{d}z_1}.$$

Now $\omega_0$ is of finite order by assumption, and the function $f_{G\mu\alpha}$ restricted to $\mathbb{Q}_p{}^\times \cong A(\mathbb{Q}_p) \subset Z(\mathbb{Q}_p)$ is the characteristic function of $p^{-\alpha}\mathbb{Z}_p^\times$. For $\alpha$ sufficiently divisible this is then, up to normalization of Haar measures, just the trace $\mathrm{Tr}(\chi_b^{Z\times G_1} f_{Z\times G_1,\mu',\alpha}, \pi_0)$. This proves that $\mathrm{Tr}(\chi_b^G f_{G,\mu,\alpha}, \Pi)$ and $\mathrm{Tr}(\chi_b^{Z\times G_1} f_{Z\times G_1,\mu',\alpha}, \pi_0)$ differ by a positive, non-zero, scalar. The proof of the theorem is now complete. □

## 6. Global extension

In this section we prove a technical proposition concerning the restriction of automorphic representations of $G$ to the subgroup $G_1 \subset G$ (the kernel of the factor of similitudes). Recall that we have the surjection $G_1 \times Z \twoheadrightarrow G$.

PROPOSITION 6.1. *Let $\Pi$ be a cuspidal automorphic representation of $G(\mathbb{A})$, then its restriction to the group $G_1(\mathbb{A}) \times Z(\mathbb{A})$ contains a cuspidal automorphic representation of $G_1(\mathbb{A}) \times Z(\mathbb{A})$.*

REMARK. The proof we give here is copied from Clozel's article [**23**, p. 137]; cf. Labesse-Schwermer [**67**, p. 391].

PROOF OF PROPOSITION 6.1. Let $A \subset Z$ be the split center. Define $\mathbb{G}_1$ to be the subset $\mathbb{G}_1 := A(\mathbb{A})G(\mathbb{Q})G_1(\mathbb{A}) \subset G(\mathbb{A})$. Then $\mathbb{G}_1$ is a subgroup because $G(\mathbb{Q})$ normalizes $G_1(\mathbb{A})$. Furthermore the subgroup $\mathbb{G}_1$ is closed in $G(\mathbb{A})$, and we have $[A(\mathbb{A})G_1(\mathbb{A})] \cap G(\mathbb{Q}) = A(\mathbb{Q})G_1(\mathbb{Q}) \subset G(\mathbb{A})$ (cf. Clozel [Lemme 5.8, *loc. cit.*]). Let $\chi$ be the central character of $\Pi$; and let $\varepsilon$ be the restriction of $\chi$ to $G_1(\mathbb{A}) \times A(\mathbb{A})$. Let $\rho_0$ be the representation of $G_1(\mathbb{A})$ on the space $L^2_0(G_1(\mathbb{Q})\backslash G_1(\mathbb{A}), \varepsilon)$ of cuspidal functions transforming under $G_1(\mathbb{A})$ via $\varepsilon$. We extend the representation $\rho_0$ to a representation of $\mathbb{G}_1$ by defining: $\rho_1(z\gamma x)f(y) = \chi(z)f(\gamma^{-1}y\gamma x)$, for $z \in A(\mathbb{A}), \gamma \in G(\mathbb{Q}), x \in G_1(\mathbb{A}), y \in G_1(\mathbb{A})$. We do not copy the verification that this representation is well-defined [*loc. cit*, 5.16]. Define the representation $\rho = \mathrm{Ind}_{\mathbb{G}_1}^{G(\mathbb{A})}(\rho_1)$ of $G(\mathbb{A})$. A computation shows that $\rho$ is isomorphic to the representation of $G(\mathbb{A})$ on the space $L^2_0(G(\mathbb{Q})\backslash G(\mathbb{A}), \chi)$ of functions on $G(\mathbb{Q})\backslash G(\mathbb{A})$ transforming via $\chi$ under the action of $A(\mathbb{A})$. Consequently, if $\Pi$ occurs in the representation $\mathrm{Ind}_{\mathbb{G}_1}^{G(\mathbb{A})}(\rho_1)$, then its restriction to $\mathbb{G}_1$ will contain irreducible $\mathbb{G}_1$-subrepresentations of $\rho_1$. □

## 7. The isolation argument

Let $\mathrm{Sh}_K$ be a Shimura variety of PEL-type of type (A), and let $G$ be the corresponding unitary group of similitudes over $\mathbb{Q}$ (cf. Equation (2.1)). We write $E$ for the reflex field and we let $p$ be a prime of good reduction [2]. Let $b \in B(G_{\mathbb{Q}_p}, \mu)$ be an admissible isocrystal. Let $\mathfrak{p}$ be a prime of the reflex field $E$ above $p$. Let $\mathbb{F}_q$ be the residue field of $E$ at $\mathfrak{p}$. Let $\mathrm{Sh}^b_{K,\mathfrak{p}}$ be the corresponding Newton stratum of $\mathrm{Sh}_{K,\mathfrak{p}}$, a locally closed subvariety of $\mathrm{Sh}_{K,\wp}$ over $\mathbb{F}_q$ [**88**].

Let $\alpha$ be a positive integer. We fix an embedding $E_\mathfrak{p} \subset \overline{\mathbb{Q}}_p$ and we write $E_{\mathfrak{p},\alpha}$ for the extension of the field $E_\mathfrak{p}$ of degree $\alpha$ inside $\overline{\mathbb{Q}}_p$.

THEOREM 7.1 (Wedhorn-Viehmann). *The variety* $\mathrm{Sh}^b_{K,\mathfrak{p}}$ *is not empty.*

REMARK. In the statement of the above theorem we have not been precise about the form of the compact open subgroup $K \subset G(\mathbb{A}_\mathrm{f})$. Note however that for any pair $(K, K')$ of compact open subgroups, hyperspecial at $p$, we have the finite étale morphisms $\mathrm{Sh}_K \leftarrow \mathrm{Sh}_{K\cap K'} \to \mathrm{Sh}_{K'}$ respecting the Newton stratification modulo $\mathfrak{p}$. Therefore, showing the Newton stratum is non-empty for one $K$ is equivalent to showing it is non-empty for all $K$.

PROOF. Fix a sufficiently divisible and even integer $\alpha$ such that the conclusion of Proposition 5.1 is true. We start with the formula of Kottwitz. We write $\phi_\alpha$ for the function $\phi_{G,\mu,\alpha}$ from the previous section [3] on $G(E_{\mathfrak{p},\alpha})$. Similarly $f_\alpha := f_{G,\mu,\alpha}$. We pick a prime $\ell \neq p$ and fix an isomorphism $\overline{\mathbb{Q}}_\ell \cong \mathbb{C}$ (and suppress it from all notations). Let $\xi$ be an irreducible complex (algebraic) representation of $G$, and write $\mathcal{L}$ for the corresponding $\ell$-adic local system on the

2. Here '*good reduction*' is in the sense of Kottwitz [**59**, §6]; in particular $K$ decomposes into a product $K = K_pK^p$ with $K_p \subset G(\mathbb{Q}_p)$ hyperspecial.

3. Where the notation $E_\alpha$ from that section should be replaced with $E_{\mathfrak{p},\alpha}$, and similarly $F^+$ of that section should be replaced by the algebra $F^+ \otimes \mathbb{Q}_p = \prod_\wp F^+_\wp$, where $\wp$ ranges over the places above $p$.

Shimura tower. Then the Kottwitz formula states:

(7.1)
$$\sum_{x'\in \mathrm{Fix}^b_{f^p\times\Phi^\alpha_{\mathfrak{p}}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(f^p\times\Phi^\alpha_{\mathfrak{p}},\iota^*(\mathcal{L})_x) = |\operatorname{Ker}^1(\mathbb{Q},G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta)O_\gamma(f^{\infty p})\mathrm{TO}_\delta(\phi_\alpha)\operatorname{Tr}\xi_{\mathbb{C}}(\gamma_0),$$

where $\mathrm{Fix}^b_{f^p\times\Phi^\alpha_{\mathfrak{p}}(\overline{\mathbb{F}}_q)}$ is the set of fixed points of the Hecke correspondence $f^p\times\Phi^\alpha_{\mathfrak{p}}$ acting on $\mathrm{Sh}^b_{K,\mathbb{F}_q}$, and where the sum ranges of the Kottwitz triples $(\gamma_0;\gamma,\delta)$ with the additional condition that the isocrystal defined by $\delta$ is equal to $b$. In Equation (7.1) the map $\iota$ is the embedding of $\mathrm{Sh}^b_{K,\mathbb{F}_q}$ into $\mathrm{Sh}_{K,\mathbb{F}_q}$.

We may rewrite the right hand side of Equation (7.1) as

$$|\operatorname{Ker}^1(\mathbb{Q},G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta)\cdot O_\gamma(f^{\infty p})\mathrm{TO}_\delta(\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma b}\phi_\alpha)\operatorname{Tr}\xi_{\mathbb{C}}(\gamma_0), \tag{7.2}$$

where now the sum ranges over *all* Kottwitz triples and where $\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma b}$ is the characteristic function on $G(E_{\mathfrak{p},\alpha})$ such for each element $\delta\in G(E_{\mathfrak{p},\alpha})$ we have $\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma b}(\delta) = 1$ if and only if the conjugacy class $\gamma = \mathcal{N}(\delta)$ satisfies $\Phi(\gamma) = \lambda\overline{\nu}$ for some positive real number $\lambda\in\mathbb{R}^\times_{>0}$. Assume the triple $(\gamma_0;\gamma,\delta)$ is such that the corresponding term $c(\gamma_0;\gamma,\delta)O_\gamma(f^{\infty p})\mathrm{TO}_\delta(\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma b}\phi_\alpha)\operatorname{Tr}\xi_{\mathbb{C}}(\gamma_0)$ is non-zero. Then, by the proof of Kottwitz [**57**], we know that the triple $(\gamma_0;\gamma,\delta)$ arises from some virtual Abelian variety with additional PEL-type structures. In particular the isocrystal defined by $\delta$ lies in the subset $B(G_{\mathbb{Q}_p},\mu)\subset B(G_{\mathbb{Q}_p})$. Thus its end point is determined. We have $\gamma=\mathcal{N}(\delta)$ and $\Phi(\gamma)=\lambda\overline{\nu}_b$ for some $\lambda$ (Proposition 1.1). Therefore the isocrystal defined by $\delta$ must be equal to $b$. Thus the above sum precisely counts Abelian varieties with additional PEL type structures over $\mathbb{F}_{q^\alpha}$ such that their isocrystal equals $b$.

We show that the sum in Equation (7.2) is non-zero. Let $\mathcal{E}$ be the (finite) set of endoscopic groups $H$ associated to $G$ and unramified at all places where the data $(G,K)$ are unramified. By the stabilization argument of Kottwitz [**58**], the expression in Equation (7.2) is equal to the stable sum

$$\sum_{\mathcal{E}} \iota(G,H)\cdot \mathrm{ST}^*_e((\chi^G_b f_\alpha)^H), \tag{7.3}$$

where $(\chi^G_b f_\alpha)^H$ are the transferred functions, whose existence is guaranteed by the fundamental lemma, the $*$ in $\mathrm{ST}^*_e$ means that one only considers stable conjugacy classes satisfying a certain regularity condition (which is empty in case $H$ is a maximal endoscopic group), and finally $\iota(G,H)$ is a constant depending on the endoscopic group (cf. [*loc. cit.*] for the definition).

We consider only functions such that the transfer $(\chi^G_b f_\alpha)^H$ vanishes for proper endoscopic groups, and therefore we may ignore the regularity condition [4]. Thus, Equation (7.3) simplifies

4. In fact, due to the form of the function $f_\infty$ we have $\mathrm{ST}^*_e = \mathrm{ST}_e$, see [**81**, thm 6.2.1] or [**25**, (2.5)].

for such functions and gives the equation:

$$\sum_{x' \in \mathrm{Fix}^b_{f^p \times \Phi^\alpha_{\mathfrak{p}}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(f^p \times \Phi^\alpha_{\mathfrak{p}}, \iota^*(\mathcal{L})_x) = \sum_{\mathcal{E}} \iota(G, H) \mathrm{ST}_e((\chi^G_b f_\alpha)^H). \tag{7.4}$$

Visibly, if we show that the left hand side of Equation (7.4) is non-zero for some Hecke operator $f^p$, then the variety $\mathrm{Sh}^b_{K,\mathbb{F}_q}$ is non-empty. We will show that the right hand side of Kottwitz's formula does not vanish for some choice of $K^p$ and some choice of $f^p$.

We write $G_0^*, G_1^*, G^*$ for the quasi-split inner forms of $G_0$, $G_1$, and $G$ respectively (we remind the reader that $G_0$ is defined over $F^+$ and that $G_1 = \mathrm{Res}_{F^+/\mathbb{Q}} G_0$). The group $G^*$ is the maximal endoscopic group of $G$. Let $\{x_1, x_2, \ldots, x_d\}$ be the set of prime numbers such that the group $G_{\mathbb{Q}_{x_i}}$ is ramified. For $v$ a prime number with $v \notin \{x_1, x_2, \ldots, x_d\}$ the local group $G_{\mathbb{Q}_v}$ is quasi-split, and therefore we may (and do) identify it with the group $G^*_{\mathbb{Q}_v}$. Below we will transfer functions from the group $G(\mathbb{A})$ to the group $G^*(\mathbb{A})$; at the places $v$ with $v \notin \{x_1, x_2, \ldots, x_d, \infty\}$ we have $G(\mathbb{Q}_v) = G^*(\mathbb{Q}_v)$ and using this identification we may (and do) take $(h_v)^{G^*(\mathbb{Q}_v)} = h_v$ for any $h_v \in \mathcal{H}(G(\mathbb{Q}_v))$.

To help the reader understand what we do below at the places $x_i$ (and *why* we do this), let us interrupt this proof with a general remark on the fundamental lemma. It is important to realise that if $v = x_i$ is one of the bad places, then the fundamental lemma guarantees the *existence* of the transferred function $h_v \rightsquigarrow (h_v)^{G^*(\mathbb{Q}_v)}$; however, in its current state, the fundamental lemma does <u>not</u> give an explicit description of a transferred function $(h_v)^{G^*(\mathbb{Q}_v)}$. The fundamental lemma only gives explicit transfer in case the group is unramified and the level is hyperspecial. In our case the transferred function $(h_v)^{G^*(\mathbb{Q}_v)}$ is not explicit, and this could introduce signs and cancellations that we cannot control. This makes it hard to show that expressions such as the one in Equation (7.10) do not vanish. In the argument below we solve the issue by taking $h_v$ to be a pseudocoefficient of the Steinberg representation. For these functions an explicit transfer is known (the transfer is again a pseudocoefficient of the Steinberg representation) and therefore we will be able to control the signs and avoid cancellations. This ends the remark, let us now continue with the proof.

We are going to construct an automorphic representation $\Pi_0$ of $G^*$ with particularly nice properties. From this point onward we take $\xi$ to be a fixed, sufficiently regular complex representation (in the sense of [**28**, Hyp. (1.2.3)]). We also assume that $\xi$ defines a coefficient system of weight 0 (cf. [**25**]), and even better that $\xi$ is trivial on the center of $G^*$. Fix three additional, different, prime numbers $p_1, p_2, p_3$ ($\neq p$) such that the group $G_{\mathbb{Q}_{p_i}}$ is split for $i = 1, 2, 3$. Let $\Pi_{0,p_1}$ be a cuspidal representation of the group $G(\mathbb{Q}_{p_1}) = G^*(\mathbb{Q}_{p_1})$. Let $A(\mathbb{R})^+$ be the topological neutral component of the set of real points of the split center $A$ of $G$. We apply a theorem of Clozel and Shin [**20, 94**] to find an automorphic representation $\Pi_0 \subset L^2_0(G^*(\mathbb{Q})A(\mathbb{R})^+ \backslash G^*(\mathbb{A}))$ of $G^*(\mathbb{A})$ with:

(1) $\Pi_{0,\infty}$ is in the discrete series and is $\xi$-cohomological;

(2) $\Pi_{0,p}$ lies in the class $\mathfrak{R}(b)$ (cf. Definition 4.12);

(3) $\Pi_{0,p_1}$ lies in the *inertial orbit*[5] $\mathcal{I}(\Pi_{0,p_1})$ of $\Pi_{0,p_1}$ at $p_1$;

(4) $\Pi_{0,p_2}$ is isomorphic to the Steinberg representation (up to an unramified twist of finite order);

(5) $\Pi_{0,x_i}$ is isomorphic to an unramified twist (of finite order) of the Steinberg representation of $G(\mathbb{Q}_{x_i})$ (for $i = 1, 2, \ldots, d$);

(6) $\Pi_{0,v}$ is unramified for all primes $v \notin \{p, p_1, p_2, p_3, x_1, x_2, \ldots, x_d\}$;

(7) The central character of $\Pi_0$ has finite order.

Because the component at $p_1$ of $\Pi_0$ is cuspidal, the representation $\Pi_0$ is a cuspidal automorphic representation. The point (7) is possible because of the condition on the weight of $\xi$.

We now choose the group $K \subset G(\mathbb{A}_\mathrm{f})$, and we will also choose a compact open group $K^*$ in $G^*(\mathbb{A}_\mathrm{f})$. Write $S = \{p, p_1, p_2, p_3\}$. Write $S' = \{p, p_1, p_2, p_3, x_1, x_2, \ldots, x_d\}$ for the union of $S$ with the set of all places where the group $G$ is ramified.

The compact open group $K \subset G(\mathbb{A}_\mathrm{f})$ is a (any) group with the following properties:

(1) $K$ is a product $\prod_v K_v \subset G(\mathbb{A}_\mathrm{f})$ of compact open groups;

(2) for all $v \notin S'$ the group $K_v$ is hyperspecial;

(3) $K_p$ is hyperspecial;

(4) $K_{p_3}$ is sufficiently small so that $\mathrm{Sh}_K$ is smooth and $(\Pi_{0,p_3})^{K_{p_3}} \neq 0$;

(5) $K_{x_i}$ is sufficiently small so that the function $f_{x_i}$ is $K_{x_i}$-spherical;

(6) for all $v \notin \{x_1, x_2, \ldots, x_d\}$ the space $(\Pi_{0,v})^{K_v}$ is non-zero.

The group $K^* \subset G^*(\mathbb{A}_\mathrm{f})$ is a (any) group with the following properties:

(1) $K^*$ is a product $\prod_v K^*_v \subset G^*(\mathbb{A}_\mathrm{f})$ of compact open groups;

(2) for any prime $v \notin \{x_1, \ldots, x_d\}$ we have $K^*_v = K_v \subset G(\mathbb{Q}_v) = G^*(\mathbb{Q}_v)$;

(3) for all $i \in \{1, 2, \ldots, d\}$ we have $(\Pi_{0,x_i})^{K_{x_i}} \neq 0$;

We now choose the Hecke function $f \in \mathcal{H}(G(\mathbb{A}_\mathrm{f}))$. Consider the function $f^{p\infty} \in \mathcal{H}(G(\mathbb{A}_\mathrm{f}))$ of the form

$$f^{p\infty} := f_{p_1} \otimes f_{p_2} \otimes f_{p_3} \otimes f_{x_1} \otimes f_{x_2} \otimes \cdots \otimes f_{x_d} \otimes f^{S'}, \tag{7.5}$$

where

- $f_{p_1}$ is a pseudo-coefficient on $G(\mathbb{Q}_{p_1})$ of the representation $\Pi_{p_1}$;
- $f_{p_2}$ is a pseudo-coefficient of the Steinberg representation of $G(\mathbb{Q}_{p_2})$;
- $f_{p_3} = \mathbf{1}_{K_{p_3}}$;
- $f_{x_i}$ is (essentially) a pseudo-coefficient of the Steinberg representation of $G(\mathbb{Q}_{x_i})$ for $i = 1, 2, \ldots, d$ (see below for the precise statement and the construction);

5. For the definition of inertial orbit, see [**89**, V.2.7].

– Before we define the function $f^{S'}$ we explain a fact: There are only *finitely many* cuspidal automorphic representations $\Pi \subset L^2_0(G^*(\mathbb{Q})A(\mathbb{R})^+\backslash G^*(\mathbb{A}))$ of $G^*$ whose component at infinity is equal to $\Pi_\infty$ and have invariant vectors under the group $K$. In particular also the set of their possible outside $S'$-components $\Pi^{S'}$ is finite. Therefore, we may find a function $f^{S'} \in \mathcal{H}(G^*(\mathbb{A}_{\mathrm{f}}^{S'})) = \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^{S'}))$ whose trace on $\Pi^{S'}$ is equal to 1 if $\Pi^{S'} \cong \Pi_0^{S'}$ and whose trace equals 0 otherwise for all $\Pi$ with $\Pi_\infty = \Pi_{0,\infty}$ and $\Pi^K \neq 0$. We fix $f^{S'}$ to be a function having this property.

We need to comment on the pseudo-coefficients $f_{x_i}$. In the literature these coefficients are usually only constructed for groups under conditions on the center [**26**, §3.4], such as the group be semi-simple, or with anisotropic center. We have neither of these conditions. Let $x = x_i$ be one of the bad places and write $H$ for the derived group of $G$, we write $Z$ for the center of $G$. We write $H^*$ for the derived group of $G^*$ (then $H^*$ is the quasi-split inner form of $H$). The center $Z$ of $G$ is canonically isomorphic with the center of $G^*$ (and the same is true for the centers of $H$ and $H^*$). Let $k$ be any smooth function on the group $H(\mathbb{Q}_x)$. Let $O_x \subset Z(\mathbb{Q}_x)$ be the maximal compact open subgroup of the center $Z(\mathbb{Q}_x)$ of $G(\mathbb{Q}_x)$. We now define a function $\widetilde{k}$ on the group $G(\mathbb{Q}_x)$. Consider first the following function on the group $H(\mathbb{Q}_x) \times Z(\mathbb{Q}_x)$:

$$(g, z) \longmapsto \int_{(H\cap Z)(\mathbb{Q}_x)} (k \times \mathbf{1}_{O_x})(gt, zt)\mathrm{d}t, \tag{7.6}$$

where $\mathrm{d}t$ is an invariant measure on the finite group $(H \cap Z)(\mathbb{Q}_x)$. The function in Equation (7.6) is $(H \cap Z)(\mathbb{Q}_x)$-invariant, and thus defines a function on the subgroup

$$\frac{H(\mathbb{Q}_x) \times Z(\mathbb{Q}_x)}{(H \cap Z)(\mathbb{Q}_x)} \subset G(\mathbb{Q}_x).$$

We extend this function by 0 to obtain the function $\tilde{k}$ on the group $G(\mathbb{Q}_x)$.

Let $H^*$ be the quasi-split inner form of $H$; then $H^*$ is also the derived group of $G^*$. By the fundamental lemma we may transfer smooth functions on the group $G(\mathbb{Q}_x)$ to functions on the group $G^*(\mathbb{Q}_x)$, and similarly functions from the group $H(\mathbb{Q}_x)$ to functions on the group $H^*(\mathbb{Q}_x)$. The formula in Equation (7.6) makes sense if we replace $H$ by its quasi-split inner form; thus we also have a construction $k \mapsto \widetilde{k}$ for smooth functions on $H^*(\mathbb{Q}_x)$. The construction in Equation (7.6) is compatible with transfer of functions, i.e. the function $(\widetilde{k})^{G^*(\mathbb{Q}_x)}$ on $G^*(\mathbb{Q}_x)$ has the same stable orbital integrals as the function $\big(\widetilde{k^{H^*(\mathbb{Q}_x)}}\big)$ for all $k \in \mathcal{H}(H(\mathbb{Q}_x))$.

We now take the function $k$ on $H(\mathbb{Q}_x)$ to be a certain sign $\varepsilon$ times a pseudocoefficient of the Steinberg representation, which exists because the center of $H$ is anisotropic. (We choose the sign $\varepsilon$ later). Define $f_x := \widetilde{k}$. In case the group has anisotropic center, the transfer of a pseudocoefficient of the Steinberg representation is again a pseudocoefficient of the Steinberg representation. Thus we may (and do) take the transferred function $(f_x)^{G^*(\mathbb{Q}_x)}$ to be the one obtained from a pseudocoefficient via the construction in Equation (7.6).

We show that the function $(f_x)^{G^*(\mathbb{Q}_x)}$ is (essentially) a pseudocoefficient of the Steinberg representation. Let us first make this statement precise. Let $\chi$ be a character of the group $G(\mathbb{Q}_x)$. The character $\chi$ induces a character $\overline{\chi}$ of the cocenter $C(\mathbb{Q}_x)$ of the group $G(\mathbb{Q}_x)$. We call the character $\chi$ *unramified* if $\overline{\chi}$ is trivial on the maximal compact open subgroup $K_C$ of $C(\mathbb{Q}_x)$. We claim that the sign $\varepsilon$ can be chosen so that the function $(f_x)^{G^*(\mathbb{Q}_x)}$ has the following two properties:

– For every unramified character $\chi$ of $G(\mathbb{Q}_x)$:

$$\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}(\chi)) \neq 0. \tag{7.7}$$

– For every smooth irreducible representation $\Pi_x$ occurring as the $x$-component of a cuspidal automorphic representation $\Pi$ of $G^*$ we have

$$\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \Pi_x) \in \mathbb{R}_{\geq 0}. \tag{7.8}$$

We first verify Equation (7.8). Let $\Pi_x$ be a smooth irreducible representation of the group $G^*(\mathbb{Q}_x)$, let $\theta_{\Pi_x}$ be its character. We assume that $\Pi_x$ is the $x$-component of a cuspidal automorphic representation $\Pi$ of the group $G^*$. Let $\pi_1, \ldots, \pi_d$ be the irreducible $H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)$-subrepresentations of $\Pi_x$, and let $\theta_1, \ldots, \theta_d$ be their characters. We have $\theta_{\Pi_x} = \sum_{i=1}^d \theta_i$. Then (modulo a positive constant depending on $\mathrm{d}t$):

(7.9)

$$\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \Pi_x) = \int_{H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)} f^{G^*(\mathbb{Q}_x)}(g) \sum_{i=1}^{d} \theta_i(g)\mathrm{d}g = \sum_{i=1}^{d} \int_{H^*(\mathbb{Q}_x)Z(\mathbb{Q}_x)} f^{G^*(\mathbb{Q}_x)}(g)\theta_i(g)\mathrm{d}g.$$

By the definition of the function $f^{G^*(\mathbb{Q}_x)}$ from Equation (7.6) the summand on the right hand side equals, up to some positive constant, the trace of the pseudocoefficient of the Steinberg representation on the group $H(\mathbb{Q}_x)$ against $\pi_i$. Such a trace is non-zero only if $\pi_i$ is isomorphic to one of the representations $V_P$ defined by Borel-Wallach [**10**, 6.2.14]. We show that $\pi_i$ must be the Steinberg representation. The representation $\Pi_x$ occurs as the component at $x$ of a cuspidal automorphic representation. Therefore $\Pi_x$ is *unitary*. Thus the representation $\pi_i$ is unitary as well. By [6.4, *loc. cit.*] the only representations $V_P$ which are unitary, are the Steinberg representation and the trivial representation. Let us exclude the trivial representation. By Clifford theory, all the representations occurring in $\Pi_x$ are conjugate under elements of the group $G(\mathbb{Q}_x)$. Consequently, if one of the occurring representations is finite dimensional, then they are all finite dimensional. This means that $\Pi_x$ is finite dimensional and thus the representation $\Pi$ is finite dimensional. Thus $\pi_i$ cannot be trivial. Therefore we can pick the sign $\varepsilon$ such that Equation (7.8) is true.

We now verify Equation (7.7). By construction the function $f^{G^*(\mathbb{Q}_x)}$ is supported on the inverse image of $K_C$ in $G$. Because $\chi$ is unramified it is constant on the support of $f^{G^*(\mathbb{Q}_x)}$. Therefore we have $\mathrm{Tr}(f_x^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}(\chi)) = \mathrm{Tr}(f_x^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)})$. We verify that the trace $\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)})$ is non-zero. Let $P_{0,x}$ be a Borel subgroup of $G^*_{\mathbb{Q}_x}$ and let $P'_{0,x}$ be the pull back of $P_{0,x}$ to $H^*_{\mathbb{Q}_x}$. Let $I$ be the space of locally constant complex valued functions on $G^*(\mathbb{Q}_x)/P_{0,x}(\mathbb{Q}_x)$ and $I'$ be the same space, but then for the group $H^*(\mathbb{Q}_x)$. We extend any function $h \in I'$ by 0 and this gives us the composition of maps $I' \rightarrowtail I \twoheadrightarrow \mathrm{St}_{G^*(\mathbb{Q}_x)}$. This composition is trivial on the subspaces $C^\infty(H^*(\mathbb{Q}_x)/P(\mathbb{Q}_x)) \subset I'$ for any proper parabolic subgroup $P$ of $H^*$ containing $P^*_{0,x}$. We obtain an $H^*(\mathbb{Q}_x)$-injection $\mathrm{St}_{H^*(\mathbb{Q}_x)} \rightarrowtail \mathrm{St}_{G^*(\mathbb{Q}_x)}$. It follows from Equation (7.9) that $\mathrm{Tr}(f^{G^*(\mathbb{Q}_x)}, \mathrm{St}_{G^*(\mathbb{Q}_x)}) \neq 0$.

We have now completed the definition of the components $f_{x_i}$, thus also the definition of the Hecke operator $f^{p\infty}$ is complete (see Equation (7.5)). We emphasize that at the primes $v \notin \{x_1, x_2, \dots, x_d\}$ we take $(f_v)^{G^*(\mathbb{Q}_v)} = f_v$ (we have $G^*(\mathbb{Q}_v) = G(\mathbb{Q}_v)$) and at the primes $v \in \{x_1, x_2, \dots, x_d\}$ we control the traces of the transferred function $(f_v)^{G^*(\mathbb{Q}_v)}$ against smooth representations via the conclusion in Equation (7.7).

Due to the cuspidal component $f_{p_1}$ of $f^p$, the trace formula simplifies. Because $f_{p_2}$ is stabilizing (Labesse [**65**]), the contribution of the proper endoscopic groups are zero, and the right hand side of Equation (7.4) becomes a sum of the form

$$\sum_{\Pi} m(\Pi)\,\mathrm{Tr}((f_\infty f^p)^{G^*(\mathbb{A}^p)}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha), \Pi), \tag{7.10}$$

where $\Pi$ ranges over cuspidal automorphic representations of $G^*(\mathbb{A})$, and $m(\Pi)$ is the multiplicity of $\Pi$ in the discrete spectrum of $G^*(\mathbb{A})$ with trivial central character on $A(\mathbb{R})^+$ ($A$ is both the split center of the group $G$ as well as the split center of the group $G^*$). Here we are applying the simple trace formula of Arthur [**3**, Cor. 23.6] (cf. proof of [**2**, thm 7.1]), the correcting term in Arthur's formula vanishes due to the pseudocoefficients in the Hecke operator. The sum in Equation (7.10) expands to the sum

$$\sum m(\Pi)\,\mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty)\,\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_{G,\mu,\alpha}, \Pi_p)\dim\left((\Pi_{p_3})^{K_{p_3}}\right)\prod_{i=1}^{d}\mathrm{Tr}(f_{x_i}^{G^*(\mathbb{Q}_{x_i})}, \Pi_{x_i}), \tag{7.11}$$

where $\Pi$ ranges over the irreducible subspaces of $L_0^2(A(\mathbb{R})^+G^*(\mathbb{Q})\backslash G^*(\mathbb{A}))$ such that

– $\Pi^{S'} \cong \Pi_0^{S'}$;
– $\Pi_{p_1}$ lies in the inertial orbit $\mathcal{I}(\Pi_{p_1})$ of the representation $\Pi_{p_1}$;
– $\Pi_{p_2}$ is, up to unramified twist, isomorphic to the Steinberg representation of $G(\mathbb{Q}_{p_2})$;
– $\Pi_{x_i}$ is such that $\mathrm{Tr}(f_{x_i}, \Pi_{x_i}) \neq 0$.

By Proposition 6.1 we may find a cuspidal automorphic representation $\pi_0$ of $G_1^*(\mathbb{A})$ contained in $\Pi_0$. Let now $\Pi$ be an automorphic representation of $G^*(\mathbb{A})$ contributing to Equation (7.11). Thus the representation $\Pi^{S'}$ is isomorphic to the representation $\Pi_0^{S'}$. Let $\pi$ be a cuspidal automorphic subrepresentation of $\mathrm{Res}_{[G_1^*\times Z](\mathbb{A})}(\Pi)$ (Proposition 6.1). Enlarge $S'$ to a larger finite set $S''$ so that the representations $\pi$ and $\Pi$ are unramified for all places outside the set $S''$. At the unramified places $v \notin S''$ the representation $\mathrm{Res}_{[G_1^*\times Z](\mathbb{Q}_v)}(\Pi_{0,v})$ contains exactly one unramified representation: $\pi_{0,v}$. Therefore we have $(\pi)^{S''} \cong (\pi_0)^{S''}$.

We now apply base change. The representation $\pi$ has the following properties:

(1) $\pi$ is cuspidal;

(2) $\pi_\infty$ is in the discrete series;

(3) $\pi_{p_1}$ is cuspidal;

(4) $\pi_{p_2}$ is an unramified twist of the Steinberg representation.

Consider the group $G_0^{*+} := \mathrm{Res}_{F/F^+}(G^*_{0,F})$. Let $\mathbb{A}_{F^+} := \mathbb{A} \otimes_{\mathbb{Q}} F^+$ and $\mathbb{A}_F := \mathbb{A} \otimes_{\mathbb{Q}} F$. Then $G_0^{*+}(\mathbb{A}_{F^+}) = G_0^*(\mathbb{A}_F)$. Because of the above properties (1), ..., (4), we may base change $\pi$ to an automorphic representation $BC(\pi)$ of $G_0^{*+}(\mathbb{A}_{F^+})$. Here we are using Corollary 5.3 from Labesse [**66**] to see that $\pi$ has a weak base change, and then the improvement of the statement at Theorem 5.9 of [*loc. cit.*], stating that [6], at the places where the unitary group is quasi-split (so in particular at $p$) the (local) base change of the representation $\pi_p$ is the representation $BC(\pi)_p$. By the same argument the base change $BC(\pi_0)$ exists as well. By strong multiplicity one for the group $G_0^{*+}$ we have $BC(\pi_\wp) \cong BC(\pi_{0,\wp})$ for all $F^+$-places $\wp$ above $p$.

We give the final argument when $F/F^+$ is inert at the $F^+$-place $\wp|p$, the case of the general linear groups being easier.

The representation $\pi_\wp$ is of the form $\mathrm{Ind}_{P(F^+_\wp)}^{G_1(F^+_\wp)}(\rho_\wp)$ because $\pi_p$ lies in the set $\mathfrak{R}_1(b)$. In this induction the parabolic subgroup $P$ has Levi component $M$ with $M(\mathbb{Q}_p) = M_{\wp,1} \times M_{\wp,2}$ with $M_{\wp,1}$ a general linear group and $M_{\wp,2}$ is a unitary group. The representation $\rho_\wp$ decomposes into $\rho_\wp \cong \rho_{\wp,1} \otimes \rho_{\wp,2}$, where $\rho_{\wp,1}$ is a generic unramified representation of $M_{\wp,1}$ and $\rho_{\wp,2}$ is an unramified twist of the Steinberg representation of $M_{\wp,2}$. The base change is compatible with parabolic induction, the base change of a generic unramified representation is again unramified [**77**] and the base change of a twist of the Steinberg representation is again a twist of the Steinberg representation [**78**]. Thus the representation $BC(\pi_\wp) \cong BC(\pi_{0,\wp})$ is an induction from a representation of the form

$$\Big(\chi_1, \chi_2, \ldots, \chi_{a_\wp}, \mathrm{St}_{\mathrm{GL}_b(F^+_\wp)}, \overline{\chi}_{a_\wp}^{-1}, \overline{\chi}_{a_\wp-1}^{-1}, \ldots, \overline{\chi}_1^{-1}\Big)$$

where $a_\wp = \mathrm{Rank}(M_{\wp,1})$ and $b_\wp = n - a_\wp$. Consequently, we have the character relations

$$\Theta_{\pi_{0,\wp}} \circ \mathcal{N} = \Theta_{\pi_\wp} \circ \mathcal{N}, \tag{7.12}$$

where $\mathcal{N}$ is the norm mapping from $G_0^{*+}(F^+_\wp)$ to $G_0^*(F^+_\wp)$. The norm mapping $\mathcal{N}$ from $\theta$-conjugacy classes in $G_0^{*+}(F^+_\wp)$ to $G_0^*(F^+_\wp)$ is surjective for the semi-simple conjugacy classes [**91**, Prop. 3.11(b)]. Thus the characters $\Theta_{\pi_\wp}$ and $\Theta_{\pi_{0,\wp}}$ coincide on $G_0(F^+_\wp)$. By Proposition 5.1 there is a positive constant $C_\Pi \in \mathbb{R}_{>0}$ such that (for $\alpha$ sufficiently divisible)

$$\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p) = C_\Pi \, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p}). \tag{7.13}$$

---

6. Labesse assumes that the extension $F^+/\mathbb{Q}$ is of degree at least 2. We we do not have this assumption. Labesse only needs his assumption to apply the simple trace formula. For our representation $\pi$ Labesse's assumption is redundant, because we have an auxiliary place ($v = p_1$) where the representation $\pi$ is cuspidal.

Remark: To find Equation (7.13) we applied Proposition (5.1) two times: first to compare $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)$ with $\mathrm{Tr}(\chi_b^{G_1 \times Z} f_\alpha^{G_1 \times Z}, \pi_p)$, and then to compare $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})$ with $\mathrm{Tr}(\chi_b^{G_1 \times Z} f_\alpha^{G \times Z_1}, \pi_{0,p})$.

We may now complete the proof. We return to Equation (7.11):

$$\sum m(\Pi) \, \mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty) \, \mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p) \dim \left( (\Pi_{p_3})^{K_{p_3}} \right) \prod_{i=1}^{d} \mathrm{Tr}(f^{G^*(\mathbb{Q}_{x_i})}, \Pi_{x_i}), \tag{7.14}$$

where $\Pi$ ranges over the irreducible subspaces of $L_0^2(A(\mathbb{R})^+ G^*(\mathbb{Q}) \backslash G^*(\mathbb{A}))$ satisfying the conditions listed below Equation (7.11). The following 6 facts have been established:

(1) The sum in Equation (7.14) is non-empty because $\Pi_0$ occurs in it (by the Propositions 4.8 and 4.10, and the Equations (7.7), (7.8), the term corresponding to $\Pi_0$ in the Sum (7.14) is non-zero).

(2) The multiplicity $m(\Pi)$ is a positive real number.

(3) For any $\Pi$ in Equation (7.14) with $\Pi \ncong \Pi_0$ we must have $\mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_\infty) = \mathrm{Tr}(f_\infty^{G^*(\mathbb{R})}, \Pi_{0,\infty})$ (here we use that $\xi$ is sufficiently regular).

(4) By Equation (7.13) the trace $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_p)$ equals $\mathrm{Tr}(\chi_b^{G(\mathbb{Q}_p)} f_\alpha, \Pi_{0,p})$ up to the positive number $C_\Pi$.

(5) The dimensions $\dim \left( (\Pi_{p_3})^{K_{p_3}} \right)$ and $\dim \left( (\Pi_{0,p_3})^{K_{p_3}} \right)$ differ by a positive real number.

(6) The product $\prod_{i=1}^{d} \mathrm{Tr}(f^{G^*(\mathbb{Q}_{x_i})}, \Pi_x)$ is a non-negative real number for all automorphic representations $\Pi$ contributing to Equation (7.14).

(facts (2) and (5) are trivial). From facts (1), (2), ..., (6) we conclude that Equation (7.14) must be non-zero. This completes the proof. □

CHAPTER 5

# Equidistribution

Let $S$ be an unitary Shimura variety of PEL type and consider a prime $\mathfrak{p}$ where $S$ has good reduction. The Newton stratification of $S$ modulo $\mathfrak{p}$ is a canonical decomposition of $S_{\mathbb{F}_q}$ into an union of locally closed subvarieties. These subvarieties are stable under the Hecke correspondences. We consider the supersingular stratum $B$ of $S_{\mathbb{F}_q}$ and work under the condition that $B$ is a finite variety and that the Shimura variety is a variety of Kottwitz (as in Chapter 2). The set of geometric points $B(\overline{\mathbb{F}}_q)$ is then a finite set, equipped with an action of the Hecke correspondences and the Frobenius element. We study the orbits of points $x \in B(\overline{\mathbb{F}}_q)$ under sequences of Hecke operators. We give an explicit description of these Hecke orbits and show, under mild conditions (§7), that the Hecke operators act inside the Hecke orbits with equidistribution. See Theorem 3.1 for the precise statement.

We would like to mention the work of Menares [**75**]. We learned the idea of equidistribution in supersingular Hecke orbits from his article. He proved that the Hecke operators $T_m$ for the group $\mathrm{GL}_2(\mathbb{Q})$ act with equidistribution on the supersingular stratum of the modular curve $X_0(p)$.

## 1. Some simple Shimura varieties

Consider the class of Shimura varieties of Kottwitz [**58**]. Such varieties are associated to a division algebra $D$ whose center is a CM field $F$. We will embed the field $F$ into the complex numbers, and we assume that $F$ splits into a compositum $F = \mathcal{K}F^+$ of a quadratic imaginary number field $\mathcal{K} \subset \mathbb{C}$ and a totally real number field $F^+$.

For any commutative $\mathbb{Q}$-algebra $R$, the group $G(R)$ is by definition the group of elements $g \in D \otimes_{\mathbb{Q}} \mathbb{R}$ such that $xx^* \in R^\times$. If $K \subset G(\mathbb{A}_f)$ is a compact open subgroup, sufficiently small, then we have a variety $\mathrm{Sh}_K$ defined over the reflex field $E$. Let $\mathfrak{p}$ be an $E$-prime where the variety $\mathrm{Sh}_K$ has good reduction in the sense defined by Kottwitz [**59**]. In particular $\mathrm{Sh}_K$ extends to a smooth and proper scheme defined over $\mathcal{O}_{E_\mathfrak{p}}$. We write $\mathbb{F}_q$ for the residue field of $E$ at $\mathfrak{p}$. Let $p$ be the rational prime number under $\mathfrak{p}$. We fix an embedding $\nu_p \colon E \to \overline{\mathbb{Q}}_p$ which is compatible with $\mathfrak{p}$.

We will always work under the assumption that the prime number $p$ is split in the field $\mathcal{K}$. Let $B$ be the supersingular locus of $\mathrm{Sh}_{K,\mathbb{F}_q}$ [**87**]. We assume that $B$ is a *finite variety*. In fact, among the set of all Kottwitz varieties, this rarely happens. However, the class of varieties for which $B$ is 0-dimensional is still quite interesting; for example it contains all the varieties considered by Harris and Taylor to prove the local Langlands conjecture [**45**].

To simplify the exposition, we also assume that the image of the group $K$ in the cocenter of the group $G$ is maximal.

The condition that $B$ is finite is a condition on the signatures of the unitary group at infinity, and the decomposition of the prime number $p$ in the field $F^+$. More explicitly, let $U \subset G$ be the subgroup of elements with trivial factor of similitude. Then $U(\mathbb{R})$ is isomorphic to a product of real, standard unitary groups $U(s_v, n - s_v)$, where $v$ ranges over the infinite $F^+$-places. We may, and do, assume that $s_v \leq \frac{1}{2}n$. The field $F^+$ is embedded into $\overline{\mathbb{Q}} \subset \mathbb{C}$ and also in $\overline{\mathbb{Q}}_p$, and therefore the group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ acts on the set of infinite $F^+$-places. In case $s_v > 1$ for some $v$, then certainly $B$ is infinite. Assume that $s_v \leq 1$ for all $v$. Then the variety $B$ is finite if and only if $s_v = 1$ for at most one infinite $F^+$-place in each $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-orbit of infinite $F^+$-places. For the proof of this statement, see § 3.4.3.

Let $A$ be the free complex vector space on the set $B(\overline{\mathbb{F}}_q)$. The Hecke algebra $\mathcal{H}(G(\mathbb{A}_{\mathrm{f}})//K)$ acts on the variety $B$ through correspondences and on the vector space $A$ via endomorphisms. Let $f_\infty$ be a function at infinity whose stable orbital integrals are prescribed by the identities of Kottwitz in [**57**]; it can be taken to be (essentially) an Euler-Poincaré function [**58**, Lemma 3.2] (cf. [**27**]). The function has the following property: Let $\pi_\infty$ be an $(\mathfrak{g}, K_\infty)$-module occurring as the component at infinity of an automorphic representation $\pi$ of $G$. Then the trace of $f_\infty$ against $\pi_\infty$ is equal to the Euler-Poincaré characteristic $\sum_{i=0}^{\infty} N_\infty (-1)^i \dim \mathrm{H}^i(\mathfrak{g}, K_\infty; \pi_\infty \otimes \xi)$, where $N_\infty$ is a certain explicit constant (cf. [**58**, p. 657, Lemma 3.2]).

By the main result of Chapter 2 we have for every Hecke operator $f^p \in \mathcal{H}(G(\mathbb{A}_{\mathrm{f}}^p)//K)$ that

$$\mathrm{Tr}(f^p \otimes \mathbf{1}_{K_p}, A) = \varepsilon \sum_{\pi \subset \mathcal{A}(G), \pi_p \text{ Steinberg type}} \mathrm{Tr}(f_\infty f, \pi^p) + \sum_{\pi \subset \mathcal{A}(G),\ \dim(\pi)=1} \mathrm{Tr}(f_\infty f, \pi^p), \tag{1.1}$$

where the sign $\varepsilon$ is equal to $(-1)^{t(n-1)}$ with $t$ the number of infinite $F^+$-places $v$ such that $p_v = 1$. We recall the definition of "Steinberg type":

DEFINITION 1.1. A smooth representation $\pi_p$ of $G(\mathbb{Q}_p)$ is of *Steinberg* type if the following two conditions hold: (1) For all $F^+$-places $\wp$ above $p$ we have

$$\pi_\wp = \begin{cases} \mathrm{St}_{\mathrm{GL}_n(F^+_\wp)} \otimes \phi_\wp & s_\wp = 1 \\ \text{Generic unramified} & s_\wp = 0, \end{cases}$$

where $\phi_\wp$ is an unramified character. (2) The factor of similitude $\mathbb{Q}_p{}^\times$ of $G(\mathbb{Q}_p)$ acts through an unramified character on the space of $\pi_p$.

We use the result in Equation (1.1) to deduce an equidistribution statement of Hecke operators acting on the basic stratum $B(\overline{\mathbb{F}}_q)$.

## 2. Hecke operators

In this section we define a sequence of Hecke operators $T_{r,m} \in \mathcal{H}(G(\mathbb{A}_\mathrm{f}))$. Consider the $\mathbb{Q}$-group $G_+ := \mathrm{Res}_{\mathcal{K}/\mathbb{Q}} G_{\mathcal{K}}$ with $G_+(\mathbb{Q}) = \mathcal{K}^\times \times D^\times$. Let $S$ be a finite set of finite, rational primes, such that:

(1) for all primes $\ell$ that do not lie in $S$, the group $G_+(\mathbb{Q}_\ell)$ is a product of general linear groups over finite, unramified extensions of $\mathbb{Q}_\ell$;

(2) $K$ splits into a product $K = K_S K^S$, where $K_S$ is a subgroup of $G(\mathbb{A}_{\mathrm{f},S})$ and $K^S$ is a subgroup of $G(\mathbb{A}_\mathrm{f}^S)$;

(3) The prime $p$ lies in $S$;

(4) The group $K^S$ is obtained by taking the $\widehat{\mathbb{Z}}^S$-points of a smooth model $\mathcal{G}$ of $G_+$ over the ring $\mathbb{Z}[\ell^{-1} | \ell \in S]$.

Let $G^+$ be the $\mathbb{Q}$-group $\mathcal{K}^\times \times \mathrm{GL}_n(F)$. Then $G^+$ is an inner form of $G_+$, and we have $G^+(\mathbb{Q}_\ell) \cong G_+(\mathbb{Q}_\ell)$ for all primes $\ell$ not in $S$. The group $G^+$ has an obvious model over $\mathbb{Z}$, and thus we have the hyperspecial subgroup $G^+(\widehat{\mathbb{Z}}) \subset G^+(\mathbb{A}_\mathrm{f})$. Let $m$ and $r$ be integers, where we have $0 \leq r \leq n$ (no condition on $m$). Then, by definition, the operator $T^+_{r,m}$ is defined to be the characteristic function:

$$T^+_{r,m} := \mathrm{char}\left( G^+(\widehat{\mathbb{Z}}) \cdot (1) \times \mathrm{diag}(\underbrace{m, m, \ldots, m}_{r}, 1, 1, \ldots 1) \cdot G^+(\widehat{\mathbb{Z}}) \right) \in \mathcal{H}(G^+(\mathbb{A}_\mathrm{f})), \tag{2.1}$$

where we should clarify the notation. We have $G^+(\widehat{\mathbb{Z}}) = \widehat{\mathcal{O}}_{\mathcal{K}}^\times \times \mathrm{GL}_n(\widehat{\mathcal{O}}_F)$, where $\widehat{\mathcal{O}}_{\mathcal{K}}^\times$ is the factor of similitude. With $(1) \times \mathrm{diag}(\ldots)$, we mean an element of $G^+(\widehat{\mathbb{Z}})$ that has trivial factor of similitude, and $\mathrm{diag}(\ldots)$ describes a diagonal matrix in the general linear group over $\widehat{\mathcal{O}}_F$.

Because the group $G_+(\mathbb{A}_\mathrm{f}^S)$ is isomorphic to $G^+(\mathbb{A}_\mathrm{f}^S)$, the operator $T^{+S}_{r,m} = \bigotimes_{\ell \notin S} T^{(\ell)}_{r,m}$ lives also in the algebra $\mathcal{H}(G_+(\mathbb{A}_\mathrm{f}^S))$. We have the base change morphism

$$\mathrm{BC} \colon \mathcal{H}(G_+(\mathbb{A}_\mathrm{f}^S) // G_+(\widehat{\mathbb{Z}}^S)) \longrightarrow \mathcal{H}(G(\mathbb{A}_\mathrm{f}^S) // K^S).$$

We define the operator $T^S_{r,m}$ to be $\mathrm{BC}(T^{+S}_{r,m})$, and we define

$$T_{r,m} := \mathbf{1}_{K^S} \otimes T^S_{r,m} \in C_c^\infty(G(\mathbb{A}_\mathrm{f}) // K). \tag{2.2}$$

We define the Hecke algebra $\mathcal{T} \subset C_c^\infty(G(\mathbb{A}_\mathrm{f}))$ to be the complex algebra generated by the operators $T_{r,m}$. The operators $T_{r,m}$ commute with each other, and satisfy no other algebraic relation. Thus the algebra $\mathcal{T}$ is isomorphic to the polynomial ring $\mathbb{C}[T_{r,m} | r, m]$ on a countable, infinite number of variables. The module $A$ is semi-simple as $\mathcal{H}(G(\mathbb{A}_\mathrm{f}^p))$-module (thus also as $\mathcal{T}$-module) because we know from our formula in Equation (1.1) that all irreducible subquotients occurring in $A$ occur in the discrete spectrum of $G$.

Using $K$ we define the *degree* of the operator $T_{r,m}$ via the integral

$$\deg(T_{r,m}) := \int_{G(\mathbb{A}_\mathrm{f})} T_{r,m}(g) \, \mathrm{d}\mu(g), \tag{2.3}$$

where the Haar measure $\mu$ on $G(\mathbb{A}_f)$ is normalized so that it gives $K$ measure 1.

## 3. Hecke orbits

The Hecke algebra $\mathcal{T}$ does not act transitively on the supersingular stratum; there are two innocent obstructions: (1) an obstruction from the cocenter of the group $G$, and (2) the Hasse invariant $\mathrm{Ker}^1(G, \mathbb{Q})$, which need not be trivial. In this section we will define certain 'candidate' orbits of $\mathcal{T}$ acting on $B$. Our main Theorem will state that $\mathcal{T}$ acts transitively and with equidistribution on these orbits.

Note that obstructions (1) and (2) are what one expects: For the first one (1): If the image of $K \subset G(\mathbb{A}_f)$ in the cocenter $C(\mathbb{A}_f)$ is not sufficiently large (and this will always be the case for many $C$, due to the presence of Abelian class groups), then the double coset space

$$G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K), \tag{3.1}$$

is not connected, and a point in one connected component will be sent by a Hecke operator to another connected component only if this operator is non-trivial on the cocenter. However, our operators in $\mathcal{T}$ all act trivially.

The second condition (2) is there because $\mathrm{Sh}_K(\mathbb{C})$ is *not* equal to the double coset space in Equation (3.1), rather it is a disjoint union

$$\mathrm{Sh}_K(\mathbb{C}) = \coprod_{\mathrm{Ker}^1(G:\mathbb{Q})} G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K), \tag{3.2}$$

of copies of this double coset space, indexed by the group $\mathrm{Ker}^1(G : \mathbb{Q})$ (this group depends only on the cocenter of $G$, and is trivial in case $n$ is even, see [**59**, p. 393]). The Hecke correspondences act on the right hand side via their natural action on the double coset spaces. Thus, clearly, over $\mathbb{C}$, all points in a Hecke orbit will have the same invariant in $\mathrm{Ker}^1(G : \mathbb{Q})$.

Let $d\colon G \twoheadrightarrow C$ be the cocenter of the group. We have the morphism $h$ from Deligne's torus $\mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m$ to $G_{\mathbb{R}}$. By composing this morphism with the natural morphism we obtain a morphism $h'\colon \mathbb{S} \to C_{\mathbb{R}}$. The couple $(C, \{h'\})$ is a zero dimensional Shimura datum. Deligne [**29**] has proved that $\mathrm{Sh}(C, \{h'\})$ parametrizes the connected components of the original variety, i.e. the natural morphism

$$\pi_0(G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K)) \longrightarrow C(\mathbb{Q})\backslash(\{h\} \times C(\mathbb{A}_f)/d(K)), \tag{3.3}$$

is an isomorphism. Via this mapping, the action of the Hecke operator $f \in \mathcal{H}(G(\mathbb{A}_f))$ on the left hand side coincides with the action of the operator $\Psi(f)$ in $\mathcal{H}(C(\mathbb{A}_f))$ on the right hand side. Here the map $\Psi\colon \mathcal{H}(G(\mathbb{A}_f)) \to \mathcal{H}(C(\mathbb{A}_f))$ is characterized by

$$\forall c \in C(\mathbb{A}_f)\ \forall f \in \mathcal{H}(G(\mathbb{A}_f)) :$$

$$[\Psi f]\,(c) = \begin{cases} \int_{G_{\mathrm{der}}(\mathbb{A}_f)} f(gh)\mathrm{d}\mu(h) & \text{if } c = \overline{g} \in \mathrm{Im}(G(\mathbb{A}_f) \to C(\mathbb{A}_f)) \\ 0 & \text{otherwise,} \end{cases}$$

where the Haar measure on $G_{\mathrm{der}}(\mathbb{A}_{\mathrm{f}})$ is the one which gives the group $K \cap G_{\mathrm{der}}(\mathbb{A}_{\mathrm{f}})$ volume 1. Let $E$ be the reflex field of the datum $(G, X)$. Deligne proved that the mapping in Equation (3.3) is $\mathrm{Aut}(\mathbb{C}/E)$-equivariant. Thus the map in Equation (3.3) descents to an isomorphism of $E$-schemes $\pi_0(\mathrm{Sh}(G : K)) \xrightarrow{\sim} \mathrm{Sh}(C : d(K))$. The variety $\mathrm{Sh}_K$ is an union of $\# \operatorname{Ker}^1(G, \mathbb{Q})$ copies of the variety $\mathrm{Sh}(G : K)$ [**59**, §6]. We obtain an $E$-isomorphism

$$\pi_0(\mathrm{Sh}_K) \xrightarrow{\sim} \coprod_{\operatorname{Ker}^1(G,\mathbb{Q})} \mathrm{Sh}(C : d(K)). \tag{3.4}$$

Both sides are finite étale $E$-schemes and the $\mathrm{Gal}(\overline{\mathbb{Q}}/E)$-action is unramified at $\mathfrak{p}$. Locally at the prime $\mathfrak{p}$ we have a natural model of $\mathrm{Sh}_K$ over the ring of integers $\mathcal{O}_{E_{\mathfrak{p}}}$, and we construct a model of the right hand side in the straightforward manner: Take the global sections $A$ of the scheme $\coprod_{\operatorname{Ker}^1(G,\mathbb{Q})} \mathrm{Sh}(C : d(K))_{E_{\mathfrak{p}}}$. Then $A$ is a $\mathbb{Q}_p$-algebra; let $A^\circ \subset A$ be the integral closure of $\mathbb{Z}_p$ in $A$. Then $\mathrm{Spec}(A^\circ)$ is our integral model. We write $Y = \mathrm{Spec}(A^\circ)$ and view it as a as scheme over $\mathcal{O}_{E,\mathfrak{p}}$. We reduce the map in Equation (3.4) modulo $\mathfrak{p}$ and compose with $B \subset \mathrm{Sh}_{K,\mathfrak{p}} \twoheadrightarrow \pi_0(\mathrm{Sh}_{K,\mathfrak{p}})$ to obtain the map

$$\psi \colon B \longrightarrow Y \tag{3.5}$$

For each point $y \in Y$ we have the fibre $B_y$ of $\psi$ above $y$. Define $A_y$ to be the free complex vector space on the set $B_y(\overline{\mathbb{F}}_q)$. Then $A$ is the direct sum of the $A_y$ with $y$ ranging over the set $Y(\overline{\mathbb{F}}_q)$. For each $y \in Y$ we have the map (of vector spaces):

$$\Psi_y \colon A_y \longrightarrow \mathbb{C}, \qquad \sum_{x \in B_y(\overline{\mathbb{F}}_q)} a_x \cdot [x] \longmapsto \sum_{x \in B_y(\overline{\mathbb{F}}_q)} a_x. \tag{3.6}$$

Write $E_y = \sum_{x \in B_y(\overline{\mathbb{F}}_q)} [x] \in A_y$. Define the endomorphism

$$\mathrm{Avg}_y \colon A_y \longrightarrow A_y, \quad v \mapsto \frac{\Psi_y(v)}{\# B_y(\overline{\mathbb{F}}_q)} \cdot E_y.$$

The fibres $B_y(\overline{\mathbb{F}}_q)$ are all of the same cardinality $\# C(\widehat{\mathbb{Z}})/d(K)$. Take the direct sum of $\mathrm{Avg}_v$ over all $y \in Y$ to obtain an endomorphism

$$\mathrm{Avg} \colon A \longrightarrow A \tag{3.7}$$

which takes the 'average' of an element $v \in A$ along the fibres of the mapping $\psi \colon B \to Y$. We will prove that any element $v \in A$ will converge to its average under the action of the sequence of Hecke operators $T_{r,m} \in \mathcal{T}$.

The complex vector space $A$ is finite dimensional and therefore carries a norm $|\cdot|$, uniquely defined up to equivalence of norms. Using this norm we may give the statement of the main Theorem:

THEOREM 3.1. *Let $v \in A$ be an element. Then there exists a constant $C \in \mathbb{R}_{>0}$ such that for any $\varepsilon > 0$ there exist an index $M$, such that for all square free integers $m > M$ and all $r$*

*with* $1 \le r \le n-1$ *we have*

$$\left| \frac{T_{r,m}(v)}{\deg(T_{r,m})} - \mathrm{Avg}(v) \right| \le C m^{\varepsilon - [F:\mathbb{Q}]\frac{r(n-r)}{2}}.$$

Remark. With the same method of proof we obtain equidistribution results also for other sequences of Hecke operators. For example, fix an operator $T \in \mathcal{T}$ and consider the sequence of its powers for the convolution product $\{T^N\}_{N \in \mathbb{Z}_{\ge 1}}$. Of course, the rate of convergence will depend on the sequence of operators you choose.

Remark. Perhaps one could relax the condition that $m$ be square free somewhat. One will then have to deal with some combinatorial issues related to the Satake transform. The condition becomes relevant at Equation 4.4 of the proof; the resulting combinatorial problem is discussed (for example) in the article [**41**].

In sections 4–5 we prove Theorem 3.1.

## 4. A vanishing statement

Observe that to the character formula for $A$ in Equation (1.1) expresses $A$ as a sum of Hecke modules of the form $(\pi_{\mathrm{f}}^p)^{K^p}$. We define $A_0 \subset A$ to be the $\mathcal{T}$-submodule generated by modules $(\pi_{\mathrm{f}}^p)^{K^p}$ for $\pi$ an infinite dimensional automorphic representation of $G(\mathbb{A})$.

The following Proposition proves the essential part of Theorem 3.1.

Proposition 4.1. *Let* $v \in A_0$*, then there exists a constant* $C \in \mathbb{R}_{>0}$ *such that for all integers* $r$ *with* $0 \le r \le n$ *and all square free integers* $m$ *coprime to* $S$ *we have*

$$\left| \frac{T_{r,m}(v)}{\deg(T_{r,m})} \right| \le C \binom{n}{r}^{c_F(m)} m^{-[F:\mathbb{Q}]\frac{r(n-r)}{2}}.$$

Notation. Let $m$ be a positive integer, unramified in $F$. We wrote $c_F(m)$ for the number of $\mathcal{O}_F$-prime ideals $\lambda$ containing the number $m$.

Proof. By our Theorem in Equation (1.1) it suffices to prove that the limit $\lim_{m\to\infty} \frac{T_{r,m}(v)}{\deg(T_{r,m})}$ vanishes[1] for each vector $v \in \pi_{\mathrm{f}}^K$ in each automorphic representation $\pi$ contributing to the character formula of $A_0$. Let $\pi$ be one of these cuspidal automorphic representations. We may use base change to send $\pi$ to an automorphic representation $BC(\pi)$ of the algebraic group $\mathcal{K}^\times \times D^\times$ (see [**4**]), and we may send the automorphic representation $BC(\pi)$ to an automorphic representation $\Pi := JL(BC(\pi))$ of the algebraic group $G^+ := \mathcal{K}^\times \times \mathrm{GL}_n(F)$ (see [**101**] and [**6**]). This automorphic representation is *discrete*. At $p$ we have $G^+(\mathbb{Q}_p) \cong G(\mathbb{Q}_p) \times G(\mathbb{Q}_p)$ and $\Pi_p$ is isomorphic to $\pi_p \otimes \pi_p$. The representation $\pi_p$ is essentially square integrable because it is an unramified twist of the Steinberg representation,

1. Here, and hereafter, when we say "limit" or "vanishes", we mean that this limit does so with the correct rate of convergence stated in the Proposition.

and therefore $\Pi_p$ also has this property. The representation $\Pi$ is then forced to be cuspidal by the classification of Moeglin-Waldspurger of the discrete spectrum [**80**].

Because $\Pi$ is cuspidal the Ramanujan conjecture applies to it. This conjecture is true for $\Pi$ because $\Pi$ is obtained by base change and Jacquet-Langlands from an automorphic representation $\pi$ of an unitary group (of similitudes). Thus $\Pi$ is conjugate self-dual. Furthermore, $\Pi$ is cohomological because $\pi$ has the property that $\mathrm{Tr}(f_\infty, \pi_\infty) \neq 0$. For such representations $\Pi$ the conjecture is proved to be true in the articles [**14**, **25**, **95**]. Thus the components $\Pi_\lambda$ are *tempered* $\mathrm{GL}_n(F_\lambda)$-representations for all primes $\lambda$ of $F$.

The non-trivial element $\theta$ of the group $\mathrm{Gal}(\mathcal{K}/\mathbb{Q})$ acts on the group $G^+$, and, 'par transport du structure', $\theta$ acts on the space of automorphic forms $\mathcal{A}(G^+)$ on $G^+$. The transferred representation $\Pi$ is $\theta$-stable. On the one hand the two isomorphic representations $\Pi$ and $\Pi^\theta$ both occur as subspaces in $\mathcal{A}(G^+)$, and on the other hand we have (strong) multiplicity one for the group $G^+$. Therefore $\Pi$ and $\Pi^\theta$ are the same subspace and we have a natural isomorphism $A_\theta \colon \Pi \xrightarrow{\sim} \Pi^\theta$ induced by $\theta$ acting on the space $\mathcal{A}(G^+)$.

We must show that the limit $\lim_{m\to\infty} T_{r,m}(v)/\deg(T_{r,m})$ vanishes for all vectors $v \in \pi_{\mathrm{f}}^K$. Let $v$ be one such vector, and assume that $v \neq 0$. We have $\pi \cong \pi_S \otimes \pi^S$ and we may assume that $v$ is an elementary tensor $v = v_S \otimes v^S$, with $v_S \in \pi_S$ and $v^S \in \pi^S$. To prove that the limit $\lim_{m\to\infty} T_{r,m}(v)/\deg(T_{r,m})$ vanishes it suffices to prove that the limit $\lim_{m\to\infty} T_{r,m}(v^S)/\deg(T_{r,m})$ vanishes. The space $\pi^{S,K^S}$ is one-dimensional and $v^S$ is a basis of this space. Therefore $\mathrm{Tr}(T^S_{r,m}, \pi^S)$ is the scalar $\lambda$ such that $T^S_{r,m}(v) = \lambda v$. Up to possibly a sign we have $\lambda = \mathrm{Tr}(T^{+S}_{r,m}, \Pi^S)$, and thus

$$\forall r, m : \quad \left|\frac{T_{r,m}(v^S)}{\deg(T_{r,m})}\right| \leq C \left|\frac{\mathrm{Tr}(T^{+S}_{r,m}, \Pi^S)}{\deg(T_{r,m})}\right|, \tag{4.1}$$

for some constant $C$ which does not depend on $r, m$.

To bound the right hand side of Equation (4.1) we will focus first on the degree $\deg(T_{r,m})$, and we will compare it with the classical notion of degree $\deg(T^+_{r,m})$ for the Hecke operators on the general linear group. It suffices to do this comparison up to a constant independent of $r, m$. The function $T_{r,m}$ on $G(\mathbb{A}_{\mathrm{f}})$ is the transfer of the function $T^+_{r,m}$ on $G^+(\mathbb{A}_{\mathrm{f}})$ via the functorialities $G \rightsquigarrow G_+ \rightsquigarrow G^+$ (base change and Jacquet-Langlands respectively). The transfer of the trivial representation along these functorialities is the again the trivial representation. Thus, up to a constant $C$ not depending on $r, m$, we have

$$\deg(T_{r,m}) = C \int_{G^+(\mathbb{A}^S_{\mathrm{f}})} T^+_{r,m}(g)\, \mathrm{d}\mu(g), \tag{4.2}$$

which is (up to a constant) the volume of the subset

$$G^+(\widehat{\mathbb{Z}}) \cdot (1) \times \mathrm{diag}(m, m, \ldots, m, 1, 1, \ldots 1) \cdot G^+(\widehat{\mathbb{Z}}) \subset G^+(\mathbb{A}_{\mathrm{f}}), \tag{4.3}$$

In turn this volume is the number of right $G^+(\widehat{\mathbb{Z}})$-cosets of the subset in Equation (4.3), and this gives back the classical notion of "degree".

Let $\ell$ be a prime divisor of $m$. Because $m$ is square free, the prime $\ell$ divides $m$ precisely once, and the $\ell$-th part of the function $T^+_{r,m}$ equals

$$\text{char}\left(G^+(\mathbb{Z}_\ell)\cdot(1)\times\text{diag}(\ell,\ell,\ldots,\ell,1,1,\ldots,1)\cdot G^+(\mathbb{Z}_\ell)\right)\in\mathcal{H}(G(\mathbb{Q}_\ell)). \tag{4.4}$$

The element $(1)\times\text{diag}(\ell,\ell,\ldots,\ell,1,1,\ldots,1)\in G^+(\mathbb{Q}_\ell)$ is the evaluation at $\ell$ of a *miniscule* cocharacter $\mu_r\in X_*(G^+)$. The field $F$ is unramified above $\ell$, and therefore $\ell$ is a prime element of the local field $F_\lambda$ for every $F$-place $\lambda$ dividing $\ell$. Because $\mu_r$ is miniscule there is a simple formula for the Satake transform of $T^{+(\ell)}_{r,m}$ (cf. [**54**]):

$$\mathcal{S}(T^{+(\ell)}_{r,m})=1\otimes\bigotimes_{\lambda|\ell}q_\lambda^{\frac{r(n-r)}{2}}\sum_{1\le i_1<i_2<\cdots<i_r\le n}X_{i_1}X_{i_2}\cdots X_{i_r}, \tag{4.5}$$

in the algebra

$$\mathbb{C}[X_*(T_{\mathbb{Q}_\ell})]=\mathbb{C}[Z]\otimes\bigotimes_{\lambda|\ell}\mathbb{C}[X_1^{\pm1},X_2^{\pm1},\ldots,X_n^{\pm1}], \tag{4.6}$$

where $T_{\mathbb{Q}_\ell}\subset G^+_{\mathbb{Q}_\ell}$ is the diagonal torus. We specify that the big tensor product in these Equations ranges over all the $F$-places $\lambda$ lying above $\ell$, and for such an $F$-place $\lambda$, we write $q_\lambda$ for the cardinality of the residue field at $\lambda$.

The degree $\deg(T_{r,m})$ is the evaluation of the polynomial $\mathcal{S}(T^{(\ell)}_{r,m})$ at the Hecke matrix of the trivial representation $\varphi_{\text{Triv}}$, and is therefore made completely explicit at this point. We may now estimate $|\mathcal{S}(T^{(\ell)}_{r,m})(\varphi_{\text{Triv}})|$. If we evaluate the symmetric polynomial

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n}X_{i_1}X_{i_2}\cdots X_{i_r},$$

at the Hecke matrix of the trivial representation of $\text{GL}_n(F_\lambda)$, then the largest monomial which appears is

$$q_\lambda^{\frac{n-1}{2}+\frac{n-3}{2}+\ldots+\frac{n-2r+1}{2}}=q_\lambda^{\frac{r(n-r)}{2}}.$$

Thus we have the following lower bound:

$$|\mathcal{S}(T^{+(\ell)}_{r,m})(\varphi_{\text{Triv}})|\ge\prod_{\lambda|\ell}q_\lambda^{\frac{r(n-r)}{2}}=\ell^{[F:\mathbb{Q}]\frac{r(n-r)}{2}}.$$

The representation $\Pi_\ell$ is tempered, and therefore the absolute values of the eigenvalues of its Hecke matrix are all equal to 1. Thus

$$|\mathcal{S}(T^{+(\ell)}_{r,m})(\varphi_{\Pi_\ell})|\le\prod_{\lambda|\ell}\binom{n}{r}=\binom{n}{r}^{c_F(\ell)}.$$

We now return to the estimation started in Equation (4.1). We have

$$\left|\frac{\text{Tr}(T^+_{r,m,S},\Pi_S)}{\deg(T_{r,m})}\right|\le C\prod_{\ell|m}\left|\frac{\mathcal{S}(T^{(\ell)}_{r,m})(\varphi_{\Pi_\ell})}{\mathcal{S}(T^{(\ell)}_{r,m})(\varphi_{\text{Triv}})}\right|\le C\binom{n}{r}^{c_F(m)}m^{-[F:\mathbb{Q}]\frac{r(n-r)}{2}},$$

where $C$ is a certain constant not depending on $r$ and $m$. This completes the proof. □

To turn the convergence rate of the above Proposition 4.1 to the convergence rate of the Theorem 3.1 we weaken our result slightly using Stirling's formula.

LEMMA 4.2. *For any $\varepsilon > 0$ there exists an integer $M > 0$ such that for all square free $m > M$ we have* $\binom{n}{r}^{c_F(m)} \leq m^\varepsilon$

REMARK. We prove this for $F = \mathbb{Q}$; we leave it to the reader to reduce to this case, or to extend the argument below.

PROOF OF LEMMA 4.2. We have $(c_{\mathbb{Q}}(m))! \leq m$. Write $m = \Gamma(x)$ for some $x \in \mathbb{R}_{\geq 0}$ where $\Gamma$ is the usual Gamma function. Then $c_{\mathbb{Q}}(m) \leq x$ and from Stirling's formula we get

$$\frac{c_{\mathbb{Q}}(m)}{\log(m)} \sim \frac{c_{\mathbb{Q}}(m)}{\log\left(\sqrt{2\pi x} e^{-x} x^x\right)} \leq \frac{1}{\log(x) - 1 + \frac{\log(\sqrt{2\pi x})}{x}}.$$

The right hand side converges to 0 for $x \to \infty$. Thus we may find (for any $\varepsilon > 0$) an $M$ such that $\exp(c_{\mathbb{Q}}(m)) \leq m^\varepsilon$ for all $m > M$. This completes the proof. □

## 5. Completion of the proof

The proof of the main theorem is now not more than a formality. Recall that in Equation (3.6) we constructed, for each point $y \in Y(\overline{\mathbb{F}}_q)$, a mapping $\Psi_y \colon A_y \twoheadrightarrow \mathbb{C}$. We may take the sum over all $y$ and obtain in this way an equivariant surjection from $A$ onto the free complex vector space $A_{\mathrm{Ab}}$ on the set $Y(\overline{\mathbb{F}}_q)$. Then $A_{\mathrm{Ab}}$ accounts precisely for the contribution of the one dimensional representations

$$\bigoplus_{\pi \subset \mathcal{A}(G), \dim(\pi)=1, \pi_\infty = 1} \pi_{\mathrm{f}}^K,$$

to the automorphic character formula for $A$ (cf. Equation (1.1)). We have an exact sequence $A_0 \rightarrowtail A \twoheadrightarrow A_{\mathrm{Ab}}$ of Hecke modules, and the 'average' mapping $\mathrm{Avg} \colon A_{\mathrm{Ab}} \to A$ from Equation (3.7) splits this sequence. For $v = v_0 + v_{\mathrm{Ab}} \in A_0 \oplus A_{\mathrm{Ab}}$ we have $v_{\mathrm{Ab}} = \mathrm{Avg}(v)$ on the one hand, and on the other hand the sequence $\frac{T_{r,m}(v)}{\deg(T_{r,m})}$ converges to $v_{\mathrm{Ab}}$ with the correct rate of convergence by Proposition 4.1 (and Lemma 4.2). This completes the proof. □

## 6. Towards the general case of unitary Shimura varieties

In this section we sketch how to extend result of our article [**63**] to a larger class of Shimura varieties which may have endoscopy and be non-compact, but satisfy a simplifying condition on the basic isocrystal.

The discussion in this section is still incomplete, because there are corrective terms in the trace formula which need to be estimated. We have not yet done this estimation.

We will consider a Shimura variety of PEL-type, of type A, as considered by Kottwitz in [**59**]. Thus we assume fixed a PEL-datum consisting of

**(A1)** A simple algebra[2] $Y$ over a CM field $F$;

**(A2)** A positive involution on the algebra $Y$ which induces complex conjugation on $F$;

**(A3)** A Hermitian $Y$-module $(V, \langle\cdot,\cdot\rangle)$, where $\langle\cdot,\cdot\rangle$ is symplectic;

**(A4)** $h\colon \mathbb{C} \to \mathrm{End}_Y(V)_\mathbb{R}$ is a morphism of $\mathbb{R}$-algebras such that $h(\overline{z}) = h(z)^*$ for all $z \in \mathbb{C}$.

Let $(G, X)$ be the Shimura datum associated to (**A1**), (**A2**), (**A3**) and the morphism $h^{-1}$. We assume that there is a quadratic imaginary extension $\mathcal{K}$ of $\mathbb{Q}$ and a totally real extension $F^+$ of $\mathbb{Q}$ such that $F = \mathcal{K}F^+$. Then the group $G_\mathcal{K}$ is isomorphic to a product of (Weil-restriction of scalars of) general linear groups. We let $p$ be a prime of good reduction in the sense of Kottwitz [**59**, §5] and we assume that $p$ splits in $\mathcal{K}/\mathbb{Q}$. We write $E$ for the reflex field of the Shimura datum. Furthermore we let $K \subset G(\mathbb{A}_\mathrm{f})$ be a compact open subgroup of the form $K = K_pK^p$, with $K_p$ hyperspecial and $K^p$ sufficiently small so that the PEL-type moduli problem of level $K$ is defined over $\mathcal{O}_E \otimes \mathbb{Z}_{(p)}$ and the variety $\mathrm{Sh}_K$ is smooth and quasi-projective.

Pick an $E$-prime $\mathfrak{p}$ above $p$ and let $B$ be the basic stratum of the variety $\mathrm{Sh}_{K,\mathbb{F}_q}$, where $\mathbb{F}_q$ is the residue field of $\mathcal{O}_E$ at $\mathfrak{p}$. We pick an embedding of $\overline{\mathbb{Q}} \to \overline{\mathbb{Q}}_p$ which extends the embedding of $E$ into $\overline{\mathbb{Q}}_p$ defined by $\mathfrak{p}$. We fix once and for all an embedding of $F$ into $\mathbb{C}$, and $\overline{\mathbb{Q}}$ will always mean the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. The field $\overline{\mathbb{F}}_q$ is the residue field of $\overline{\mathbb{Q}}_p$ and the field $\mathbb{F}_q$ is the residue field of $E$ at $\mathfrak{p}$.

Because we have the embeddings $F \subset \overline{\mathbb{Q}} \subset \overline{\mathbb{Q}}_p$, the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ acts on the set of infinite $F^+$-places $V(F^+)$ and we may identify any $\wp|p$ with a Galois orbit $V(\wp)$ of infinite places. Let $U \subset G$ be the subgroup of elements with trivial factor of similitude. Then $U(\mathbb{R})$ is a product of standard real groups: $U(\mathbb{R}) = \prod_{v \in V(F^+)} U(s_v, n - s_v)$ for certain numbers $s_v$. We assume that $s_v \leq \frac{1}{2}n$ so that these numbers are well defined. The additional technical condition that we make is the following:

HYPOTHESIS 6.1. *There exists an $F^+$-prime $\wp$ such that the number $s_\wp$ is coprime to $n$.*

Let $\alpha$ be a integer. Consider the function $f = f_\infty f_\alpha f^p$ in the Hecke algebra of $G$, where $f_\infty$ is a Clozel-Delorme function for the trivial complex representation of $G_\mathbb{C}$ and $f^p \in \mathcal{H}(G(\mathbb{A}_\mathrm{f}^p))$ is any $K^p$-spherical Hecke operator. Let $\ell$ be a prime number different from $p$ and fix an isomorphism $\overline{\mathbb{Q}}_\ell \cong \mathbb{C}$ of abstract fields. Without further mention, we will use this isomorphism to turn the complex valued function $f^{p\infty}$ into a function which is $\overline{\mathbb{Q}}_\ell$-valued in the cases where this is necessary. Write $\iota$ for the inclusion $B \hookrightarrow \mathrm{Sh}_{K,\mathbb{F}_q}$. Recall that the article [**59**] gives the result:

(6.1)
$$\sum_{x' \in \mathrm{Fix}_{\Phi_\mathfrak{p}^\alpha \times f^{\infty p}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(\Phi_\mathfrak{p}^\alpha \times f^{\infty p}, \iota^*(\overline{\mathbb{Q}}_\ell)_x) = |\ker^1(\mathbb{Q}, G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\phi_\alpha),$$

2. The notation $Y$ is nonstandard and questionable; we use it because the usual notation $B$ conflicts with our notation for the basic stratum.

where the notations are from [§19, *loc. cit*]. We restrict this formula to the basic stratum $B$ by considering on the right hand side only *basic* Kottwitz triples. The equation then becomes

$$\sum_{x' \in \mathrm{Fix}^B_{\Phi^\alpha_\mathfrak{p} \times f^{\infty p}}(\overline{\mathbb{F}}_q)} \mathrm{Tr}(\Phi^\alpha_\mathfrak{p} \times f^{\infty p}, \iota^*(\overline{\mathbb{Q}}_\ell)_x) =$$

$$= |\ker^1(\mathbb{Q}, G)| \sum_{(\gamma_0;\gamma,\delta)} c(\gamma_0;\gamma,\delta) O_\gamma(f^{\infty p}) TO_\delta(\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma c} \phi_\alpha), \tag{6.2}$$

where $E_{\mathfrak{p},\alpha}$ is the unramified extension of degree $\alpha$ of $E_\mathfrak{p}$ (in $\overline{\mathbb{Q}}_p$). The function $\chi^{G(E_{\mathfrak{p},\alpha})}_{\sigma c}$ is the characteristic function of the subset of $\sigma$-compact elements in $G(E_{\mathfrak{p},\alpha})$, and $\mathrm{Fix}^B_{\Phi^\alpha_\mathfrak{p} \times f^{\infty p}}$ is the fibre product $\mathrm{Fix}_{\Phi^\alpha_\mathfrak{p} \times f^{\infty p}} \times_\Delta B$, where $\Delta$ is the diagonal variety in $\mathrm{Sh}_{K,\mathbb{F}_q} \times \mathrm{Sh}_{K,\mathbb{F}_q}$. By the stabilization argument of Kottwitz in [**57**] the right hand side of Equation (6.2) simplifies to

$$\sum_{\mathcal{E}} \iota(G,H) \cdot \mathrm{ST}^*_e(\chi^{G(\mathbb{Q}_p)}_c f^H), \tag{6.3}$$

where $\mathcal{E}$ is the set of isomorphism classes of elliptic endoscopic triples of $G$, and $\mathrm{ST}^*_e$ is a sum of stable integral orbitals on the elliptic $(G,H)$-regular elements in $H(\mathbb{Q})$ [**57**]. If $H$ is the maximal endoscopic group, then this regularity condition is empty, and we have $\mathrm{ST}^*_e(\chi^{G(\mathbb{Q}_p)}_c f) = \mathrm{ST}_e(\chi^{G(\mathbb{Q}_p)}_c f)$. We also mention that the notation $\chi^{G(\mathbb{Q}_p)}_c f^H$ is slightly abusive, because $f^H$ is a global function, while $\chi^{G(\mathbb{Q}_p)}_c$ is a function at $p$. When we write the product $\chi^{G(\mathbb{Q}_p)}_c f$ we actually mean the function $f^p \otimes (\chi^{G(\mathbb{Q}_p)}_c f_\alpha)$, so the truncation only occurs at $p$.

Lemma 6.2. *Let $P \subset G(\mathbb{Q}_p)$ be a proper standard parabolic subgroup of $G(\mathbb{Q}_p)$. Then the truncated constant term $\chi^{G(\mathbb{Q}_p)}_c f^{(P)}_\alpha$ vanishes.*

Proof. Let $P$ be a parabolic subgroup of $G(\mathbb{Q}_p)$. We have $f^{(P)}_\alpha = \mathbf{1}_{q^{-\alpha}} \otimes \bigotimes_{\wp|p} \prod f^{(P_\wp)}_{n\alpha s}$, where $P_\wp$ is the $\wp$-th component of $P$. If $P$ is proper, then $P_\wp$ is proper as well. Pick some $\wp|p$ such that $s_\wp$ is coprime to $n$ (Hypothesis 6.1). We look at the $\wp$-th component $f^\wp_\alpha$ of the function $f_\alpha \in \mathcal{H}_0(G(\mathbb{Q}_p))$ via the isomorphism $\mathcal{H}_0(G(\mathbb{Q}_p)) \cong \mathcal{H}_0(\mathbb{Q}_p{}^\times) \otimes \bigotimes_{\wp|p} \mathcal{H}_0(\mathrm{GL}_n(F^+_\wp))$. In the notation of [**63**], we have $f^\wp_\alpha = f_{n\alpha_v s_v}$ [Prop. 3.3, *loc. cit.*]. By the explicit description in [Lem. 1.9, *loc. cit.*] of the truncated constants terms of $f_{n\alpha_v s_v}$ we see that these constant terms vanish for the proper parabolic subgroups in case $s$ is coprime to $n$. □

Proposition 6.3. *For any proper endoscopic group $H$ of $G$ we have $(\chi^{G(\mathbb{Q}_p)}_c f)^H_\alpha = 0$.*

Proof. The transfer $f \rightsquigarrow f^H$ from the function on $G(\mathbb{A})$ to the endoscopic group $H(\mathbb{A})$ factors through the transfer from $G$ to its quasi-split inner form $G^*$. At $p$, the group $G(\mathbb{Q}_p)$ is quasi split and therefore $G(\mathbb{Q}_p) = G^*(\mathbb{Q}_p)$ and we take the transfer from functions on $G(\mathbb{Q}_p)$ to functions on $G^*(\mathbb{Q}_p)$ to be trivial. Thus we must transfer the function $\chi^{G(\mathbb{Q}_p)}_c f_\alpha$ on $G^*(\mathbb{Q}_p)$ to $H(\mathbb{Q}_p)$. We first consider the function $f_\alpha \in \mathcal{H}_0(G^*(\mathbb{Q}_p))$ ($\mathcal{H}_0$ denotes the spherical Hecke algebra). In section 3.4, case 2 on page 1668 of [**95**], Sug Woo Shin describes explicitly the transfer for quasi-split similitudes unitary groups. He starts by describing the endoscopic groups, and explains that any $H$ can be identified with a group of the form $G(GU^*(n_1) \times GU^*(n_2))$ with $n = n_1 + n_2$ (there are some conditions on the possible partitions $n = n_1 + n_2$ here, but they are of no importance to us). In particular we assume $H$ is the Levi-component of a maximal standard parabolic subgroup $P_H$ of $G^*$. By the second last displayed formula on page 1668 of [*loc. cit*] the transfer of $f_\alpha$ to a function on $H(\mathbb{Q}_p)$ is given by $f_\alpha^{(P_H(\mathbb{Q}_p))} \cdot \chi^+_{\varpi,u}$, where $\chi^+_{\varpi,u}$ is some function which we will not need to specify for our argument. The transfer of a conjugacy class in $H(\mathbb{Q}_p)$ to a conjugacy class in $G^*(\mathbb{Q}_p)$ is the obvious construction (i.e. induced from the inclusion $H(\mathbb{Q}_p) \subset G^*(\mathbb{Q}_p)$). Consequently the function

$$\left(\chi_c^{G(\mathbb{Q}_p)}|_{H(\mathbb{Q}_p)}\right) f_\alpha^{(P_H(\mathbb{Q}_p))} \chi^+_{\varpi,u} \in \mathcal{H}(H(\mathbb{Q}_p)), \tag{6.4}$$

is a transfer of the function $\chi_c^{G(\mathbb{Q}_p)} f_\alpha$ to $H(\mathbb{Q}_p)$. Therefore the transfer vanishes by Lemma 6.2. □

A function $h$ is called cuspidal if for every non-elliptic semi-simple conjugacy class $\gamma$ the orbital integral $O_\gamma(h)$ vanishes.

Lemma 6.4. *The truncated function* $\chi_c^{G(\mathbb{Q}_p)} f_\alpha$ *is cuspidal.*

Proof. Any non-elliptic conjugacy class of $G(\mathbb{Q}_p)$ is conjugated to an element of $M$ for some proper standard Levi-subgroup $M$ of $G(\mathbb{Q}_p)$. Let $P$ be the corresponding standard parabolic subgroup of $G$. Then the orbital integral $O_\gamma(\chi_c^{G(\mathbb{Q}_p)} f)$ is the product of a certain Jacobean factor with the $M$-orbital integral of $\gamma$ of the function $\chi_c^{G(\mathbb{Q}_p)} f^{(P)} = 0$ (Proposition 6.3). Thus the function is cuspidal. □

We are thus left with the term $\mathrm{St}_{G^*}((\chi_c^{G(\mathbb{Q}_p)} f)^{G^*})$ in Equation (6.3), which can be treated by base change as in [**28**, §4.3]. The final result is, as above, that the dominant term is given by the trivial representation (or Abelian characters).

APPENDIX A

# Existence of cuspidal representations of $p$-adic reductive groups

We prove the following Theorem:

THEOREM 0.5. *Let $G$ be a connected reductive group over $F$. Then $G(F)$ has a cuspidal complex representation.*

This theorem is "folklore", but we have not found a proof for it in the literature. After some reduction steps the proof consists of finding certain characters in general position of elliptic maximal tori of $G$. In case the cardinal of the residue field of $F$ is "large with respect to $G$", then there are quick arguments to show that characters in general position exist; see for example [**15**, lemma 8.4.2]. It is the small groups over small fields and big Weyl groups that might cause problems, and in this chapter we show that such problems do not occur.

This appendix is independent of the rest of this thesis.

## 1. Reduction to a problem of classical finite groups of Lie type

Let $P \subset G(F)$ be a maximal proper parahoric subgroup with associated reductive quotient $M$ over $k$. We claim that $M(k)$ has an irreducible cuspidal representation $\sigma$. When $\sigma$ is proved to exist, then we may construct a cuspidal representation of $G$ as follows, see [**83**], [**82**] and [**83**]. Inflate $\sigma$ to obtain a $P$-representation. We may compactly induce the $P$-representation $\sigma$ to a representation of $G(F)$. This $G(F)$-representation need not be irreducible, but its irreducible subquotients are all cuspidal. Therefore Theorem 0.5 reduces to the next proposition.

PROPOSITION 1.1. *Let $G$ be a connected reductive group over the finite field $k$. The group $G(k)$ has a cuspidal complex representation.*

PROOF. We will first reduce to $G$ simple and adjoint. Consider the morphism $G(k) \to G_{\mathrm{ad}}(k)$. If $\pi$ is a irreducible representation of $G_{\mathrm{ad}}(k)$, then, when restricted to a representation of $G(k)$ it will decompose as a finite direct sum $\pi = \bigoplus_i \pi_i$ of irreducible representations. Recall that $\pi$ is cuspidal if and only if $\mathrm{H}^0(N(k), V) = 0$ for all rational parabolic subgroups $P \subset G$ with Levi decomposition $P = MN$. The map $G \to G_{\mathrm{ad}}$ is an isomorphism on its image when restricted to $N$. For any parabolic subgroup $P = MN \subset G_{\mathrm{ad}}$ the inverse image of $P$ in $G$ is a parabolic subgroup with the same unipotent part. Thus, if $\pi$ is cuspidal as $G_{\mathrm{ad}}(k)$-representation, then the $\pi_i$ are cuspidal representations of $G(k)$. Therefore, we may assume that $G$ is adjoint. But then $G$ is a product of $k$-simple adjoint groups. If the theorem is true for all the factors, then the theorem is true for $G$. So we may assume that $G = \mathrm{Res}_{k'/k} G'$ where $G'$ is (absolutely) simple and defined over some finite extension $k'$ of $k$. We have $G(k) = G'(k')$, and under this equality cuspidal representations correspond to cuspidal representations. Therefore, we may assume that $G$ is simple and adjoint.

The simple reductive groups $G$ over $k$ are classified by their root system. We will distinguish cases between the possible root systems. Let us first assume that the root system of $G$ is exceptional, i.e. of the form ${}^3D_4$, $E_6$, ${}^2E_6$, $E_7$, $E_8$, $F_4$, ${}^2F_4$, $G_2$ or ${}^2G_2$. In Carter's book [**15**, §13.9] one finds for each exceptional group the complete list of its unipotent irreducible complex trace characters. He also mentions for each group how many of these characters are cuspidal. As it turns out, in each of the exceptional cases, this number is $> 0$ and so in particular all the exceptional groups have a cuspidal representation. Some of the classical groups do not have cuspidal *unipotent* characters. So unfortunately for those groups we cannot find a cuspidal representation in Carter's list.

It remains to verify Proposition 1.1 for the simple adjoint groups $G/k$ which are *classical*. Thus if $G$ is split, then it is of type $A_n, B_n, C_n$ or $D_n$, and if it is non-split, then it is of type ${}^2A_n$ or ${}^2D_n$. To do this we will use Deligne-Lusztig theory in Section 2 to reduce the problem to finding characters in general position. In section 3 we will then verify that all split groups have such a character. In sections 4 and 5 we will then find characters in general position for the remaining non-split root systems. The proof of Proposition 1.1 will then be complete. □

## 2. Characters in general position

Let $G/k$ be a reductive group with connected center. We will apply results of Deligne-Lusztig [**33**]. Pick $\ell$ a prime number different from $p$. Suppose that we are given the following data: $T \subset G$ a maximal torus and $\theta\colon T(k) \to \overline{\mathbb{Q}}_\ell^\times$ a rational character. Then, to this data Deligne and Lusztig associate a virtual character $R_T^\theta$ of $G(k)$ with $\overline{\mathbb{Q}}_\ell$-coefficients [**33**, p. 114].

Let $\sigma(G)$ be the $k$-rank of $G$ and let $\sigma(T)$ be the $k$-rank of $T$. Proposition [**33**, Prop. 7.4] states that the character $(-1)^{\sigma(G)-\sigma(T)} R_T^\theta$ comes from an actual irreducible $G(k)$-representation $\pi_T^\theta$ if the character $\theta$ is in *general position*, ie if the rational Weyl group of $T$ acts freely on it. Theorem [**33**, thm 8.3] states that if, additionally, $T$ is elliptic, then $\pi_T^\theta$ is cuspidal. Assume for the moment that we have such a pair $(T, \theta)$. Pick an isomorphism $\iota\colon \overline{\mathbb{Q}}_\ell \xrightarrow{\sim} \mathbb{C}$; then the $G(k)$-representation $\pi_T^\theta \otimes_\iota \mathbb{C}$ is complex cuspidal and irreducible. Therefore, the proof of Proposition 1.1 is reduced to Proposition 3.2, Proposition 4.1 and Proposition 5.1.

## 3. The split classical groups

Before continuing with the proof, we recall some generalities. Let $G/k$ be a reductive group. Let $(T_0, B_0)$ be a pair consisting of a maximal torus and a Borel subgroup which contains $T_0$, both defined over $k$. Let $W_0/k$ be the Weyl group of $T_0 \subset G$. The Frobenius $\mathrm{Frob}_q = (x \mapsto x^q) \in \mathrm{Gal}(\overline{k}/k)$ acts on the root datum of $G$ by a diagram automorphism. By abuse of notation this diagram automorphism is also denoted $\mathrm{Frob}_q$.

We carry out the following construction. Let $\chi\colon T_{0,\overline{k}} \to \mathbb{G}_{m,\overline{k}}$ be a character. Restrict to $T_0(k)$ to get a morphism $T_0(k) \hookrightarrow T_0(\overline{k}) \to \overline{k}^\times$. From this construction we obtain a map $X^*(T_0) \to \mathrm{Hom}(T_0(k), \overline{k}^\times)$, and this map fits in the exact sequence

$$0 \longrightarrow X^*(T_0) \xrightarrow{\Phi-1} X^*(T_0) \longrightarrow \mathrm{Hom}(T_0(k), \overline{k}^\times) \longrightarrow 0, \tag{3.1}$$

of $\mathbb{Z}[W_0(k)]$-modules (see [**33**, §5]). Here $\Phi$ is the relative $q$-Frobenius of $T_{0,\overline{k}}$ over $\overline{k}$, i.e. given by $f \otimes \lambda \mapsto f^q \otimes \lambda$ on the global sections $\mathcal{O}_{T_0}(T_0) \otimes_k \overline{k}$ of $T_{0,\overline{k}}$. Recall that we write $\mathrm{Frob}_q$ for the Frobenius $f \otimes \lambda \mapsto f \otimes \lambda^q$ on $\mathcal{O}_{T_0}(T_0) \otimes_k \overline{k}$.

Definition 3.1. Two elements $w, w'$ in $W_0(\overline{k})$ are *Frobenius conjugate*, or $\mathrm{Frob}_q$-*conjugate*, if there exists an $x \in W_0(\overline{k})$ such that $w' = xw\mathrm{Frob}_q(x)^{-1}$.

The $G(k)$-conjugacy classes of rational maximal tori in $G_{\overline{k}}$ are parametrized by the Frobenius conjugacy classes of $W_0(\overline{k})$ in the following manner. Let $N_0$ be the normalizer of $T_0$ in $G$. We have a surjection from $G(\overline{k})$ to the set of maximal tori in $G_{\overline{k}}$ by sending $g \in G(\overline{k})$ to the torus ${}^gT_0 := gT_0g^{-1}$. The torus ${}^gT_0 \subset G_{\overline{k}}$ is rational (i.e. $\mathrm{Gal}(\overline{k}/k)$-stable) if and only if $g^{-1}\mathrm{Frob}_q(g) \in N_0(\overline{k})$.

Assume that we have two elements $g, g' \in G(\overline{k})$ such that the tori ${}^gT_0$, ${}^{g'}T_0$ in $G_{\overline{k}}$ are rational. Then, $g^{-1}\mathrm{Frob}_q(g)$ and $g'^{-1}\mathrm{Frob}_q(g')$ lie in $N_0(\overline{k})$ so we map them to elements of the Weyl group $W_0(\overline{k})$ via the canonical surjection $\pi\colon N_0(\overline{k}) \to W_0(\overline{k})$. The torus ${}^gT_0 \subset G_{\overline{k}}$ is equal to the torus ${}^{g'}T_0 \subset G_{\overline{k}}$ if and only if

$$\pi(g^{-1}\mathrm{Frob}_q(g)) \equiv \pi(g'^{-1}\mathrm{Frob}_q(g')) \in W_0(\overline{k})/_{\text{Frobenius conjugacy}},$$

(for the proof of this fact, see [**36**, III.3.23]). This completes the description how Frobenius conjugacy classes in $W_0(\overline{k})$ parametrize $G(k)$-conjugacy classes of maximal tori in $G_{\overline{k}}$.

Notation. We will write $T_0(w)$ for the torus ${}^gT_0$.

Proposition 3.2. *Let $G/k$ be a classical simple adjoint group. Then $G$ has an anisotropic maximal torus $T \subset G$ together with a character $\theta\colon T(k) \to \mathbb{C}^\times$ in general position.*

Proof. To prove this proposition we will translate it to an explicit combinatorial problem on Dynkin diagrams. We will then use the classification of such diagrams and calculate to obtain the desired result.

Let $(T_0, B_0)$ be a pair consisting of a split maximal torus and a Borel subgroup which contains $T_0$, both defined over $k$. Let $w \in W_0(\overline{k})$ be a Coxeter element and let $T = T_0(w) \subset G$ be the maximal torus corresponding to the Frobenius conjugacy class $\overline{w} \subset W_0(\overline{k})$ generated by $w$.

Pick $g \in G(\overline{k})$ such that $g^{-1}\mathrm{Frob}_q(g) \in N_0(\overline{k})$ and $\pi(g^{-1}\mathrm{Frob}_q(g)) = w \in W_0(\overline{k})$. The conjugation-by-$g$-map $G_{\overline{k}} \to G_{\overline{k}}$ induces an isomorphism from $T_{0,\overline{k}}$ to ${}^gT_{0,\overline{k}} = T_{\overline{k}}$, and in turn an isomorphism $X^*(T) \xrightarrow{\sim} X^*(T_0)$. Under this isomorphism, the Frobenius $\mathrm{Frob}_q$ on $X^*(T)$ corresponds to the automorphism $w\mathrm{Frob}_q$ on $X^*(T_0)$, and similarly $\Phi$ on $X^*(T)$ corresponds to $w\Phi$ on $X^*(T_0)$.

To see that the torus $T_0(w)$ is anisotropic it suffices to prove that $X^*(T_0(w))^{\mathrm{Frob}_q} = 0$. We will verify this in each individual case below.

The rational Weyl group $W_T(k)$ of the torus $T$ is equal to the set of those elements $w \in W_T(\overline{k})$ in the absolute Weyl group whose action on the characters $X^*(T)$ is equivariant for the Frobenius $\mathrm{Frob}_q$. Therefore, under the bijection $W_T(\overline{k}) \xrightarrow{\sim} W_0(\overline{k})$, the image of $W_T(k)$ in $W_0(\overline{k})$ is equal to the set of all $t \in W_0(\overline{k})$ such that $t(w\mathrm{Frob}_q) = (w\mathrm{Frob}_q)t$. Because $T_0$ is split, the automorphism $\mathrm{Frob}_q$ acts trivially on $X^*(T_0)$. Therefore, the image of $W_T(k)$ in $W_0(\overline{k})$ is the centralizer of $w \in W_0(\overline{k})$. Because $w$ is a Coxeter element this centralizer is equal to the subgroup generated by $w \in W_0(\overline{k})$.

Choose an embedding of groups $\iota\colon \overline{k}^\times \hookrightarrow \mathbb{C}^\times$. Then, using $\iota$, we may identify $\mathrm{Hom}(T(k), \overline{k}^\times)$ with $\mathrm{Hom}(T(k), \mathbb{C}^\times)$. The set $\mathrm{Hom}(T(k), \mathbb{C}^\times)$ is the set of characters of $T(k)$. We are interested in the subset of $\mathrm{Hom}(T(k), \mathbb{C}^\times)$ consisting of those characters which are in general position. Under the bijection $\frac{X^*(T_0)}{(w\Phi-1)X^*(T_0)} \xrightarrow{\sim} \mathrm{Hom}(T(k), \mathbb{C}^\times)$ the action of the group $W_T(k)$ on the right corresponds to the action of the subgroup $\langle w \rangle \subset W_0(\overline{k})$ on the set on the left. The problem of finding an elliptic torus together with a character in general position is thus translated into a problem of the root system of $(G, B_0, T_0)$: Pick any Coxeter element $w$ in the Weyl group of the root system, and find an element $v$ in $\frac{X^*(T_0)}{(w\Phi-1)X^*(T_0)}$ which is such that $w^r v \neq v$ for all $r = 1 \dots h$, where $h = \#\langle w \rangle$ is the Coxeter number of $G$.

Before starting the computations, let us make the following 3 remarks to clarify. First, the relative $q$-Frobenius $\Phi$ acts on $X^*(T_0)$ by $\chi \mapsto \chi^q$ ($T_0$ is split). And second, because the group $G$ is adjoint, the root lattice of $G$ is equal to the weight lattice $X^*(T_0)$. Finally, the facts on Dynkin diagrams that we state below come from Bourbaki [**11**, chap 6, §4 – §13].

- $G$ is split of type $B_n$ with $n \in \mathbb{Z}_{\geq 2}$. The root system of $G$ may be described as follows. Let $V = \mathbb{R}^n$ with its canonical basis $e_1, \dots, e_n$ and the standard inner product. Define $\alpha_1 = e_1 - e_2, \alpha_2 = e_2 - e_3, \dots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_n$. The elements $\alpha_1, \dots, \alpha_n \in \mathbb{Z}^n$ are the simple roots, and the root lattice is equal to $\mathbb{Z}^n \subset \mathbb{R}^n$. The element $w = w_{\alpha_1} w_{\alpha_2} \cdots w_{\alpha_n}$ is a Coxeter element of the Weyl group; it acts on $\mathbb{R}^n$ by $(x_1, \dots, x_n) \mapsto (-x_n, x_1, \dots, x_{n-1})$. It is clear that there are no elements in the root lattice invariant under the action of $w\mathrm{Frob}_q$. This implies that $T_0(w)$ is anisotropic.

We claim that the element $e_1 \in \mathbb{Z}^n$ reduces to an element of $\mathbb{Z}^n/(w\Phi - 1)\mathbb{Z}^n$ in general position. The order of $w$ is equal to $2n$, so $\#\mathrm{stab}_{\langle w \rangle}(v)$ divides $2n$. Therefore, it suffices to check that for all $r \in \{1, \dots, n\}$ we have $w^r(e_1) - e_1 \notin (w\Phi - 1)\mathbb{Z}^n$.

We distinguish cases. Assume first $r = n$. Then $w^r$ acts on $V$ by $v \mapsto -v$. We have $w^n(e_1) - e_1 = (-2, 0, \dots, 0)$. Assume that we have an $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$ with $(w\Phi - 1)x = (-2, 0, 0, \dots, 0)$. Then

$$(3.2) \qquad -qx_n - x_1 = -2,\ qx_1 - x_2 = 0,\ qx_2 - x_3 = 0,\ \dots,\ qx_{n-1} - x_n = 0.$$

From this we get $x_n = q^{n-1}x_1$, and $-2 = -q^n x_1 - x_1 = -(1 + q^n)x_1$ which is not possible. So we have dealt with the case $r = n$.

Now assume that $r \in \{1, \dots, n-1\}$. Then $w^r(e_1) - e_1 = e_{r+1} - e_1$. Assume that we have an $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$ such that

$$(3.3) \qquad -qx_n - x_1 = -1, \quad qx_r - x_{r+1} = 1, \quad \text{and} \quad qx_{i-1} - x_i = 0 \quad (\forall i \notin \{1, r+1\}).$$

We find

$$x_n = q^{n-r-1}x_{r+1} = q^{n-r-1}(qx_r - 1) = q^{n-r}x_r - q^{n-r-1} = q^{n-1}x_1 - q^{n-r-1},$$

and $x_1 - 1 = -qx_n = -q(q^{n-1}x_1 - q^{n-r-1})$, which implies

$$x_1 = \frac{q^{n-r} + 1}{q^n + 1},$$

but $|q^{n-r} - 1|_\infty < |q^n + 1|_\infty$, so $x_1$ is not integral: contradiction. This completes the proof that $e_1 \in X(T)$ is a character in general position in case $G$ is of type $B_n$.

- $G$ is split of type $C_n$ with $n \in \mathbb{Z}_{\geq 2}$. The root system of $G$ may be described as follows. Let $V = \mathbb{R}^n$ with its canonical basis $e_1, \dots, e_n$ and the standard inner product. Define $\alpha_1 = e_1 - e_2, \alpha_2 = e_2 - e_3, \dots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = 2e_n$. The elements $\alpha_1, \dots, \alpha_n \in \mathbb{Z}^n$ are the simple roots, and the root lattice $\Lambda$ is equal to the set of $(x_1, \dots, x_n) \in \mathbb{Z}^n \subset \mathbb{R}^n$ with $\sum_{i=1}^n x_i \equiv 0 \mod 2$. The element $w = w_{\alpha_1}w_{\alpha_2} \cdots w_{\alpha_n}$ is a Coxeter element of the Weyl group; it acts on $\mathbb{R}^n$ by $(x_1, \dots, x_n) \mapsto (-x_n, x_1, \dots, x_{n-1})$. It is clear that there are no elements in the root lattice invariant under the action of $w\mathrm{Frob}_q$. This implies that $T_0(w)$ is anisotropic.

We claim that the element $2e_1 \in \Lambda$ reduces to an element of $\Lambda/(w\Phi - 1)\Lambda$ in general position. It suffices to verify that $w^r(2e_1) - 2e_1 \notin (w\Phi - 1)\Lambda$ for all $r \in \{1, \dots, n\}$. Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ be the vector satisfying the equations in Equation 3.3. Then the vector $x' := 2x$ satisfies $w^r(2e_1) - 2e_1 = (w\Phi - 1)x'$. Therefore,

$$x'_1 = 2 \cdot \frac{q^{n-r} + 1}{q^n + 1}.$$

For $q \neq 2$ we have $2|q^{n-r} + 1|_\infty < |q^n + 1|_\infty$, and for $q = 2$ the numerator and denominator are coprime. Therefore $x_1$ is not integral.

- $G$ is split of type $A_n$ with $n \in \mathbb{Z}_{\geq 1}$. Consider inside $\mathbb{R}^{n+1}$ the hyperplane $V$ with equation $\sum_{i=1}^{n+1} \xi_i = 0$. Define $\alpha_1 = e_1 - e_2, \alpha_2 = e_2 - e_3, \dots, \alpha_n = e_n - e_{n+1}$ (simple roots), $\Lambda = \mathbb{Z}^{n+1} \cap V$ (root lattice), and $w = w_{\alpha_1}w_{\alpha_2} \cdots w_{\alpha_n}$ (Coxeter element). The element $w$ acts on $V \subset \mathbb{R}^{n+1}$ by rotation of the coordinates: $(x_1, x_2, \dots, x_n, x_{n+1}) \mapsto (x_{n+1}, x_1, x_2, \dots, x_n)$. We have $(w\Phi-1)(x_1,\ldots,x_{n+1}) = (qx_{n+1}-x_1, qx_1-x_2, qx_2-x_3,\ldots,qx_n-x_{n+1})$. It is clear that there are no elements in the root lattice invariant under the action of $w\mathrm{Frob}_q$. This implies that $T_0(w)$ is anisotropic.

We claim that the element $v := e_1 - e_{n+1} \in \Lambda$ reduces to an element of $\Lambda/(w\Phi-1)\Lambda$ which is in general position. The order of $w$ equals $n+1$. Let $r \in \{1,\ldots,n\}$. Suppose for a contradiction that $w^r(v) - v = (e_{r+1}-e_r) - (e_1 - e_{n+1}) \in (w\Phi-1)\Lambda$. Then we have an element $(x_1,\ldots,x_{n+1}) \in \Lambda$ such that

$$\begin{aligned} &qx_{n+1}-x_1=-1,\ qx_{r-1}-x_r=-1,\ qx_r-x_{r+1}=1,\ qx_n-x_{n+1}=1,\\ &qx_{i-1}-x_i=0 \quad (\forall i\notin\{r+1,r,1,n+1\}). \end{aligned}$$

By substitution we deduce from this $q^{n+1}x_{n+1} = x_{n+1} - q^n - q^{n+1-r} + q^{n-r} + 1$. But, $q^{n+1} - 1 > q^n + q^{n-r+1} - q^{n-r} - 1$, so $x_{n+1}$ cannot be integral: contradiction.

- $G$ is split of type $D_n$ with $n \in \mathbb{Z}_{\geq 4}$. Define $\alpha_1 = e_1 - e_2$, $\alpha_2 = e_2 - e_3$, …, $\alpha_{n-1} = e_{n-1} - e_n$, $\alpha_n = e_{n-1} + e_n$ (simple roots), $\Lambda$ the set of $(x_1,\ldots,x_n) \in \mathbb{Z}^n$ such that $\sum_{i=1}^n x_i \equiv 0 \mod 2$ (root lattice).

Unfortunately the above procedure to produce anisotropic tori and characters in general position does not work for this group $G$ for the following reason. Let $w = w_{\alpha_1}\cdots w_{\alpha_n}$ be the Coxeter element of the Weyl group which is the product of the reflections in the simple roots. Then $w$ acts on $V$ by

$$(x_1,x_2,\ldots,x_n) \mapsto (-x_n, x_1,\ldots,x_{n-2},-x_{n-1}).$$

This implies that the vector $(2,\ldots,2,-2) \in \Lambda$ is stable under the action of Frobenius and thus the corresponding torus is not anisotropic.

Let $W_0$ be the Weyl group of the system $D_n$. We have a split exact sequence

$$1 \longrightarrow \{-1\}^n_{\det=1} \longrightarrow W_0 \longrightarrow \mathfrak{S}_n \longrightarrow 1, \tag{3.4}$$

where $\mathfrak{S}_n$ acts on $\mathbb{Z}^n$ via the natural action and an $\varepsilon = (\varepsilon_i) \in \{-1\}^n_{\det=1}$ acts on a vector $e_i \in \mathbb{Z}^n$ of the standard basis by $\varepsilon e_i = \varepsilon_i e_i$.

Write $n = m+1$. Let $w = (123\ldots m) \in \mathfrak{S}_n$. Write $t_k \in \{-1\}^n$ for the element with $-1$ on the $k$-th coordinate, and with 1 on all other coordinates. Define $w' = t_n t_m w \in W_0$. We consider the maximal torus $T$ in $G$ of type $w'$. The action of $\mathrm{Frob}_q$ on the character group of this torus is given by

$$\mathbb{Z}^n \ni (x_1,\ldots,x_m,x_n) \mapsto (x_m, x_1,\ldots,-x_{m-1},-x_n).$$

We see that there are no non-zero vectors in $\mathbb{Z}^n$ which are invariant under this action. Therefore the torus $T$ is anisotropic.

The rational Weyl group of $T$ is the set of $s \in W_0$ which commute with $w'$. Let us compute this group. Write $\varphi\colon W_0 \twoheadrightarrow \mathfrak{S}_n$ the natural surjection (see Equation 3.4). Let $s \in W_T(k)$, then $w = \varphi(w') = \varphi(sws^{-1})$. Therefore $\varphi(s)$ commutes with $w$. This implies that $\varphi(s)$ is a power of $w$. Write $s = \varepsilon w^k$ for some $\nu \in \{-1\}^n_{\det=1}$. We have

$$st_n s^{-1} = t_n, \quad st_m s^{-1} = t_{w^k(m)}, \quad \text{and} \quad sws^{-1}\varepsilon w^k w w^{-k}\varepsilon = \varepsilon w \varepsilon.$$

Therefore

$$s(w')s^{-1} = s(t_n t_m w)s^{-1} = t_n t_{w^k(m)} \varepsilon w \varepsilon,$$

which is equivalent to

$$(\varepsilon_{w(i)}\varepsilon_i) = t_{w^k(m)} t_m.$$

A priori there are 4 solutions $\varepsilon \in \{-1\}^n$ of this equation. When we add the condition $\det(\varepsilon) = 1$, then precisely 2 of those solutions remain.

Let $\varepsilon \in \{-1\}^n_{\det=1}$ be such that $(\varepsilon_{w(i)}\varepsilon_i) = t_{w(m)} t_m$. We have an exact sequence

$$1 \longrightarrow \{1, \nu\} \longrightarrow W_T(k) \longrightarrow \langle \varepsilon w \rangle \longrightarrow 1,$$

where $\nu \in \{-1\}^n_{\det=1}$ is given by $\nu_i = -1$ for $i \le m$ and $\nu_n = (-1)^m$.

We claim that $v = 2e_m \in \Lambda$ reduces to an element of $\Lambda/(w'\Phi - 1)\Lambda$ in general position. Assume that

$$(w'\Phi - 1)(x_1, \dots, x_m, x_n) = (qx_m - x_1, qx_1 - x_2, \dots, -qx_{m-1} - x_m, -qx_n - x_n).$$

We ignore the last coordinate, and only work with the vector $(x_1, \dots, x_m)$. By substitution we deduce that $q^m x_m = -2 - x_m \pm 2q^{m-r}$. This implies

$$x_m = 2\frac{q^{m-r} \pm 1}{q^m + 1}. \tag{3.5}$$

For $(q, r) \ne (2, 1)$ we have $|q^m + 1|_\infty > 2|q^{m-r} \pm 1|_\infty$, and for $(q, r) = (2, 1)$ the numerator and denominator have a gcd which divides 3, so then $q^m + 1 = 3$ and we must have $m = 1$, but we assumed $m \ge 2$. Therefore $x_m$ is not integral. □

## 4. The unitary groups

PROPOSITION 4.1. *Let $n \ge 3$. The simple adjoint group over $k$ with root system ${}^2A_{n-1}$ has an anisotropic maximal torus $T$ together with a character $T(k) \to \mathbb{C}^\times$ in general position.*

PROOF. Let $E \subset \overline{k}$ be the quadratic extension of $k$, and let $\sigma\colon E \xrightarrow{\sim} E$ be the unique non-trivial $k$-automorphism of $E$. The unitary group $\mathrm{U}_n$ over $k$ is the group of matrices $g \in \mathrm{Res}_{E/k}\mathrm{GL}_{n,E}$ such that $\sigma(g)^{\mathrm{t}} g = 1$. The adjoint group $\mathrm{U}_{n,\mathrm{ad}}$ of $\mathrm{U}_n$ is the group $\mathrm{PU}_n$ and this group has root system ${}^2A_{n-1}$.

We will distinguish cases between $n$ odd and $n$ even. Assume first that $n$ is odd. Let $T_0$ be the torus $(U_1)^n$ embedded diagonally in $U_n$. Then $\mathrm{Frob}_q$ acts on $X^*(T_0)$ by $x \mapsto -x$. We have $X^*(T_0) = \mathbb{Z}^n$ and under this equality, the Weyl group $W_{T_0}(\overline{k})$ is identified with $\mathfrak{S}_n$. Let

$w = (123\dots n) \in \mathfrak{S}_n = W_{T_0}(\bar{k})$ and let $T$ be the torus $T_0(w)$. The relative Frobenius $\Phi$ acts on $X^*(T) = \mathbb{Z}^n$ by

$$(x_1, \dots, x_n) \mapsto (-x_n, -x_1, \dots, -x_{n-1}). \tag{4.1}$$

We claim that the torus $T$ is anisotropic over $k$. To see this, let $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$ be $w\mathrm{Frob}_q$-invariant. Then $(x_1, \dots, x_n) = (-x_n, -x_1, \dots, -x_{n-1})$, it follows $x_1 = (-1)^n x_1$, and because $n$ is odd, this implies $x_1 = 0$. The same argument applies to the other $x_i$, and therefore $x = 0$. We proved $X^*(T)^{\mathrm{Frob}_q} = 0$, and thus $T$ is anisotropic.

The center of $U_n$ is equal to $U_1$ embedded diagonally. Let $T_{\mathrm{ad}}$ be the image of the torus $T$ in the adjoint group of $\mathrm{U}_n$. Then $\Lambda = X^*(T_{\mathrm{ad}})$ is the subset of $\mathbb{Z}^n$ consisting of those vectors $x \in \mathbb{Z}^n$ such that $\sum_{i=1}^n x_i = 0$. The Weyl group is $\mathfrak{S}_n$ and it acts on $\Lambda$ via the restriction of the natural action $\mathfrak{S}_n \circlearrowleft \mathbb{Z}^n$ to $\Lambda$. The rational Weyl group $W_{T_{\mathrm{ad}}}(k) \subset \mathfrak{S}_n$ is the set of elements $w$ commuting with $\mathrm{Frob}_q$. The rational Weyl group is equal to $\langle w \rangle \subset \mathfrak{S}_n$ because all elements of the Weyl group commute with $-1$.

To find an element in general position we must find a vector $v \in \Lambda$ which is such that $w^r(v) - v \notin (\Phi - 1)\Lambda$ for all $r = 1 \dots n-1$. We claim that $v = e_1 - e_n \in \Lambda$ is such a vector.

Assume for a contradiction that $(w\Phi - 1)x = w^r v - v$ for some $x \in \Lambda$. Then

$$(-qx_n - x_1, -qx_1 - x_2, \dots, -qx_{n-1} - x_n) = (e_{r+1} - e_r) - (e_1 - e_n).$$

By substitution we deduce from this $(-q)^n x_n = x_n - (-q)^{n-1} - (-q)^{n-r} + (-q)^{n-1-r} + 1$, and thus

$$x_n = -\frac{(-q)^{n-1} + (-q)^{n-r} - (-q)^{n-1-r} - 1}{(-q)^n - 1} \in \mathbb{Z}.$$

We show that this is not possible. We will distinguish cases. Assume first that the pair $(q, r)$ is such that the inequality $|(-q)^r + (-q) + 1| < q^{r+1} - 2$ holds. We may then estimate

$$\begin{aligned} |(-q)^{n-1} + (-q)^{n-r} - (-q)^{n-1-r} - 1|_\infty &= |(-q)^{n-1-r}((-q)^r + (-q) - 1) - 1| \\ &\le q^{n-1-r} \cdot |(-q)^r + (-q) - 1| + 1 \\ &< q^n - 2q^{n-1-r} + 1 \le q^n - 1 \le |(-q)^n - 1|_\infty. \end{aligned}$$

This proves that $x_n$ cannot be integral.

Let us determine the pairs $(q, r)$ for which the above inequality is not true. We have $|(-q)^r + (-q) + 1| \le q^r + q + 1$. The inequality $q^r + q + 1 < q^{r+1} - 2$ does not hold for $(q, r) \in \{(2,1), (2,2), (3,1)\}$. To see that it holds in all other cases, observe first that if the inequality holds for $(q, r)$ then it holds also for $(q, r+1)$. By direct verification we see that it holds for $(2,3)$, $(3,2)$, and for $(q,1)$ in case $q > 3$.

For $(q, r) \in \{(2,2), (3,1)\}$ we have the inequality $|(-q)^r + (-q) + 1| < q^{r+1} - 2$, so the above proof also applies to these cases. In case $(q, r) = (2,1)$, then we obtain $x_1 = -1 + \frac{(-2)^{n-2}}{(-2)^n - 1}$, which is not integral. This completes the proof for $n$ odd.

Now assume that $n$ is even. Write $n = m + 1$, so that $m$ is odd. Let $T_0 \subset U_n$ be the torus $(U_1)^n$ embedded on the diagonal of $U_n$. Let $w = (123\dots m) \in \mathfrak{S}_n = W_{T_0}(\bar{k})$, and consider the torus $T := T_0(w)$. We have $X^*(T) = \mathbb{Z}^n$ on which the Frobenius acts by $-w$. The rational Weyl group $W_T(k) \subset \mathfrak{S}_n$ is the set of $s \in \mathfrak{S}_n$ which commute with $-w$. Therefore $W_T(k) = \langle w \rangle$, and in particular $W_T(k) \subset \mathfrak{S}_m$. Let $T_{\rm ad}$ be the image of the torus $T$ in the adjoint group of $\mathrm{U}_n$. The lattice $X^*(T_{\rm ad}) = \Lambda \subset \mathbb{Z}^n$ is the set of vectors $(x_1, \ldots, x_n)$ with $\sum_{i=1}^n x_i = 0$. The tori $T$ and $T_{\rm ad}$ are anisotropic. Write $T = T_m \times U_1$, where the torus $T_m$ is the maximal torus in the group $U_m$ that we considered in the odd case. The rational Weyl group $W_T(k)$ preserves this decomposition of $T$. We have the map $X^*(T_{m,\rm ad}) \to X^*(T_{\rm ad})$, $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0)$. This map is $\mathfrak{S}_m$-equivariant, and it induces a map

$$\frac{X^*(T_{m,\rm ad})}{(w\Phi - 1)X^*(T_{m,\rm ad})} \rightarrowtail \frac{X^*(T_{\rm ad})}{(w\Phi - 1)X^*(T_{\rm ad})},$$

which is $w$-equivariant. Therefore characters in general position are send to characters in general position. By the argument above we know that $T_m$ has characters in general position, so this completes the proof for $n$ even. □

## 5. The non-split orthogonal groups

Proposition 5.1. *Let $n \in \mathbb{Z}_{\geq 4}$. The simple adjoint group $G$ over $k$ with root system ${}^2D_{n-1}$ has a maximal torus $T \subset G$ with a character $T(k) \to \mathbb{C}^\times$ in general position.*

*Proof.* Let $J$ be the $2n \times 2n$-matrix consisting of the blocks $\left(\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right)$ on the diagonal, and all other entries 0. The group $\mathrm{O}_{2n}$ over $k$ is the set of matrices $g \in \mathrm{GL}_{2n,k}$ which are such that $g^{\rm t} J g = J$. The group $\mathrm{SO}_{2n}$ is the group of matrices $g \in \mathrm{O}_{2n}$ such that $\det(g) = 1$. The non-split form $\mathrm{SO}'_{2n}$ over $k$ is obtained from $\mathrm{SO}_{2n}$ by twisting the action of $\mathrm{Frob}_q$ with the matrix $s \in \mathrm{GL}_{2n}$ consisting of the blocks $\left(\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right)$ on the diagonal, except for the last block on the diagonal which is $\left(\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right)$. This corresponds to replacing the matrix $J$ with the matrix $sJs^{-1} = sJs$ in the definition of the orthogonal group.

In characteristic $p \neq 2$, the group $\mathrm{SO}_{2n}$ (resp. $\mathrm{SO}_{2n}$) is connected and has root system $D_{n-1}$ (resp. ${}^2D_{n-1}$). For $p = 2$ it is the connected component of identity, $\mathrm{SO}_{2n}^\circ$ (resp. $\mathrm{SO}_{2n}'^\circ$), that has root system $D_{n-1}$ (resp. ${}^2D_{n-1}$).

The torus $(\mathrm{SO}_2^\circ)^n$ on the diagonal in $\mathrm{SO}_{2n}^\circ$ is a maximal torus, and the torus $T_0 = (\mathrm{SO}_2^\circ)^{n-1} \times U_1$ is a maximal torus of $\mathrm{SO}_{2n}'^\circ$. We have $X^*(T_0) = \mathbb{Z}^n$ and $\mathrm{Frob}_q$ acts on $X^*(T_0)$ by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-1}, -x_n)$. Let $W_0$ be the absolute Weyl group of $T_0$. We have a split exact sequence

$$1 \longrightarrow \{-1\}^n_{\det=1} \longrightarrow W_0 \longrightarrow \mathfrak{S}_n \longrightarrow 1, \tag{5.1}$$

where $\mathfrak{S}_n$ acts on $\mathbb{Z}^n$ by permuting the standard basis vectors, and where an $\varepsilon = (\varepsilon_i) \in \{-1\}^n_{\det=1}$ acts on a vector $e_i \in \mathbb{Z}^n$ of the standard basis by $\varepsilon e_i = \varepsilon_i e_i$.

Let $w \in \mathfrak{S}_n \subset W_0$ be the $n$-cycle $(123\ldots n)$ and consider the torus $T := T_0(w)$. Then $X^*(T) \xrightarrow{\sim} \mathbb{Z}^n$ via which the action $\mathrm{Frob}_q \circlearrowleft X^*(T)$ corresponds to the action of $w\mathrm{Frob}_q$ on $\mathbb{Z}^n$.

We verify that this torus is anisotropic. Let $x = (x_1, \ldots, x_n) \in X^*(T_0)^{w\mathrm{Frob}_q}$. Then

$$(x_1, \ldots, x_n) = w\mathrm{Frob}_q(x_1, \ldots, x_n) = (-x_n, x_1, \ldots, x_{n-1}),$$

which implies $x = 0$. Therefore $T$ is anisotropic.

The rational Weyl group $W_T(k) \subset W_0$ is the set of $s \in W_0$ which commute with $w\mathrm{Frob}_q$. Let us determine this group. Write $\varphi$ for the map $W_0 \twoheadrightarrow \mathfrak{S}_n$ (see Equation 5.1), and write $t_j \in \{-1\}^n$ for the element with $-1$ on the $j$-th coordinate, and with 1 on all other coordinates. Then $\mathrm{Frob}_q = t_n$.

If $s \in W_t(k)$, then $s(wt_n)s^{-1} = wt_n$. We apply $\varphi$ to this equality to obtain $w = \varphi(w) = \varphi(sws^{-1})$, and thus $\varphi(s) \in \mathfrak{S}_n$ commutes with $w$. This implies that $\varphi(s)$ is a power of $w$. Write $s = \varepsilon w^k$ where $\varepsilon \in \{-1\}^n_{\det=1}$. We have

$$st_ns^{-1} = \varepsilon w^k(t_n)w^{-k}\varepsilon = t_k,$$

and

$$sws^{-1} = \varepsilon w^k w w^{-k}\varepsilon = \varepsilon w \varepsilon.$$

Therefore,

$$wt_n = s(wt_n)s^{-1} = \varepsilon w \varepsilon \cdot t_k.$$

This is equivalent to,

$$\varepsilon w^{-1} \varepsilon w = t_k t_n.$$

Write $\varepsilon = (\varepsilon_i) \in \{-1\}^n_{\det=1}$. Then we have,

$$\varepsilon w^{-1}\varepsilon w = (\varepsilon_i) \cdot (\varepsilon_{w(i)}) = (\varepsilon_i \varepsilon_{i+1}) = t_k t_n. \tag{5.2}$$

We will now distinguish cases between $n$ is odd and $n$ is even. Assume first that $n$ is odd. Return to Equation 5.2, we have $(\varepsilon_i\varepsilon_{i+1}) = t_kt_n$. After the choice of $\varepsilon_n$, the $\varepsilon_i$ for $i < n$ are uniquely determined by this equation. If $\varepsilon$ is one of the solutions, then $-\varepsilon$ is the other solution. We have $\det(-\varepsilon) = (-1)^n \det(\varepsilon) = -\det(\varepsilon)$. Therefore, precisely one of the two solutions has determinant 1. We conclude that the rational Weyl group $W_T(k)$ is equal to $\langle \varepsilon w \rangle$, where $\varepsilon \in \{-1\}^n_{\det=1}$ is the unique element such that $(\varepsilon_i\varepsilon_{i+1}) = t_1t_n$.

Let $\mathrm{SO}^\circ_{2n,\mathrm{ad}}$ be the adjoint group of $\mathrm{SO}^\circ_{2n}$ and let $T_{\mathrm{ad}}$ be the image of the torus $T$ in $\mathrm{SO}^\circ_{2n,\mathrm{ad}}$. Then $X^*(T_{\mathrm{ad}}) \subset X^*(T) = \mathbb{Z}^n$ is the sublattice of elements $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ such that $\sum_{i=1}^n x_i = 0$.

We claim that the element $v = 2e_n \in \Lambda$ reduces to an element $\overline{v} \in \Lambda/(w\Phi - 1)\Lambda$ in general position. We have $(\varepsilon w)^r 2e_n = \pm 2e_{n-r}$, for all $r = 1, \ldots, n-1$. We left the sign unspecified, but we mention that it depends on $r$.

Suppose that there exists an $x = (x_1, \ldots, x_n) \in \Lambda$ such that

$$(w\Phi - 1)(x_1, \ldots, x_n) = (-qx_n, qx_1, \ldots, qx_{n-1}) - (x_1, \ldots, x_n) = 2e_n \pm 2e_r.$$

This implies $-q^nx_n = -2q^{n-r} \pm 2 + x_n$, and thus

$$x_n = 2\frac{q^{n-r} \pm 1}{q^n + 1}.$$

We have already verified in Equation 3.5 that $x_n$ cannot be integral. This completes the proof for $n$ odd.

Assume now that $n$ is even. We have $-1 \in \{-1\}^n_{\det=1}$ in case $n$ is even. In Equation 5.2 we found that $(\varepsilon_i\varepsilon_{i+1}) = t_k t_n$. We obtain from this, $\varepsilon_i = \varepsilon_1$ for $i \leq k$ and $\varepsilon_i = -\varepsilon_1$ for $i > k$. Therefore $\det(\varepsilon) = (-1)^k$, independently of $\varepsilon_1$. Therefore, $\varepsilon \in \{-1\}^n_{\det=1}$ only if $k$ is even, and if this is the case, then the equation $(\varepsilon_i\varepsilon_{i+1}) = t_k t_n$ has exactly 2 solutions for $\varepsilon \in \{-1\}^n_{\det=1}$.

We conclude that $\#W_T(k) = n$, but the group is not cyclic: Pick an $\varepsilon \in \{-1\}^n_{\det=1}$, such that $(\varepsilon_i\varepsilon_{i+1}) = t_2 t_n$ holds. Then the rational Weyl group $W_T(k)$ is equal to $\{-1\} \times \langle \varepsilon w^2 \rangle$. We have

$$\forall \nu_1 \in \{-1\} \exists \nu_2 \in \{-1\} : \qquad \nu_1(\varepsilon w^2)^r(2e_n) = \nu_2 2e_{2r}.$$

We show that $2e_n \pm 2e_{n-2r} \notin (w\Phi - 1)\Lambda$ for all $r = \left\{1, \ldots, \frac{n}{2}\right\}$, and all signs, so that the element $\overline{2e_n} \in \Lambda/(w\Phi - 1)\Lambda$ is in general position. Assume for a contradiction that $x \in \Lambda$ is such that

$$(w\Phi - 1)(x_1, \ldots, x_n) = 2e_n \pm 2e_{2r}.$$

Then we may proceed as in Equation 3.5 to find that $x_n$ is not integral.

All possible cases are now verified and the proof of Theorem 0.5 is completed. $\square$

APPENDIX B

# Jacquet modules (joint with Erez Lapid)

Let $F$ be a non-Archimedean local field with residue characteristic $p$ and consider the locally compact, totally disconnected group $G_n := \mathrm{GL}_n(F)$. Let $P = M \ltimes N$ be the standard, block upper triangular, parabolic subgroup of type $(n_1, n_2, \dots, n_k)$ with the standard Levi decomposition. Thus $M \simeq \prod_{i=1}^k G_{n_i}$. The *normalized Jacquet functor* $J_P$ is a functor from the category of smooth admissible complex representations of $G_n$ to those of $M$. It is defined as the space of coinvariants for the action of the unipotent group $N$ on $\pi$, twisted by a certain normalizing character. More precisely,

$$J_P(\pi) := \pi_N[\delta_P^{-1/2}], \quad \text{where} \ \ \delta_P(m) := |\det(\mathrm{Ad}(m)|\mathrm{Lie}(N))|\,, m \in M.$$

In general, it is a difficult problem to compute $J_P(\pi)$, or even its semisimplification, for an arbitrary irreducible $\pi$. In this appendix we will give an explicit formula for $J_P(\pi)$ for a certain class of irreducible representations, namely the *ladder representations* introduced in [**71**]. The case where $P$ is the minimal parabolic subgroup for which $J_P(\pi) \neq 0$ was considered in [ibid.]. Here we will extend it to any $P$.

The class of ladder representations contains the class of Speh representations. The main result of [**71**] is to extend the determinantal formula of Tadić for Speh representations [**99**] (cf. also [**18**]) to ladder representations (see (1.2) below). Speh representations are important in the representation theory of the general linear group, because they form the building blocks for the unitary dual of $G_n$. More precisely, it was shown by Tadić that any irreducible unitary representation is isomorphic to the parabolic induction of Speh representations twisted by certain (explicit, but not necessarily unitary) characters [**98**]. In particular, this is the case for the local components of representations occurring in the discrete automorphic spectrum of $G_n$ over a global field.

We prove that the Jacquet module of a ladder representation is semisimple, multiplicity free, and that its irreducible constituents are themselves tensor products of ladder representations. In contrast, the class of Speh representations is not stable under taking the Jacquet module. In other words, (non Speh) ladder representations are encountered in the Jacquet module of Speh representations. Hence, ladder representations are important for global applications.

Our result has an application to Shimura varieties. In Chapters 2 and 3 we computed the Hasse-Weil zeta function of the basic stratum of certain simple Shimura varieties at split primes of good reduction following the method of Langlands and Kottwitz [**59**]. Apart from the basic stratum, these varieties admit additional Newton strata (cf. [**87**]). In order to compute the zeta function of a given stratum $S$ one may proceed as in [**63**] provided that one knows the Jacquet modules of the representations occurring in the cohomology of $S$. These representations turn out to be (essentially) Speh representations, and hence the problem reduces to the one considered in this appendix. Details will be given elsewhere.

## 1. The Jacquet modules of a Ladder representation

We first introduce some more notation. We write $\mathcal{R} = \bigoplus_{n\in\mathbb{Z}_{\geq 0}} \mathrm{Groth}(G_n)$ where $\mathrm{Groth}(G_n)$ is the Grothendieck group of the category $\mathrm{Rep}(G_n)$ of smooth complex representations of $G_n$ of finite length. The group $\mathcal{R}$ has a structure of a graded ring (introduced by Zelevinsky in [**105**]) with multiplication given by

$$\pi_1 \times \pi_2 := \mathrm{Ind}_{P_{n_1,n_2}}^{G_{n_1+n_2}}(\pi_1 \otimes \pi_2) \in \mathrm{Rep}(G_{n_1+n_2}),$$

(normalized induction) for $\pi_1 \in \mathrm{Rep}(G_{n_1}), \pi_2 \in \mathrm{Rep}(G_{n_2})$, $n_1, n_2 \in \mathbb{Z}_{\geq 0}$ where $P_{n_1,n_2}$ is the standard parabolic subgroup of $G_{n_1+n_2}$ of type $(n_1, n_2)$. The unit element of $\mathcal{R}$ is the one-dimensional representation of $G_0$.

Fix an integer $d > 0$ and a cuspidal representation $\rho$ of $G_d$. For our purposes, a segment $[a, b]$ is a set of integers of the form $\{a, a+1, \ldots, b\}$ with $b \geq a$. For any segment $[a, b]$ the representation $\rho[|\det\cdot|^a] \times \cdots \times \rho[|\det\cdot|^b]$ admits a unique irreducible quotient $\delta([a, b])$, the so-called *generalized Steinberg* representation. A *ladder* is a finite sequence of segments $[a_1, b_1], \ldots, [a_t, b_t]$ such that $a_1 > a_2 > \cdots > a_t$ and $b_1 > b_2 > \cdots > b_t$. Given a ladder of segments, we may form the representation $\delta([a_1, b_1]) \times \cdots \times \delta([a_t, b_t])$. This representation admits a unique irreducible quotient, $\mathrm{LQ}(\delta([a_1, b_1]) \times \cdots \times \delta([a_t, b_t]))$ which is the *Langlands quotient* in the case at hand. The representations which arise in this manner are by definition the *ladder representations*. The subclass of *Speh representations* (up to twists) is obtained by taking $a_{i+1} = a_i - 1$ and $b_{i+1} = b_i - 1$ for all $i = 1, \ldots, t-1$.

The ring $\mathcal{R}$ is actually a bi-algebra (and in fact has an additional structure of a Hopf-algebra) with respect to the comultiplication $\Delta\colon \mathcal{R} \to \mathcal{R} \otimes \mathcal{R}$ defined by $\pi \mapsto \sum_{i=0}^{n} J_{P_{i,n-i}}(\pi)$, $\pi \in \mathrm{Rep}(G_n)$. In particular we have

$$\Delta(\delta([a, b])) = \sum_c \delta([c+1, b]) \otimes \delta([a, c]), \tag{1.1}$$

where we have used the convention that $\delta([a, b]) = 0$ if $b < a - 1$ and $\delta([a, a-1]) = 1 \in \mathcal{R}$.

THEOREM 1.1. *Suppose that $a_1 > \cdots > a_t$ and $b_1 > \cdots > b_t$. Then*

$$\Delta(\mathrm{LQ}(\delta([a_1, b_1]), \ldots, \delta([a_t, b_t]))) = \sum_{c_1 > \cdots > c_t \in \mathbb{Z}} \mathrm{LQ}(\delta([c_1+1, b_1]), \ldots, \delta([c_t+1, b_t])) \otimes \mathrm{LQ}(\delta([a_1, c_1]), \ldots, \delta([a_t, c_t])).$$

REMARK. Note the similarity between this formula and the formula

$$\Delta(\delta([a_1,b_1])\times\cdots\times\delta([a_t,b_t]))= \sum_{c_1,\dots,c_t\in\mathbb{Z}}\delta([c_1+1,b_1])\times\cdots\times\delta([c_t+1,b_t])\otimes\delta([a_1,c_1])\times\cdots\times\delta([a_t,c_t]).$$

Let us now prove Theorem 1.1. By the determinantal formula of Tadić [**71**] we have

$$\mathrm{LQ}(\delta([a_1,b_1]),\dots,\delta([a_t,b_t]))=\det(\delta([a_i,b_j]))_{i,j=1,\dots,t}. \tag{1.2}$$

Therefore,

$$\Delta(\mathrm{LQ}(\delta([a_1,b_1]),\dots,\delta([a_t,b_t])))=\det(\Delta(\delta([a_i,b_j])))_{i,j=1,\dots,t}.$$

By (1.1) and using the multi-linearity of the determinant we get

$$\sum_{c_1,\dots,c_t\in\mathbb{Z}}\det(\delta([c_j+1,b_j])\otimes\delta([a_i,c_j]))=\sum_{c_1,\dots,c_t\in\mathbb{Z}}\big(\prod_{j=1}^{t}\delta([c_j+1,b_j])\big)\otimes\det(\delta([a_i,c_j])).$$

Write $S_t$ for the symmetric group on the set $\{1,2,\dots,t\}$. Observe that if $c_j=c_k$ for some $j\neq k$ then $\det(\delta([a_i,c_j]))=0$ since two columns in the matrix are identical. Therefore, only distinct $c_1,\dots,c_t$ contribute to the right hand side of the above equation, and we can write the sum as

$$\begin{aligned}&\sum_{c_1>\cdots>c_t\in\mathbb{Z}}\sum_{s\in S_t}\big(\prod_{j=1}^{t}\delta([c_{s(j)}+1,b_j])\big)\otimes\det(\delta([a_i,c_{s(j)}]))\\&=\sum_{c_1>\cdots>c_t\in\mathbb{Z}}\sum_{s\in S_t}\operatorname{sgn}s\,\big(\prod_{j=1}^{t}\delta([c_{s(j)}+1,b_j])\big)\otimes\det(\delta([a_i,c_j]))\\&=\sum_{c_1>\cdots>c_t\in\mathbb{Z}}\det(\delta([c_i+1,b_j]))\otimes\det(\delta([a_i,c_j])).\end{aligned}$$

Applying (1.2) once more, we obtain Theorem 1.1.

COROLLARY 1.2. *Suppose that $a_1>\cdots>a_t$ and $b_1>\cdots>b_t$. Then the Jacquet module of $\mathrm{LQ}(\delta([a_1,b_1]),\dots,\delta([a_t,b_t]))$ with respect to the parabolic subgroup of type $(n_1,\dots,n_k)$ is*

$$\bigoplus_f \mathrm{LQ}(f^{-1}(1))\otimes\cdots\otimes\mathrm{LQ}(f^{-1}(k)) \tag{1.3}$$

*where the sum is over all $k$-colorings $f:\cup_{i=1}^t([a_i,b_i]\times\{i\})\to\{1,\dots,k\}$ such that*

(1) *$j\mapsto f(j,i)$ is (weakly) monotone decreasing for all $i=1,\dots,t$,*

(2) *$n_l=d\cdot\left|f^{-1}(l)\right|$ for all $l=1,\dots,k$,*

(3) *for any $l=1,\dots,k$ and $i=1,\dots,t$, let $m_{i,l}=\min\{j\in[a_i,b_i+1]:f(j,i)\le l\}$ (with $f(b_i+1,i)=-\infty$) and $n_{i,l}=\max\{j\in[a_i-1,b_i]:f(j,i)\ge l\}$ (with $f(a_i-1,i)=\infty$). Then $m_{i,l}>m_{i+1,l}$ and $n_{i,l}>n_{i+1,l}$ for all $i=1,\dots,t-1$, $l=1,\dots,k$.*

FIGURE 1. An example of a 4-coloring of 3 segments satisfying the conditions above.

The corollary extends the result of [**71**] (i.e., the case $n_1 = \cdots = n_t = d$). Up to semisimplification, the corollary follows from Theorem 1.1 by induction on $k$. To show that the Jacquet module is semisimple it suffices to note that the summands in (1.3) have distinct supercuspidal supports. This follows from the fact that given $b_1 > \cdots > b_t$ and a multiset $A$ of integers, there is at most one sequence $a_1 > \cdots > a_t$ such that $a_i \leq b_i + 1$ for all $i$ and $A = \cup[a_i, b_i]$. We apply this inductively on $l$ to show that $m_{i,l}$ and $n_{i,l}$, $i = 1, \ldots, t$ are determined by the supercuspidal support.

# Bibliography

[1] J. Arthur – "A trace formula for reductive groups. I. Terms associated to classes in $G(\mathbf{Q})$", *Duke Math. J.* **45** (1978), no. 4, p. 911–952.

[2] ———, "The invariant trace formula. II. Global theory", *J. Amer. Math. Soc.* **1** (1988), no. 3, p. 501–554.

[3] ———, "An introduction to the trace formula", in *Harmonic analysis, the trace formula, and Shimura varieties*, Clay Math. Proc., vol. 4, Amer. Math. Soc., Providence, RI, 2005, p. 1–263.

[4] J. Arthur & L. Clozel – *Simple algebras, base change, and the advanced theory of the trace formula*, Annals of Mathematics Studies, vol. 120, Princeton University Press, Princeton, NJ, 1989.

[5] A.-M. Aubert – "Dualité dans le groupe de Grothendieck de la catégorie des représentations lisses de longueur finie d'un groupe réductif $p$-adique", *Trans. Amer. Math. Soc.* **347** (1995), no. 6, p. 2179–2189.

[6] A. I. Badulescu – "Global Jacquet-Langlands correspondence, multiplicity one and classification of automorphic representations", *Invent. Math.* **172** (2008), no. 2, p. 383–438, with an appendix by Neven Grbac.

[7] ———, "On p-adic Speh representations", (2011).

[8] A. Beĭlinson & J. Bernstein – "Localisation de $g$-modules", *C. R. Acad. Sci. Paris Sér. I Math.* **292** (1981), no. 1, p. 15–18.

[9] I. N. Bernstein & A. V. Zelevinsky – "Induced representations of reductive $\mathfrak{p}$-adic groups. I", *Ann. Sci. École Norm. Sup. (4)* **10** (1977), no. 4, p. 441–472.

[10] A. Borel & N. Wallach – *Continuous cohomology, discrete subgroups, and representations of reductive groups*, second éd., Mathematical Surveys and Monographs, vol. 67, American Mathematical Society, Providence, RI, 2000.

[11] N. Bourbaki – *Lie groups and Lie algebras. Chapters 4–6*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2002, Translated from the 1968 French original by Andrew Pressley.

[12] J.-F. Boutot, L. Breen, P. Gérardin, J. Giraud, J.-P. Labesse, J. S. Milne & C. Soulé – *Variétés de Shimura et fonctions L*, Publications Mathématiques de l'Université Paris VII [Mathematical Publications of the University of Paris VII], vol. 6, Université de Paris VII U.E.R. de Mathématiques, Paris, 1979.

[13] C. J. Bushnell & G. Henniart – *The local Langlands conjecture for* GL(2), Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 335, Springer-Verlag, Berlin, 2006.

[14] A. Caraiani – "Local-global compatibility and the action of monodromy on nearby cycles", *Duke Math. J.*, à paraître.

[15] R. W. Carter – *Finite groups of Lie type*, Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1985, Conjugacy classes and complex characters, A Wiley-Interscience Publication.

[16] W. Casselman – "The unramified principal series of p-adic groups. I. The spherical function", *Compositio Math.* **40** (1980), no. 3, p. 387–406.

[17] C.-L. Chai – "Newton polygons as lattice points", *Amer. J. Math.* **122** (2000), no. 5, p. 967–990.

[18] G. Chenevier & D. Renard – "Characters of Speh representations and Lewis Caroll identity", *Represent. Theory* **12** (2008), p. 447–452.

[19] N. Chriss & V. Ginzburg – *Representation theory and complex geometry*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2010, Reprint of the 1997 edition.

[20] L. Clozel – "On limit multiplicities of discrete series representations in spaces of automorphic forms", *Invent. Math.* **83** (1986), no. 2, p. 265–284.

[21] ______, "Orbital integrals on $p$-adic groups: a proof of the Howe conjecture", *Ann. of Math. (2)* **129** (1989), no. 2, p. 237–251.

[22] ______, "The fundamental lemma for stable base change", *Duke Math. J.* **61** (1990), no. 1, p. 255–302.

[23] ______, "Représentations galoisiennes associées aux représentations automorphes autoduales de $\mathrm{GL}(n)$", *Inst. Hautes Études Sci. Publ. Math.* (1991), no. 73, p. 97–145.

[24] ______, "Nombre de points des variétés de Shimura sur un corps fini (d'après R. Kottwitz)", *Astérisque* (1993), no. 216, p. Exp. No. 766, 4, 121–149, Séminaire Bourbaki, Vol. 1992/93.

[25] ______, "Purity reigns supreme", *Int. Math. Res. Notices* (2011), À paraître.

[26] L. Clozel & G. Chenevier – "Corps des nombres peu ramifiés et formes automorphes autoduales", *Journal of the A.M.S.* **22** (2009), p. 467–519.

[27] L. Clozel & P. Delorme – "Pseudo-coefficients et cohomologie des groupes de Lie réductifs réels", *C. R. Acad. Sci. Paris Sér. I Math.* **300** (1985), no. 12, p. 385–387.

[28] L. Clozel, M. Harris & J.-P. Labesse – "Construction of automorphic Galois representations, I", in *On the stabilization of the trace formula*, Stable Trace Formula, Shimura Varieties and Arithmetic Applications, vol. 1, Int. Press, Somerville, MA, 2011, p. 497–527.

[29] P. Deligne – "Travaux de Shimura", in *Séminaire Bourbaki, 23ème année (1970/71), Exp. No. 389*, Springer, Berlin, 1971, p. 123–165. Lecture Notes in Math., Vol. 244.

[30] ______, "Formes modulaires et représentations de GL(2)", in *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, Springer, Berlin, 1973, p. 55–105. Lecture Notes in Math., Vol. 349.

[31] ______, "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, Amer. Math. Soc., Providence, R.I., 1979, p. 247–289.

[32] P. Deligne, D. Kazhdan & M.-F. Vigneras – *Représentations des groupes réductifs sur un corps local*, Travaux en Cours. [Works in Progress], Hermann, Paris, 1984.

[33] P. Deligne & G. Lusztig – "Representations of reductive groups over finite fields", *Ann. of Math. (2)* **103** (1976), no. 1, p. 103–161.

[34] P. Deligne & M. Rapoport – "Les schémas de modules de courbes elliptiques", in *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, Springer, Berlin, 1973, p. 143–316. Lecture Notes in Math., Vol. 349.

[35] M. Demazure – *Lectures on p-divisible groups*, Lecture Notes in Mathematics, vol. 302, Springer-Verlag, Berlin, 1986, Reprint of the 1972 original.

[36] F. Digne & J. Michel – *Representations of finite groups of Lie type*, London Mathematical Society Student Texts, vol. 21, Cambridge University Press, Cambridge, 1991.

[37] G. van Dijk – "Computation of certain induced characters of p-adic groups", *Math. Ann.* **199** (1972), p. 229–240.

[38] M. Duflo & J.-P. Labesse – "Sur la formule des traces de Selberg", *Ann. Sci. École Norm. Sup. (4)* **4** (1971), p. 193–284.

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

[40] K. Fujiwara – "Rigid geometry, Lefschetz-Verdier trace formula and Deligne's conjecture", *Invent. Math.* **127** (1997), no. 3, p. 489–533.

[41] B. H. Gross – "On the Satake isomorphism", in *Galois representations in arithmetic algebraic geometry (Durham, 1996)*, London Math. Soc. Lecture Note Ser., vol. 254, Cambridge Univ. Press, Cambridge, 1998, p. 223–237.

[42] A. Grothendieck – *Groupes de Barsotti-Tate et cristaux de Dieudonné*, Les Presses de l'Université de Montréal, Montréal, Que., 1974, Séminaire de Mathématiques Supérieures, No. 45 (Été, 1970).

[43] T. C. Hales – "Unipotent representations and unipotent classes in $\mathrm{SL}(n)$", *Amer. J. Math.* **115** (1993), no. 6, p. 1347–1383.

[44] M. Harish-Chandra – "Harmonic analysis on reductive $p$-adic groups", in *Harmonic analysis on homogeneous spaces (Proc. Sympos. Pure Math., Vol. XXVI, Williams Coll., Williamstown, Mass., 1972)*, Amer. Math. Soc., Providence, R.I., 1973, p. 167–192.

[45] M. Harris & R. Taylor – *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, NJ, 2001, with an appendix by Vladimir G. Berkovich.

[46] Y. Ihara – "The congruence monodromy problems", *J. Math. Soc. Japan* **20** (1968), p. 107–121.

[47] ———, *On congruence monodromy problems*, MSJ Memoirs, vol. 18, Mathematical Society of Japan, Tokyo, 2008, Reproduction of the Lecture Notes: Volume 1 (1968) [MR0289518], Volume 2 (1969) [MR0289519] (University of Tokyo), with author's notes (2008).

[48] H. Jacquet & R. P. Langlands – *Automorphic forms on* GL(2), Lecture Notes in Mathematics, Vol. 114, Springer-Verlag, Berlin, 1970.

[49] C. Jantzen – "Jacquet modules of $p$-adic general linear groups", *Represent. Theory* **11** (2007), p. 45–83 (electronic).

[50] M. Kashiwara & T. Tanisaki – "The Kazhdan-Lusztig conjecture for Kac-Moody Lie algebras", *Sūrikaisekikenkyūsho Kōkyūroku* (1995), no. 895, p. 44–66, Noncommutative analysis on homogeneous spaces (Japanese) (Kyoto, 1994).

[51] N. M. Katz & B. Mazur – *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies, vol. 108, Princeton University Press, Princeton, NJ, 1985.

[52] R. E. Kottwitz – "Orbital integrals on $\mathrm{GL}_3$", *Amer. J. Math.* **102** (1980), no. 2, p. 327–384.

[53] ———, "Rational conjugacy classes in reductive groups", *Duke Math. J.* **49** (1982), no. 4, p. 785–806.

[54] ———, "Shimura varieties and twisted orbital integrals", *Math. Ann.* **269** (1984), no. 3, p. 287–300.

[55] ———, "Isocrystals with additional structure", *Compositio Math.* **56** (1985), no. 2, p. 201–220.

[56] ———, "Base change for unit elements of Hecke algebras", *Compositio Math.* **60** (1986), no. 2, p. 237–250.

[57] ———, "Shimura varieties and $\lambda$-adic representations", in *Automorphic forms, Shimura varieties, and L-functions, Vol. I (Ann Arbor, MI, 1988)*, Perspect. Math., vol. 10, Academic Press, Boston, MA, 1990, p. 161–209.

[58] ———, "On the $\lambda$-adic representations associated to some simple Shimura varieties", *Invent. Math.* **108** (1992), no. 3, p. 653–665.

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

[60] ———, "Isocrystals with additional structure. II", *Compositio Math.* **109** (1997), no. 3, p. 255–339.

[61] ———, "Dimensions of Newton strata in the adjoint quotient of reductive groups", *Pure Appl. Math. Q.* **2** (2006), no. 3, Special Issue: In honor of Robert D. MacPherson. Part 1, p. 817–836.

[62] ———, "The number of points on the modular curve over finite fields", 2011.

[63] A. Kret – "The basic stratum of some simple Shimura varieties", *Math. Ann.*, À paraître.

[64] ———, "The basic stratum of Kottwitz's arithmetical varieties", *Preprint* (2012).

[65] J.-P. Labesse – "Cohomologie, stabilisation et changement de base", *Astérisque* (1999), no. 257, p. vi+161, Appendix A by Laurent Clozel and Labesse, and Appendix B by Lawrence Breen.

[66] ———, "Changement de base CM et séries discrètes", in *On the stabilization of the trace formula*, Stab. Trace Formula Shimura Var. Arith. Appl., vol. 1, Int. Press, Somerville, MA, 2011, p. 429–470.

[67] J.-P. Labesse & J. Schwermer – "On liftings and cusp cohomology of arithmetic groups", *Invent. Math.* **83** (1986), no. 2, p. 383–401.

[68] J.-P. Labesse & J.-L. Waldspurger – "La formule des traces tordue d'apres le Friday morning seminar", *Preprint* (2012).

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

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

[71] E. Lapid & A. Mínguez – "On a determinantal formula of Tadić", *Amer. J. Math.* **À paraître**, `http://www.ma.huji.ac.il/~erezla/publications.html`.

[72] G. Laumon – *Cohomology of Drinfeld modular varieties. Part I*, Cambridge Studies in Advanced Mathematics, vol. 41, Cambridge University Press, Cambridge, 1996, Geometry, counting of points and local harmonic analysis.

[73] G. Laumon & M. Rapoport – "The Langlands lemma and the Betti numbers of stacks of $G$-bundles on a curve", *Internat. J. Math.* **7** (1996), no. 1, p. 29–45.

[74] E. Mantovan – "The Newton stratification", *Paris book project* (2011), à paraître.

[75] R. Menares – "Equidistribution of Hecke points on the supersingular module", *Proc. Amer. Math. Soc.*, à paraître.

[76] J. S. Milne – "Points on Shimura varieties mod $p$", in *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, p. 165–184.

[77] A. Mínguez – "Unramified representations of unitary groups", in *On the stabilization of the trace formula*, Stab. Trace Formula Shimura Var. Arith. Appl., vol. 1, Int. Press, Somerville, MA, 2011, p. 389–410.

[78] C. Mœglin – "Classification et changement de base pour les séries discrètes des groupes unitaires $p$-adiques", *Pacific J. Math.* **233** (2007), no. 1, p. 159–204.

[79] C. Mœglin & J.-L. Waldspurger – "Sur l'involution de Zelevinski", *J. Reine Angew. Math.* **372** (1986), p. 136–177.

[80] ______, "Le spectre résiduel de GL($n$)", *Ann. Sci. École Norm. Sup. (4)* **22** (1989), no. 4, p. 605–674.

[81] S. Morel – *On the cohomology of certain noncompact Shimura varieties*, Annals of Mathematics Studies, vol. 173, Princeton University Press, Princeton, NJ, 2010, with an appendix by Robert Kottwitz.

[82] L. Morris – "$P$-cuspidal representations of level one", *Proc. London Math. Soc. (3)* **58** (1989), no. 3, p. 550–558.

[83] ______, "Tamely ramified intertwining algebras", *Invent. Math.* **114** (1993), no. 1, p. 1–54.

[84] F. Oort – "Newton polygons and formal groups: conjectures by Manin and Grothendieck", *Ann. of Math. (2)* **152** (2000), no. 1, p. 183–206.

[85] ______, "Newton polygon strata in the moduli space of abelian varieties", in *Moduli of abelian varieties (Texel Island, 1999)*, Progr. Math., vol. 195, Birkhäuser, Basel, 2001, p. 417–440.

[86] D. Ramakrishnan – "Pure motives and automorphic forms", in *Motives (Seattle, WA, 1991)*, Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, p. 411–446.

[87] M. Rapoport – "A guide to the reduction modulo $p$ of Shimura varieties", *Astérisque* (2005), no. 298, p. 271–318, Automorphic forms. I.

[88] M. Rapoport & M. Richartz – "On the classification and specialization of $F$-isocrystals with additional structure", *Compositio Math.* **103** (1996), no. 2, p. 153–181.

[89] D. Renard – *Représentations des groupes réductifs p-adiques*, Cours Spécialisés [Specialized Courses], vol. 17, Société Mathématique de France, Paris, 2010.

[90] F. Rodier – "Représentations de GL($n$, $k$) où $k$ est un corps $p$-adique", in *Bourbaki Seminar, Vol. 1981/1982*, Astérisque, vol. 92, Soc. Math. France, Paris, 1982, p. 201–218.

[91] J. D. Rogawski – *Automorphic representations of unitary groups in three variables*, Annals of Mathematics Studies, vol. 123, Princeton University Press, Princeton, NJ, 1990.

[92] P. Scholze – "The langlands-kottwitz method and deformation spaces of p-divisible groups", *Journal of the AMS* (2011), À paraître.

[93] J. A. Shalika – "The multiplicity one theorem for $\mathrm{GL}_n$", *Ann. of Math. (2)* **100** (1974), p. 171–193.

[94] S. W. Shin – "Automorphic Plancherel density theorem", *Israel J. Math* (2011), À paraître.

[95] ______, "Galois representations arising from some compact Shimura varieties", *Ann. of Math. (2)* **173** (2011), no. 3, p. 1645–1741.

[96] ______, "On the cohomological base change for unitary simulitude groups", *Preprint* (2011).

[97] R. P. Stanley – *Enumerative combinatorics. Volume 1*, second éd., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012.

[98] M. Tadić – "Classification of unitary representations in irreducible representations of general linear group (non-Archimedean case)", *Ann. Sci. École Norm. Sup. (4)* **19** (1986), no. 3, p. 335–382.

[99] ______, "On characters of irreducible unitary representations of general linear groups", *Abh. Math. Sem. Univ. Hamburg* **65** (1995), p. 341–363.

[100] J. Tits – "Reductive groups over local fields", in *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 1*, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, p. 29–69.

[101] M.-F. Vigneras – "On the global correspondence between GL($n$) and division algebras", *Lecture notes from IAS* (1984).

[102] J.-L. Waldspurger – "La formule de Plancherel pour les groupes $p$-adiques (d'après Harish-Chandra)", *J. Inst. Math. Jussieu* **2** (2003), no. 2, p. 235–333.

[103] C.-H. Wang – “Representations de Weil pour les groupes de similitudes et changement de base”, *thesis, Paris-sud* (2012).

[104] T. Wedhorn & E. Viehmann – “Ekedahl-Oort and Newton strata for Shimura varieties of PEL type”, *Preprint* (2011).

[105] A. V. Zelevinsky – “Induced representations of reductive p-adic groups. II. On irreducible representations of GL($n$)”, *Ann. Sci. École Norm. Sup. (4)* **13** (1980), no. 2, p. 165–210.