J. Tilouine  
LAGA, CNRS, UMR 7539, Institut Galilée, Université Paris 13,  
93430 Villetaneuse, France.  
E-mail : tilouine@math.univ-paris13.fr

H. Carayol  
IRMA, Université Louis Pasteur, 7, rue René Descartes, 67084 Strasbourg Cedex,  
France.  
E-mail : carayol@math.u-strasbg.fr

M. Harris  
Université Paris VII, Institut de Mathématiques de Jussieu, 2, place Jussieu,  
75251 Paris Cedex 05 France.  
E-mail : harris@math.jussieu.fr

M.-F. Vignéras  
Université Paris VII, Institut de Mathématiques de Jussieu, 2, place Jussieu,  
75251 Paris Cedex 05, France.  
E-mail : vigneras@math.jussieu.fr

Classification mathématique par sujets (2000). — 11-04, 11F03, 11F11, 11F30, 11F33,  
11F41, 11F60, 11F70, 11F72, 11F75, 11G15, 11G18, 11G40, 11R39, 11S37, 14G15,  

Mots clefs. — Algèbre de Hecke, cohomologie d’intersection, compactifications,  
conjecture de Gouvêa-Mazur, conjecture de Grothendieck, conjecture de Langlands  
locale, correspondance fonctorielle, correspondance de Jacquet-Langlands, déforma-  
tions de groupes formels, ensembles µ-admissibles et µ-permis, espaces analytiques  
rigides, espaces de modules, espace localement symétrique, fonctions L, fonction zêta  
semi-simple, forme automorphe, forme automorphe p-adique, formes modulaires,  
formule des traces, formule des traces de Lefschetz, groupes de Barsotti-Tate, groupe  
p-divisible, méthode de descente, modèle de Whittaker, pentes, polygones de Newton,  
séries d’Eisenstein, sous-groupes parahoriques, tour d’Igusa, variétés abéliennes,  
variétés de Deligne-Lusztig affines, variété de Shimura.
FORMES AUTOMORPHES (I)

ACTES DU SEMESTRE DU CENTRE ÉMILE BOREL,
PRINTEMPS 2000

édité par Jacques Tilouine, Henri Carayol, Michael Harris,
Marie-France Vignéras

Résumé. — Ce volume est le premier d’une série de deux consacrés aux formes automorphes sous leurs aspects géométrique et arithmétique et à certains points du programme de Langlands. Les thèmes abordés dans ce volume concernent les formes modulaires p-adiques, la correspondance locale de Langlands pour $GL(n)$, la cohomologie des variétés de Shimura, leur réduction modulo $p$ et leurs stratifications associées aux polygones de Newton.

Abstract (Automorphic forms (I), Proceedings of the Semester of the Émile Borel Center, Spring 2000)

This volume is the first of a series of two devoted to Automorphic Forms, in a geometric and arithmetic point of view. They also deal with certain parts of Langlands program. The themes treated in this volume include $p$-adic modular forms, the local Langlands correspondence for $GL(n)$, the cohomology of Shimura varieties, their reduction modulo $p$ and their stratification by Newton polygons.
TABLE DES MATIÈRES

Résumés des articles ................................................................. ix
Abstracts ........................................................................... xiii
Préface ............................................................................ xvii

K. BUZZARD — Questions about slopes of modular forms .......... 1
   Introduction ............................................................................. 1
   1. $\Gamma_0(N)$-regularity .......................................................... 5
   2. The conjecture ...................................................................... 7
   3. Remarks on the conjecture .................................................. 10
   4. Consequences of the conjecture and related questions .......... 11
   References ............................................................................ 14

M. HARRIS — The Local Langlands correspondence: Notes of (half) a course at the IHP Spring 2000 ........................................................... 17
   Introduction ............................................................................. 18
   1. Galois representations attached to automorphic representations of $GL(n)$ . 24
   2. Shimura varieties as moduli varieties ..................................... 35
   3. $p$-divisible $O$-modules and Drinfel’d bases ............................ 46
   4. Stratification and vanishing cycles ....................................... 58
   5. Construction of a local correspondence ............................... 74
   6. The second basic identity and Isogeny classes in the special fiber .... 92
   7. Comparison of trace formulas ............................................. 108
   8. Strata in Shimura varieties of PEL type ................................ 125
   Appendices ............................................................................ 137
   References ............................................................................ 142

H. HIDA — p-Adic automorphic forms on reductive groups ........... 147
   1. Introduction ......................................................................... 148
   2. Elliptic Curves ..................................................................... 155
   3. Vertical Control for Elliptic Modular Forms .......................... 166
### TABLE DES MATIÈRES

4. Hecke Equivariance of the Eichler-Shimura Map ........................................... 170  
5. Moduli of Abelian Schemes ................................................................. 182  
6. Shimura Varieties ............................................................................. 197  
7. Formal Theory of Automorphic Forms .................................................. 207  
8. Vertical Control for Projective Shimura Varieties ................................. 219  
9. Hilbert Modular Forms ....................................................................... 230  
10. Igusa Towers ...................................................................................... 241  
References .............................................................................................. 252  

F. OORT — *Newton polygons and $p$-divisible groups: a conjecture by Grothen-dieck* ................................................................. 255  
Introduction .............................................................................................. 255  
1. Notations ............................................................................................... 256  
2. Results : deformations of $p$-divisible groups ....................................... 259  
3. Results : deformations of polarized $p$-divisible groups and of abelian varie-
ties ......................................................................................................... 260  
4. Methods : deformations to $a \leq 1$ ....................................................... 261  
5. Methods : Cayley-Hamilton ................................................................... 263  
6. Polarized abelian varieties and quasi-polarized $p$-divisible groups .......... 263  
7. Some questions and conjectures ................................................................ 265  
References .............................................................................................. 267  

M. RAPPORT — *A guide to the reduction modulo $p$ of Shimura varieties* ...... 271  
1. Motivation : The elliptic modular curve ............................................... 273  
Part I. Local theory  
2. Parahoric subgroups ............................................................................. 277  
3. $\mu$-admissible and $\mu$-permissible set ................................................ 280  
4. Affine Deligne-Lusztig varieties ............................................................ 288  
5. The sets $X(\mu, b)_K$ ............................................................................ 291  
6. Relations to local models ....................................................................... 297  
Part II. Global theory  
7. Geometry of the reduction of a Shimura variety ..................................... 299  
8. Pseudomotivic and quasi-pseudomotivic Galois gerbs ............................ 302  
9. Description of the point set in the reduction .......................................... 305  
10. The semi-simple zeta function ................................................................ 308  
References .............................................................................................. 311  
Notes added June 2003 ............................................................................ 314  
Additional references ............................................................................. 318  

L. SAPER — *$\mathcal{L}$-modules and the Conjecture of Rapoport and Goresky-
MacPherson* ............................................................................................ 319  
1. Introduction ............................................................................................ 320  
2. Compactifications .................................................................................. 320  
3. The conjecture ....................................................................................... 322  

ASTÉRISQUE 298
TABLE DES MATIÈRES

4. A generalization ........................................................... 323
5. \(\mathcal{L}\)-modules .......................................................... 324
6. Examples of \(\mathcal{L}\)-modules .............................................. 326
7. Micro-support of \(\mathcal{L}\)-modules ......................................... 326
8. A vanishing theorem for \(\mathcal{L}\)-modules .............................. 328
9. Micro-purity of intersection cohomology .................................. 329
10. Functoriality of micro-support and proof of the Main Theorem .......... 330
11. Example/application: ordinary cohomology ............................ 331
References .......................................................................... 333

D. Soudry — On Langlands functoriality from classical groups to \(\text{GL}_n\) ....... 335
Introduction ............................................................................. 335
1. The weak lift from \(\text{SO}_{2n+1}\) to \(\text{GL}_{2n}\) ........................................ 337
2. \(L\)-functions for \(G \times \text{GL}_k\), where \(G\) is a quasi-split classical group (generic representations) ......................................................... 347
3. On the weak lift from a quasi-split classical group to \(\text{GL}_N\). ........... 361
4. Illustrations of Proofs in Low Rank Examples .......................... 368
References .............................................................................. 388

M. Strauch — On the Jacquet-Langlands correspondence in the cohomology of the Lubin-Tate deformation tower ............................... 391
1. Introduction ........................................................................... 392
2. Deformation spaces and their cohomology groups ........................ 394
3. The approach via a Lefschetz trace formula ............................... 400
4. Fixed points and the period morphism ..................................... 404
References ............................................................................... 408
RÉSUMÉS DES ARTICLES

Questions about slopes of modular forms
Kevin Buzzard ............................................................... 1

Nous formulons une conjecture prédissant, dans de nombreux cas, les valeurs $p$-adiques exactes des valeurs propres de l’opérateur de Hecke $T_p$ agissant sur les espaces de formes modulaires classiques. Cette conjecture a des conséquences très concrètes sur la théorie classique, mais elle suggère aussi de nombreuses symétries inexpliquées concernant les courbes de Coleman-Mazur.

The Local Langlands correspondence : Notes of (half) a course at the IHP Spring 2000
Michael Harris .............................................................. 17

L’article contient une description assez complète des principaux résultats du livre de l’auteur avec Richard Taylor, qui décrit les représentations galoisiennes réalisées dans la cohomologie de certaines variétés de Shimura associées aux groupes unitaires, et qui obtient la conjecture locale de Langlands pour $GL(n)$ d’un corps $p$-adique comme conséquence. Les principales étapes de la démonstration de la conjecture locale de Langlands y sont présentées, parfois simplifiées. Le gros de l’article concerne la géométrie de la variété de Shimura aux places de mauvaise réduction, où l’on dispose néanmoins de bons modèles locaux, et la description des points dans la fibre spéciale à la manière de Langlands et Kottwitz. La dernière section de l’article décrit les extensions éventuelles de ces résultats aux variétés de Shimura plus générales, ainsi qu’un compte rendu des travaux de Laurent Fargues sur ces questions.

$p$-Adic automorphic forms on reductive groups
Haruzo Hida ............................................................... 147

Nous démontrons le contrôle vertical pour les formes automorphes ordinaires $p$-adiques et l’irreductibilité de la tour d’Igusa pour les variétés de Shimura symplectique et unitaire.

Newton polygons and $p$-divisible groups : a conjecture by Grothendieck
Frans Oort ............................................................... 255

En 1970 Grothendieck a formulé une conjecture concernant les déformations de groupes $p$-divises (groupes de Barsotti-Tate). Nous décrivons une
démonstration de cette conjecture. Cela donne une information sur des strates définies par le polygone de Newton dans les espaces de modules des variétés abéliennes en caractéristique positive.

_A guide to the reduction modulo_ \( p \) _of Shimura varieties_

**MICHAEL RAPOPORT**  ........................................................... 271

Cet article est un survol de résultats sur la réduction des variétés de Shimura à structure de niveau parahorique.

**\( \mathcal{L} \)-modules and the Conjecture of Rapoport and Goresky-MacPherson**

**LESLIE SAPER**  ................................................................. 319

Considérons les groupes de cohomologie d’intersection (de perversité intermédiaire) de diverses compactifications d’un espace localement hermétiqne symétrique. Rapoport et, indépendamment, Goresky et MacPherson ont conjecturé que ces groupes coïncident pour la compactication de Borel-Serre réductive et la compactification de Baily-Borel-Satake. Cet article décrit la théorie des \( \mathcal{L} \)-modules et la façon dont elle peut s’employer pour résoudre la conjecture. Plus généralement, nous traitons une compactification de Satake pour laquelle toutes les composantes réelles à la frontière sont de « rang égal ». Les détails en seront disponibles ailleurs [26]. Comme application supplémentaire de la théorie des \( \mathcal{L} \)-modules, nous prouvons un théorème d’annulation sur le groupe de cohomologie ordinaire d’un espace localement symétrique. Ceci répond à une question soulevée par Tilouine.

**On Langlands functoriality from classical groups to \( GL_n \)**

**DAVID SOUDRY**  ................................................................. 335

Cet article est une exposition de la méthode de descente de Ginzburg, Rallis et Soudry. Cette méthode construit, pour une représentation irréductible, automorphe et cuspidale \( \tau \) telle que \( \tau = \tau^* \), une représentation irréductible, automorphe, cuspidale et générique \( \sigma(\tau) \) d’un groupe classique quasi-deployé \( G \) (qui dépend de \( GL_n \) et \( \tau \) ), telle que \( \tau \) corresponde à \( \sigma(\tau) \) par la correspondance functorielle faible (« weak lifting »). Cette construction est valable aussi pour toutes les représentations de \( GL_n(\mathbb{A}) \) qui appartiennent à la partie « tempérée » de l’image de la correspondance functorielle de Langlands de \( G \) à \( GL_n \).

**On the Jacquet-Langlands correspondence in the cohomology of the Lubin-Tate deformation tower**

**MATTHIAS STRAUCH**  ........................................................... 391

Soient \( F \) un corps local non-archimédiens et \( X \) un \( \mathfrak{O}_F \)-module formel de hauteur \( n \) sur \( \overline{F}_p \). Les schémas de déformations de \( X \) munis de structures de niveau de Drinfeld fournissent un système projectif d’espaces analytiques rigides \( (M_K)_K \), où \( K \) parcourt l’ensemble des sous-groupes compacts ouverts de \( G = GL_n(F) \). La limite inductive \( H^*_\mathcal{L} \) des espaces \( H^*_\mathcal{L}(M_K \otimes \overline{F}_\ell, \mathbb{Q}_\ell) \) (\( \ell \neq p \))
constitue une représentation virtuelle lisse du groupe $G \times B^\times$, $B$ étant une algèbre à division sur $F$ d’invariant $1/n$. Si $\pi$ est une représentation super-cuspidale de $G$, les travaux de Boyer et Harris-Taylor impliquent que dans le groupe de Grothendieck des représentations admissibles de $B^\times$ on a la relation $\text{Hom}_G(H^\times_\pi, \pi) = n \cdot (-1)^{n-1} J^L(\pi)$, $J^L$ désignant la correspondance de Jacquet-Langlands. Dans cet article nous proposons une approche de ce résultat fondé sur une formule des traces à la Lefschetz conjecturale, et nous calculons la contribution venant des points fixes.
ABSTRACTS

Questions about slopes of modular forms
KEVIN BUZZARD ............................................................... 1

We formulate a conjecture which predicts, in many cases, the precise $p$-adic valuations of the eigenvalues of the Hecke operator $T_p$ acting on spaces of classical modular forms. The conjecture has very concrete consequences in the classical theory, but can also be thought of as saying that there is a lot of unexplained symmetry in many of the Coleman-Mazur eigencurves.

The Local Langlands correspondence: Notes of (half) a course at the IHP Spring 2000
MICHAEL HARRIS .............................................................. 17

The article provides a reasonably self-contained account of the main results of the author’s book with Richard Taylor, containing a description of the Galois representations obtained in the cohomology of certain Shimura varieties attached to unitary groups, and obtaining the local Langlands conjecture for $GL(n)$ of $p$-adic fields as a consequence. The main steps in the proof of the local Langlands conjecture are presented, and in some cases simplified. The bulk of the paper concerns the geometry of the Shimura variety at places of bad reduction, where good local models are nevertheless available, and the description of points in the special fiber in the manner of Langlands and Kottwitz. The article concludes with a section describing possible extensions of these results to other Shimura varieties, and an account of some of the work of Laurent Fargues along those lines.

$p$-Adic automorphic forms on reductive groups
HARUZO HIDA ................................................................. 147

In these lecture notes, we will prove vertical control theorems for ordinary $p$-adic automorphic forms and irreducibility of the Igusa tower over unitary and symplectic Shimura varieties.
Newton polygons and \(p\)-divisible groups: a conjecture by Grothendieck

Frans Oort ................................................................. 255

In my talk in 2000 I discussed a conjecture in 1970 by Grothendieck concerning deformations of \(p\)-divisible groups; a proof of this conjecture gives access to finding properties of Newton polygon strata in the moduli spaces of polarized abelian varieties in positive characteristic.

A guide to the reduction modulo \(p\) of Shimura varieties

Michael Rapoport ........................................................... 271

This is a survey of recent work on the reduction of Shimura varieties with parahoric level structures.

\(L\)-modules and the Conjecture of Rapoport and Goresky-MacPherson

Leslie Saper ................................................................. 319

Consider the middle perversity intersection cohomology groups of various compactifications of a Hermitian locally symmetric space. Rapoport and independently Goresky and MacPherson have conjectured that these groups coincide for the reductive Borel-Serre compactification and the Baily-Borel-Satake compactification. This paper describes the theory of \(L\)-modules and how it is used to solve the conjecture. More generally we consider a Satake compactification for which all real boundary components are equal-rank. Details will be given elsewhere [26]. As another application of \(L\)-modules, we prove a vanishing theorem for the ordinary cohomology of a locally symmetric space. This answers a question raised by Tilouine.

On Langlands functoriality from classical groups to \(GL_n\)

David Soudry ................................................................. 335

This article is a survey of the descent method of Ginzburg, Rallis and Soudry. This method constructs, for an irreducible, automorphic, cuspidal, self-conjugate representation \(\tau\) on \(GL_n(\mathbb{A})\), an irreducible, automorphic, cuspidal, generic representation \(\sigma(\tau)\), on a corresponding quasi-split classical group \(G\), which lifts weakly to \(\tau\). This construction works well also for all representations of \(GL_n(\mathbb{A})\), which are in the so-called “tempered” part of the expected image of Langlands functorial lift from \(G\) to \(GL_n\).

On the Jacquet-Langlands correspondence in the cohomology of the Lubin-Tate deformation tower

Matthias Strauch ........................................................... 391

Let \(F\) be a local non-archimedean field, and let \(X\) be a one-dimensional formal \(\sigma_F\)-module over \(\overline{\mathbb{F}}_p\) of height \(n\). The formal deformation schemes of \(X\) with Drinfeld level structures give rise to a projective system of rigid-analytic spaces \((M_K)_K\), where \(K\) runs through the compact-open subgroups of \(G = GL_n(F)\). On the inductive limit \(H^*_c\) of the spaces \(H^*_c(M_K \otimes \overline{F}^\ell, \mathbb{Q}_\ell)\) (\(\ell \neq p\))
there is a smooth action of $G \times B^\times$, $B$ being a central division algebra over $F$ with invariant $1/n$. For a supercuspidal representation $\pi$ of $G$ it follows from the work of Boyer resp. Harris-Taylor that in the Grothendieck group of admissible representations of $B^\times$ one has $\text{Hom}_G(H^*_c, \pi) = n \cdot (-1)^{n-1} \mathcal{JL}(\pi), \mathcal{JL}$ denoting the Jacquet-Langlands correspondence. In this paper we propose an approach that is based on a conjectural Lefschetz trace formula for rigid-analytic spaces, and we calculate the contribution coming from the fixed points.
Un semestre spécial consacré aux formes automorphes sous leurs différents aspects (analyse harmonique, théorie des représentations, géométrie algébrique et arithmétique) s’est déroulé du 17 février au 11 juillet 2000, dans le cadre du centre Émile Borel à Paris dans les locaux de l’Institut Henri Poincaré.

Les principaux thèmes abordés ont été :

- la décomposition spectrale de l’espace des formes automorphes, la formule des traces globale, la fonctorialité de Langlands
- la réalisation géométrique des correspondances de Langlands locales et globales pour $GL(n)$ y compris pour des représentations galoisiennes modulo $\ell$, l’étude de la cohomologie de certaines variétés de Shimura et de leurs modèles locaux,
- les congruences entre formes modulaires, les familles $p$-adiques de telles formes, leurs représentations galoisiennes et leurs fonctions $L$ $p$-adiques,
- les fonctions $L$ automorphes, la théorie réciproque de Hecke et ses applications à la fonctorialité de Langlands,
- l’analyse harmonique et la théorie des représentations des groupes $p$-adiques.

Les activités ont été organisées autour de douze cours, d’un séminaire hebdomadaire comptant un ou deux exposés, et de trois mini-colloques.

a) Les titres des cours étaient les suivants :
- L. Clozel (Paris 11) : Lemme fondamental et comptage des points sur les variétés de Shimura,
- M. Harris Paris 7), G. Henniart (Paris 11) : Correspondance de Langlands locale,
- H. Hida (UCLA) : Formes modulaires $p$-adiques,
- J.-P. Labesse (Paris 7), C. Moeglin (Paris 7) : Spectre des formes automorphes et formule des traces,
- L. Lafforgue (CNRS et IHES) : Correspondance de Langlands globale pour $GL(n)$ sur les corps de fonctions,
- M. Rapoport (Bonn U.) : Sur la réduction modulo $p$ des variétés de Shimura,
- J. Rogawski (UCLA) : Formule des traces relatives,
– D. Soudry (Tel Aviv) : On Langlands functoriality from classical groups to $GL(n)$,
– G. Stevens (Boston U.) : $p$-adic families of modular forms with positive slopes,
– R. Taylor (Harvard U.) : On Artin conjecture for odd icosahedral degree two Artin representations,
– J. Tilouine (Paris 13), E. Urban (Paris 13 et Columbia) : Représentations galoisiennes et familles de Hida pour les groupes symplectiques,
– M.-F. Vignéras (Paris 7) : Correspondance de Langlands modulo $\ell$.

b) Les titres des mini-colloques étaient :
– Variétés de Shimura et arithmétique des formes automorphes du 25 au 29 avril.
– Aspects géométriques des formes automorphes,(Journées Solstice d’Été) du 19 au 24 juin,

celui du mini-colloque était :
– Analyse harmonique et représentations des groupes $p$-adiques, le 16 juin.

Les organisateurs ont demandé aux orateurs de rédiger des textes portant sur leur cours, exposé ou sur des sujets voisins et tiennent à remercier les auteurs qui ont accepté de le faire, pour leur contribution d’abord, mais aussi pour leur patience; nous espérons qu’ils voudront bien nous excuser du retard avec lequel ces volumes paraissent.

Les articles ont été examinés par des rapporteurs que nous remercions pour leur aide aussi désintéressée qu’utile. Nous tenons à exprimer la reconnaissance des organisateurs et, nous l’espérons, des participants, pour l’atmosphère cordiale dans laquelle s’est déroulé le semestre. Y ont contribué à leur énergie et leur compétence :

– les directeurs successifs de l’Institut Henri Poincaré et leur équipe : Joseph Oesterlé qui nous a conseillé et aidé pendant la phase préparatoire, et Michel Broué, nouveau directeur au moment de la conférence proprement dite qui a toujours réagi avec célérité et efficacité aux divers problèmes que nous lui posions.

– tout le personnel de l’Institut Henri Poincaré, en particulier Mme Nocton, bibliothécaire, qui a accepté d’accroître les heures d’ouverture de la bibliothèque pour faciliter le travail des mathématiciens visitant l’Institut pendant la durée du semestre.

Enfin, deux mentions spéciales reviennent à Monique Douchez, secrétaire de l’équipe Formes Automorphes de l’Université de Paris 7, pour un travail d’organisation considérable effectué avec bonne humeur et diligence, et à la secrétaire du Centre E. Borel, Annie Touchant, qui a assuré avec une gentillesse et une amabilité indéfectibles la gestion quotidienne des problèmes administratifs et matériels qui se posaient aux très nombreux participants.

Les organisateurs
QUESTIONS ABOUT SLOPES OF MODULAR FORMS

by

Kevin Buzzard

Abstract. — We formulate a conjecture which predicts, in many cases, the precise $p$-adic valuations of the eigenvalues of the Hecke operator $T_p$ acting on spaces of classical modular forms. The conjecture has very concrete consequences in the classical theory, but can also be thought of as saying that there is a lot of unexplained symmetry in many of the Coleman-Mazur eigencurves.

Introduction

Let $N \geq 1$ be a fixed integer, and let $p$ denote a fixed prime not dividing $N$. If $k \in \mathbb{Z}$ then there is a complex vector space $S_k(\Gamma_0(Np))$ of cusp forms of weight $k$ and level $Np$. This space is finite-dimensional over the complex numbers and comes equipped with an action of the Hecke operator $U_p$, an endomorphism whose eigenvalues are non-zero complex numbers. The characteristic polynomial of $U_p$ has integer coefficients, which implies that the eigenvalues are algebraic integers. Hence we can consider the eigenvalues as lying in $\mathbb{C}$ or in $\mathbb{Q}_l$ for any prime $l$.

The $U_p$-eigenvalues fall naturally into two classes, $p$-old ones and $p$-new ones. The $p$-old eigenvalues are the roots of $X^2 - a_p X + p^{k-1}$, where $a_p$ runs through the eigenvalues of $T_p$ acting on $S_k(\Gamma_0(N))$. A deep theorem of Deligne says that the $p$-old eigenvalues all have complex absolute value $p^{(k-1)/2}$. The $p$-new eigenvalues are what is left, and it is well-known that these eigenvalues are square roots of $p^{k-2}$. Hence the complex valuations of these $U_p$-eigenvalues are known in every case. Moreover, from these definitions it is clear that if $l \neq p$ is a prime then the $U_p$-eigenvalues are all $l$-adic units.

2000 Mathematics Subject Classification. — 11-04, 11F11, 11F30, 11F33.

Key words and phrases. — Modular forms, slopes, Gouvêa-Mazur conjecture.
From this point of view, the question that remains about valuations of eigenvalues is:

**Question.** — *What can one say about the p-adic valuations of the eigenvalues of $U_p$?*

The term “slopes” is used nowadays to refer to these valuations. A study of the simplest special case, namely $N = 1$ and $p = 2$, shows that the answer is nowhere near as simple as the other cases. The forms which are 2-new at weight $k$ will all have slope $\frac{k-2}{2}$ and this leaves us with the oldforms, whose slopes we can easily compute from the theory of the Newton Polygon, if we know the 2-adic valuations of the eigenvalues of $T_2$ acting on cusp forms of level 1. The smallest $k$ for which non-zero level 1 cusp forms exist is $k = 12$; the space $S_{12}(\text{SL}_2(\mathbb{Z}))$ is one-dimensional, and $T_2$ acts as multiplication by $-24$. Hence the 2-old eigenvalues of $U_2$ at weight 12 and level 2 are the two roots of $X^2 + 24X + 2^{11}$, and these two roots have 2-adic valuations equal to 3 and 8. Note that $3 \neq 8$, and so the story is already necessarily different to the complex and $l$-adic cases. We include a short table of valuations and slopes for small weights.

<table>
<thead>
<tr>
<th>Weight</th>
<th>2-adic valuations of $T_2$-eigenvalues at level 1</th>
<th>Slopes of $U_2$ at level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>3</td>
<td>3,8</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>6,6</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>3,7,12</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>4,8,13</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>3,9,9,16</td>
</tr>
<tr>
<td>22</td>
<td>5</td>
<td>5,10,10,16</td>
</tr>
<tr>
<td>24</td>
<td>3,7</td>
<td>3,7,11,16,20</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>4,12,12,12,21</td>
</tr>
</tbody>
</table>

From this table, one wonders whether there is any structure at all in the slopes. However, the purpose of this paper is to suggest that in fact there is a very precise structure here. In fact, in this paper we explain a completely elementary conjectural combinatorial recipe, recursive in the weight $k$, for generating the above table line by line. In fact, for a large class of pairs $(N, p)$ (including $(1, p)$ for all primes $p < 100$ apart from 59 and 79) we give a conjectural recipe for the valuations of the $T_p$-eigenvalues at level $N$, and hence the slopes of $U_p$ at level $Np$. We strongly believe that there should be a recipe for generating the slopes of $U_p$ at level $Np$ for any $N$ and $p$, given as an input the slopes for level $N$ and weights at most $p + 2$. However we have not yet managed to formulate such a recipe at the present time. In this paper, we offer a recipe only in the case where $p$ is $\Gamma_0(N)$-regular, a term that we shall define later.
Before we explain our conjectural recipe, we shall explain what is known about the slopes of $U_p$, and what has been conjectured before. The first observation, hinted at by the apparent randomness in the table above, is that to find structure in the slopes of $U_p$ one should, contrary to the complex and $l$-adic cases, not consider the slopes at one fixed weight, but let the weight vary. There are well-known concrete examples of this phenomenon. For example, a theorem of Hida says that for fixed level, the number of $U_p$-eigenvalues with slope zero is bounded, and indeed for $k \geq 2$ this number depends only on $k$ modulo $p - 1$ (resp. modulo 2) for $p$ odd (resp. $p = 2$).

As an example of this, we note that there are no slope zero forms in the table above, and we deduce from Hida’s theorem that in fact for $N = 1$ and $p = 2$ there will never be any slope zero forms, however high the weight gets.

These theorems about $U_p$-eigenvalues of slope 0 were generalised by Gouvea and Mazur to an explicit conjecture in [11] about the number of eigenvalues of arbitrary slope as the weight varies. The Gouvea-Mazur conjecture says that if $M \geq 0$ is any integer, then for $k$ and $k'$ sufficiently large (which nowadays means at least $M + 2$) and congruent modulo $(p - 1)p^M$, the number of $U_p$-eigenvalues of slope $\alpha$ at weight $k$ and weight $k'$ should be the same, for any $\alpha \leq M$. Experimental evidence for this conjecture was supplied by Mestre in the case where $N = 1$ and $p$ is small. A few years after this conjecture was made, ground-breaking work of Coleman in [6] showed that cuspidal eigenforms naturally lay in $p$-adic analytic families, and an analysis by Wan [17] of Coleman’s methods showed that one could deduce a weaker version of the Gouvea-Mazur conjectures, namely that for $k$ and $k'$ sufficiently large, and congruent modulo $(p - 1)p^M$, the number of eigenvalues with slope $\alpha$ at these two weights were equal, if $\alpha \leq O(\sqrt{M})$. The constants here are all explicit.

Note added in proof: For a few years the gap between the conjecture and the theoretical results was a mystery, but in some sense the mystery was resolved when a counterexample to the Gouvea-Mazur conjecture was found by the author and F. Calegari in the case $N = 1$ and $p = 59$. This paper was written before the counterexample was found and in fact it was the results in this paper which led the author and Calegari to a study of the particular case $p = 59$, which is the smallest prime for which (at level 1) the results of this paper do not apply. Note that for $N = 1$, although the Gouvea-Mazur conjecture is false for $p = 59$, it may well still be true for $N = 1$ and $p < 59$, and indeed perhaps the results of this paper are an indication that it is true if $p$ is $\Gamma_0(N)$-regular (see later for the definition). This paper is not about the counterexample at $p = 59$ but about the extra structure discovered for $p < 59$. The counterexample at $p = 59$ is explained in [4].

The families in Coleman’s work were beautifully interpolated into a mysterious geometric object, constructed by Coleman and Mazur, called an “eigencurve”, whose very existence implies deep results about modular forms. One can compute what are essentially local equations for small pieces of these eigencurves for explicit $p$ and $N$. 

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
and computations of this nature have been undertaken by Emerton in [9] and Coleman, Stevens and Teitelbaum in [7], where for \( N = 1 \) and \( p = 2, 3 \) respectively the authors manage to compute the majority of the part of the eigencurve with smallest slope. Computations like this have concrete consequences in the theory—for example, Emerton deduced from his computations that when \( N = 1 \), the smallest slope of \( U_2 \) was periodic as the weight increased, repeating the pattern 3, 6, 3, 4, 3, 5, 3, 4 indefinitely (one can see the first instance of this pattern in the table above, which already indicates that the table is much too small to be able to indicate what is going on).

The computations of Mestre concerning the Gouvêa-Mazur conjecture were done about ten years ago, and because computers are currently increasing vastly in speed, it was clear that one could go much further nowadays. The author’s motivations for actually going further were several—firstly, Wan’s results, and unpublished analogous theorems of the author for automorphic forms on definite quaternion algebras, both gave a version of the Gouvêa-Mazur conjecture with \( \alpha \leq O(\sqrt{M}) \) rather than \( \alpha \leq M \), and this led us to believe that perhaps the Gouvêa-Mazur conjectures were too optimistic. Hence we thought we would make a concerted effort to search for a counterexample (Note added in proof: see [4] for the counterexample that we ultimately found). Secondly, several years ago we had come up with an (again unpublished) fast algorithm for computing a matrix representing \( T_2 \) on \( S_k(\text{SL}_2(\mathbb{Z})) \) and we felt that this would help us with the project. Thirdly, it seemed that a serious computation would be a way to get a “feeling” for the Coleman-Mazur eigencurves. Finally, William Stein has recently written a package that computes spaces of modular forms, and a serious computation seemed like a good way of testing his programs. We should remark that Gouvêa also did many computations since [11] was written, and the reader that wants to see the current status of things is strongly recommended to refer to [10] or to [16].

Our extensive numerical calculations did not (initially) reveal any counterexamples to the Gouvêa-Mazur conjecture (Note added in proof: however they did lead us to the observation that \( p = 59 \) was somehow different to other primes \( p < 59 \) and this is what ultimately led to the counterexample). On the contrary, to our surprise, they revealed what in many cases seemed to be far more structure. The Gouvêa-Mazur conjectures predict local constancy of slopes, in some sense, whereas, with the help of the numerical data, we were able to formulate in many cases a new conjecture, which predicted all slopes precisely.

Our investigations of the phenomenon of patterns in slopes were inspired by the aforementioned computations of Emerton, and also by results in Lawren Smithline’s 1999 UC Berkeley thesis. We are grateful to both Smithline and Emerton for several helpful remarks. Smithline proves in his thesis that there is some structure to the set of slopes of weight zero 3-adic overconvergent modular forms of level 1, and this structure was one of the reasons why we were inspired to do these computations. We
are also grateful to William Stein and Tamzin Cuming for providing many spare CPU cycles, and to the referee for several helpful comments.

Although the conjectural formula that is the heart of this paper is of a purely elementary nature, it seems very complicated to explain. The structure of this paper is as follows. In the first section we explain what we mean by the notion of $\Gamma_0(N)$-regularity above. In the second, we formulate the conjecture. The third section is an attempt to explain heuristically our motivation behind the precise details of the conjecture. Finally, the fourth section raises further related questions. In particular, the finiteness questions 4.4 and 4.5 do not apparently appear to have been raised before.

Although we shall not mention overconvergent forms in the main body of this paper, we should perhaps mention that the original reason we were motivated to do these computations was to try and understand the geometry of the Coleman-Mazur eigencurves in some specific cases. Closely related to conjectures about the values of slopes of classical modular forms are analogous conjectures about the values of slopes of overconvergent forms, as one can see by a simple continuity argument and the theorem of Coleman that overconvergent forms of small slope are classical. In fact these conjectures below about slopes of classical forms could be entirely rephrased in terms of overconvergent forms. On the other hand, this rephrasing seemed equally complicated, if not more, and so we have not mentioned it below. However, in the specific case of $p = 2$ and $N = 1$, the author and F. Calegari have managed to come up with a conjecture for both classical and overconvergent forms that is much simpler to state, and have furthermore have succeeded in proving it for overconvergent forms of weight 0—we can prove that the valuation of the $n$th slope of $U_2$ is $1 + 2v_2\left(\frac{(2n)!}{n!}\right)$. In particular, all slopes are positive odd integers, which could perhaps be regarded as some very weak evidence towards Question 4.2 below. See the forthcoming [5] for more details.

Note added in proof: the forthcoming Northwestern thesis [12] of Graham Herrick attempts to explain the main conjecture of this paper in a much more conceptual manner.

1. $\Gamma_0(N)$-regularity

Let $N$ be a fixed positive integer and let $p$ be a fixed prime not dividing $N$. For $k \geq 2$ an even integer, define $v_k$ to be the sequence of $p$-adic valuations of the eigenvalues of $T_p$, acting on $\tilde{S}_k(\Gamma_0(N))$, with multiplicities, arranged in increasing order. For example, if $N = 1$ and $p = 2$ then we see from the table above that $v_{24} = [3, 7]$, where we use square brackets to denote a sequence.

We firstly remark that there is probably no elementary combinatorial formula for predicting $v_k$ in general. For example, when $k = 12$ and $N = 1$ one finds that $T_p$ acts
as a \( p \)-adic unit for most primes, but occasionally (for example for \( p = 2, 3, 5, 7 \) and 2411) the eigenvalue of \( T_p \) is divisible by \( p \). Our goals are thus slightly more modest. Define \( k_p = \frac{p+3}{2} \) if \( p > 2 \), and define \( k_2 = 4 \).

**Question 1.1.** — Is there an elementary combinatorial recipe which, given \( v_k \) for all \( k \leq k_p \), predicts \( v_k \) for all \( k \geq 2 \)?

We have substantial numerical evidence suggesting that the answer to the question above is “yes”, although we have not really made the question precise. Indeed, we shall not make this question precise in general, but only in the case where the prime \( p \) is \( \Gamma_0(N) \)-regular. We now give a definition of \( \Gamma_0(N) \)-regularity.

**Definition 1.2 (\( \Gamma_0(N) \)-regularity: \( p \) odd).** — If \( p > 2 \) then we say that \( p \) is \( \Gamma_0(N) \)-regular if the eigenvalues of \( T_p \) acting on \( S_k(\Gamma_0(N)) \) are all \( p \)-adic units, for all even integers \( 2 \leq k \leq k_p \).

If \( p = 2 \) then this definition is not a good idea in general, because by Hida theory we see that the number of unit eigenvalues of \( T_p \) at weight 4 is bounded above by \( \dim(S_2(\Gamma_0(2N))) - \dim(S_2(\Gamma_0(N))) \), which is almost always less than \( \dim(S_4(\Gamma_0(N))) \).

**Definition 1.3 (\( \Gamma_0(N) \)-regularity: \( p = 2 \)).** — We say that the prime \( p = 2 \) is \( \Gamma_0(N) \)-regular if

1. The eigenvalues of \( T_2 \) on \( S_2(\Gamma_0(N)) \) are 2-adic units.
2. There are exactly \( \dim(S_2(\Gamma_0(2N))) - \dim(S_2(\Gamma_0(N))) \) eigenvalues of \( T_2 \) on \( S_4(\Gamma_0(N)) \) which are 2-adic units, and all the others (if any) have 2-adic valuation equal to 1.

The reader who would like a uniform definition should perhaps think of the definition as saying that \( p \) is \( \Gamma_0(N) \)-regular if the valuations of the eigenvalues of \( T_p \) for \( k \leq k_p \) are as small as possible. This definition for \( p = 2 \) is a little ad-hoc, and is based on the fact that a computation in the case of \( p = 2 \) and \( N = 5 \) showed that we did not want 2 to be \( \Gamma_0(5) \)-regular. The modification is motivated by the following consequence of (one form of) the Gouvêa-Mazur conjecture: if \( p = 2 \) and all eigenvalues of \( T_2 \) on \( S_2(\Gamma_0(N)) \) are units, then there should be no eigenvalues of \( T_2 \) on \( S_4(\Gamma_0(N)) \) with valuation strictly between 0 and 1. This justifies the phrase “as small as possible” above.

Assume for the rest of this section that \( p > 3 \). Then any continuous odd irreducible Galois representation \( \rho : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{F}_p) \) which has determinant equal to an integer power of the cyclotomic character, and which is modular, has a twist coming from a characteristic zero form of weight at most \( p + 1 \), level equal to the conductor of \( \rho \), and trivial character. Furthermore, one can read off whether the eigenvalue of \( T_p \) on such a form is a \( p \)-adic unit by the local behaviour of \( \rho \) at \( p \). Details of these results can be found for example in [15] and [8]. Finally, by the theory of theta cycles, if
there is a mod $p$ eigenform of level $N$ and weight $k$ with $\frac{p+3}{2} < k \leq p + 1$ which is in the kernel of $T_p$, then there is another such form of weight $p + 3 - k \leq \frac{p+3}{2}$. From these facts, one can easily deduce

**Lemma 1.4.** — $p > 3$ is $\Gamma_0(N)$-regular if and only if any irreducible modular Galois representation $\rho : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{F}_p)$ with conductor dividing $N$ and determinant a power of the mod $p$ cyclotomic character is necessarily reducible when restricted to a decomposition group at $p$.

The restriction to $p > 3$ is because of technical problems lifting mod $p$ forms with trivial character to characteristic zero forms with trivial character, and could perhaps be avoided if we worked with $\Gamma_1(N)$, or with mod $p$ modular forms.

If one now assumes Serre’s conjecture on modularity of continuous irreducible odd mod $p$ Galois representations, then one can deduce a purely representation-theoretic criteria for $\Gamma_0(N)$-irregularity, because the word “modular” in the lemma above can then be replaced by “continuous and odd”. This formulation of $\Gamma_0(N)$-regularity can perhaps be thought of as an analogue to the representation-theoretic criteria for irregular (in the classical sense) primes—if a prime is irregular in the classical sense then there is an upper-triangular 2-dimensional mod $p$ Galois representation which is non-split, unramified away from $p$ and tamely ramified at $p$.

The first few $\text{SL}_2(\mathbb{Z})$-irregular primes are 59, 79, 107, 131, 139, 151, 173, . . . .

### 2. The conjecture

Recall that we have fixed $N$ and $p$, where now $p \geq 2$ is back to being an arbitrary prime not dividing $N$. Below, we shall give a recipe for constructing sequences $s_2, s_4, s_6, \ldots$. These sequences depend only on $k, p$ and the dimension of various space of cusp forms of level $N$ and $Np$. The main conjecture of this paper is

**Conjecture 2.1.** — Assume that $p$ is $\Gamma_0(N)$-regular. Then the sequences $s_2, s_4, \ldots$ of integers are precisely the sequences $v_2, v_4, \ldots$ of $p$-adic valuations of $T_p$ acting on $S_k(\Gamma_0(N))$.

The recipe defining the $s_k$ is messy, and it seems to us that its ideal presentation is as a computer program. The recipe depends on the dimension of various spaces of cusp forms, sometimes with non-trivial character, and the only package of which we are currently aware that has these things built in is the MAGMA package [1]. We have implemented our conjecture in MAGMA, and the source is available at the author’s web page [2]. We have also implemented our conjecture in pari-gp [14], but this was a little messier because we also had to implement some of the theory of Dirichlet characters, and also routines for computing dimensions of spaces of cusp forms with non-trivial character. Again, the source is available at [2]. The reader
may well find playing with these programs a lot more enjoyable than attempting to read the description of the conjecture below.

Firstly, some notation. A sequence denotes a finite sequence \([a_1, a_2, \ldots, a_n]\) of integers. The square brackets are merely notational. If \(s = [a_1, a_2, \ldots, a_n]\) is a sequence, then we let \(l(s) := n\) denote the length of \(s\) and we let \(s[i] := a_i\) denote its \(i\)th term. We say that a sequence \(s\) is increasing if \(s[i] \leq s[i + 1]\) for all \(i\) with \(1 \leq i < l(s)\). The union \(a \cup b\) of two sequences is the sequence of length \(l(a) + l(b)\) defined as the sequence \(a\) followed by the sequence \(b\). Note that this is of course not commutative in general. If \(a\) and \(b\) are sequences of the same length \(l\), then \(\text{Min}(a, b)\) denotes the sequence of length \(l\) whose \(i\)th term is \(\text{Min}(a[i], b[i])\).

For \(k\) an integer, write \(d(k)\) for the dimension of \(S_k(\Gamma_0(N))\), write \(d_p(k)\) for the dimension of \(S_k(\Gamma_0(Np))\), and for \(\varepsilon\) a Dirichlet character of level \(p\), write \(d_{p, \varepsilon}(k)\) for the dimension of \(S_k(\Gamma_0(N)) \cap \Gamma_1(p); \varepsilon\). For \(n, r \geq 0\) we define \(\kappa(n, r)\) to be the constant sequence \([r, r, \ldots, r]\) of length \(n\). If \(v\) is a sequence of length \(l\) and \(e\) is an integer, we define \(v + e\) to be the sequence \([v[1] + e, v[2] + e, \ldots, v[l] + e]\) and \(v - e\) to be the sequence \([e - v[l], e - v[l-1], \ldots, e - v[1]]\) (note the reversal of order). If \(v\) is a sequence and \(0 \leq \delta \leq l(v)\), we define \(\sigma(v, \delta)\) to be the truncation \([v[1], v[2], \ldots, v[\delta]]\) of \(v\). More generally, if \(1 \leq \delta_1, \delta_2 \leq l(v)\), we define \(\sigma(v, \delta_1, \delta_2)\) to be \([v[\delta_1], v[\delta_1 + 1], \ldots, v[\delta_2]]\), where this is interpreted as the empty sequence if \(\delta_2 < \delta_1\). For \(\alpha \in \mathbb{Q}\), we write \([\alpha]\) for the largest integer which is at most \(\alpha\).

We begin by defining sequences \(t_2, t_4, \ldots\) of integers; note that later on we will define \(s_k\) to be \(t_k\) for \(k > 2\), and hence in particular for \(k > 2\) we are conjecturing that \(t_k\) is going to be the sequence of slopes at weight \(k\). We will define the first few \(t_k\) “by hand”, and then proceed recursively to define \(t_k\) for all positive even integers \(k\).

We set \(t_2 = \kappa(d_p(2) - d(2), 0)\). If \(p = 2\) then we set \(t_4 = t_2 \cup \kappa(d(4) - l(t_2), 1)\) and define \(k_{\text{min}} = 6\). If \(p > 2\) then for \(4 \leq k \leq p + 1\) even we set \(t_k = \kappa(d(k), 0)\) and set \(k_{\text{min}} = p + 3\). For \(p > 2\) what we are doing here is assuming that all slopes are 0 for all weights \(k \leq p + 3\), and in particular for all \(k \leq k_p\), which is our \(\Gamma_0(N)\)-regularity condition.

Now let us assume that \(k \geq k_{\text{min}}\) is even and that we have defined \(t_l\) for all even \(l < k\). We will now define \(t_k\). The definition depends on three parameters \(a, b\) and \(c\), defined thus. Let \(a\) be the unique element of \(\mathbb{Z}_{\geq 1}\) such that \(p^a < k - 1 - \leq p^{a+1}\). Let \(b\) be the unique integer with \(1 \leq b \leq p - 1\) such that \(p^ab < k - 1 - \leq p^a(b + 1)\). Set \(c = 1 + \left\lfloor \frac{(k - 2 - p^a b)}{p^a} \right\rfloor\). Then \(1 \leq c \leq p\). Also, let \(m\) be the number of cusps on \(X_0(N)\).

We will firstly define a sequence \(V\) which will be the “first few slopes” of \(t_k\). The algorithm used for the definition of \(V\) will depend on \(b\) and \(c\). More precisely, the definition of \(V\) will depend on which of the following cases we are in: \(b + c \leq p - 1, b < p - 1 < b + c\) or \(b = p - 1\). Note that if \(p = 2\) then the third case is the only one that can arise. We will attempt to give some explanation of what is happening...
at this point in the algorithm, in particular the motivation behind the definitions of the $k_1$, in the next section.

**Case 1:** $b+c \leq p-1$. We set $k_1 = k - b(p-1)p^{a-1}$ and $k_2 = k - b(p-1)p^{a-1} - 2(b+c-1)p^{a-1}$. We set $v_1 = t_{k_1}$ and $v_2 = t_{k_2}$. Define $B = p^a b + p^{a-1} (c-1) + 1$, set $e = k - B$ and let $\varepsilon$ denote $\chi^{B-1}$, where $\chi$ is any Dirichlet character of conductor $p$ and order $p-1$ (note that $p > 2$). Finally set $s = 1 + d_p, \varepsilon (1+\varepsilon)$.

If $l(v_1) \geq s - 1$ then we set $V_1 = \sigma(v_1, s - 1)$. Otherwise we set $V_1 = v_1 \cup (e - \sigma(v_2, s - 1 - l(v_1)))$. Finally, we set $V = V_1 \cup \kappa(m, e)$.

**Case 2:** $b < p - 1 < b + c$. We set $k_1 = k - ((b+1)p^{a-1}(p-1))$ and $k_2 = k - p^{a-1}(p-1)$. We set $v_1 = t_{k_1}$ and $v_2 = t_{k_2}$. We define $B = (b+1)p^{a-1}(p-1) + 1$ and set $e = k - B$. Finally, set $s = 1 + d_p(1+\varepsilon)$, let $s_2 = \lfloor (s-1)/2 \rfloor$ and let $e_2 = \lfloor e/2 \rfloor$.

If $l(v_1) \geq s - 1$ then we set $V_1 = \sigma(v_1, s - 1)$. If $s - 1 \leq 2l(v_1) < 2(s - 1)$ then we set $V_1 = v_1 \cup (e - \sigma(v_1, s - 1 - l(v_1)))$. If however $2l(v_1) < s - 1$ then define $w = \sigma(v_2, l(v_1) + 1, s_2)$, and our definition of $V_1$ depends on the parity of $s$. If $s$ is even then we set $V_1 = v_1 \cup w \cup [e_2] \cup (e - 1 - w) \cup (e - v_1)$ and if $s$ is odd then we set $V_1 = v_1 \cup w \cup (e - 1 - w) \cup (e - v_1)$. Note here that $[e_2]$ denotes the sequence with one element, $e_2$.

Finally, if $e = 1$ then we define $V = V_1 \cup \kappa(m - 1, 1)$ and otherwise we set $V = V_1 \cup \kappa(m, e)$.

**Case 3:** $b = p - 1$. This is the only case that occurs when $p = 2$. It is similar to case 2 but $w$ is slightly modified. We set $k_1 = k - p^a(p - 1)$ and $k_2 = k - p^{a-1}(p-1)$. We set $v_1 = t_{k_1}$ and $v_2 = t_{k_2}$. We set $B = p^a(p - 1) + 1$ and $e = k - B$. Next, set $s = 1 + d_p(1+\varepsilon)$, set $s_2 = \lfloor (s-1)/2 \rfloor$ and set $e_2 = \lfloor e/2 \rfloor$.

Again, if $l(v_1) \geq s - 1$ then we set $V_1 = \sigma(v_1, s - 1)$, and if $s - 1 \leq 2l(v_1) < 2(s - 1)$ then we set $V_1 = v_1 \cup (e - \sigma(v_1, s - 1 - l(v_1)))$. If however $2l(v_1) < s - 1$ then define $w_0 = \sigma(v_2, l(v_1) + 1, s_2)$ and set $w = \min(w_0 + 1, \kappa(l(w_0), e_2))$ (recall that this minimum is taken pointwise). Now we proceed as in case 2. If $s$ is even then we set $V_1 = v_1 \cup w \cup [e_2] \cup (e - 1 - w) \cup (e - v_1)$, and if $s$ is odd then we set $V_1 = v_1 \cup w \cup (e - 1 - w) \cup (e - v_1)$.

Finally, if $e = 1$ then we set $V = V_1 \cup \kappa(m - 1, 1)$ and otherwise we set $V = V_1 \cup \kappa(m, e)$.

We are finally ready to define $t_k$. If $l(V) \geq d(k)$ then we simply let $t_k$ be $\sigma(V, d(k))$.

Otherwise, we set $k_3 = 2B - k$, $v_3 = t_{k_3}$, and define $t_k = \sigma(V \cup (e + v_3), d(k))$.

This gives us an infinite sequence of sequences $t_2$, $t_4$, ..., The definition of $s_k$ is now simple: $s_k = t_k$ if $k > 2$, and $s_2 = \kappa(d(2), 0)$. Having now defined $s_k$ we remind the reader that the conjecture is that $s_k$ should be the slopes of $T_p$ on modular forms of level $N$.

Note that although the definition of $s_k$ is messy, it is elementary to implement on a computer, and in particular it is much easier to compute $s_k$ than to compute
actual slopes of modular forms. For example when $p = 2$ it takes under a second to compute $s_{1,000,000}$ (note that one does not need to compute $s_k$ for all $k < 1,000,000$ to compute $s_{1,000,000}$; indeed one only needs to compute 49 other $s_k$) but computing the characteristic polynomial of a matrix acting on weight 1,000,000 modular forms would be beyond modern computers.

We remark finally that the fact that $s_2$ differs from $t_2$ indicates that perhaps one should work with slopes of $U_p$ at level $Np$ rather than $T_p$ at level $N$.

3. Remarks on the conjecture

Although the conjecture made above has some interesting consequences (see the next section) and raises some related interesting questions, the author feels that the precise form of the conjecture itself is deeply unsatisfactory. The conjecture is basically saying that there is a very precise and unproven structure amongst slopes, but it seems to us that when this structure is discovered and proved, it will probably not prove the conjecture as it stands—it is much more likely to explain how the conjecture should have been formulated. F. Calegari and the author in fact have a much more readable form of the conjecture in the special case $p = 2$ and $N = 1$, and have proved several cases of it (see [3]).

There was a lot of motivation behind the recipe in the conjecture. The recipe was formulated by firstly assuming that the Gouvêa-Mazur conjectures were in fact much too weak, and seeing what kind of consequences this assumption had. We take the time here to explain a little about the motivation behind the details that we understand.

Let us consider, for example, the definition of $V_1$ in Case 1. What is going on is that $V_1$ should in fact be the vector of $U_p$-slopes on $S_{1+e}(\Gamma_0(N) \cap \Gamma_1(p); \varepsilon)$. The number $e$ is chosen so that $e + 1$ is congruent to $k$ modulo $p^{r-1}$. The power of the Teichmüller character chosen is to ensure that the weight-characters $x \mapsto x^k$ and $x \mapsto x^{1+e}(x)$ are in the same component of weight space. Hence one should expect small slopes in $S_{1+e}(\Gamma_0(N) \cap \Gamma_1(p); \varepsilon)$ and $S_k(\Gamma_0(Np))$ to be close, and we are predicting that many of them coincide. The $w_p$ operator will send a form of slope $s$ in $S_{1+e}(\Gamma_0(N) \cap \Gamma_1(p); \varepsilon)$ to a form of slope $e - s$ in $S_{1+e}(\Gamma_0(N) \cap \Gamma_1(p); \varepsilon)$. So to explain the higher slopes in $V_1$ we should look at small slopes in $S_{1+e}(\Gamma_0(N) \cap \Gamma_1(p); \varepsilon^{-1})$. The weight-character corresponding to these forms is close to $x \mapsto x^r$ if $r$ is an integer which is congruent to $k + 2 - 2B$ modulo $(p - 1)$ and to $1 + e$ modulo a high power of $p$. The integer $k_2$ has this property, because $B$ is congruent to $b + c$ modulo $p - 1$. This then completely explains the motivation behind the definition of $V_1$ in case 1.

As another example, we explain the motivation for the final $e + v_3$. Let us assume that we are in case 3. If $f$ is an eigenform of weight $k_3$ and slope $\sigma$ then the Hodge-Tate weights of the associated Galois representation are 0 and $k_3 - 1$. Tate twisting this representation $e$ times gives a Galois representation with Hodge-Tate weights
QUESTIONS ABOUT SLOPES

\[ e = k - 1 - p^a(p - 1) \] and \[ k_3 - 1 + e = p^a(p - 1) \]. Hence \( p \)-adically these Hodge-Tate weights are close to \( k - 1 \) and 0 respectively, and the conjecture is predicting that there is a modular Galois representation which does have Hodge-Tate weights 0 and \( k - 1 \) and which is highly congruent to this Tate twist. The associated modular form will have slope \( e + \sigma \) and will presumably be highly congruent to the \( p \)-adic modular form \( \vartheta^e f \) of weight \( k_3 + 2e = k \).

It is a pleasant exercise, if one really wants to understand the nuts and bolts of the conjecture, to try and explain all the combinatorics involved in this way. However there is one step that the author cannot explain in this conceptual manner, and that is the construction of \( w \) in the middle of case 3. The fact that one sometimes has to add precisely 1 to an entry of \( w_0 \) seems to say geometrically that the eigencurve looks less “flat” near a \( p \)-newform, but is varying in a very precise way. What seems to be happening is that families of overconvergent eigenforms that do not contain any classical \( p \)-newforms seem to be a lot flatter in general than families containing newforms. Here we use the word flat in a non-technical sense, to mean that the slope tends to vary a lot less as one moves through the family.

4. Consequences of the conjecture and related questions

In this last section we raise some consequences and probable consequences of the conjecture, and the thoughts behind it. We start by emphasizing that we strongly believe that \( \Gamma_0(N) \)-regularity is a red herring, and that there should be a recipe which gives either the valuations of the eigenvalues of \( T_p \) at level \( N \), or the slopes of \( U_p \) at level \( Np \), in all cases. This recipe should take as input the slopes at weight \( k \leq k_p \), or perhaps the slopes at weight \( k \leq p+1 \). However, if \( p \) is not \( \Gamma_0(N) \)-regular the situation is more complicated. As an example of why it is more complicated we present the first consequence of our conjecture:

**Consequence 4.1.** — If \( p \) is \( \Gamma_0(N) \)-regular, and Conjecture 2.1 is true, then for any \( k \geq 2 \) the eigenvalues of \( T_p \) on \( S_k(\Gamma_0(N)) \) are all non-zero, and the valuation of any such eigenvalue is an integer.

Wan has asked whether in the general case the denominators of the valuations are bounded by a constant depending on \( N \) and \( p \), but independent of \( k \). One may ask a stronger question (recall that \( k_2 = 4 \) and \( k_p = \frac{p^{2a}}{2} \) for \( p > 2 \), and also that \( p \) is prime to \( N \)):

**Question 4.2.** — Let \( M \) be the lowest common multiple of the denominators of the slopes of \( U_p \) on forms of level \( Np \) and weight \( k \), with \( 2 \leq k \leq k_p \). Does the denominator of any slope at level \( Np \) at any weight divide \( M \)?

Related to these questions are questions about fields of definitions of modular forms. Let \( f_k \) denote the characteristic polynomial of \( T_2 \) acting on the space \( S_k(\text{SL}_2(\mathbb{Z})) \).

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
Maeda has conjectured that this polynomial is always irreducible over $\mathbb{Q}$, and various authors have checked both this statement and the stronger statement that the Galois group of its splitting field is the full symmetric group. For example, the author checked this for all $k \leq 2048$. On the other hand, if one looks at the factorization of $f_k$ over $\mathbb{Q}_2$ for small values of $k$, one cannot help but notice that the irreducible factors are always linear or quadratic. This is related to the fact that our conjectures are frequently forcing slopes to be spread out, making it more difficult for them to repeat. We remark that the corresponding extensions of $\mathbb{Q}_2$ are sometimes ramified, even though we are conjecturing that the valuations are always integral. This raises the specific question

**Question 4.3.** — Let $f$ be a normalised eigenform of level 1. Does the extension of $\mathbb{Q}_2$ generated by the coefficients of $f$ always have degree at most 2 over $\mathbb{Q}_2$?

More generally, we have

**Question 4.4.** — Let $N$ be a positive integer, and let $p$ be a prime not dividing $N$. Is there a bound $B = B(N, p)$ such that if $f$ is any normalised eigenform of level $N$, then the coefficients of $f$ generate an extension of $\mathbb{Q}_p$ of degree at most $B$? Equivalently, is there a subfield $F \subset \mathbb{Q}_p$, finite over $\mathbb{Q}_p$, and depending only on $N$ and $p$, such that any normalised eigenform $f \in S_k(\Gamma_0(N); \mathbb{Q}_p)$ has $q$-expansion in $F[[q]]$?

One might even consider the case where $p$ divides $N$ but we have not done any computations at all in this case.

A remark related to these questions: it is a recent theorem of Kilford (see [13]) that if $f$ is a normalised eigenform of level $\Gamma_1(4)$ and any odd weight, the coefficients of $f$ necessarily lie in $\mathbb{Q}_2$. This result could have been noticed over 50 years ago, and the author finds it interesting that it was proved before it was conjectured. This might reflect on the current ease with which one can compute spaces of forms, thanks to Stein. Kilford’s proof relies strongly on Coleman’s theory of overconvergent modular forms, and explicit computations of matrices related to the $U_2$ operator. Note added in proof: these results have now been generalised to level $\Gamma_1(2^n)$ in [5].

Motivated by the Fontaine-Mazur conjecture, one can move completely away from the theory of modular forms. If $\rho : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\mathbb{Q}_p)$ is continuous, irreducible, unramified at a finite set of primes, and crystalline at $p$, then it has a conductor $N(\rho)$, which is a positive integer prime to $p$. As before, let $N$ be any positive integer and let $p$ be a prime not dividing $N$.

**Question 4.5.** — Is there a subfield $F \subset \overline{\mathbb{Q}}_p$, finite over $\mathbb{Q}_p$ and depending only on $N$ and $p$, such that if $\rho : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(\overline{\mathbb{Q}}_p)$ is irreducible, crystalline at $p$, and has conductor $N$, then the trace of $\rho(g)$ lies in $F$ for all $g \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$?

One could relax the crystalline condition to a potentially semi-stable one, and let $p$ divide $N$, for an even stronger question—here one has to define the $p$-part of
the conductor of such a representation using Fontaine theory. One could even ask $n$-dimensional analogues of this question, but we shall leave this to the optimistic reader.

We now move onto a rather amusing consequence of Conjecture 2.1. The dependence of the recipe in the conjecture on $N$ is only via the dimension of various spaces of cusp forms of level $N$ and $Np$. There are cases where these dimensions happen to coincide for different $N$. For example, $\dim(S_k(\Gamma_0(6))) = \dim(S_k(\Gamma_0(8)))$ for all $N$, as can be seen from classical formulae for these dimensions. In these cases, the recipe might produce the same results for different $N$. For example, $p = 5$ is both $\Gamma_0(6)$-regular and $\Gamma_0(8)$-regular, and as a consequence one gets the following rather strange result:

**Consequence 4.6.** — If Conjecture 2.1 is true, then the $5$-adic valuations of the eigenvalues of $T_5$ on $S_k(\Gamma_0(6))$ coincide, with multiplicities, with the $5$-adic valuations of $T_5$ on $S_k(\Gamma_0(8))$. Similarly the slopes of $U_5$ on $S_k(\Gamma_0(30))$ coincide with the slopes of $U_5$ on $S_k(\Gamma_0(40))$.

The author has checked the above consequence in MAGMA for $k \leq 60$. The reader who knows about Coleman’s results and the overconvergent theory will realise that as another consequence of the conjecture, the overconvergent (finite) slopes of $U_5$ at tame levels 6 and 8 must coincide for any weight-character in the closure of $\mathbf{Z}$. This result is surely not explained by a morphism between the two eigencurves, and the author has no idea of a more conceptual explanation of this phenomenon. Perhaps it is just a numerical coincidence. Even more unnerving is that it is very easy to find many more examples where coincidences at small weight imply equalities at all weights. The author does not know, unfortunately, of an example where the set of slopes coming from two levels are the same at all small weights but where $p$ is not regular (regularity in the sense of this paper).

We next raise some combinatorial problems, which can presumably be attacked using only elementary techniques, and are almost certainly accessible.

**Question 4.7.** — Is the conjecture well-defined, in the sense that every time a sequence is implicitly assumed to have at least a given length, it does have this length?

**Question 4.8.** — The sequences $v_k$ are by definition increasing. Are the sequences $s_k$ produced by the conjecture always increasing?

One baulks at the combinatorics behind these questions, although they are surely both accessible. We believe that the answers are affirmative, in both cases, but have only checked the details in the case $p = 2$ and $N = 1$.

**Question 4.9.** — Does Conjecture 2.1 imply that if $p$ is $\Gamma_0(N)$-regular then the valuation of any eigenvalue of $T_p$ on $S_k(\Gamma_0(N))$ is at most $\frac{k-1}{p+1}$?
This phenomenon, that slopes tend to be very small, was explicitly noted by Gouvêa. The author again convinced himself that the conjecture did indeed imply that all slopes were at most $\frac{k-1}{3}$ in the case $p = 2$ and $N = 1$.

Gouvêa also considered the following: if one divides the sequence of slopes of $U_p$ at weight $k$ by a factor of $k - 1$, one gets a sequence of rationals in the closed interval $[0, 1]$, and this sequence can be thought of as giving a (finite) probability measure on this closed interval.

**Question 4.10.** — Does Conjecture 2.1 imply that, as $k$ increases, the measures tend to a limit, and if so then what is this limit?

Numerical experiments with $p = 2$ and $N = 1$ suggest to the author that in this case measures were tending to a limit, which gave the point $\frac{1}{2}$ a mass of $\frac{1}{3}$, and which distributed the remaining mass of $\frac{2}{3}$ uniformly on $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. This points to a natural conjecture in the general case.

There remains the very natural

**Question 4.11.** — Does Conjecture 2.1 imply the Gouvea-Mazur conjectures in the $\Gamma_0(N)$-regular case?

Again, the author convinced himself that this was the case when $p = 2$ and $N = 1$. We will not trouble the reader with the excruciating details.

**References**


QUESTIONS ABOUT SLOPES


K. Buzzard, Department of mathematics, Imperial College, 180 Queen’s Gate, London, SW7 2AZ, UK • E-mail: buzzard@imperial.ac.uk
Url: http://www.ma.imperial.ac.uk/~buzzard/
THE LOCAL LANGLANDS CORRESPONDENCE:
NOTES OF (HALF) A COURSE AT THE IHP SPRING 2000

by

Michael Harris

Abstract. — The article provides a reasonably self-contained account of the main results of the author’s book with Richard Taylor, containing a description of the Galois representations obtained in the cohomology of certain Shimura varieties attached to unitary groups, and obtaining the local Langlands conjecture for $GL(n)$ of $p$-adic fields as a consequence. The main steps in the proof of the local Langlands conjecture are presented, and in some cases simplified. The bulk of the paper concerns the geometry of the Shimura variety at places of bad reduction, where good local models are nevertheless available, and the description of points in the special fiber in the manner of Langlands and Kottwitz. The article concludes with a section describing possible extensions of these results to other Shimura varieties, and an account of some of the work of Laurent Fargues along those lines.

Résumé (La correspondance de Langlands locale). — L’article contient une description assez complète des principaux résultats du livre de l’auteur avec Richard Taylor, qui décrit les représentations galoisiennes réalisées dans la cohomologie de certaines variétés de Shimura associées aux groupes unitaires, et qui obtient la conjecture locale de Langlands pour $GL(n)$ d’un corps $p$-adique comme conséquence. Les principales étapes de la démonstration de la conjecture locale de Langlands y sont présentées, parfois simplifiées. Le gros de l’article concerne la géométrie de la variété de Shimura aux places de mauvaise réduction, où l’on dispose néanmoins de bons modèles locaux, et la description des points dans la fibre spéciale à la manière de Langlands et Kottwitz. La dernière section de l’article décrit les extensions éventuelles de ces résultats aux variétés de Shimura plus générales, ainsi qu’un compte rendu des travaux de Laurent Fargues sur ces questions.

2000 Mathematics Subject Classification. — Primary 11G18; Secondary 11F70, 11S37, 14G35, 14L05, 11F72.

Key words and phrases. — Local Langlands conjecture, Shimura variety, automorphic form, $p$-divisible group, trace formula.

Introduction

The present notes cover 50% of the material presented in a course given jointly with Guy Henniart during the special semester “Formes Automorphes”, held at the Institut Henri Poincaré in Paris between February and June 2000, as well as a little more material I didn’t have time to present. The purpose of the course was to explain two proofs of the local Langlands conjecture for \( p \)-adic fields, due respectively to Richard Taylor and myself [HT], and to Henniart [He5]. My lectures were naturally concerned with [HT], the main burden of which is to construct a candidate for a local Langlands correspondence, and to prove that this putative correspondence is (nearly) compatible with the global correspondence realized on the cohomology of certain specific Shimura varieties. The techniques applied derive mainly from arithmetic algebraic geometry: we study the bad reduction of the Shimura varieties in question by interpreting them locally/infinitesimally as formal deformation spaces for \( p \)-divisible groups with additional structure of a kind already studied by Drinfel’d. This yields a stratification of the special fiber, with particularly nice properties, in terms of \( p \)-rank of the universal \( p \)-divisible group. The cohomology of the Shimura varieties is then calculated by means of vanishing cycles on the bad special fiber. Thanks to Berkovich’s work on étale cohomology of (rigid) analytic spaces, the vanishing cycles can be computed infinitesimally, which permits determination of their stalks in terms of certain universal representation spaces. An extension, to our situation of bad reduction, of the trace formula techniques perfected by Langlands and Kottwitz for calculating zeta functions of Shimura varieties at places of good reduction, provides the necessary compatibility of local and global correspondences.

My goal in the course was to present a self-contained account of the main results of [HT]. In so doing, I chose to sacrifice the description of the global structure of the strata in the special fiber, and of the vanishing cycles sheaves on the strata, in the first place because this would have been impossible in the eight three-hour sessions available, but also because no such description seemed likely to be available for other Shimura varieties.\(^{(1)}\) My presentation therefore diverged from that of [HT], in that I studied the vanishing cycles by means of formal completions along points in the special fiber, following the approach of Rapoport and Zink in [RZ], rather than along the strata. This was nearly successful: the geometric material was covered in detail, but I ran out of time and was unable to do justice to the detailed comparison of trace formulas. This was just as well, because I did not find a satisfactory approach to the latter – an approach likely to extend to other groups – until long after the semester had ended and all the visitors had gone home.\(^{(2)}\) This is the approach presented in the present notes.

\(^{(1)}\)In the meantime, Elena Mantovan’s Harvard Ph.D. thesis [Ma] has revealed this expectation to be unduly pessimistic.

\(^{(2)}\)To be honest, talking to the visitors was much more interesting than perfecting the final stages of the argument.
We introduce the notation that will be used throughout these notes. Let $p$ be a rational prime number. For any finite extension $K$ of $\mathbb{Q}_p$ and any positive integer $n$, we let $\mathcal{A}(n, K)$ denote the set of equivalence classes of irreducible admissible representations of $GL(n, K)$, $\mathcal{A}_0(n, K)$ the subset of supercuspidal representations. Let $\mathcal{G}(n, K)$ denote the set of equivalence classes of $n$-dimensional complex representations of the Weil-Deligne group $WD_K$ on which Frobenius acts semisimply, $\mathcal{G}_0(n, K)$ the subset of irreducible representations. We will frequently write $G_n$ for $GL(n)$.

A local Langlands correspondence for $p$-adic fields is the following collection of data:

1. For every $p$-adic field and integer $n \geq 1$, a bijection $\pi \mapsto \sigma(\pi)$ between $\mathcal{A}(n, K)$ and $\mathcal{G}(n, K)$ that identifies $\mathcal{A}_0(n, K)$ with $\mathcal{G}_0(n, K)$.

2. Let $\chi$ be a character of $K^\times$, which we identify with a character of $WD_K$ via the reciprocity isomorphism of local class field theory. Then $\sigma(\pi \otimes \chi \circ \det) = \sigma(\pi) \otimes \chi$.

3. If $\pi \in \mathcal{A}(n, K)$ with central character $\xi_\pi \in \mathcal{A}(1, K)$, then $\xi_\pi = \det(\sigma(\pi))$.

4. $\sigma(\pi^\vee) = \sigma(\pi)^\vee$, where $^\vee$ denotes contragredient.

5. Let $\alpha : K \to K_1$ be an isomorphism of local fields. Then $\alpha$ induces bijections $\mathcal{A}(n, K) \to \mathcal{A}(n, K_1)$ and $\mathcal{G}(n, K) \to \mathcal{G}(n, K_1)$ for all $n$, and we have $\sigma(\alpha(\pi)) = \alpha(\sigma(\pi))$. In particular, if $K$ is a Galois extension of a subfield $K_0$, then the bijection $\sigma$ respects the $\text{Gal}(K/K_0)$-actions on both sides.

6. Let $K'/K$ denote a cyclic extension of prime degree $d$. Let $BC : \mathcal{A}(n, K) \to \mathcal{A}(n, K')$ and $AI : \mathcal{A}(n, K') \to \mathcal{A}(nd, K')$ denote the local base change and automorphic induction maps [AC, HH]. Let $\pi \in \mathcal{A}(n, K)$, $\pi' \in \mathcal{A}(n, K')$. Then

\begin{align*}
\sigma(BC(\pi)) &= \sigma(\pi)|_{WD_{K'}} , \\
\sigma(AI(\pi')) &= \text{Ind}_{K'/K} \sigma(\pi'),
\end{align*}

where $\text{Ind}_{K'/K}$ denotes induction from $WD_{K'}$ to $WD_K$.

Let $n$ and $m$ be positive integers, $\pi \in \mathcal{A}(n, K)$, $\pi' \in \mathcal{A}(m, K)$. Then

- $L(s, \pi \otimes \pi') = L(s, \sigma(\pi) \otimes \sigma(\pi'))$.

- For any additive character $\psi$ of $K$, $\varepsilon(s, \pi \otimes \pi', \psi) = \varepsilon(s, \sigma(\pi) \otimes \sigma(\pi'), \psi)$.

Here the terms on the left of (0.7) and (0.8) are as in [JPSS, Sh] and are compatible with the global functional equation for Rankin-Selberg $L$-functions. The right-hand terms are given by Artin and Weil (for (0.7)) and Langlands and Deligne (for (0.8)) and are compatible with the functional equation of $L$-functions of representations of the global Weil group. In particular both sides have Artin conductors and (0.8) implies that $a(\sigma(\pi)) = a(\pi)$.

The local Langlands conjecture, established in [HT] and in [He5], is the assertion that a local Langlands correspondence exists. The existence of some family of bijections $\mathcal{A}(n, K) \leftrightarrow \mathcal{G}(n, K)$, identifying $\mathcal{A}_0(n, K)$ with $\mathcal{G}_0(n, K)$, preserving conductors and satisfying weakened versions of properties (0.2)-(0.5), had been proved by Henniart a number of years before [He2]. Henniart’s main tools are a counting
argument for local fields of positive characteristic, based on Laumon’s theory of the \(\ell\)-adic Fourier transform (the subsets of \(A_0(n, K)\) and \(G_0(n, K)\) with fixed conductor are finite) and an “approximation” of local fields of characteristic zero by local fields of positive characteristic. The properties established in [He2] do not suffice to characterize the correspondence uniquely. However, another theorem of Henniart ([He4]; cf. (A.2.5), below) guarantees that properties (0.1)-(0.8) do suffice to determine a unique correspondence. Nevertheless, the “numerical local Langlands correspondence” of [He2] is a necessary ingredient of all proofs to date of the local Langlands correspondence in mixed characteristic. In the present notes, it is invoked in (5.3). (3)

The notes are divided into eight more or less fictitious lectures, following my original plan which proved too ambitious; even the first seven lectures did not fit in the time allotted. The first lecture covers the arguments common to [HT] and [He5]: the construction of special families of cohomological automorphic representations of \(GL(n)\) of CM fields, corresponding to certain cases of non-Galois automorphic induction of Hecke characters. These arguments are mostly taken from [H2], which uses these special automorphic representations to reduce the local Langlands conjecture – more precisely, property (0.8), the others being established by geometric means – to the local/global compatibility, asserted as Main Theorem 1.3.6.

The next three lectures present an attenuated version of the geometric part of [HT]. The main object of these notes is the Shimura variety attached to the unitary (similitude) group \(G\) of a division algebra of dimension \(n^2\) over a CM field \(F\), with involution of the second kind fixing the real subfield \(F^+\) of \(F\). As complex analytic varieties, they are compact quotients of the unit ball of dimension \(n-1\). Lecture 2 introduces these Shimura varieties as moduli spaces of abelian varieties with PEL type. Their regular integral models in ramified level, over a \(p\)-adic place \(w\) of \(F\) split over \(F^+\), are defined by means of Drinfel’d bases. The main properties of the latter are recalled in Lecture 3, which also carries out the thankless task of explaining how Hecke operators act on Drinfel’d bases. The stratification by \(p\)-rank of the special fiber at a split place is defined in Lecture 4: it is shown that there is one stratum, a union of locally closed smooth subvarieties, in each dimension \(h = 0, 1, \ldots, n-1\). Infinitesimal uniformization, as in [RZ], is then combined with the results of Berkovich to show that the stalks of the vanishing cycles sheaves are constant along strata, and are isomorphic on the \(h\)-dimensional stratum to a standard space \(\Phi_{n-h, w}\) with canonical action of \(GL(n-h, F_w) \times J_{n-h, w}\), where \(J_{n-h, w}\) is a specific anisotropic inner form of \(GL(n-h)\) over \(F_w\).

(3) Of course the local Langlands conjecture is originally due to Langlands! The form presented here became standard after it was understood that conditions (0.7) and (0.8) for \(m = 1\) do not suffice to characterize the correspondence.

(4) In his IH\(P\) lectures, Henniart replaced the counting argument in positive characteristic by a reference to Lafforgue’s theorem which establishes the global Langlands correspondence for function fields, with the local Langlands correspondence in positive characteristic as a corollary. The original proof [LRS] of the local Langlands correspondence in positive characteristic used [He2].
In Lecture 4, the course begins to diverge from [HT]. In the first place, we work directly with the strata, rather than with the Igusa varieties of the first kind of [HT]. These are modular varieties defined (in characteristic $p$) independently of the Shimura variety. The Igusa variety of the first kind is isomorphic as ringed space to the stratum, but not as a scheme over the (finite) base field. More importantly, we do not introduce the Igusa varieties of the second kind. These are pro-étale covers of the Igusa varieties of the first kind – for the $h$-dimensional stratum, the covering group is the maximal compact subgroup of $J_{n-h}$ – and their existence is combined in [HT] with a theorem of Berkovich to prove that the vanishing cycle sheaves are locally constant along the strata. Igusa varieties were first defined for the special fibers of integral models of elliptic modular curves, and were studied in detail in the book of Katz and Mazur [KM]. Their properties have been at the heart of many of the most important developments of arithmetic algebraic geometry of the last 30 years. It is likely that the more general Igusa varieties described in [HT], and their generalizations constructed by Mantovan [Ma], will also find applications to arithmetic. However, for applications to automorphic forms (with coefficients in characteristic zero!) the infinitesimal structure at points in the special fiber appears to suffice.

The space $\Phi_{n-h}$ is the “fundamental local representation,” which also carries an action of the Weil group of $F_w$. It is universal in the sense that $\Phi_g$ occurs as the stalk of the vanishing cycles along the codimension $g - 1$-stratum for any of the Shimura varieties we study. In Lecture 5, we use a comparison of trace formulas to prove a conjecture of Carayol, showing that, for $h = 0$, $\Phi_n$ simultaneously realizes the Jacquet-Langlands correspondence between representations of $J_n$ and the discrete series of $GL(n, F_w)$ and a bijection $\mathcal{A}_0(n, F_w) \leftrightarrow \mathcal{G}_0(n, F_w)$ that satisfies properties (0.1)-(0.7). Here again we depart, slightly, from [HT]. In [HT], comparisons of trace formulas are established for all strata simultaneously, and in each case the comparison is between a Lefschetz trace formula for the action of Hecke operators on the special fiber and Arthur-Selberg trace formula, in its cohomological version [A], for the action of Hecke operators on the cohomology of the generic fiber. This comparison is carried out in Lectures 6 and 7, where it is called the Second Basic Identity. However, an alternative comparison is available for the minimal (0-dimensional) stratum, one that provides slightly stronger information for supercuspidal representations. Indeed, one can use the infinitesimal uniformization to derive Carayol’s conjecture from a comparison of the trace formula for $G$ with that of an inner form attached to the (unique) isogeny class contributing to the minimal stratum. Such an argument was already used in [H1], in the setting of $p$-adic uniformization of the generic fiber, where it took the form of a Hochschild-Serre spectral sequence for rigid étale cohomology, since vastly generalized in the thesis of Laurent Fargues [Fa]. A more immediate precursor is to be found in the thesis of P. Boyer [Bo], which also contributed the fundamental observation, used here and in [HT], that the cohomology of the strata of positive
dimension is a sum of induced $GL(n, F_w)$-modules, hence has no intertwining with the supercuspidal part of the cohomology. However, the simplifications obtained in this way (arising from the degeneration of the supercuspidal part of the vanishing cycles spectral sequence (5.1.3) and from Clozel’s purity lemma, cf. (5.1.6)) are not strictly necessary; the trace identities and dévissage suffice. Indeed, [HT] treats the more general case, not considered here, of discrete series representations.

As mentioned above, the Second Basic Identity is stated and proved in Lectures 6 and 7. But first it is shown that the Second Basic Identity, combined with the First Basic Identity – a summary of the geometric information contained in Lectures 2-4 – suffices to prove Main Theorem 1.3.6. The strategy used in [HT] to prove the Second Basic Identity roughly follows Kottwitz’ approach in [K5] to the zeta functions of Shimura varieties. One uses a version of Honda-Tate theory adapted to PEL types to “count” the points in the special fiber in a rough way, then one applies techniques from Galois cohomology to rewrite the result of this “count” in a form suited to comparison with the cohomological trace formula. However, our approach in [HT] differs from that of Kottwitz in three particulars. First, and most obviously, Kottwitz only considers the case of good reduction (hyperspecial level), which give rise to unramified local Galois representations, whereas the point of [HT] is to study ramification. Thus [HT] considers the cohomology of individual strata, rather than the full special fiber, with coefficients given by the vanishing cycle sheaves. Next, Kottwitz counts fixed points of Hecke correspondences, twisted by powers of Frobenius, over finite fields, and obtains formulas in terms of twisted orbital integrals. These fixed point formulas are then interpreted as traces in $\ell$-adic cohomology of the special fiber by means of Grothendieck’s version of the Lefschetz trace formula. In [HT] we also use an $\ell$-adic Lefschetz trace formula, specifically the one proved by Fujiwara [F], designed to apply to non-proper varieties such as the strata of our Shimura varieties. However, instead of counting points over finite fields we count fixed points of Hecke correspondences over the algebraic closure of the residue field of $F_w$ – on a fixed stratum, a sufficiently regular Hecke correspondence already incorporates a twist by a power of Frobenius – and obtain formulas in terms of orbital integrals involving an inner twist of a Levi subgroup of $GL(n, F_w)$. Finally, Kottwitz’ formalism leads to an expression of the result of the point count as a sum over rational conjugacy classes in $G$ modulo stable conjugacy, an expression well-adapted for comparison with the stable trace formula. The formalism in [HT] leads naturally to an expression as a sum of rational conjugacy classes in $G$ modulo adelic conjugacy, adequate for application to local questions, at least for inner forms of $GL(n)$ where the problem of local instability does not arise.

The present version of the counting argument of [HT] features several technical simplifications, mainly in the treatment of inertial equivalence. The formulas in [HT] are complicated by the need to take into account the reducibility of the restriction of an irreducible representation of $J_{n-h}$ to its maximal compact subgroup. The present
account avoids these complications by exploiting invariance properties of the fundamental local representation (cf. Proposition 5.5.9). This approach also eliminates the need for an intermediate expression of the point count in terms of orbital integrals on $J_{n-h}$.

Lecture 8, for which there was no time at the IHP, contains some new material. The article [H3] outlines a possible extension of some of the techniques and results of [HT] to general Shimura varieties. Since it is not known how to generalize Drinfeld’s bases, nor even whether such a generalization is possible, it is proposed in [H3] to work directly on the rigid analytic space associated to the Shimura variety in characteristic zero, decomposing it into rigid analytic subspaces according to a stratification of the special fiber in minimal (hyperspecial) level by isocrystal type. For Shimura varieties of PEL type, L. Fargues has carried out much of this program and more in his thesis [Fa]. As mentioned above, he has constructed a Hochschild-Serre spectral sequence, as in [H1], to determine the cohomological contribution of an isogeny class, and proved, as in [Bo] and [HT], that only the basic isogeny class intertwines non-trivially with the supercuspidal representations. Lecture 8 proves the assertions stated without proof in [H3] and provides an introduction to Fargues’ results.

Rather than provide complete proofs – one can find these in [HT] – the present notes aim to provide some understanding of the techniques used in [HT]. Generally speaking, when concepts give way to calculation, I have preferred to cut short the discussion and refer to [HT] or to the literature. Exceptions are made where the approach followed here diverges from that of [HT]: in such instances, I have tried to give enough details to convince the reader that the present approach is correct, or at least has avoided obvious pitfalls! On the other hand, I have included material not in [HT] that seemed appropriate at the time of the course. In particular, §2 and §3 contain a brief review of the deformation theory of one-dimensional formal $O$-modules, following Lubin-Tate and Drinfeld’s.

It remains to thank the audience at the IHP for having put up with my many blunders; Guy Henniart for planning the course with me (and making no blunders whatsoever); Ariane Mézard and especially Laurent Fargues for having read and pointed out some of the errors in earlier drafts (all copies of which should be immediately destroyed!); and the fellow organizers of the automorphic semester – Henri Carayol, Jacques Tilouine, and Marie-France Vignéras – as well as the directors of the IHP, Joseph Oesterlé and Michel Brion, and especially Annie Touchant of the Centre Emile Borel, for having made the semester an unqualified success. Finally, I am deeply grateful to the referees for their meticulous reading of the manuscript.

(5) The notes distributed during the semester, on which the present text is based, occasionally referred to the courses Clozel and Labesse gave during the automorphic semester. Although the notes of their courses are not being published, some of these references have been retained, as have other references to the time and place of my own lectures, as a reminder of the original context.
1. Galois representations attached to automorphic representations of $GL(n)$

1.1. Cohomological, conjugate self-dual representations. — Fix a prime $p$. Let $E$ be an imaginary quadratic field in which $p$ splits, $F^+$ a totally real field of degree $d$, $F = E \cdot F^+$. Complex conjugation is denoted $c$. Choose a distinguished complex embedding $\tau_0$ of $F$, and let $\Sigma$ denote the set of complex embeddings of $F$ with the same restriction to $E$ as $\tau_0$. This $\Sigma$ is a CM type, and is in bijection with the set of real embeddings of $F^+$. We consider automorphic representations $\Pi$ of $GL(n, F)$, or more precisely of $GL(n, \mathbb{A}_F)$. Any such representation can be factored $\Pi = \Pi_\infty \otimes \Pi_f$, where $\Pi_f$ is an admissible irreducible representation of $GL(n, \mathbb{A}_{F,f})$ and $\Pi_\infty$ is a Harish-Chandra module for $GL(n, \mathbb{C})^\Sigma$; i.e., an admissible irreducible $(\mathfrak{g}(n, \mathbb{C})^d, U(n)^d)$-module.

We will only be concerned with $\Pi$ such that $\Pi_\infty$ is cohomological. We will also restrict attention to cuspidal $\Pi$, though general discrete cohomological $\Pi$ also play a role in the more detailed results of [HT]. Then $\Pi$ is generic, by Shalika’s theorem. Let $(\xi, W_\xi)$ denote a finite-dimensional irreducible representation of $GL(n, F)$. This is equivalent to giving a pair of finite-dimensional irreducible representations $(\xi_\sigma, \xi_\sigma^\vee)$ of $GL(n, \mathbb{C})$ for each $\sigma \in \Sigma$. For any representation $\tau$, we let $\tau^\vee$ denote its contragredient.

(1.1.1) Fact. — For every irreducible finite-dimensional representation $(\Xi, W_\Xi) = (\xi \otimes \xi_c, W_\xi \otimes W_{\xi_c})$ of $GL(n, \mathbb{C}) \times GL(n, \mathbb{C})$ such that

$$\xi_c \sim \xi^\vee$$

there is a unique generic $(\mathfrak{g}(n, \mathbb{C}), U(n))$-module $\Pi_\Xi$ such that the relative Lie algebra cohomology $H^*(\mathfrak{g}(n, \mathbb{C}), U(n); \Pi_\Xi \otimes W_\Xi)$ is non-trivial. Moreover, $\Pi_\Xi \circ c \sim \Pi_\Xi^\vee$.

The above fact is a special case of the construction in [C2, 3.5], which covers nearly all generic cohomological $(\mathfrak{g}(n, \mathbb{C}), U(n))$-modules. It will suffice for the purposes of the present notes. The relative Lie algebra cohomology is relevant to calculating the cohomology of the adelic locally symmetric space attached to $GL(n, F)$, via Matsushima’s formula. Here and in what follows, we denote by $\mathcal{A}_0(G)$ the set of cuspidal automorphic representations of a reductive algebraic group $G$.

(1.1.3) Matsushima’s formula. — Let $G$ be a reductive algebraic group over $\mathbb{Q}$ (e.g., $GL(n, F)$, via restriction of scalars), and $(\Xi, W_\Xi)$ a finite-dimensional algebraic representation of $G$. For any open compact subgroup $K$ of $G(\mathbb{A}_f)$, let

$$\mathcal{M}_K(G) = G(\mathbb{Q}) \backslash G(\mathbb{A}) / Z_G(\mathbb{R}) \cdot K_\infty \cdot K,$$

where $K$ runs over open compact subgroups of $G(\mathbb{A}_f)$. Let $\mathcal{L}_\Xi$ be the local system

$$G(\mathbb{Q}) \backslash G(\mathbb{A}) \times W_\Xi / Z_G(\mathbb{R}) \cdot K_\infty \cdot K$$

over $\mathcal{M}_K(G)$ (this is a local system provided the central character of $\Xi$ is trivial on the Zariski closure of a sufficiently small congruence subgroup of the global units, which
we assume to be the case). Then there is a $G(A_f)$-equivariant subspace $H^\ast_{\text{cusp}}(\mathcal{L}_\Xi)$ of $\lim_{\rightarrow K} H^\ast(M_K(G), \mathcal{L}_\Xi)$ such that

$$H^\ast_{\text{cusp}}(\mathcal{L}_\Xi) = \bigoplus_H H^\ast(\mathfrak{g}, Z_G(\mathbb{R}) \cdot K_\infty ; \Pi_\infty \otimes W_\Xi) \otimes \Pi_f$$

as $G(A_f)$-modules, where $\Pi$ runs through $\mathcal{A}_0(G)$.

(1.1.4) — When $G = GL(n)_F$, we assume $\Xi = \otimes \Xi_\sigma$, where each $\Xi_\sigma$ satisfies condition (1.1.2); recall that $\Sigma$ indexes embeddings of $F^+$.

Fact (1.1.1) shows that the sum runs over $\Pi$ such that

$$\Pi_\infty \simeq \Pi_\Xi := \otimes \Xi_\sigma,$$

and this implies that $\Pi_\infty \circ c \sim \Pi'_\infty$. We also make an analogous global restriction:

(1.1.5) $$\Pi' \sim \Pi'.'$$

This is necessary in order to attach compatible families of $\ell$-adic representations to $\Pi$, following Clozel’s construction.

1.2. Fake unitary (similitude) groups, descent and base-change. — The relation to cohomology of the symmetric spaces attached to $GL(n)_F$ plays no role in what follows. The Galois representations are instead constructed on the $\ell$-adic cohomology of Shimura varieties attached to certain unitary groups. This is the next theme.

Let $B$ be a central division algebra of dimension $n^2$ above $F$, and let $t_B : B \to F$ and $n_B : B \to F$ denote the reduced trace and reduced norm, respectively. Suppose $B$ admits an involution of the second kind, i.e., an anti-automorphism $*: B \to B$ restricting to $c$ on the center $F$. This is a purely local hypothesis; i.e., it depends only on the completions $B_v$ of $B$ at places of $F$. Let $S_B$ denote the set of places of $F^+$ above which $B$ is ramified. If $v \in S_B$, we assume that $v$ splits in $F$. Then the existence of the involution implies that $B$ ramifies at both places of $F$ dividing $B$. We assume that, at every place $v'$ dividing a $v \in S_B$, $B$ is a division algebra. We will later be fixing a rational prime $p$ and a place $w$ of $F$ dividing $p$. We assume $B$ split at $w$ but make no hypothesis regarding the invariants of $B$ at the remaining divisors of $p$. We choose a maximal order $O_B \subset B$ such that the involution $*$ restricts to an involution of $O_B, p = O_B \otimes F \mathbb{Z}_p$.

Let $B^{op}$ denote the opposite algebra, and let $V$ be the $F$-vector space $B$, viewed as a $B \otimes_F B^{op}$-module. The involution $*$ is assumed to be positive; i.e., $\text{Tr}_{F/Q}(t_B(g \cdot g^*)) > 0$ for all nonzero $g \in B$. Let $B^- \subset B$ denote the $(-1)$-eigenspace for the involution $*$. For any $\beta \in B^-$, we define an involution of the second kind $\#_\beta$ by $x^\#_\beta = \beta x^* \beta^{-1}$ and a $B - *$-hermitian alternating pairing (i.e., alternating upon restriction of scalars to $\mathbb{Q}$, and hermitian in the sense that $(\beta v, w) = (v, \beta^* w)$) $V \times V \to \mathbb{Q}$ by

$$(x_1, x_2)_\beta = \text{Tr}_{F/Q}(t_B(x_1 \beta x_2^*)).$$
Then for \( b \in B, b_{\text{op}} \in B^{\text{op}}, \) we have

\[
((b \otimes b_{\text{op}})x_1, x_2)_{\beta} = (x_1, (b^* \otimes b_{\text{op}}^* \beta) x_2)_{\beta}.
\]

Let \( G_\beta \) be the algebraic group over \( \mathbb{Q} \) whose group of \( R \)-points, for any \( \mathbb{Q} \)-algebra \( R \), is given by the set of \( g \in (B^{\text{op}} \otimes \mathbb{Q} R)^{\times} \) such that, for some \( \lambda \in R^\times \), the following equation is satisfied:

\[
g \cdot g_{\# \beta} = \lambda.
\]

Then \( G_\beta \) is connected and reductive, and \( g \to \lambda \) defines a map \( \nu : G_\beta \to \mathbb{G}_m \). The kernel \( G_{\beta,1} \) of \( \nu \) is the restriction of scalars to \( \mathbb{Q} \) of a group \( G^+_\beta \) over \( F^+ \).

We identify

\[
B \otimes \mathbb{Q} \mathbb{R} \xrightarrow{\sim} \prod_{\tau \in \Sigma} M(n, F_\tau) \xrightarrow{\sim} M(n, \mathbb{C})^d.
\]

For each \( \tau \in \Sigma, (\bullet, \bullet)_{\beta,\tau} \) thus defines a \( * \)-hermitian form \( (\bullet, \bullet)_{\beta,\tau} \) on \( M(n, \mathbb{C}) \). If \( n \) is even, we assume \( 1 + dn/2 \) has the same parity as \( |S_B| \). Then a calculation in Galois cohomology (cf. [C2, §2]; [HT, Lemma 1.1]) shows that \( \beta \) can be chosen in \( B^{-} \) such that \( G_\beta \) is quasi-split at all rational primes that do not split in \( E/\mathbb{Q} \), and such that the form \( (\bullet, \bullet)_{\beta,\tau} \) is of signature \( (n, n(n-1)) \) (resp. \( (0, n^2) \)) for \( \tau = \tau_0 \) (resp. \( \tau \neq \tau_0 \)). Thus \( G_{\beta,\tau_0}^+ \) is isomorphic to \( U(1, n-1) \) but \( G_{\beta,\tau}^+ \) is a compact unitary group for all real places \( \sigma \neq \tau_0 \). We fix such a \( \beta \) and drop it henceforth from the notation.

We write \( K \text{Sh}(G) \) for the locally symmetric space denoted \( M_K \) above. It is in fact a hermitian locally symmetric space of (complex) dimension \( n-1 \), hence a quasi-projective variety by the theorem of Baily-Borel. Because \( B \) is a division algebra, \( K \text{Sh}(G) \) is in fact projective for all \( K \), and smooth if \( K \) is sufficiently small (which we assume). Thus there is no distinction between \( H^*_{\text{cusp}} \) and \( H^* \) in Matsushima’s formula.

The representation \( (\xi, W_\xi) \) defined above gives rise to a representation of \( G \) (take the factors in \( \Sigma \), and regard \( G(R) \) as a subgroup of unitary similitudes in \( GL(n, \mathbb{C})^\Sigma \)). We denote by \( L_\xi \) the corresponding local system on \( \text{Sh}(G) = \lim_K K \text{Sh}(G) \).

If \( p \) splits in \( E \), we can identify

\[
G(\mathbb{Q}_p) \xrightarrow{\sim} \prod_{v|p} B^{\text{op},\times}_v \times \mathbb{Q}_p^\times
\]

where the map \( G(\mathbb{Q}_p) \to \mathbb{Q}_p^\times \) is given by \( \nu \). Thus if

\[
\pi = \pi_{\infty} \otimes_p \pi_p \in \mathcal{A}_0(G),
\]

we can further factor \( \pi_p \) as

\[
\pi_p = \otimes_{v|p} \pi_v \otimes \psi_p,
\]

where \( \psi_p \) is a character of \( \mathbb{Q}_p^\times \).
(1.2.5) Remark. — In practice, we will arrange that $\psi_p$ always be an unramified character. We will moreover make a habit of suppressing the effect of $\psi_p$, which merely complicates the formulas while adding nothing of substance.

The following theorem was originally considered by Clozel. The first complete proof of the base change in both directions was published in the appendix by Clozel and Labesse to Labesse’s book in Astérisque \([CL, L]\). This book contains a much more general framework for proving theorems of this kind, by comparison of stable trace formulas.

(1.2.6) Stable Base Change Theorem (Clozel, Labesse). — Let $(\Xi, W_\Xi)$ be as in (1.1.4). Let $\Pi \subset A_0(GL(n)_F)$ be a cuspidal automorphic representation with central character $\psi_\Pi$, and let $\psi$ be a Hecke character of $E$. Suppose

(a) $\Pi_\infty \simeq \Pi_\xi$,
(b) $\Pi^c \sim \Pi^\vee$;
(c) For every place $v \in S(B)$, $\Pi_v$ is in the discrete series.
(d) $\psi_\Pi |_{A_\xi^c} = \psi/\psi$.
(e) $(\xi |_{E_\xi})^{-1} = \psi_{E_\xi}$.

Then there exists an automorphic representation $\pi$ of $G$ whose base change to $GL(n)_F \times E^\times$ equals $(\Pi, \psi)$. Moreover, $\pi_f$ occurs in the cohomology of $Sh(G)$ with coefficients in $L_\xi$.

Conversely, given $\pi \in \mathcal{A}_0(G)$, cohomological for $\xi$, there exists a pair $(\Pi, \psi)$ satisfying (a)-(c), with $\psi = \psi_{E_\xi}^{c} |_{A_\xi^c}$, such that $(\Pi, \psi)$ is the base change of $\pi$ at all unramified places and at all places that split in $E$. Moreover, if $\pi_v$ is supercuspidal (or corresponds via Jacquet-Langlands to a supercuspidal if $v \in S(B)$) for some $v$ dividing some $p$ that splits in $E$, then $\Pi$ is cuspidal.

(1.2.7) Remark. — Say $\Pi \in CU(n, F)$ if it satisfies (a)-(c). Starting from $\Pi \in CU(n, E)$, one sees easily that there is no obstruction to finding $\psi$ satisfying (d) and (e).

I need to explain the meaning of base change. The group $R_{E/Q}\Gamma_E$ is naturally an inner form of the quasi-split group $R_{E/Q}GL(1)_E \times R_{F/Q}GL(n)_F$. Then the base change of $\pi$ to an automorphic representation $\pi_E$ of $R_{E/Q}GL(1)_E \times R_{F/Q}GL(n)_F$ can be regarded as a pair consisting of an automorphic representation $\Pi$ of $GL(n)_F$ and a Hecke character $\psi$ of $E^\times$; this explains the notation above. To simplify, I assume all $\psi$’s are trivial, but denote them as (?). At places $p$ that split as $yy\mathbb{F}$ in $E$, the base change is simple. Choose one $y$ and write

$$G(\mathbb{Q}_p) \simeq \mathbb{Q}_p^\times \times B_y^{op, x},$$

where $B_y$ is of course a product of central simple algebras over the completions of $F$ at places dividing $y$. Thus given $\pi$, we can write $\pi_p = ? \otimes \pi_y$, and define $\Pi_p = \pi_y \otimes \pi_y^\#$, 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
where $\pi_y^\#(g) = \pi_y((g^\#)^{-1})$. This doesn’t depend on the choice of $y$. Moreover, we recover $\pi_y$ from $\Pi_p$.

If $p$ is inert, then $G(\mathbb{Q}_p)$ is a product of quasi-split unitary groups (up to the center). Local base change for representations of unitary groups is not known in general. But if $\pi_p$ is unramified, and if $G$ is split over an unramified extension of $\mathbb{Q}_p$, let $B \subset G$ be a Borel subgroup, $T \subset B$ the Levi factor. We can identify

$$T(\mathbb{Q}_p) = \{(d_0; d_1, \ldots, d_n) \mid d_0 = d_i \cdot d_{n+1-i}, i = 1, \ldots, n\}.$$ 

If $\alpha$ is a character of $T(\mathbb{Q}_p)$, let

$$BC(\alpha)(d_0; d_1, \ldots, d_n) = \alpha(d_0 \cdot d_0^c; d_0 \cdot d_1^c, \ldots, d_0 \cdot d_n^c/d_1^c).$$

If $\pi_p$ is the unramified representation $\pi(\alpha)$ corresponding to $\alpha$, then $(\Pi_p, \psi_p) = BC(\pi_p) = \pi(BC(\alpha))$. We leave it as an exercise to the reader to determine the Satake parameter of $\Pi_p$ (as opposed to the unramified character $\psi_p$). We have thus defined $\Pi_p$ for almost all $p$, and by strong multiplicity one, this suffices to determine $\Pi$.

Henceforward, to simplify the exposition and minimize notation, we assume $\Xi$ to be the trivial representation. No essential elements of the proof are lost under this assumption. However, for applications to the local Langlands conjecture, we need to be able to consider more general $\Xi$.

1.3. Kottwitz’ theorem and its refinements. — We can identify complex cohomology with $\ell$-adic cohomology, for example, by choosing an isomorphism $\mathbb{C} \sim \mathbb{Q}_\ell$. Thus if $\Pi$ is as above, we can choose a character $\psi$ and define $\pi$ such that $\pi_f \subset H^*(\text{Sh}(G), \mathbb{Q}_\ell)$. It is known that $\text{Sh}(G)$ admits a canonical model over $F$ (recalled next week), and thus there is a virtual representation of $\Gamma_F = \text{Gal}(\overline{F}/F)$ on the $\pi_f$-isotypic subspace

$$R_\ell(\pi, G) = \sum_i (-1)^{n-1+i} \text{Hom}_{\Gamma_F}(A_f, H^i(\text{Sh}(G), \mathbb{Q}_\ell)).$$

(Warning: the sign $(-1)^{n-1}$ will disappear later in the course.) Define

$$R_\ell(\pi) = R_\ell(\pi, G) \otimes \psi^c |_{\Gamma_F}.$$ 

Here is the relation between $R_\ell(\pi)$ and $\Pi = BC(\pi)$:

(1.3.1) Theorem (Kottwitz, [K4]). — There is a constant $a(\pi)$ such that for almost all places $p$ not dividing $\ell$, such that $\pi_p$ is unramified, and for all $v$ dividing $p$, the local representation $R_\ell(\pi)_v$ is isomorphic to $a(\pi)$ copies of

$$\bigoplus_{i=1}^n \alpha_i(\Pi_v)^{-1},$$

where the $v$-component $\Pi_v$ of $\Pi$ is the unramified representation attached to the $n$-tuple of characters $(\alpha_i(\Pi_v))$.

(Note: sign conventions differ in the literature.)

Here is an argument, based on ideas of Clozel, to show that $\pi_f$ occurs only in the middle degree $n - 1$. Kottwitz’ theorem uses the theory of the zeta function of the
reduction mod $v$ of $\text{Sh}(G)$. In particular, the $\alpha_i(\Pi_v)$ are eigenvalues of Frobenius on $H^*(\text{Sh}(G), \mathbb{Q}_\ell)$ (up to sign). In particular, they are algebraic numbers whose complex absolute values are determined by the degree of cohomology in which they occur. However, $\Pi$ is cuspidal, hence $\Pi_v$ is unitary for all $v$. It follows from the classification of unitary representations of $GL(n, F_v)$ (Tadic) and Deligne’s purity theorem that all $\alpha_i(\Pi_v)$ have the same complex absolute values. Thus $\pi_f$ can only occur in one dimension of cohomology. By the hard Lefschetz theorem, this can only be the middle dimension.

Taylor has given an argument (cf. [HT, §vii 1.8]) to show that the constant $a(\pi)$, an uncontrolled multiplicity that arises in the comparison of trace formulas, can be factored out of $R_\ell(\Pi)$; i.e., we can write

$$R_\ell(\Pi)_{ss} = R_{\ell,0}(\Pi)^{a(\pi)}$$

for some $n$-dimensional semisimple representation $R_{\ell,0}(\Pi)$ of $\Gamma_F$; here the subscript $ss$ denotes semisimplification, which is all we can understand via traces. The argument requires that $R_\ell(\Pi)$ be Hodge-Tate. In our case, $R_\ell(\Pi)$ is realized as a subquotient of the cohomology of a Kuga fiber variety, by the Leray spectral sequence. Since the Kuga fiber variety is smooth and projective over some number field, its cohomology is potentially semi-stable at all places, by Tsuji’s theorem. Then any subquotient is also pst, hence is Hodge-Tate. So there is no problem. But even if there were, by controlling the local ramification at inert places, one can arrange to have $a(\pi) = 1$ (joint work in progress with Labesse). We define

$$r_\ell(\Pi) = (R_{\ell,0}(\Pi)_{ss})^\vee$$

Applying Kottwitz’ theorem, we obtain

$$(1.3.4) \text{Theorem (Clozel [C2] + Taylor). — Let } \Pi \in CU(n, E). \text{ Then there is a compatible family } (r_\ell(\Pi)) \text{ of } n\text{-dimensional } \lambda\text{-adic representations such that, for all finite places } v \text{ outside a finite set } S \text{ containing all ramified places, Kottwitz’ theorem yields}$$

$$[r_\ell(\Pi) |_{\Gamma_v}]_{ss} \simeq \sigma_\ell(\Pi_v).$$

**Remark.** — We are always working with the unitary normalization of the Langlands correspondence. So the $L$-function of $\Pi$ has to be shifted by $(n-1)/2$ in order to obtain the $L$-function of a Galois representation:

$$(1.3.5) L^S(s, \Pi) = L^S(s + \frac{n-1}{2}, \sigma_\ell(\Pi)).$$

where the left hand side is automorphic (in the unitary normalization) and the right hand side is the partial Galois-theoretic $L$-function, with the factors at the finite set $S$ of bad primes removed.

Here $\sigma_\ell$ is the local Langlands correspondence, so far defined for unramified representations. A compatible family $(r_\ell(\Pi))$ as above is called weakly associated to $\Pi$. The goal of my lectures is to present a proof of the generalization of this theorem,
contained in my article with Taylor, to all places \( v \); in other words, to show that 
\((r_\ell(\Pi))\) is strongly associated to \( \Pi \) (after semi-simplification locally). The remainder of today’s lecture explains how to reduce the local Langlands conjecture to the statement that the representations 
\((r_\ell(\Pi))\) are strongly associated to \( \Pi \).

But first, to make sense of this, we need to have constructed a family of local bijections \( \pi \leftrightarrow \sigma(\pi) \) between \( \mathcal{A}(n, F_v) \) and \( \mathcal{G}(n, F_v) \) for all places \( v \) that are candidates for the local Langlands correspondence. We need to know that \( \sigma \) comes from \( \mathcal{A}_0(n, F_v) \rightarrow \mathcal{G}_0(n, F_v) \), that \( \sigma(\pi^\vee) = \sigma(\pi)^\vee \), that \( \sigma \) commutes with character twists, Galois automorphisms, base change, automorphic induction: in short, that \( \sigma \) satisfies all hypotheses enumerated in the introduction except, perhaps, compatibility with local \( \varepsilon \)-factors. In later lectures I will explain how to construct such a correspondence by algebraic geometry, such that

\[(1.3.6) \text{Main Theorem (HT)}. \quad \text{Let} \ \Pi \in \mathcal{C}U(n, E). \ \text{Then for all places} \ v \ \text{not dividing} \ \ell, \]

\[ [r_\ell(\Pi)]|_{\mathfrak{r}_v} \simeq [\sigma_\ell(\Pi_v)].(6) \]

Now I have to explain how this theorem implies compatibility with \( \varepsilon \) factors.

1.4. Non-Galois automorphic induction. — The arguments in this section are taken from the article [H2], and were extended slightly in [HT]. Here and in what follows, we use the notation \( \boxplus \) to denote Langlands sum. Let \( K \) be a local field, \( \pi \in \mathcal{A}(n, K) \), \( \pi' \in \mathcal{A}(m, K) \), and let \( P \subset G_{n+m} \) be the standard parabolic subgroup with Levi subgroup \( G_n \times G_m \); the representation \( \pi \otimes \pi' \) defines by inflation a representation of \( P \). We define the representation

\( \pi \boxplus \pi' \in \mathcal{A}(n + m, K) \)

to be the Langlands subquotient of the normalized induction

\[ \text{Ind}_{P}^{G_{n+m}} \pi \otimes \pi'. \]

We now consider a triple of CM fields \( F_3 \supset F_2 \supset F_1 \) as before, all containing \( E \), with totally real subfields \( F_i, i = 1, 2, 3 \). We assume \( F_3/F_1 \) is Galois and \( \Gamma = \text{Gal}(F_3/F_1) \) is solvable. The goal is to show that certain algebraic Hecke characters \( \chi \) of \( F_2 \) define by induction automorphic representations \( I_{F_3/F_1}(\chi) \) of \( GL(d, F_1) \), where \( d = [F_2:F_1] \). The meaning is clear. If \( \chi \) is associated to an \( \ell \)-adic character \( r_\ell(\chi) \) of \( \Gamma_{F_2} \), then \( I_{F_3/F_1}(\chi) \) should be weakly associated to \( \text{Ind}_{F_3/F_1} r_\ell(\chi) \), where \( \text{Ind}_{F_3/F_1} \) denotes induction from \( \Gamma_{F_2} \) to \( \Gamma_{F_1} \). Concretely, let \( v \) be a place of \( F_1 \), unramified in \( F_2 \), such that \( \chi_v \) is unramified at all \( w \) dividing \( v \). For each \( w \mid v \) \( F_{2,w} \) is a cyclic extension of \( F_{1,v} \) of degree \( f_w \), and we define the representation \( I_{w/v} \chi_w \) of \( GL(f_w, F_{1,v}) \) by cyclic automorphic induction: it is the unramified representation associated to the \( f_w \)-tuple.

\[(6)\text{Note added in proof: Taylor and T. Yoshida have just proved the expected strengthening of this theorem, in which semisimplification is replaced by Frobenius’ semisimplification.}\]
of characters \( \chi_w \circ \gamma \) as \( \gamma \) runs through \( \text{Gal}(F_{2,w}/F_{1,w}) \). Then the \( v \) component of \( I_{F_{2}/F_{1}}(\chi) \) must be the (Langlands) sum of the \( I_{w/v}\chi_w \) for \( w \mid v \). The problem is to show that these local components fit together into an automorphic representation.

\textbf{(1.4.1) Remark.} — Regarding the archimedean constituents, recall that we are always working with the unitary normalization of the \( L \)-function. So in fact, \( \chi \cdot |s|^{(d-1)/2} \), rather than \( \chi \), is an algebraic Hecke character.

Suppose we can do this for quite general \( \chi \): just how general will become clear in a moment. Suppose moreover that \( I_{F_{2}/F_{1}}(\chi) \in \text{CU}(d, F_{1}) \), so that we can apply the Main Theorem. Let \( F_{2}' \supset F_{2} \supset F_{1} \) be another triple as above, with \( |F_{2}' : F_{1}| = d' \), and suppose we have a second character \( \chi' \). Write \( \Pi(\chi) = I_{F_{2}/F_{1}}(\chi), \Pi(\chi') = I_{F_{2}'/F_{1}}(\chi') \).

Consider the Rankin-Selberg \( L \)-function, with its functional equation \([\text{JPSS, Sh}]\)

\begin{equation}
L(s, \Pi(\chi) \otimes \Pi(\chi')) = \prod_{v} \varepsilon_{v}(s, \Pi(\chi)_v \otimes \Pi(\chi')_v, \psi_v)L(1 - s, \Pi(\chi) \vee \Pi(\chi') \vee).
\end{equation}

The local \( \varepsilon \) factors are those of the automorphic theory. On the other hand, we have \( \Pi(\chi) \) weakly associated to \( \text{Ind}_{F_{2}'/F_{1}} r_{\ell}(\chi') \), and likewise for \( \Pi(\chi') \). The tensor product \( \text{Ind}_{F_{2}/F_{1}} r_{\ell}(\chi) \otimes \text{Ind}_{F_{2}'/F_{1}} r_{\ell}(\chi') \) is an \( \ell \)-adic representation corresponding to a complex representation of the global Weil group of \( F_{1} \), hence there is a functional equation

\begin{equation}
L(s, \text{Ind}_{F_{2}/F_{1}} r_{\ell}(\chi) \otimes \text{Ind}_{F_{2}'/F_{1}} r_{\ell}(\chi'))
= \prod_{v} \varepsilon_{v}(s, \text{Ind}_{F_{2,v}/F_{1,v}} r_{\ell}(\chi_v) \otimes \text{Ind}_{F_{2,v}'/F_{1,v}} r_{\ell}(\chi'_v), \psi_v)
L(1 - s, \text{Ind}_{F_{2}/F_{1}} r_{\ell}(\chi) \vee \text{Ind}_{F_{2}'/F_{1}} r_{\ell}(\chi') \vee),
\end{equation}

where the local \( \varepsilon \) factors are those of Langlands and Deligne. By the technique explained by Henniart in his course, and recalled in Appendix (A.2) (cf. [He1]), this implies identities of local \( \varepsilon \) factors for all \( v \):

\begin{equation}
\varepsilon_{v}(s, \Pi(\chi)_v \otimes \Pi(\chi')_v, \psi_v) = \varepsilon_{v}(s, \text{Ind}_{F_{2,v}/F_{1,v}} r_{\ell}(\chi_v) \otimes \text{Ind}_{F_{2,v}'/F_{1,v}} r_{\ell}(\chi'_v), \psi_v).
\end{equation}

Fix \( v \) of residue characteristic \( p \) split in \( E \), as before, \( K = F_{1,v} \) and assume \( v \) is inert in \( F_{3} \) and \( F_{3}' \). Let \( \sigma \) and \( \sigma' \) be two representations of \( \Gamma_v \) factoring through \( \text{Gal}(F_{3,v}/K) \) and \( \text{Gal}(F_{3}'_{v}/K) \), respectively, and let \( \pi \) and \( \pi' \) be the corresponding elements of \( A_0(n, K) \) (resp. \( A_0(n', K) \)). Note that for any \( \sigma \in G_0(n, K) \) we can choose a local extension \( F_{3,v}/K \) which is solvable and such that \( \sigma \) comes from \( \text{Gal}(F_{3,v}/K) \) up to an unramified twist, which we ignore. By Brauer’s theorem, there are intermediate fields \( K \subset F_{2,j,v} \subset F_{3,v} \), characters \( \chi_{j,v} \) of \( \text{Gal}(F_{3,v}/F_{2,j,v}) \), and integers \( e_j \) such that

\[ \sigma = \sum_{j} e_j \text{Ind}_{F_{2,j,v}/K} \chi_{j,v}; \]

likewise for \( \sigma' \). Applying the above identity of \( \varepsilon \) factors, and ignoring \( \psi_v \), one obtains

\[ \varepsilon_{v}(s, \sigma \otimes \sigma') = \prod_{j,j'} \varepsilon_{v}(s, \Pi(\chi_{j,v}) \otimes \Pi(\chi'_{j',v}))^{e_{j'} \epsilon_{j'}}. \]
Now we apply the Main Theorem. Say

$$\Pi(\chi_j)_v = \boxtimes_i \pi_{i,j}; \quad \Pi(\chi'_j)_v = \boxtimes_{i'} \pi'_{i',j'},$$

with the notation $\boxtimes$ defined as above. The Main Theorem states that

$$\sigma = \sum_{i,j} e_{ij} \sigma(\pi_{i,j}); \quad \sigma' = \sum_{i',j'} e_{i'j'} \sigma(\pi'_{i',j'})$$

Write $\sigma_{i,j} = \sigma(\pi_{i,j})$, etc. On the other hand, by the additive properties of the automorphic $\varepsilon$ factors (cf. (A.2.2), (A.2.4)),

$$\prod_{j,j'} \varepsilon_v(s, \Pi(\chi_j)_v \otimes \Pi(\chi'_j)_v)^{e_{i,j} \cdot e_{i',j'}} = \prod_{i,j,i',j'} \varepsilon_v(s, \pi_{i,j} \otimes \pi'_{i',j'})^{e_{i,j} \cdot e_{i',j'}}$$

$$= \prod_{i,j,i',j'} \varepsilon_v(s, \sigma^{-1}(\pi_{i,j}) \otimes \sigma^{-1}(\pi'_{i',j'}))^{e_{i,j} \cdot e_{i',j'}}$$

$$= \varepsilon_v(s, \sigma^{-1}(\sigma)) \otimes \sigma^{-1}(\sigma')) \quad (0.8)$$

This yields the identity (0.8) of $\varepsilon$ factors, which, assuming the Main Theorem (1.3.6), completes the proof of the local Langlands conjecture. Since $\sigma$ commutes with twists by characters, one sees that it doesn’t matter if we only get $\sigma$ and $\sigma'$ up to unramified twists.

By simple approximation arguments, we see that any local extension $K'/K$ can be realized as an $F_3v/F_1v$ as above. In this way, we find that it suffices to prove that, for any intermediate field $F_2$ in a solvable extension and any character $\chi_v$ of $F_2v$, then (up to unramified twists) $\chi_v$ can be realized as the local component of a Hecke character $\chi$ for which $I_{F_2/F_1}(\chi)$ exists as an automorphic representation of $GL(d, F_1)$.

If $F_2/F_1$ is cyclic the existence of a global $I_{F_2/F_1}(\chi)$ is guaranteed by the base change theory of Arthur-Clozel. By induction, $I_{F_2/F_1}(\chi)$ exists when $F_2$ is solvable over $F_1$, without any additional hypothesis on the fields or characters. On the other hand, if $F_2/F_1$ is cyclic and $\Pi_2$ is an automorphic representation of $GL(d, F_2)$ invariant under $\text{Gal}(F_2/F_1)$, then Arthur and Clozel prove the existence of $\Pi_1$ on $GL(d, F_1)$ whose base change to $F_2$ is $\Pi_2$: this is the descent of $\Pi_2$ to $F_1$.

So one might argue as follows: let $\Gamma_2 = \text{Gal}(F_3/F_2)$, and, motivated by the usual restriction/induction formula on the Galois side, replace $\chi$ by

$$\Pi_3(\chi) := \boxtimes_{\Gamma' \Gamma_2} \chi \circ N_{F_3/F_2}.$$ 

The result is invariant under $\Gamma$, by construction, so one should be able to descend to fixed fields of successive cyclic Galois groups of prime order.

The problem is that descent is ambiguous. To simplify, assume there is an intermediate field field $F_1 \subset E \subset F_3$ with $F_3/E$ and $E/F_1$ cyclic of prime order; let $C = \text{Gal}(F_3/E)$. Let $J(\chi, E) = \text{Res}_{F_3} \text{Ind}_{F_2/F_1} r_\ell(\chi)$. Then

$$\text{Res}_{F_3} J(\chi, E) = \text{Res}_{F_3} (J(\chi, E) \otimes \beta)$$

ASTÉRISQUE 298
for any character $\beta$ of $C$. There is a similar ambiguity in descent. Consider the first step: let $\Pi$ be an automorphic representation of $GL(d, F)$ invariant under $C$, $\Pi_C$ a descent to $GL(d, \bar{E})$. For any character $\beta$ of $C$, the twist $\Pi_C \otimes \beta$ ($\beta$ viewed as a Hecke character of $GL(1, \bar{E})$) is another descent of $\Pi$. So the total number of descents is in the order of $|C|$ (some twists may be isomorphic). (Actually, there is more ambiguity: if $\Pi_C$ is not cuspidal, say $\Pi_C = \bigoplus_j \Pi_j$, then each $\Pi_j$ can be twisted separately by a character $\beta_j$ of $C$.) On the other hand, locally everywhere $\Pi_{C,p}$ can be identified only up to twist(s) by character(s) $\alpha_{p,j}$ of the decomposition group $C_{v,p} \subset C$ at $p$. The general theory thus gives that for each $p$ there is a character $\alpha_p$ of $C_p$ such that, for almost all $p$,

$$\sigma_\ell(\Pi_{C,p}) = J(\chi, E)_p \otimes \alpha_p.$$ 

A priori, there is no way to prove that the local characters $\alpha_p$ fit together to a global character $\beta$ of $C$.

On the other hand, if it is known that $\Pi_C$ has a weakly associated $\ell$-adic representation $r_\ell(\Pi_C)$ of $\Gamma_{\bar{E}}$, then $r_\ell(\Pi_C)$ is a descent to $E$ of $\text{Res}_{E_3} \text{Ind}_{F_2/F_1} r_\ell(\chi)$. If moreover $r_\ell(\Pi_C)$ is irreducible, then the only ambiguity in descents comes from twists by characters of $C$:

$$r_\ell(\Pi_C) = J(\chi, E) \otimes \alpha$$

for a global character $\alpha$. So one can replace $\Pi_C$ by $\Pi_C \otimes \alpha^{-1}$ and continue with the descent. More generally, the analogous argument works if $\Pi_C = \bigoplus_j \Pi_j$ with a weakly associated $r_\ell(\Pi_j)$ for each $j$.

So it suffices by induction to show that there is a sequence of intermediate fields $F_1 = E_0 \subset E_1 \subset \cdots \subset E_r = F_3$ with each $E_i/E_{i-1}$ Galois and cyclic of prime degree, and a sequence of descents $\Pi(\chi, E_i)$ of $\Pi_3(\chi)$ such that each $\Pi(\chi, E_i) = \bigoplus_{i,j} \Pi_{i,j}$ with each $\Pi_{i,j} \in CU(d_j, E_i)$.

Recall (1.2.7) that $CU(d_j, E_i)$ involves 3 conditions: conjugate self-duality, regularity at $\infty$, and a local condition at some prime. We assume $v$ is inert in $F_3/F_1$ and we assume there is a second prime $w$ of $F_1$, also inert in $F_3$ and dividing a rational prime $p(w)$ split in $E$. In practice, we have to allow $p(w) = p$, which is not a problem. We want $\chi_w$ to be general, $\chi_{\infty}$ such that $\Pi_{i,j,\infty}$ is cohomological for all $i$ and $j$, and $\chi_w$ such that every $\Pi_{i,j,w}$ is supercuspidal. Since $p$ splits in $E$, $v \neq w^c$, so the condition $\chi^{-1} = \chi^c$ imposes no restriction on $\chi_v$. The condition at $w$ is a bit more subtle. By Mackey’s theorem, the restriction $\text{Res}_{E_i} \text{Ind}_{F_2/F_1} r_\ell(\chi)$ breaks up as a sum of constituents that may not necessarily be irreducible:

$$\text{Res}_{E_i} \text{Ind}_{F_2/F_1} r_\ell(\chi) = \bigoplus_a \text{Ind}_{E_i,a/E_i} a(\chi)$$

where $a$ runs through the double cosets $\Gamma_{E_i} \backslash \Gamma/\Gamma_2$, $E_i,a$ is the fixed field of $a(\Gamma_2)a^{-1} \cap \Gamma_{E_i}$, and $a(\chi)$ is the restriction to $a(\Gamma_2)a^{-1} \cap \Gamma_{E_i}$ of $\chi$ (conjugated by $a$). We need to choose $\chi_w$ so that each of these Mackey constituents is locally irreducible at $w$. 

---

SOCIÈTÉ MATHÉMATIQUE DE FRANCE 2005
This is true generically (exercise, cf. [H2, Lemma 4.7]). Then the base change theory implies that \( \Pi(\chi, E_i) = \mathbb{H}_{\Pi, a} \), with \( \Pi_{i,a} \) supercuspidal at \( w \).

Supposing we have \( \chi_w \), we choose a global Hecke character \( \chi_0 \), trivial at \( \infty \), such that \( \chi_{0,v} = \chi_v, \chi_{0,w} = \chi_w, \chi_{0,v^c} = 1 = \chi_{0,w^c} \). Let \( \chi_1 = \chi_0/\chi_0^c \). This has the right properties at \( v \) and \( w \) and satisfies \( \chi_1^c = \chi_1^{-1} \).

To obtain \( \chi_\infty \), we work backwards. For any complex place \( \tau \) of \( F_1 \), let \( \tau(k), k = 1, \ldots, d \) be the primes of \( F_2 \) dividing \( \tau \). Then

\[
\Pi(\chi)_\infty = \mathbb{B}_k\chi_\tau(k).
\]

The regularity hypothesis requires that all \( \chi_\tau(k) \) be distinct for fixed \( \tau \), and conjugate self-duality requires that

\[
\{ \chi_\tau^{-1}(k) \} = \{ \chi_{c\tau}(k) \}.
\]

The coefficient system \( \Xi \) determines the set \( \{ \chi_\tau^{-1}(k) \} \) for each \( \tau \). This is again not a restriction. Let \( \chi_2 \) be any algebraic Hecke character with \( \chi_2^c = \chi_2^{-1} \), with \( \chi_2,\tau(k) \) as just described for all \( \tau \), and such that \( \chi_{2,v} \) and \( \chi_{2,w} \) are unramified. By allowing sufficient ramification elsewhere, we can easily construct such \( \chi_2 \). Then we let \( \chi = \chi_1 \cdot \chi_2 \). This has the right behavior at \( w \) and \( \infty \), is conjugate self-dual, and is the desired \( \chi_v \) up to unramified twist. It suffices to show that \( I_{F_2/F_1} \chi \) is automorphic for such \( \chi \).

Here is the induction step. Suppose we have \( \Pi(\chi, E_i) = \mathbb{H}_{\Pi_i, a} \), with \( \Pi_{i,a} \) supercuspidal at \( w \), and with each \( \Pi_{i,a} \in CU(d_i, a); d_i \) is the dimension of the corresponding Mackey constituent. We further assume that the set \( \{ \Pi_{i,a} \} \) is invariant under \( \text{Gal}(E_i/F_1) \). Finally, we assume that the sum of the corresponding Galois representations is \( \text{Res}_{E_i} \text{Ind}(\tau(\chi)) \). Let \( C_i = \text{Gal}(E_i/E_{i-1}) \), of prime order \( q \). We let \( C_i \) act on the set \( \{ \Pi_{i,a} \} \). The orbits are either fixed points or of order \( q \). If \( \Pi_{i,a} \) is a fixed point, it descends to a \( \Pi_{i-1,a',0} \), with local component supercuspidal at \( w \) (since this is true after base change). The cohomological condition is automatic, though the relevant coefficient system depends on the orbit of \( C_i \) on the \( \tau(k) \). We need to know that

\[
\Pi_{i-1,a',0} = \Pi_{i-1,a',0}^\prime.
\]

This is true after base change, so

\[
\Pi_{i-1,a',0}^\prime = \Pi_{i-1,a',0}^\prime \cdot \eta
\]

for some character \( \eta \) of \( C_i \). But \( E_i/E_{i-1} \) comes from a cyclic extension of totally real fields \( E_i^+/E_{i-1}^+ \), so \( \eta = \eta^+ \circ N_{E_i/E_{i-1}} \). This implies that \( \eta^+ \) extends to a finite Hecke character \( \alpha \) of the ideles of \( E_{i-1} \). Replacing \( \Pi_{i-1,a',0} \) by \( \Pi_{i-1,a'} := \Pi_{i-1,a',0} \otimes \alpha \), we find that

\[
\Pi_{i-1,a'} = \Pi_{i-1,a'}^\prime.
\]

Next, if \( \{ \Pi_{i,a}^t \mid t \in C_i \} \) is a non-trivial orbit, the Langlands sum \( \mathbb{B}_i\Pi_{i,a}^t \) descends to a \( \Pi_{i-1,a'} \). Such a descent is unique, hence the conjugate-self duality is automatic.
as is the cohomological condition. The supercuspidality at \( w \) follows from the choice of generic \( \chi_w \).

Finally, we need to know that the set \( \{ \Pi_{i-1,a'} \} \) gives the right set of Galois representations. But each corresponding \( \sigma(i-1,a') \) is locally irreducible at \( w \), hence is globally irreducible, hence is determined up to a twist by a character of \( C_i \). Looking at the Mackey decomposition, we see that by choosing the right twist, we get a constituent of \( \text{Res}_{E_{i-1}} \text{Ind}(r_e(\chi)) \). In particular, the set of constituents is invariant under \( \text{Gal}(E_{i-1}/F_i) \), and this completes the induction step.

(1.4.5). — Lectures 2-7 are devoted to the proof of Main Theorem (1.3.6) in the special case where \( \Pi_v \) is the full induced representation from a supercuspidal representation of the Levi subgroup of a parabolic subgroup of \( GL(n,F_v) \). This case suffices to establish the compatibility of local \( \varepsilon \) factors, as one verifies immediately by inspecting the arguments presented above. Strangely, the more general version of the Main Theorem appears to be required to prove the modular local Langlands conjecture, due to Vignéras [V].

2. Shimura varieties as moduli varieties

2.1. Shimura varieties attached to fake unitary groups: canonical models

A Shimura datum is a pair \((G,X)\), where \( G \) is a connected reductive group over \( \mathbb{Q} \) and \( X \) is a \( G(\mathbb{R})\)-conjugacy class of homomorphisms \( h: R_{C/\mathbb{R}}(G_{m,C}) \to G_{\mathbb{R}} \), satisfying a familiar list of axioms [De1]. We will always assume the weight morphism \( w_h \), the restriction of \( h \) to \( G_{m,\mathbb{R}} \), is rational over \( \mathbb{Q} \). The centralizer of \( h \) contains the real points of the center \( Z_G \) of \( G \), as well as a maximal compact subgroup \( K_\infty \), and the axioms imply that the connected components of \( X \) are hermitian symmetric spaces homogenous under the identity component of the group of real points of the derived subgroup \( G_{\text{der}} \) of \( G \). Upon extension of scalars to \( \mathbb{C} \), an \( h \in X \) defines a homomorphism \( G_{m,\mathbb{C}} \times G_{m,\mathbb{C}} \to \hat{G}_{\mathbb{C}} \) whose first coordinate is a cocharacter denoted \( \mu = \mu_h \).

The \( G \)-conjugacy class of \( \mu \) is independent of \( h \) and its field of definition is a number field denoted \( E(G,X) \); we will write \( \mu_X \) for any point in this conjugacy class. This is a cocharacter of some maximal torus of \( G \), hence a character of a maximal torus \( \hat{T} \subset \hat{G} \). The Shimura variety \( \text{Sh}(G,X) \), whose set of complex points is given by

\[
\text{Sh}(G,X)(\mathbb{C}) = \lim_{U \subset \hat{G}(A_f)} \text{Sh}_U(G,X)(\mathbb{C}),
\]

where

\[
\text{Sh}_U(G,X)(\mathbb{C}) = G(\mathbb{Q}) \backslash (X \times G(A_f))/U \simeq \mathcal{M}_U(G)
\]

(notation as in (1.1.3)), has a canonical model over the field \( E(G,X) \). This is a general fact that will be derived for our specific Shimura varieties by interpreting them as solutions to a moduli problem.
Notation is as before: $F = F^+ E$, $B$, $\tau_0$, $\sigma_0$, $p = wu^c$, $w$, $u$, $\Sigma$, etc.

Choose an $\mathbb{R}$-algebra homomorphism $h_0 : C \to B^{\text{op}} \otimes_{\mathbb{Q}} \mathbb{R}$ such that $h_0(z)^\# = h_0(\overline{z})$ for all $z \in C$. The image is contained in $G$ and is centralized by a maximal compact subgroup of $G(\mathbb{R})$ if and only if the map $x \mapsto h_0(i)^{-1}x^\# h_0(i)$ is a positive involution. Since $\#$ is conjugate to
\[
g \mapsto \text{diag}(-1, 1, \ldots, 1)^g \text{diag}(-1, 1, \ldots, 1)
\]
this means that $h_0(z)$ must be conjugate to $\text{diag}(z, \overline{z}, \overline{z}, \ldots, \overline{z})$ (in the $\tau_0$ coordinate) and $\overline{z} \cdot I_n$ (in the remaining coordinates). Let $(G, X)$ be the Shimura datum for which $X$ is the $G(\mathbb{R})$-conjugacy class containing $h_0$. Then the reflex field $E(G, X)$ is isomorphic to $F$, identified with its image in $C$ under $\tau_0$.

Recall the Hodge-theoretic interpretation of $h \in X$. Any irreducible representation of $R_{C/\mathbb{R}}(G_m)$ is of the form $z \mapsto h_{p,q}(z) = z^{-p} \overline{z}^{-q}$ for $p, q \in \mathbb{Z}$. If $h \in X$ and $(\rho, V)$ is a representation of $G$, then $\rho \circ h$ decomposes $V \mathbb{C}$ as a sum of eigenspaces $V^{p,q}$ for $h_{p,q}$.

(Scholium: $V^{p,q} = H^0(Y, \Omega^p)$ if $V$ is the complex cohomology of a smooth complex variety $Y$.)

Let $V = B$ itself. Then
\[
B_{\mathbb{C}} = B \otimes_{\mathbb{Q}} \mathbb{C} = \oplus_{\tau \in \Sigma} B_\tau \oplus B_{\text{cor}};
\]
moreover, $B_\tau$ is the $\tau$-eigenspace for the action of $F$ for $\tau \in \text{Hom}(F, \mathbb{C})$. On the other hand, via $h_0$, we have $B(\mathbb{C}) = B(\mathbb{C})^{-1,0} \oplus B(\mathbb{C})^{0,-1}$. Now $B_{\mathbb{R}}$ has a positive involution $\ast$, defining (via the trace) a bilinear form that takes rational values on $B(\mathbb{Q})$. Hence, choosing a lattice $\Lambda \subset B(\mathbb{Q})$, we find that
\[
\Lambda \backslash B(\mathbb{R}) = \Lambda \backslash B(\mathbb{C}) / B(\mathbb{C})^{0,-1}
\]
is a polarized abelian variety $A_0$, with Lie algebra $B(\mathbb{C})^{-1,0}$. Decomposing $\text{Lie}(A_0) = B(\mathbb{C})^{-1,0}$ as a sum of $\tau$-eigenspaces, we find
\[
(2.1.3) \quad \dim \text{Lie}(A_0)_{\tau_0} = n, \quad \dim \text{Lie}(A_0)_\tau = 0, \tau \in \Sigma, \tau \neq \tau_0;
\]
and for all $\tau$,
\[
(2.1.4) \quad \dim \text{Lie}(A_0)_\tau + \dim \text{Lie}(A_0)_{\text{cor}} = n^2 = \dim_{\mathbb{F}} B.
\]
This justifies the relation to moduli explained in the next section.

2.2. The moduli problem. — If $A$ is an abelian scheme over a base scheme $S$ over $\mathbb{Q}$, let $T_j(A)$ denote the direct product of the Tate modules $T_j(A)$ over all primes $\ell$, $V_f(A) = \mathbb{Q} \otimes T_j(A)$. Let $U \subset G(A_f)$ be a compact open subgroup. Consider the functor $A_U(B, \ast)$ on schemes over $F$, which to $S$ associates the set of equivalence classes of quadruples $(\lambda, A, i, \eta)$, where $A$ is an abelian scheme of dimension $dn^2$, $\lambda : A \to \hat{A}$ is a polarization, $i : B \hookrightarrow \text{End}(A) \otimes \mathbb{Q}$ is an embedding, and $\eta : V \otimes_{\mathbb{Q}} A_f \overset{\sim}{\longrightarrow} V_f(A)$ an isomorphism of skew-hermitian (see below) $B \otimes_{\mathbb{Q}} A_f$-modules, modulo $U$; here $V$ is the $B \otimes_{\mathbb{Q}} B^{\text{op}}$-module $B$, as in (1.2).
Here is the precise meaning of “modulo $U$” following Kottwitz \cite[p. 390]{K5}). We may assume $S$ connected. The Tate module $T_f(A)$ is a smooth $A_f$-sheaf on $S$. Fixing a geometric point $s \in S$, it is thus the $A_f$-sheaf associated to the representation of $\pi_1(S, s)$ on $T_f(A_s)$. Then a level structure modulo $U$ is a $U$-orbit of isomorphisms $\eta : V \otimes \mathbb{Q} A_f \sim V_f(A_s)$ that is stable under the action of $\pi_1(S, s)$ on the right. It can be checked that this condition is independent of the choice of geometric point.

We assume the Rosati involution on $\text{End}(A) \otimes \mathbb{Q}$ restricts to the involution $\ast$ on $i(B)$ and $\eta$ takes the standard pairing on $V$ to an $A_f$-multiple of the Weil pairing for $\lambda$ on $V_f(A)$. Most importantly, $i$ induces an action $i_F$ of the center $F$ of $B$ on the $O_S$-module $\text{Lie}(A)$. For each embedding $\tau : F \rightarrow \mathbb{C}$, we let $O_{S, \tau} = O_S \otimes F, \tau \mathbb{C}$, and let $\text{Lie}(A)_\tau = \text{Lie}(A) \otimes F, \tau \mathbb{C}$. We then assume that

(2.2.1) $\text{Lie}(A)_\tau = 0$, $\tau \in \Sigma$, $\tau \neq \tau_0$;
(2.2.2) $\text{Lie}(A)_{\text{cor}}$ is a projective $O_{S, \text{cor}}$ module of rank $n^2$, $\tau \neq \tau_0$;
(2.2.3) $\text{Lie}(A)_{\tau_0}$ is a projective $O_{S, \tau_0}$ module of rank $n$;
(2.2.4) $\text{Lie}(A)_{\text{cor} \tau_0}$ is a projective $O_{S, \text{cor} \tau_0}$ module of rank $n(n - 1)$.

Note – this is important – that the action of $F$ on $\text{Lie}(A)_\tau$ is via the embedding $\tau$. Two quadruples $(A, \lambda, i, \eta)$ and $(A', \lambda', i', \eta')$ are equivalent if there is an isogeny $A \to A'$ taking $\lambda$ to a $\mathbb{Q}^\times$-multiple of $\lambda'$ and preserving the other structures. In particular, we may always assume $|\text{Ker}(\lambda)|$ prime to $p$.

We assume $U$ is sufficiently small; then $\mathcal{A}_U(B, \ast)$ is represented by a smooth projective scheme over $F$, also denoted $\mathcal{A}_U(B, \ast)$. For $B = \mathbb{Q}$ this was proved over any base prime to the level of $U$ by Mumford, using geometric invariant theory. The problem with $B$ is relatively representable over the one without $B$ “by the theory of the Hilbert scheme,” as one says at this point. In fact, the complete proof is written down nowhere, except in Shimura’s papers of the early 60s, which use the language of Weil’s algebraic geometry. (However see \cite{Hida,}.)

2.3. Points over $\mathbb{C}$. Hasse principle and connected components. — Using Riemann matrices, we show that $\mathcal{A}_U(B, \ast)$ is isomorphic to $|\text{ker}^1(\mathbb{Q}, G)|$ copies of the canonical model of $\mathcal{U} \text{Sh}(G, X)$. I begin by explaining the source of the invariant $\text{ker}^1(\mathbb{Q}, G) = \text{ker}[H^1(\mathbb{Q}, G) \to \prod_v H^1(\mathbb{Q}_v, G)]$. Recall that $G$ is the group of automorphisms of the $R_{F/\mathbb{Q}}B$-module $V$ that preserve the $\ast$-skew-hermitian pairing $(x_1, x_2)_\beta$ up to a scalar. Here and below, a $\ast$-skew-hermitian form on a $B$-module is only considered fixed up to a ($\mathbb{Q}$-rational) scalar. If $V'$ is a second skew-hermitian $B$-module of the same dimension, then $V'_{\mathbb{Q}} \sim_{\mathbb{Q}} V''_{\mathbb{Q}}$ as skew-hermitian $B$-modules, and this gives rise to a class in $H^1(\mathbb{Q}, G)$. Now suppose we have a point $x = (A, \lambda, i, \eta) \in \mathcal{A}_U(B, \ast)(\mathbb{C})$, and let $V' = H_x(A, \mathbb{Q})$. This defines a class $c(x) \in H^1(\mathbb{Q}, G)$. Now $\eta$ defines isomorphisms $V_{\mathbb{Q}_p} \sim_{\mathbb{Q}_p} V'_{\mathbb{Q}_p}$ for all finite primes $p$, so $c(x)$ becomes trivial in $H^1(\mathbb{Q}_p, G)$ for all finite $p$. Moreover, the conditions (2.2.1-4) imply that $c(x)$ becomes trivial.
in $H^1(\mathbb{R}, G)$ as well. Thus $c(x) \in \ker^1(\mathbb{Q}, G)$. Note that in any case, $V$ and $V'$ are isomorphic as $B$-modules, so only the polarization makes a difference.

There is no reason to assume the class $c(x) \in \ker^1(\mathbb{Q}, G)$ vanishes. One can determine $\ker^1(\mathbb{Q}, G)$ explicitly: it is a finite group, trivial when $n$ is even, and isomorphic to

$$\ker\left[\mathbb{F}^+, \mathbb{Q}^N_{F/F^+}(\mathbb{F}^\times) \longrightarrow \mathbb{A}^X_{F^+}/\mathbb{A}^X_{F/F^+}(\mathbb{A}^\times_F)\right].$$

when $n$ is odd. This is an elementary calculation (found on p. 394 of [K5]). We index the elements of $\ker^1(\mathbb{Q}, G)$ by $c_i, i = 1, \ldots, \kappa$, with $c_1 = 0$, and let

$$S^1_U(B, *)(\mathbb{C}) = \{x \in A_U(B, *)(\mathbb{C}) \mid c(x) = c_i\}.$$

I will show that each $S^1_U(B, *)(\mathbb{C})$ is the set of complex points of a canonical model of $\text{Sh}_U(G, X)$.

Indeed, suppose $x \in S^1_U(B, *)(\mathbb{C})$. First set $i = 1$. One thus has $H_1(A) \xrightarrow{\sim} V$ as skew-hermitian $B$-modules, and we choose an isomorphism $\iota : H_1(A) \xrightarrow{\sim} V$. Via $\iota$, the datum $\eta$ defines a point in $G(A_f)/U$. On the other hand, the complex structure on $H_1(A, \mathbb{R}) = \text{Lie}(A)$ defines a map $h' : R_{C/B}G_m \xrightarrow{\sim} GL(V)$. Since the complex structure commutes with $B$, $h'$ takes values in $B^{\text{op}, \times}$. Again, the conditions on the $B$-action on $\text{Lie}(A)$ and the positivity of the Rosati involution imply that $h' \in X$ (= the set of polarized Hodge structures of a certain type). Thus we obtain a point $\tilde{x}(\iota) \in X \times G(A_f)/U$. The choice of $\iota$ is well-defined up to an element of $G(\mathbb{Q})$, thus $x$ gives a well-defined point in $G(\mathbb{Q})\backslash(X \times G(A_f)/U) = \text{Sh}_U(G, X)(\mathbb{C})$. We thus have a map

$$S^1_U(B, *)(\mathbb{C}) \longrightarrow \text{Sh}(G, X)(\mathbb{C}).$$

By the theory of Riemann matrices, this map is a bijection. Indeed, one can recover the abelian variety $A$ from the vector space $V$ and the complex structure $h'$, at least up to isogeny; then the point in $G(A_f)/U$ gives $A$ in terms of a correct choice of lattice. On the other hand, every point in $X \times G(A_f)/U$ corresponds to a polarization on $V(\mathbb{R})$ and a lattice (with level structure) with respect to which the polarization is integral, hence to a complex abelian variety, and the additional structures are automatic, by the discussion above.

For general $i$, we have to start with an isomorphism $i^i : H_1(A) \xrightarrow{\sim} V^i$; then the same argument goes through with $G$ replaced by $G^i = \text{Aut}_B(V^i, (\cdot)_i)$. Note that nothing changes except the set of $i^i$. The procedure for relating abelian varieties over $\mathbb{C}$ to pairs consisting of an archimedean datum (in $X$) and a finite-adelic datum (in $G^i(A_f)/U = G(A_f)/U$), modulo a global datum (in $G^i(\mathbb{Q})$) is worth recalling here, since it is the model for what will be used to study the points over finite fields.

We note that in fact $G^i = G$ for all $i$. This is a consequence of the following lemma:

**(2.3.1) Lemma.** — The natural map $\ker^1(\mathbb{Q}, Z_G) \rightarrow \ker^1(\mathbb{Q}, G)$ is surjective.

Indeed, this implies that the twist of the hermitian space $V$ is induced by a twist coming from $Z_G$, hence one that has trivial image in $G^\text{ad} = \text{Aut}(G)^0$, hence defines a
trivial twist of \( G \). To prove that \( G^i = G \), we could also appeal to the Hasse principle for adjoint groups. However, lemma (2.3.1) will be used repeatedly in the second half of the course, so I sketch a proof here, due to Kottwitz. First, let \( D = G/G^{\text{der}} \). Since \( G^{\text{der}} \) is an inner form of \( SL(n) \), it satisfies the Hasse principle, and it follows from a simple diagram chase that \( \ker^1(Q, G) \to \ker^1(Q, D) \) is injective. Surjectivity is a bit trickier. Let \( T \subset G \) be a maximal torus, elliptic at some finite place, \( T_{\text{sc}} = T \cap G^{\text{der}} \). Then the short exact sequence

\[
1 \to T_{\text{sc}} \to T \to D \to 1
\]

yields a commutative diagram of long exact sequences

\[
\begin{array}{c}
\cdots \\
H^1(Q, T_{\text{sc}}) \to H^1(Q, T) \to H^1(Q, D) \to H^2(Q, T_{\text{sc}}) \\
\cdots \\
H^1(A, T_{\text{sc}}) \to H^1(A, T) \to H^1(A, D) \to H^2(A, T_{\text{sc}}) \\
\end{array}
\]

(2.3.2)

Now we have the following

\[\textbf{(2.3.3) Lemma ([K5, pp. 421-422])}. \quad \text{Let } T \text{ be a torus over } \mathbb{Q}. \text{ The group }\]

\[\ker^2(Q, T) = \ker[H^2(Q, T) \to H^2(A, T)]\]

vanishes if \( T \) is anisotropic locally at one place.

In our situation, \( T_{\text{sc}} \) is elliptic at some finite place, hence the Lemma applies. It follows that any \( y \in \ker^1(Q, D) \) comes from an \( x \in H^1(Q, T) \) whose image in \( H^1(A, T) \) comes from \( H^1(A, T_{\text{sc}}) \). Since \( H^1(Q, T_{\text{sc}}) \) maps onto \( H^1(\mathbb{R}, T_{\text{sc}}) \) (another well-known general fact, cf. [Ha, Thm. A.12]), we can replace \( x \) by \( x' \) with trivial image in \( H^1(\mathbb{R}, T) \). Let \( z \) denote the image of \( x' \) in \( H^1(Q, G) \). Clearly it maps onto \( y \), and it remains to show \( z \in \ker^1(Q, G) \). By construction, it has trivial component at \( \mathbb{R} \), and since \( H^1(Q, G^{\text{der}}) = 1 \) (by Kneser’s theorem, since \( G^{\text{der}} \) is simply connected) it is in fact in \( \ker^1(Q, G) \).

On the other hand, forgetting the \( B \) action yields a map of Shimura data \((G^i, X) \to (GSp(V^i), \mathfrak{S}^{\pm})\), hence realizes \( \text{Sh}(G^i, X) \) as a canonical model defined over its reflex field \( F \) by the general theory of Shimura varieties. In particular, the subvarieties \( S_{ij}(B, \ast) \) of \( A_{ij}(B, \ast) \) are defined over \( F \).

2.4. Discussion of the moduli problem in étale level. — Now choose a prime \( u \) of \( E \) above \( p \), and let \( w = w_1, w_2, \ldots, w_r \) be the primes of \( F \) above \( u \). Write \( K = F_w \). Since \( p \) splits in \( E \), we can identify

\[G(Q_p) \xrightarrow{\sim} GL(n, K) \times \prod_{i>1} B^\text{opp}_{w_i} \times \mathbb{Q}_p^\times\]

(cf. (1.2.3)) where the map \( G(Q_p) \to \mathbb{Q}_p^\times \) is given by \( \nu \).

Henceforward, we write \( \mathcal{O} = \mathcal{O}_w \). We assume \( U \) factors as \( U_p \times U_p \), with \( U_p \) sufficiently small, and we further assume \( U_p = \prod_i U_{w_i} \times \mathbb{Z}_p^\times \), with respect to (2.4.1).
Assume \( U_w = U_{w_1} = GL(n, \mathcal{O}) \). Then \( \mathcal{A}_U(B, \ast) \) is representable, hence has a model over \( \text{Spec}(\mathcal{O}) \), also denoted \( \mathcal{A}_U(B, \ast) \), that represents a slightly modified version of the functor considered above. First, we always take \( \lambda \) to be a prime-to-\( p \)-polarization, and the equivalence is up to prime-to-\( p \)-isogenies. More importantly, \( \text{Lie}(A) \) becomes a module over \( \mathcal{O}_S \otimes \mathbb{Z}_p \mathcal{O}_{B, p} \), hence over \( \mathcal{O}_S \otimes \mathbb{Z}_p \mathcal{O}_{F, p} \). Then conditions (2.2.1-4) are replaced by

1. \( \text{Lie}(A) \otimes \mathcal{O}_{F, p} \mathcal{O}_{w, i} = 0, i > 1 \);  
2. \( \text{Lie}(A) \otimes \mathcal{O}_{F, p} \mathcal{O} \) is a projective \( \mathcal{O}_S \) module of rank \( n \), on which \( \mathcal{O} \) acts via the structural morphism \( \mathcal{O} \to \mathcal{O}_S \).

The remaining ranks are automatically determined by the polarization condition. One verifies easily that on the generic fiber we recover the moduli problem defined in (2.2).

As above, \( \mathcal{A}_U(B, \ast) \) is the union of \( |\ker^1(F, G)| \) copies of a \( \mathcal{O} \)-model \( S_U(G, X) \) of \( \mathbb{K} \text{Sh}(G, X) \).

**Theorem.** — The scheme \( \mathcal{A}_U(B, \ast) \) is smooth and projective over \( \mathcal{O} \).

**Proof.** — We follow Carayol [Ca1]. First, \( \mathcal{A}_U(B, \ast) \) is projective: since there is an embedding in the moduli space of polarized abelian varieties, it suffices to show it is proper. We prove this by the valuative criterion. Let \( R \) be a discrete valuation ring over \( \mathcal{O} \), \( S = \text{Spec}(R) \), and suppose we have a quadruple \((A, \lambda, i, \eta)\) over the generic point \( \text{Spec}(\mathbb{K}) \). We need to extend it to a quadruple over \( S \). Let \( A_R \) denote the Néron model of \( A \) over \( R \). It makes no difference if we replace \( S \) by a finite cover, so by the semi-stable reduction theorem of Grothendieck the special fiber \( A_k \) is an extension of an abelian variety by a torus \( T \). Now there is an isomorphism \( \text{End}(A_R) \cong \text{End}(A_K) \) (functoriality of Néron models). Thus \( \mathcal{O}_B \) acts on \( A_k \), hence necessarily on \( T \), hence on the character group \( X_\ast(T) \). This group has \( \mathbb{Z} \)-rank at most equal to the \( \text{dim} A = dn^2 \), whereas \( \mathcal{O}_B \) is of \( \mathbb{Z} \)-rank \( 2dn^2 \). Since \( B \) is a division algebra, any \( \mathcal{O}_B \)-module must have rank a multiple of \( 2dn^2 \), which implies \( T \) is trivial. Thus \( A_R \) is an abelian scheme, and we have already extended \( i \). The extension of \( \lambda \) follows similarly by functoriality.

Finally, there is the question of extending \( \eta \). The components of \( \eta \) away from \( p \) extend, because the \( \ell \)-division points are étale over \( S \). So we need only worry about extending \( \eta_p \). This is the right time to introduce the theme of \( p \)-divisible groups, which will occupy the next two lectures and will recur in those that follow. Let \( A_R[p^\infty] \) denote the \( p \)-divisible group associated to \( A_R \); it is the direct limit of the finite flat group schemes \( A_R[p^n] \). The maximal order \( \mathcal{O}_B \) acts on \( A_R[p^\infty] \), and this extends to an action of \( \mathcal{O}_B \otimes Z_p \simeq \prod_i \mathcal{O}_B \otimes \mathcal{O}_{w_i} \otimes \mathbb{O}_{w_i} \). Let \( \mathcal{O}_B_i \) denote the corresponding factor. This is a direct product, hence we have a decomposition

\[
A_R[p^\infty] \simeq \oplus_i (A_R[w_i^\infty] \oplus A_R[w_i^{c, \infty}]),
\]
where $A_R[w_i^∞]$ is a $p$-divisible group with $O_B$, action. The condition that the Rosati involution restricts to the involution $\ast$ of the second kind on $O_B$ implies that the polarization identifies
\[ A_R[w_i^∞] \sim \widehat{A}_R[w_i^∞], \]
where $\widehat{}$ denotes Cartier dual.

In general the data “polarization + $\eta_p$ (mod $U_p$)” is equivalent to “level structure on $A_R[w_i^∞] \pmod{U_w}$ for all $i$ + trivialization of the Tate module of $G_m$” (mod a subgroup of $Q_p^\times$). The fact that $\eta$ is invariant under the factor $\mathbb{Z}_p^\times \subset U_p$ implies that we only have to consider level structures on $A_R[w_i^∞]$. The condition on the Lie algebras implies $A_R[w_i^∞]$ is étale for $i > 1$, so the factor $\eta_{w_i}$ extends over $S$ for $i > 1$. Finally, we have chosen $U_w$ maximal, so there is no level structure at $w$, hence nothing to extend.

Now to prove smoothness, we use Grothendieck’s infinitesimal criterion. We let $\mathfrak{S} = \mathfrak{S}_U(G, X)$ denote the special fiber of our model, and let $x = (A_x, \lambda_x, i_x, \eta_x) \in \mathfrak{S}(\mathbb{F})$ be a geometric point. Let $S$ be an Artinian local $O$-algebra with residue field $\mathbb{F}$. We need to show

(2.4.5) Deformation property. — Let $I \subset S$ be an ideal, and let $x'$ be a lifting of the geometric point $x \in \mathfrak{S}(\mathbb{F})$ to an $S/I$-valued point of $\mathcal{A}_U(B, \ast)$. Then $x'$ lifts to an $S$-valued point of $\mathcal{A}_U(B, \ast)$.

The deformation property (2.4.5) is in fact a property of the formal completion $\mathcal{A}_U(B, \ast)_\wedge$ of $\mathcal{A}_U(B, \ast)$ at $x$, a formal scheme over $\text{Spec}(O)$. It therefore suffices to prove that $\mathcal{A}_U(B, \ast)_\wedge$ is formally smooth. In fact, we will prove that $\mathcal{A}_U(B, \ast)_\wedge$ is isomorphic to the formal spectrum of a power series ring. The construction of this isomorphism will occupy the rest of the section.

We reformulate the problem as follows. We consider the functor $\mathcal{F}_1$ on $\text{Art}(O, \mathbb{F})$:
\[ S \mapsto (A, \lambda, i, \eta) + j : (A, \lambda, i, \eta) \mathbb{F} \sim (A_x, \lambda_x, i_x, \eta_x) \]
This is represented by the formal completion $\mathcal{A}_U(B, \ast)_\wedge$. Consider the second functor $\mathcal{F}_2$
\[ S \mapsto (G, \lambda, i, \eta^w) + j : (G, \lambda, i, \eta^w) \mathbb{F} \sim (A_x[p^\infty], \lambda_x[p^\infty], i_x[p^\infty], \eta_x^w) \]
The terms need to be explained. Here $G$ is a $p$-divisible group scheme over $S$, $\lambda$ an isomorphism $\widehat{G} \sim \widehat{G}$, $i$ and inclusion $O_B \otimes_{\mathbb{Z}} \mathbb{Z}_p \to \text{End}(G)$, and $\eta^w$ a $U$-level structure on the prime-to-$w$ Tate module of $A$ away from $w$ (this makes sense because $T^w(A)$ extends uniquely to any $S \in \text{Art}(O, \mathbb{F})$ as an étale sheaf).

(2.4.6) Serre-Tate Theorem. — The morphism $\mathcal{F}_1 \to \mathcal{F}_2$ is an isomorphism of functors.

Thus to determine the infinitesimal structure of $\mathcal{A}_U(B, \ast)$, it suffices to study the functor $\mathcal{F}_2$. Obviously, the structure $\eta^w$ is étale, as is the deformation of the data $A_x[w_i^∞]$ for $i > 1$. So it suffices to study the deformation of $A_x[w_1^∞]$. Now $O_{B_w} \sim \wedge$.
$M(n, \mathcal{O})$, by our original hypothesis. Thus there are $n$ orthogonal idempotents in $\mathcal{O}_B$, which decompose $A_x[w^n]$ into $n$ mutually isomorphic $p$-divisible groups with $\mathcal{O}$ action; this argument is called "Morita equivalence". Let $G_x$ denote any of these divisible $\mathcal{O}$-modules, $\iota_x : \mathcal{O} \to \mathrm{End}(G_x)$ the action and let $\mathcal{F}_3$ be the functor

$$S \mapsto (G, \iota) + j : (G, \iota) \to (G_x, \iota_x).$$

Since the remaining data are étale, the natural map $\mathcal{F}_2 \to \mathcal{F}_3$ is again an isomorphism.

Now the projective $\mathcal{O}_S$-module $\mathrm{Lie}(A_x) \otimes_{\mathcal{O}_B} \mathcal{O}_B$ is isomorphic to the sum of $n$ copies of $\mathrm{Lie}(G_x)$ (by Morita equivalence). It follows from the definition of the moduli problem that $\mathrm{Lie}(G_x)$ is a projective (i.e. free) rank 1 $\mathcal{O}_S$-module. On the other hand, the height of the $p$-divisible group $G_x$ is $n[K : \mathbb{Q}_p]$ (because $A_x[p]$ is a finite flat $p$-group scheme of rank $2 \dim A_x$). Indeed, the polarization breaks up $A_x[p]$ as $A_1[p] \times A_2[p]$ each of height $\dim(A)$, with $A_1[p] = A_x[p] \cap \prod A_x[w^n_i]$. Since all but one of these is étale, the height of $A_x[w]$ is determined, and one computes directly that the height is precisely $n[K : \mathbb{Q}_p]$.

(2.4.7) **Definition.** — Let $S$ be a scheme over $\mathcal{O}$, and choose a uniformizer $\varpi$ of $\mathcal{O}$. A $p$-divisible $\mathcal{O}$-module of height $h$ is a $p$-divisible group scheme $G$ over $S$ with an action $i : \mathcal{O} \to \mathrm{End}(G)$ such that

(i) for every pair of integers $m_1 > m_2$, the natural sequence

$$0 \to G[\varpi^{m_2}] \to G[\varpi^{m_1}] \to G[\varpi^{m_1 - m_2}] \to 0$$

is an exact sequence of finite flat group schemes;

(ii) the action of $\mathcal{O}$ on $\mathrm{Lie}(G)$ is given by the structural morphism $\mathcal{O} \to \mathcal{O}_S$.

The height of the $p$-divisible $\mathcal{O}$-module $G$ is defined to be the $h$ such that $G[\varpi]$ is a finite flat $k(w)$-vector group scheme of rank $h$.

Thus the height of $A_x[w]$ as $\mathcal{O}$-module is just $n$, and (2.4.7) is the functor classifying deformations of $(G_x, \iota_x)$ as a 1-dimensional height $n$ divisible $\mathcal{O}$-module.

We consider the canonical exact sequence

$$0 \to G_x^0 \to G_x \to G_x^{\text{ét}} \to 0$$

and let $h$ denote the height of $G_x^0$, so $G_x^0$ is a formal $\mathcal{O}$-module of height $n - h$.

For $\mathcal{O} = \mathbb{Z}_p$, and when $h = 0$, the deformation problem was solved by Lubin-Tate in 1966 [LT]. The general problem was solved by Drinfel’d [Dr]. I will follow his account and that of Hopkins-Gross [HG] (Equivariant vector bundles on the Lubin-Tate moduli space, Contemporary Math., 158 (1994), p. 23-88), skipping many details as the argument progresses. We begin with the case $h = 0$, and consider a 1-dimensional formal $\mathcal{O}$-module $F$ over $\mathcal{O}$ of height $n$. Consider the category $\mathrm{Art}(\mathcal{O}, \mathcal{F})$ of Artinian local $\mathcal{O}$-algebras $R$ with maximal ideal $m = m_R$ (containing $\varpi$) and residue field $\mathcal{F}$, and consider the functor of deformations of $F$ on $\mathrm{Art}(\mathcal{O}, \mathcal{F})$; i.e., $p$-divisible $\mathcal{O}$-modules $G$ over $\text{Spec}(R)$ given with isomorphisms $j : G_R \isom F$. Because $\text{Spec}(R)$ is
infinitesimal, $G$ is in fact a formal group, hence is given by power series: the addition law $G(X, Y)$ and multiplication $a_G(X)$ for $a \in \mathcal{O}$. To say that $G$ is a deformation of $F$ is to say that $G \equiv F \pmod{m}$ and $a_G \equiv a_F \pmod{m}$ for all $a \in \mathcal{O}$.

The difference is given by a 2-cocycle $(\Delta(X, Y), \delta_a(X))$. First, a cochain is just a collection of power series as above without constant terms. They form a (symmetric) 2-cocycle for $F$ if

$$\Delta(X, Y) = \Delta(Y, X)$$

$$\Delta(Y, Z) + \Delta(X, Y + F Z) = \Delta(X + F Y, Z) + \Delta(X, Y).$$

(Here the symbol $Y +_F Z$ means $F(Y, Z)$, etc.)

$$\delta_a(X) + \delta_a(Y) + \Delta(a_F(X), a_F(Y)) = a\Delta(X, Y) + \delta_a(X +_F Y)$$

$$\delta_a(X) + \delta_b(X) + \Delta(a_F(X), b_F(X)) = \delta_{a+b}(X)$$

$$a\delta_b(X) + \delta_a(b_F(X)) = \delta_{ab}(X)$$

Given a $\psi \in R[[X]]$ with $\psi(0) = 0$, we define the coboundary

$$\Delta(\psi)(X, Y) = \psi(Y) - \psi(F(X, Y)) + \psi(X)$$

$$\delta_a(\psi(X)) = a\psi(X) - \psi(a_F(X))$$

Then $H^2(F, R)$, the symmetric 2-cocycles with values in $R$, modulo coboundaries, classify isomorphism classes of deformations of $F$ to $R$, by

$$(\Delta, \delta_a) \mapsto G(X, Y) = (F(F(X, Y), \Delta(X, Y)), a_G(X) = F(a_F(X), \delta_a(X))).$$

The verification is by direct calculation, just as in Lubin-Tate.

The problem is then to find an explicit basis for $H^2(F, R)$.

(2.4.10) Theorem (Drinfel’d). — There is a functorial bijection between $m_R^{n-1}$ and the set of deformations of $F$ to $R$.

Note that $m_R^{n-1}$ is naturally equal to the set of continuous $\mathcal{O}$-algebra homomorphisms from the power series ring $R_{n, \mathcal{O}} = \mathcal{O}[[t_1, \ldots, t_{n-1}]]$ to $R$. Thus the functor of deformations of $F$ is prorepresented (on the category of complete (noetherian) local $\mathcal{O}$-algebras – by passage to the limit) by $\text{Spf}(\mathcal{O}[[t_1, \ldots, t_{n-1}]])$; i.e., $\mathcal{F}_3$ is prorepresented by a power series ring. In particular, taking $F$ to be the formal group $G_x$ above, we see that $\mathcal{F}_i$ is formally smooth for $i = 1, 2, 3$, which implies that $A_U(B, \star)$ is smooth at any point $x$ where $G^*_x = 0$.

The proof of Drinfel’d’s theorem, like that of Lubin-Tate, is also a direct calculation. One shows by hand that any deformation can be written in such a way that $\Delta$ and $\delta_a$ have the form

$$(\Delta, \delta_a) = \sum_{i=1}^{n-1} (\Delta_i, \delta_{a,i}) + (\deg q^{n-1} + 1)$$

where $\Delta_i$ and $\delta_{a,i}$ are homogeneous of degree $q^i$; then one shows that each $(\Delta_i, \delta_{a,i})$ is unique up to an (arbitrary) scalar in $m_R$, for $i = 1, \ldots, n-1$, and that they determine
the remainder of the deformation. Explicitly, if \((\Delta, \delta_a)\) is a cocycle, congruent to \((\text{mod } \deg n + 1)\), then
\[
(\Delta, \delta_a) \equiv (c(X + Y)^n - X^n - Y^n, c(a^n - a)X^n) \pmod{\deg n + 1}
\]
if \(n\) is not a power of \(q\) (and hence is cohomologous to 0 \((\text{mod } \deg n + 1)\)), whereas
\[
(\Delta, \delta_a) \equiv \left(c \frac{p}{\omega}[(X + Y)^n - X^n - Y^n], c \frac{a^n - a}{\omega}X^n\right) \pmod{\deg n + 1}
\]
if \(n\) is a power of \(q\). This is exactly as in Lubin-Tate except for the presence of the \(\delta_a\).

By writing down the power series, we obtain a universal deformation over \(\text{Spf}(\mathcal{O}_n, \mathcal{O})\). Taking successive subgroups of \(\omega^m\)-division points, we obtain a \(p\)-divisible \(\mathcal{O}\)-module \(\tilde{\Sigma}_{K,n}\) over \(\text{Spf}(\mathcal{O}_n, \mathcal{O})\). It is not hard to see, because it is a direct limit of finite flat group schemes over \(\text{Spf}(\mathcal{O}_n, \mathcal{O})\), that in fact \(\tilde{\Sigma}_{K,n}\) is actually a \(p\)-divisible \(\mathcal{O}\)-module over \(\text{Spec}(\mathcal{O}_n, \mathcal{O})\), and not merely over the formal completion. However, \(\tilde{\Sigma}_{K,n}\) is no longer formal (consider the pullback of the universal elliptic curve to the formal completion at a supersingular point). This is an elementary, but striking illustration of the difference between formal and algebraic geometry that creates most of the difficulty in the study of the bad reduction of the moduli space.

It is known that up to isomorphism, there is a unique 1-dimensional \(p\)-divisible \(\mathcal{O}\)-module \(\Sigma_{K,n}\) of height \(n\) over \(\mathbb{F}\), with endomorphism ring isomorphic to \(\mathcal{O}_{D_1/n}\), the maximal order in the central division algebra \(D_1/n\) over \(K\) with invariant \(1/n\). This can be proved by explicit power series calculations, using the techniques of the Lubin-Tate theory; see Drinfeld’s paper for such a proof. One construction is by taking the reduction \(\text{mod } \omega\) of (any) Lubin-Tate formal group for \(\mathcal{O}_{n}\), the unramified extension of \(\mathcal{O}\) of degree \(n\). Another construction will be discussed next week.

So much for the case \(h = 0\). Now suppose \(h\) arbitrary. We have the universal deformation \(\tilde{\Sigma}_{K,n-h}\) over \(\text{Spf}(\mathcal{O}_n, \mathcal{O})\). Let \(R \in \text{Art}(\mathcal{O}, \mathbb{F})\). Over \(\mathbb{F}\), we have an isomorphism
\[
F \cong F^0 \times (K/\mathcal{O})^h
\]
where \(F^0 \cong \Sigma_{K,n-h}\).

**Theorem (Drinfel’d).** — The functor of deformations of \(\Sigma_{K,n-h} \times (K/\mathcal{O})^h\) is prorepresented by a power series ring in \(n - 1 = (n - h - 1) + h\) variables, and canonically by the \(h\)-fold fiber product of \(\tilde{\Sigma}_{K,n-h}\) over \(\text{Spf}(\mathcal{O}_n, \mathcal{O})\).

Theorem 2.4.11, combined with the Serre-Tate Theorem 2.4.6, completes the proof of the deformation property 2.4.5, hence of the smoothness assertion of Theorem 2.4.4. As for Theorem 2.4.11, its proof is based on an argument due to Messing, that goes as follows. Evidently, any deformation \(G\) of \(\Sigma_{K,n-h} \times (K/\mathcal{O})^h\) to \(R\) has an exact sequence as in (2.4.9):
\[
0 \rightarrow G^0 \rightarrow G \rightarrow G^{\text{ét}} \rightarrow 0.
\]
Here \(G^0\) is a deformation of \(\Sigma_{K,n-h}\) and \(G^{\text{ét}}\) is a deformation of \((K/\mathcal{O})^h\), hence is isomorphic to \((K/\mathcal{O})^h\) since the latter admits no deformations. So we have to
classify extensions of deformations $G^0$ of $\Sigma_{K,n-h}$ by $(K/O)^h$. Formally, the short exact sequence

$$0 \rightarrow \mathcal{O}^h \rightarrow K^h = \lim_{\rightarrow} \mathcal{O}^h \rightarrow (K/O)^h \rightarrow 0$$

of sheaves yields a long exact sequence (of sheaves) with terms

$$\lim_{\rightarrow} \text{Hom}_{\mathcal{O}}(\mathcal{O}^h, G^0) \rightarrow \text{Hom}_{\mathcal{O}}(\mathcal{O}^h, G^0) \xrightarrow{\delta_R} \text{Ext}^1(\mathcal{O}^h, G^0) \rightarrow \text{Ext}^1(\lim_{\rightarrow} \mathcal{O}^h, G^0).$$

Now \(\text{Hom}_{\mathcal{O}}(\mathcal{O}^h, G^0)\) is represented by \(\text{Hom}_{\mathcal{O}}(\mathcal{O}^h, m_R)\). Since multiplication by \(p\) is contracting on \(m_R\) and \(m_R\) is nilpotent, the inverse limit is zero.

To conclude, it suffices to show that the map

$$\text{Ext}^1(K/O, G^0) \rightarrow \text{Ext}^1(\lim_{\rightarrow} \mathcal{O}^h, G^0)$$

is zero, in other words, that any extension \(G\) of \(G^0\) by an étale \(\mathcal{O}\)-module that is split at the closed point of \(R\) splits over \(R\) upon multiplication by a sufficiently high power of \(p\). (Here and above, the arguments, apparently merely heuristic, can be made rigorous, as in the proof of Proposition 2.5 of the Appendix of [Me].) It suffices to show that the map

$$\text{Hom}_{\mathcal{O}}(G, K/O \times G^0) \otimes K \rightarrow \text{Hom}_{\mathcal{O}}(G_{\mathbb{F}}, K/O \times \Sigma_{K,n-h}) \otimes K$$

(restriction to the closed point) is an isomorphism.

More generally, we have

\section{Theorem (Drinfel’d’s theorem on rigidity of quasi-isogenies)}

Let \(S\) be a scheme on which \(p\) is locally nilpotent, and let \(S_0\) be the subscheme defined by a nilpotent sheaf of ideals. Let \(G_1\) and \(G_2\) be two \(p\)-divisible groups over \(S\). Then restriction to \(S_0\) defines an isomorphism (of sheaves):

$$\text{Hom}(G_1, G_2) \otimes \mathbb{Q}_p \xrightarrow{\sim} \text{Hom}(G_{1,S_0}, G_{2,S_0}) \otimes \mathbb{Q}_p.$$

In other words, any map from \(G_1\) to \(G_2\) over \(S_0\) lifts uniquely to \(S\) after multiplication by a sufficiently high power of \(p\).

This theorem, which we will use repeatedly, is also the basis of Drinfel’d’s simple proof of the Serre-Tate theorem. There is a very readable proof by Katz in LNM 868, Surfaces Algébriques, pp. 141-143.

The above discussion is based on the uniqueness of \(\Sigma_{K,g}\) up to isomorphism over \(\mathbb{F}\), and the isomorphisms of formal completions are so far only rational over \(\mathbb{F}\). Next week I will explain how to descend to \(\mathbb{F}_q\).
2.5. **Hecke correspondences away from** $p$. — We continue to work over $\text{Spec}(\mathcal{O})$. Suppose $U \supset U'$ are two open compact subgroups of $G(A_f)$ with $U_w = U'_w = \text{GL}(n, \mathcal{O})$ as before and $U_w = U^p \times \prod_i U_{w_i} \times \mathbb{Z}_p^\times \supset U'^w$ (again $U'^p \supset \mathbb{Z}_p^\times$). Then there is a finite morphism $\mathcal{A}_U(B, *) \to \mathcal{A}_U(B, *)$. Since the prime-to-$p$ torsion subgroups are étale and since level structures at $w_i$ are also étale, this projection is étale.

Define

\[ G(A_f^w) = G(A_f^p) \otimes \prod_{i > 1} B_{w_i}^{\text{op}, X}, \]

so that

\[ G(A_f) = G_w \times \mathbb{Q}_p^\times \times G(A_f^w) = \text{GL}(n, K) \times \mathbb{Q}_p^\times \times G(A_f^p). \]

Thus any admissible irreducible representation $\pi$ of $G(A_f)$ can be factored

\[ \pi = \pi_w \otimes \psi \otimes \pi^w, \]

where $\pi_w \in \mathcal{A}(n, K)$, $\psi$ is a character of $\mathbb{Q}_p^\times$, and $\pi^w$ is an admissible irreducible representation of $G(A_f^w)$. In what follows, we will try to ignore $\psi$.

Now suppose $g \in G(A_f^p)$. For $U'$ we take $U \cap gUg^{-1}$. Then there are two maps, $p_1, p_2 : \mathcal{A}_{U'}(B, *) \to \mathcal{A}_U(B, *)$, with $p_1$ given by the inclusion $U' \subset K$ and $p_2$ the composition

\[ \mathcal{A}_{U'}(B, *) \to \mathcal{A}_{gUg^{-1}}(B, *) \xrightarrow{g} \mathcal{A}_U(B, *) \]

There is then a map (Hecke correspondence)

\[ T(g) = p_2 \circ p_1^*: H^*(\mathcal{A}_U(B, *), \mathbb{Q}_\ell) \to H^*(\mathcal{A}_{U'}(B, *), \mathbb{Q}_\ell). \]

The goal of next week’s lecture will be to explain how to extend this to allow level structures at $w$ and Hecke operators with non-trivial components at $w$.

3. **$p$-divisible $\mathcal{O}$-modules and Drinfel’d bases**

Today I will deal with the most tiresome part of the construction (le point le plus fastidieux du manuscrit, as Carayol wrote in his Bourbaki report), namely the explanation of the models of Shimura varieties with bad reduction and the definition of the group actions. Rather than give all the details, I will try to explain why it works. For these varieties, the construction is relatively explicit and uses strongly that we are dealing with 1-dimensional formal $\mathcal{O}$-modules.

3.1. **Dieudonné modules and formal $\mathcal{O}$-modules.** — Let $K$ be a finite extension of $\mathbb{Q}_p$, $\mathcal{O} = \mathcal{O}_K$ its ring of integers with maximal ideal $\mathfrak{p}_K$ and residue field $k = \mathbb{F}_q$, with algebraic closure $\mathbb{F}$. Let $W$ be the ring of Witt vectors of $\mathbb{F}$, $\mathcal{K}$ the fraction field of $W$; i.e. $\mathcal{K}$ is what is denoted $\mathbb{K}'$ in [HT], and let $\sigma$ denote the Frobenius (relative to $\mathbb{Q}_p$) acting on $\mathcal{K}$. We let $\mathcal{K}_K = \mathcal{K} \cdot \mathbb{K}$ and $W_K$ be the integral closure of $W$ in $\mathcal{K}_K$. For any positive integer $g$, we choose a one-dimensional formal $\mathcal{O}$-module
$\Sigma_{K,g}$ over $\mathbb{F}$ of height $g$. We write $K/\mathcal{O}$ for the étale height one $p$-divisible $\mathcal{O}$-module, and for any non-negative integer $h$ we let $\Sigma_{K,g,h} = \Sigma_{K,g} \times (K/\mathcal{O})^h$.

The uniqueness of $\Sigma_{K,g}$ up to isogeny, at least, follows from the classification of Dieudonné modules up to isogeny (isocrystals) over $\mathbb{F}$. For future reference (cf. (8.1)), we define an isocrystal to be a pair $(N, \phi)$ where $N$ is a finite-dimensional $K$-vector space and $\phi$ is a $\sigma$-linear bijection $N \to N$, in the sense that

$$\phi(av) = \sigma(a)\phi(v), \quad a \in K, \ v \in N.$$ 

The category of isocrystals is semisimple; i.e., every isocrystal is isomorphic to a sum of simple objects. Moreover, the simple objects are classified by rational numbers $r/s$ where $s = \dim N_{r/s}$ and $\phi^s(M) = p^rM$ for some $W$-lattice $M \subset N_{r/s}$. If $N \xrightarrow{\sim} \oplus N_{r_i/s_i}$ then the $r_i/s_i$ are the slopes of $N$.

To any $p$-divisible group $G$ over $\mathbb{F}$ one can associate its (contravariant) Dieudonné module $D(G)$, and its isocrystal $N(G) = D(G) \otimes_W K$. $D(G)$ is a $W$-free module of finite type over the non-commutative ring $W[F,V]$ with relations

$$Fa = \sigma(a)F; \ aV = V\sigma(a); \ FV = VF = p.$$ 

An isocrystal $N$ is attached to a $p$-divisible group if and only if all its slopes are in the interval $[0, 1]$. More precisely, the main theorem of Dieudonné theory (over perfect fields of characteristic $p$) is that the functor $G \mapsto D(G)$ is an anti-equivalence of categories with the category of $W[F,V]$-modules as above, and $N(G) \xrightarrow{\sim} N(G')$ as isocrystals if and only if $G$ and $G'$ are isogenous.

There is a similar classification of divisible $\mathcal{O}$-modules. Let $\sigma_q$ denote the lift of the Frobenius $\text{Frob}_q \in \text{Gal}(\mathbb{F}/k)$ to $\text{Gal}(K_K/K)$, and fix a uniformizer $\varpi \in \mathcal{O}$. Then a $\mathcal{O}$-Dieudonné module (resp. a $K$-isocrystal) is a $W_K[F,V]$-module where now the relations are

$$Fa = \sigma_q(a)F; \ aV = V\sigma_q(a); \ FV = VF = \varpi.$$ 

(resp. a $K$-vector space $N$ with $\sigma_q$-linear bijective morphism $\Phi$). If $G$ is a divisible $\mathcal{O}$-module, then $\mathbb{D}(G)$ has a natural $\mathcal{O}$-structure, and in this way it becomes a $\mathcal{O}$-Dieudonné module. The simple objects are again classified by slopes $r/s$; here $N_{r,s}$ has $F^s(M) = \varpi^rM$ for an appropriate lattice.

We have the relations

$$(3.1.1) \quad \text{height}(G) = \text{rank}_\mathcal{O} \mathbb{D}(G); \dim G = \dim_W V\mathbb{D}(G)/\varpi \mathbb{D}(G).$$ 

In particular, if $G$ is simple of slope $r/s$, then $s = \text{height}(G)$ and $r = \dim G$.

We can thus construct the 1-dimensional height $g$ formal $\mathcal{O}$-module as follows. Its slope is $1/g$. We take $N = K_K^g$. For any linear map $b \in GL(N)$, one can define a
σ_q-linear map φ_b = b · σ_q. We take

\[
\begin{pmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1 \\
\varpi & 0 & 0 & \ldots & 0
\end{pmatrix}
\]

Then (bσ_q)^g = \varpi · σ_q^g which implies that the slope is 1/g. This already shows uniqueness of Σ_K,g up to isogeny, and it is not hard to show that any two φ-invariant lattices are actually isomorphic. Or this can be done directly with power series, as in Drinfel’d, and we can arrange that

\[
f_\varpi(X) = X^{q^g}, \quad f_\zeta(X) = \zeta \cdot X
\]

for \(\zeta \in \mu_{q-1}\).

Drinfel’d proved the following results, generalizing the results of Lubin for one-dimensional formal groups:

**Proposition (Dr, Prop. 1.7)**

(i) The algebra \(\text{End}(N, \phi)^{op} = \text{End}(\Sigma_{K,g}) \otimes \mathbb{Q}_p\) is isomorphic to the central division algebra \(D_g = D_{K,g}\) over \(K\) with invariant 1/g.

(ii) This isomorphism identifies \(\text{End}(\Sigma_{K,g})\) with the maximal order \(\mathcal{O}_{D_{K,g}} \subset D_{K,g}\).

Let \(N : D_{K,g}^x \to K^x\) be the reduced norm, and let \(\Pi \in \mathcal{O}_{K,g} := \mathcal{O}_{D_{K,g}} = \text{End}(\Sigma_{K,g})\) be an element such that \(v(N(\Pi)) = 1\); we may even assume \(\Pi^g = \varpi\). Then there is an isogeny \(\Pi : \Sigma_{K,g} \to \Sigma_{K,g}\) with kernel a one-dimensional \(\mathcal{O}/\varpi\) vector space scheme. (The existence of such an isogeny, equivalent to (ii) of the Proposition, shows that any formal \(\mathcal{O}\)-module isogenous to \(\Sigma_{K,g}\) is isomorphic to \(\Sigma_{K,g}\).) On the other hand, because the action of \(\mathcal{O}\) on the Lie algebra is just the natural map \(\mathcal{O} \to \mathcal{O}/\varpi = \mathbb{F}_q \to \mathbb{F}\), we see that the morphism \(\text{Frob}_q : \Sigma_{K,g} \to \Sigma_{K,g}^{(q)}\) is a map of \(\mathcal{O}\)-modules. Thus \(\ker \text{Frob}_q = \ker \Pi\), which means that

\[
\Sigma_{K,g}^{(q)} \xrightarrow{\sim} \Sigma_{K,g}/(\ker \Pi) \xrightarrow{\sim} \Sigma_{K,g}.
\]

We have already defined quasi-isogenies: If \(A \in \text{Art}(\mathcal{O}, \mathbb{F})\) and \(H_1, H_2\) are two \(p\)-divisible \(\mathcal{O}\)-modules over \(A\), a quasi-isogeny between \(H_1\) and \(H_2\) is a global section \(f\) of the sheaf \(\text{Hom}_A(H_1, H_2) \otimes \mathbb{Q}\) such that \(p^a f\) is an isogeny for some \(f\). If \(\ker p^a f\) is a group of order \(p^b\), the height of \(f\) is then the integer \(b - a\). For any non-negative integer \(h\), the group of self-quasi-isogenies of \(\Sigma_{K,g,h}\) is isomorphic to \(D_x \times GL(h, K)\), where \(D_g\) is the central division algebra over \(K\) with invariant 1/g. A self-quasi-isogeny of height 0 of \(\Sigma_{K,g}\) is an invertible element of \(\mathcal{O}_{K,g}\), hence an automorphism of \(\Sigma_{K,g}\). Alternatively, every isogeny factors as a product of \(\Pi^a\) and an isomorphism (isomorphism on Lie algebras, hence isomorphism), for some \(a\). Here again, we are strongly using the one-dimensionality of \(\Sigma_{K,g}\).
We consider the functor $\text{QDef}(\Sigma_{K,g,h})$ from $\text{Art}(\mathcal{O}, \mathbb{F})$ to $\{\text{Sets}\}$:

$$A \mapsto (H/A, j : \Sigma_{K,g,h} \rightarrow H_{\mathbb{F}})$$

where $j$ is a quasi-isogeny. This functor is representable, as in [RZ], by a formal scheme $\hat{M}_{g,h}$ with infinitely many connected components. When $h = 0$ the components are indexed by the height of the quasi-isogeny in $\mathbb{Z}$, and indeed

$$\hat{M}_g \xrightarrow{\sim} \hat{M}_g(0) \times \mathbb{Z}$$

where $\hat{M}_g(0) = \text{Spf}(R_{g,K})$ represents pairs $(H, j)$ where $j$ is of height 0, hence an isomorphism.

This is the functor we studied in Lecture 2, represented by $\mathcal{O}[[u_1, \ldots, u_{g-1}]]$. The additional étale part adds first $h$ more variables to the power series ring; the connected components are indexed by $\mathbb{Z} \times \text{GL}(h, K)/\text{GL}(h, \mathcal{O})$ (quasi-isogenies of $(K/\mathcal{O})^h$ are indexed by lattices in $K^h$). We let $(\hat{\Sigma}_{K,g,h}, \tilde{j})$ denote the universal pair over $\hat{M}_{g,h}$.

We need something slightly more general: Let $\sim$ denote Cartier dual, and consider $\Sigma^+_{K,g,h} = \Sigma_{K,g,h} \times \hat{\Sigma}_{K,g,h}$. This $p$-divisible group has a canonical polarization $\psi : \Sigma^+_{K,g,h} \times \hat{\Sigma}^+_{K,g,h} \rightarrow \mu_{p^\infty}$, where $\mu_{p^\infty}$ denotes the $p$-divisible group of $\mathbb{G}_m$. The functor $\text{QDef}(\Sigma^+_{K,g,h})$ classifies pairs $(H^+/A, j : \Sigma^+_K \rightarrow H^+_{\mathbb{F}})$ where $j$ is required to respect the polarizations on the two sides up to a multiple in $\mathbb{Z}_p^{\times}$. It is represented by a formal scheme $\hat{M}^+_{g,h}$, which can be split canonically as $\hat{M}_{g,h} \times \mathbb{Q}_p^\times / \mathbb{Z}_p^\times$, with the second factor for the polarization. The universal pair over $\hat{M}^+_{g,h}$ is denoted $(\hat{\Sigma}^+_{K,g,h}, \tilde{j}^+)$. By analogy with (3.1.6), there is a non-canonical isomorphism

$$\hat{M}^+_{g,h} \xrightarrow{\sim} \hat{M}^+_{g,h}(0) \times \mathbb{Z} \times \text{GL}(h, K)/\text{GL}(h, \mathcal{O}) \times \mathbb{Q}_p^\times / \mathbb{Z}_p^\times,$$

where $\hat{M}^+_{g,h}(0)$ represents pairs $(H^+, j)$ such that $j$ is an isomorphism and such that the polarization is exact.

We have seen that, the formal completion of $\mathcal{A}_U(B, \ast)$ at a point $x$ of the special fiber is isomorphic to $\hat{M}_{g,h}$ for some $g + h = n$. This was proved over $\mathbb{F}$. Today we will carry out three additional steps:

1. First, we show how this descends to $\bar{\mathbb{F}}_q$. This can be done first on the special fiber; the Galois action lifts uniquely.
2. Next, we add (Drinfeld’s) level structures at $w$ and obtain a local uniformization with these level structures.
3. Finally, we show how the Hecke correspondences at $w$ extend to these integral models.

### 3.2. Uniformization of isogeny classes

We denote by $\mathfrak{S}_U$ the special fiber of $\mathcal{A}_U(B, \ast)$. Let $\mathfrak{S}_U^{(h)}$, or just $\mathfrak{S}^{(h)}$, be the set of points $x \in \mathfrak{S}_U(\bar{\mathbb{F}})$ such that $\mathfrak{g}_x^{(h)}$ is of height $h$. It is easy to see that this is a (reduced) subscheme; next time we’ll see it is smooth of dimension $h$. Fix $x \in \mathfrak{S}_U^{(h)}$, and consider the set $\mathfrak{S}(x) = \mathfrak{S}_U(x)$ of points in the isogeny class of $x$. It is obviously contained in $\mathfrak{S}^{(h)}$. Thus $\mathfrak{S}(x)$
consists of quadruples \( x' = (A', \lambda', i', \eta') \) such that there exists an isogeny \( \phi : A \to A' \) respecting the other structures. The kernel of \( \phi \) breaks up into the \( w \)-component and the prime-to-\( w \) component. The latter is a lattice in \( V_f^w(A) \), isomorphic via \( \eta^w \) to \( V(A_f^w) \). Since \( \phi \) respects the other structures, this gives a well-defined point in \( G(A_f^w)/U^w \). The \( w \)-component is the same as an isogeny of \( O \)-modules \( G_x \mapsto G_{x'} \). Taking \( G_x \) as our model for \( \Sigma_{K,g,h} \), with \( g = n - h \), we thus obtain a point of \( \hat{M}_{g,h}(\mathbb{F}) \).

Thus \( x' \) corresponds to a pair \((m, g^w) \in \hat{M}_{g,h}(\mathbb{F}) \times G(A_f^w)/U^w \). But this pair is only well-defined up to an element of \( I_x(\mathbb{Q}) = I(A, \lambda, i)(\mathbb{Q}) \), where \( I_x = I_{A, \lambda, i} \) is the group of self-(quasi)isogenies of the triple \((A, \lambda, i)\), acting diagonally on the two data. In this way, we obtain a bijection (uniformization of an isogeny class):

\[
\Theta : I_x(\mathbb{Q}) \setminus \hat{M}^+_g,h(\mathbb{F}) \times G(A_f^w)/U^w \longrightarrow \overline{\Sigma}(x).
\]

Injectivity is almost obvious: if two pairs \((m, g^w)\) give the same point \( x' \), then the composition of one isogeny with the inverse of the other gives a self-isogeny of \( A \) respecting all the data, hence an element of \( I_x(\mathbb{Q}) \), by definition.

The Serre-Tate theorem (2.4.6) then shows that this extends to an isomorphism of formal completions:

\[
\Theta : I_x(\mathbb{Q}) \setminus \hat{M}^+_g,h(\mathbb{F}) \times G(A_f^w)/U^w \longrightarrow \mathcal{A}_U(B, *) \overline{s}(x).
\]

The meaning of this formal completion along an isogeny class in the special fiber is explained in Rapoport-Zink [RZ, 6.22]; it is something like the formal disjoint union of the formal completions at the individual points.

Let me explain how this works on functors. Let \( R \in \text{Art}(O, \mathbb{F}) \), and \((m, g^w)\) a point in \( \hat{M}_{g,h}(R) \times G(A_f^w)/U^w \). Thus \( m \) corresponds to a pair \((H/R, j : \Sigma_{K,g,h} \to H_f) \). Recall that \( \Sigma_{K,g,h} \) is identified with \( G_x \) for the fixed basepoint. Lift \((A, \lambda, i)\) to \((A_1, \lambda_1, i_1) \in \mathcal{A}_U(B, *)(W)\) (any lifting). This is possible; indeed, we can even arrange that \((A_1, \lambda_1, i_1)\) comes from a certain CM type. Let \( G_{x,1} \) be the corresponding lifting of \( G_x \).

By rigidity of quasi-isogenies (2.4.13), the map \( j \) lifts to a quasi-isogeny \( j_1 : G_{x,1} \to H \). The kernel of this quasi-isogeny defines a (virtual) subgroup scheme \( S_m \subseteq A_1[[w]] \), whereas \( g^w \) defines a lattice \( T_{g^w} \subseteq V_f^w(A_1) \). Suppose for simplicity that \( S_m \) is a genuine subgroup scheme and \( T_{g^w} \supseteq T_f^w(A_1) \). Then the quotient by \( S_m \times (T_{g^w}/T_f^w(A_1)) \) is a new abelian scheme over \( R \), and this is the image of the point \((m, g^w)\). In general, one has to modify the construction to account for virtual subgroup schemes, but this is not difficult. At the end I will explain how this works on level \( w \)-structures in characteristic 0.

We want \( \Theta \) to be rational over \( \mathbb{F}_q \) in a certain sense. Rapoport and Zink construct a “Weil descent datum” on \( \hat{M}_{g,h} \), as follows. Let \( \sigma_q \) denote the (arithmetic) Frobenius automorphism in \( \text{Gal}(\overline{\mathbb{F}}/\mathbb{F}_q) \), and let \( \text{Frob}_q : \Sigma_{K,g,h} \to (\sigma_q)^*(\Sigma^+_{K,g,h}) \) denote the Frobenius morphism of the (polarized) \( p \)-divisible \( O \)-module as above. (We will need the polarization in what follows.) Let \( R \in \text{Art}(O, \mathbb{F}) \), with structure map \( \phi : R \to \mathbb{F} \);
let $R_{[\sigma_q]}$ be the same algebra $R$ with structure map $\sigma_q \circ \phi$. Let $(H, j)$ be an $R$-valued point of $\tilde{M}_{g,h}$. Define $H^\alpha = H$, and let $j^\alpha$ be the morphism

$$j \circ \phi^*(\text{Frob}_q^{-1}) : (\sigma_q)^* (\tilde{\Sigma}_{K,g,h})_g \rightarrow (\tilde{\Sigma}_{K,g,h})_g \rightarrow \Sigma_{K,g,h}.$$  

(Note that $\text{Frob}_q^{-1}$ is a quasi-isogeny, not a genuine morphism.) This gives rise to an isomorphism of functors $\alpha : \tilde{M}_{g,h} \xrightarrow{\sim} \sigma_q^* (\tilde{M}_{g,h})$ via

$$(3.2.3) \quad \tilde{M}_{g,h}(A) \rightarrow \tilde{M}_{g,h}(A_{[\sigma_q]}); (H, j) \longmapsto (H^\alpha, j^\alpha).$$

This morphism breaks up as a product of two factors: one on $\tilde{\Sigma}_{K,g,h}$ and one on the polarization; the second factor is just multiplication by $q$ (the action of $\sigma_q$ on roots of unity).

If $\tilde{M}_{g,h}$ had a $F_q$-rational structure, then $\alpha$ would correspond to the action of $\sigma_q$ on points (say over $F$). The fact that $\Theta$ commutes with the action of $\sigma_q$ comes down, after verification, to the fact that Frobenius on $(A, \lambda)$ corresponds to Frobenius on $G_\ell \times \check{G}_\ell$ together with the polarization.

To any such deformation problem, Rapoport and Zink associate a pair of groups $(G, J)$ over $Q_p$, with $J$ an inner form of a Levi factor of a rational parabolic subgroup of $G$. For $\tilde{M}_{n-h}^+$, the group $G$ is $GL(n, K) \times Q_p^\times$, and $J = J_{n-h,+} = D_{1/(n-h)} \times GL(h, K) \times Q_p^\times$. In any case, $J$ is the group of self-quasi-isogenies of the relevant divisible $O$-module $\Sigma_{K,g,h}$ preserving all additional structure (in the case of $\Sigma_{K,g,h}^+$, $J$ is the group that preserves the polarization). Thus it acts on the moduli problem by sending $j$ to $j \circ \delta$. These actions commute with the Weil descent datum because Frobenius commutes with everything.

3.3. Drinfel’d level structures: properties. — As before, $n = g + h$. Let $m \geq 0$, $\varpi \in O$ a uniformizing parameter, and consider Drinfel’d $\varpi^m$-level structures.

(3.3.1) Definition. — Let $R \in \text{Art}(O, F)$, and $(H, j) \in \tilde{M}_{g,h}(R)$. A map of groups

$$p : (\varpi^{-m}O/O)^n \rightarrow H[\varpi^m](R)$$

is a Drinfel’d level structure if and only if there is a free rank $g$ $O/\varpi^m O$-direct summand $M \subset (\varpi^{-m}O/O)^{g+h}$ such that

(3.3.2) $\prod_{x \in M} (T - T(p(x)))$ divides $f_{\varpi^m}(T)$, the power series representing multiplication by $\varpi^m$ on $H$;

(3.3.3) The induced map $(\varpi^{-m}O/O)^n/M \rightarrow H[\varpi^m](R)/H^0[\varpi^m](R)$ is an isomorphism.

Drinfel’d level structures were introduced in [Dr]. Another approach, developed by Katz and Mazur in [KM], is described in the following section. The present notes can only sketch the their basic properties. A complete discussion, with proofs of all properties used implicitly below, can be found in Chapter III.2 of [HT].
The functor on Art(\(O, \mathbb{F}\)) that takes \(A\) to the set of \((H, j, p)\), with \(H\) and \(j\) as before and \(p\) a Drinfel’d level structure, is relatively representable over \(\hat{M}_{g,h}\) by a formal scheme \(\check{M}_{g,h,m}\). We can do the same with +; however, we always consider polarizations only up to \(\mathbb{Z}_p^*\)-multiples. One of the main results of [Dr] is that the formal scheme \(\check{M}_{g,h,m}\) is flat over \(\hat{M}_{g,h}\) and is regular; however, it has bad singularities in characteristic \(p\). Its rigid generic fiber is precisely \(\check{M}^\gamma_{g,h,m}\), with \(U(m) \subset GL(g + h, O)\) the principal congruence subgroup of level \(\varpi^m\). Note that the free rank \(g\) summand \(M\) in the previous paragraph is a discrete invariant of the triple \((H, j, p)\); thus

\[
\hat{M}_{g,h,m} = \prod_M \check{M}_{g,h,m}
\]

(the index \(m\) is implicit in \(M\)). When necessary, we say \(p\) is of “type \(M\”

The Weil descent datum on \(\hat{M}_{g,h}\) lifts trivially to each \(\check{M}_{g,h,m}\), and stabilizes each component \(\check{M}_{g,h;m}\). Indeed, since \(H^\alpha = H\), we can define \(\alpha : \check{M}_{g,h;m} \to Fr_k(\check{M}_{g,h;m})\) (on \(A\)-valued points, as above) by sending \((H, j, p)\) to \((H, j^\alpha, p)\). Again, all these constructions go through in the variants with +. Similarly, the action of \(J\) on \(\hat{M}_{g,h}\) lifts to each \(\check{M}_{g,h;m}\) and each \(\check{M}_{g,h,M}\), inducing the action already defined on the rigid generic fiber.

The action of \(G\), previously defined on \(\check{M}_{g,h}\), also extends to the family of integral models \(\check{M}_{g,h,m}\). Here is the construction. Let \((H, j, p) \in \check{M}_{g,h;m}(A)\) for some test scheme \(A\). Suppose moreover that \(p\) is of type \(M\), and lift \(M\) to a rank \(g\) direct summand \(M_0\) of \(\mathcal{O}^{g+h}\); let \(P_{M_0} \subset G\) be the stabilizer of the \(K\)-subspace spanned by \(M_0\). First suppose that \(\gamma^{-1} \in M_0(\mathcal{O})\) and that \(\gamma \cdot \mathcal{O}^{g+h} \subset \varpi^{-m} \mathcal{O}^{g+h} \subset \varpi^{-m} \mathcal{O}^{g+h}\). Suppose in addition that \(\gamma \in P_{M_0}\), and let \((\gamma_g, \gamma_h)\) denote its projection on \(GL(g) \times GL(h)\). Then \(\gamma\) takes the triple \((H, j, p)\) over a test scheme \(A\), where \(p\) is a Drinfel’d level \(m\)-structure, to a triple \((H^\gamma, j^\gamma, p^\gamma)\), with \(p^\gamma\) a Drinfel’d level \(m^\prime\)-structure. Here

\[
H^\gamma = H/p(\gamma \cdot \mathcal{O}^{g+h})
\]

where \(p(\gamma \cdot \mathcal{O}^{g+h})\) is viewed as a finite flat subgroup scheme of \(H[\varpi^m]\) with “full set of sections” \(p(\gamma \cdot \mathcal{O}^{g+h}) < H[\varpi^m]\) (A). (This notion will be defined more generally in the global setting.) Now let \(\Sigma(\gamma_g) = \Sigma_{K,g,h}/\ker(Frob^K_{\mathcal{O}}(\det(\gamma_g)))\). Then \(j^{-1}\) identifies \(H^\gamma_{\mathcal{K}}\) with

\[
\Sigma^\gamma_{K,g,h} = \Sigma_{K,g,h}/\ker(Frob^K_{\mathcal{O}}(\det(\gamma_g))) \times (K^h/\gamma_h \cdot \mathcal{O}^h)
\]

where \(v_K\) is the valuation on \(K\). Indeed, this follows upon comparing orders from the fact that every finite flat subgroup of \(\Sigma_{K,g}\) is of the form \(\ker(Frob^K_{\mathcal{O}})^d\) for some \(d\). We obtain \(j^\gamma\) by composing

\[
\Sigma_{K,g} \times (K/\mathcal{O})^h \xrightarrow{(\det(\gamma_g), \gamma_h)} \Sigma^\gamma_{K,g,h} \times (K^h/\gamma_h \cdot \mathcal{O}^h) \xrightarrow{j} H^\gamma_{\mathcal{K}}.
\]
Finally, \( p^\gamma \) is just \( p \circ \gamma \) where \( \gamma \) is viewed as an embedding
\[
\varpi^{-m'} \mathcal{O}^{g+h}/\mathcal{O}^{g+h} \subset \varpi^{-m} \mathcal{O}^{g+h}/\varpi \cdot \mathcal{O}^{g+h}.
\]

Letting \( m \) and \( m' \) vary, we obtain an action of \( \gamma \) satisfying the above properties on the tower \( \{ \tilde{M}_{g,h;m} \} \). This action obviously commutes with the action of \( J \times W_K \), and it is easy to see that it coincides with the usual action on the relevant subset of the rigid generic fiber. We note that if \( x \in \mathcal{O} \) and \( x \neq 0 \) then the element \((x^{-1}, x^{-1}) \in P_{M_0} \) acts trivially. Thus we may extend the partially defined action to obtain an action of \( P_{M_0} \times J \times W_K \) on the tower \( \{ \tilde{M}_{g,h;m} \} \), factoring through \( (P_{M_0} \times J)/K^x \times W_K \), where \( K^x \) is embedded diagonally. Finally, we have the Iwasawa decomposition \( G = P_M \cdot GL(g + h, \mathcal{O}) \). There is no problem defining an action of \( GL(g + h, \mathcal{O}) \) on Drinfel’d level structures (by the standard action on \( (\varpi^{-m} \mathcal{O}/\mathcal{O})^{g+h} \)) for all \( m \); thus we can extend the action to \( (G \times J)/K^x \times W_K \) on the tower \( \{ \tilde{M}_{g,h;m} \} \), which we denote \( \hat{M}_{g,h;m} \) and view as the projective limit of the \( \hat{M}_{g,h;m} \). Note that covering the isomorphism (3.1.7) (and ignoring the +) we have an isomorphism of (ind-profinite) schemes over \( k \):

\[
(3.3.4) \quad \hat{M}_{g,h;m,0} \cong \tilde{M}_{g,h,m,0} \rightarrow \mathbb{Z} \times GL(h, K)
\]

Again, all these constructions go through with additional structures +. In the next construction we will include these structures, just for a change. For any \( m \) and any type \( M \), there are natural morphisms \( \pi : \hat{M}_{g,h;m} \rightarrow \tilde{M}_{g,m} \). Here if \( (H, j, p) \in \tilde{M}_{g,h,m}(A) \) for some test scheme \( A \), we let \( \pi(H, j, p) = (H^0, j^0, p^0) \), where \( H^0 \) is the connected part of \( H \), \( j^0 \) the restriction of \( j \) to \( \Sigma_{K,g} \), whose image is \( H^0_{K,g} \), and \( p^0 \) the restriction of \( p \) to \( M \). We can factor \( \pi = \pi_3 \circ \pi_2 \circ \pi_1 \). Here, letting \( \hat{M}_{g,h;m,0} \) denote the moduli space over \( \hat{M}_{g,h} \) of Drinfel’d level \( \varpi^m \)-structures on the connected subgroup of the variable height \( g + h \)-divisible \( \mathcal{O} \)-module, we have

\[
\pi_1 = \pi_{1,M} : \hat{M}_{g,h;m} \rightarrow \hat{M}_{g,h;m,0} \times \hat{M}_{0,h}\n\]
takes \( (H, j, p) \) to \( (H, j, p^0, p^\text{ét}) \), with \( p^\text{ét} \) the induced level structure
\[
(\varpi^{-m} \mathcal{O}/\mathcal{O})^{g+h}/M \rightarrow \hat{H}/H^0[\varpi^{-m}].
\]

Moreover, \( \pi_2(H, j, p^0, p^\text{ét}) = (H, j, p^0) \) (forget \( p^\text{ét} \)), and \( \pi_3 \) is the base change to \( \hat{M}_{g,h,0} \) of an analogous map \( \pi'_3 : \hat{M}_{g,h,0} \rightarrow \tilde{M}_{g,0} \). (i.e., forget the étale part altogether).

\[(3.3.5) \text{Proposition.} \quad \text{The map } \pi_2 \text{ (resp. } \pi'_3, \text{ resp. } \pi_1 \text{) is étale, (resp. smooth, resp. radicial over the special fiber).} \]

\textbf{Proof.} — The statement concerning \( \pi_2 \) is easy. The smoothness of \( \pi'_3 \) follows directly from Theorem 2.4.11 (Drinfel’d’s theorem has no + and \( K' \), but these just add profinite limits of discrete parameters). The assertion regarding \( \pi_1 \) is left as an exercise. \qed
3.4. Drinfel’d level structures: global construction. — We will need a more general definition of Drinfel’d level structures on a 1-dimensional divisible \( O \)-module \( H \) of height \( n \) over a general base scheme \( S \). In particular, we do not know that the connected part is of constant height. The Katz-Mazur definition is as follows. Consider the finite flat group scheme \( H[\varpi^m] \) over \( S \), and consider homomorphisms of abelian groups \( p : (\varpi^{-m}O/O)^n \to H[\varpi^m](S) \). Every point \( p(x) \) is then an \( S \)-subscheme of \( H[\varpi^m](S) \). The set \( \{ p(x), x \in (\varpi^{-m}O/O)^n \} \) is a “full set of sections” of \( H[\varpi^m] \) if, for any affine \( S \)-scheme \( \text{Spec}(R) \) and every function \( \phi \in B = H^0(H[\varpi^m]_R, O_{H[\varpi^m]_R}) \), there is the equality of characteristic polynomials in \( R[T] \):

\[
\det(T - \phi) = N_{B[T]/R[T]}(T - \phi) = \prod_{x \in (\varpi^{-m}O/O)^n} (T - \phi(p(x))).
\]

(Equivalently: if \( N(\phi) = \prod_x \phi(p(x)) \in R \) these are equivalent by replacing \( R \) by \( R[T] \).)

(3.4.1) Definition. — \( p \) is a Drinfel’d basis if and only if the set \( \{ p(x) \} \) is a full set of sections.

We need three properties of this definition:

(3.4.2) The functor \( S \mapsto p \), where \( H \) is a fixed 1-dimensional divisible \( O \)-module of height \( n \) over \( S \), is representable.

(3.4.3) When \( H \) is étale, it is just the usual level \( m \) structure.

(3.4.4) It coincides with Drinfel’d’s definition when \( H \) is a formal group.

The first property implies that it applies to \( S = \mathcal{A}_U(B, \ast) \), defining a moduli scheme \( \mathcal{A}_{U(m)}(B, \ast) \) over \( O \). The second property implies that the generic fiber of \( \mathcal{A}_{U(m)}(B, \ast) \) (over \( \text{Spec}(K) \)) is isomorphic to the moduli space for level \( U(m) \) structure, where \( U(m) = U^n \times U_w(m) \), with \( U_w(m) \) the principal congruence subgroup of \( GL(n, O) \) of level \( \varpi^m \). Thus the notation is consistent. The third property implies that, in order to determine local properties of \( \mathcal{A}_{U(m)}(B, \ast) \), it suffices to study Drinfel’d bases over \( \tilde{M}_{g,h} \) for general \( g \) and \( h \). In particular, the results of Drinfel’d quoted above imply that \( \mathcal{A}_{U(m)}(B, \ast) \) is a regular scheme, flat over \( \mathcal{A}_U(B, \ast) \) for all \( m \).

We prove properties (3.4.2-3.4.4) in turn.

(3.4.5) Lemma. — The functor is representable.

Proof (as in Katz-Mazur). — The functor

\[ T \mapsto \{ p : (\varpi^{-m}O/O)^n \to H[\varpi^m](T) \} \]

is represented by \( S = H[\varpi^m]\tilde{\varpi}^m \). So we need to show that the condition of being a full set of sections is represented by a closed subscheme. We may localize on \( S \) to assume that \( S = \text{Spec}(R) \), \( H[\varpi^m] = \text{Spec}(B) \) with \( B \) free of rank \( M \) over \( R \). Let \( b_1, \ldots, b_M \) be an \( R \) basis of \( B \). Let \( P_1, \ldots, P_M \) be the tautological sections of \( H[\varpi^m] \) over \( S \). The condition that they form a full set of sections depends on the choice of
a variable \( R \)-algebra \( R' \), but in fact any function over any algebra \( R' \) is of the form 
\[ \sum_i t_i b_i, \] 
with \( t_i \in R' \). So it suffices to look at the universal case \( R' = R]\[T_1, \ldots, T_M\], and the universal function \( \Phi = \sum_i T_i b_i \). The condition that \( P_1, \ldots, P_M \) is a full set of sections is the condition
\[
\text{Norm}_{B[T_1,\ldots,T_M]/R'}(\Phi) = \prod_i f(P_i)
\]
i.e.
\[
\text{Norm}\left( \sum_i T_i b_i \right) = \prod_i \left( \sum_i T_i b_i(P_i) \right).
\]
Both sides are homogeneous forms of degree \( M \) in \( T_1, \ldots, T_M \), with coefficients in \( R \).

The equality comes down to equality of coefficients, and this is given by a set of equations in \( R \), i.e. a closed condition on \( S \).

(3.4.6) Lemma. — If \( C = H[\varpi^m] \) is étale over a scheme \( Z \), then a Drinfel’d basis is just a level structure.

Proof. — This is easy. We can trivialize \( C \) (by base change to \( C \)). Then \( P_1, \ldots, P_M \) is a level structure if and only if there is a basis \( b_i \) of \( B \) with \( b_i(P_j) = \delta_{ij} \). Then \( b_i \cdot b_j = \delta_{ij} b_i \), and for this basis, the equality of norms in Lemma 1 is obvious. (Taking as basis \( b_i \) for \( B[T_1, \ldots, T_M]/R' \), the matrix of \( \sum T_i b_i \) is diagonal with entries \( T_i \).)

(3.4.7) Lemma. — The above definition coincides with Drinfel’d’s when \( H \) is formal.

Proof. — We admit the following elementary lemma ([KM], p. 42, Lemma 1.10.2):

(3.4.8) Lemma. — Let \( R \) be a ring, \( F(X) \in R[X] \) a monic polynomial of degree \( M \geq 1 \), \( a_1, \ldots, a_M \) elements of \( R \). Let \( B = R[X]/(F) \). Then the following two conditions are equivalent:

(a) We have the factorization \( F(X) = \prod_i (X - a_i) \).

(b) For every \( \phi \in B \), we have the factorization
\[
\det(T - \phi) = \prod_i (T - \phi(a_i)).
\]

Sketch of proof. — The determinant is relative to the free extension \( B/R \). Then \( (b) \Rightarrow (a) \) because in \( B \) the characteristic polynomial of \( X \) is \( F \), i.e. \( \det(T - X) = F(T) \).

Applying (b) to \( \phi = X \), we thus get \( F(T) = \prod_i (T - a_i) \) which is (a). In the other direction, we can regard the coefficients of \( \phi \) and the \( a_i \) as independent variables in a big field \( K \), and
\[
K[X]/\prod_i(X - a_i) \cong \prod K[X]/(X - a_i)
\]
so the relation of characteristic polynomials is clear.
Now condition (3.3.2) of Drinfel’d’s definition is the one that applies to a formal group:

\[ f_{w^n}(T) = g(T) \prod_x (T - T(p(x))), \]

for some power series \( g \). Over \( \mathbb{F} \), we may assume \( f_{w^n}(T) = T^{w^n g} \) (the height = \( g \)). Comparing degrees and leading coefficients, this implies that \( g(T) \) is a unit in \( R[[T]]^\times \) with constant term 1. Now in the lemma we may take

\[ B = R[[X]]/(f_{w^n}) \xrightarrow{\sim} R[X]/(F) \]

for some monic polynomial, by Weierstrass preparation. Drinfel’d’s condition is (a) of the lemma; the Katz-Mazur condition is (b).

Now by putting together the uniformization morphism \( \Theta \) with Lemma 3.4.7, we obtain morphisms of all levels. Let \( x \in \mathcal{S}_{U}^{(h)} \), and let \( \mathcal{S}(x, m) \) denote the inverse image of the isogeny class \( \mathcal{S}(x) \) in \( \mathcal{A}_{U(m)}(B, \ast) \). Then because the Drinfel’d basis depends only on the \( p \)-divisible group, we can lift \( \Theta \) to

\[ \Theta_m : I_x(Q) \backslash \tilde{M}_{n-h,h;m}^+ \times G(A_f^w)/U^w \xrightarrow{\sim} \mathcal{A}_{U(m)}(B, \ast) \tilde{\mathcal{S}}(x_m). \]

This uniformization depends on \( U^w \) and on \( m \), but they fit together in the limit to yield

\[ \Theta_\infty : I_x(Q) \backslash \tilde{M}_{n-h,h;\infty}^+ \times G(A_f^w) \xrightarrow{\sim} \lim_{U_w, m} \mathcal{A}_{U(m)}(B, \ast) \tilde{\mathcal{S}}(x_m). \]

This commutes with the Weil group action, as before. Note that the action of \( I_x(Q) \) on \( \tilde{M}_{n-h,h;\infty}^+ \) is given by associating to a self-quasi-isogeny of \( A_x \) a self-quasi-isogeny of \( G_x \). In other words, it factors through a homomorphism \( I_x(Q) \rightarrow J = J_{n-h,h} = D_{1/(n-h)}^x \times GL(n, K) \). Write \( G^{(h)}(A_f) = G(A_f^w) \times J_{n-h,h} \) (an abuse of notation, because \( G^{(h)}(A_f) \) is not the group of \( A_f \)-points of something called \( G^{(h)} \)). Then (3.4.9) can be rewritten

\[ [\tilde{M}_{n-h,h;\infty}^+ \times (I_x(Q) \backslash G^{(h)}(A_f))]/J_{n-h,h} \xrightarrow{\sim} \lim_{U_w, m} \mathcal{A}_{U(m)}(B, \ast) \tilde{\mathcal{S}}(x_m). \]

where the \( J_{n-h,h} \)-action on the left hand side is diagonal (on the left on \( \tilde{M}_{n-h,h;\infty} \) and on the right on the adelic group).

### 3.5. Action of adelic group with Drinfel’d level structures.

It is not difficult to define an action of \( G(A_f^w) \) on the right-hand side of (3.4.9) so that it coincides with the obvious action on the left-hand side; this is standard in the theory of Shimura varieties. On the other hand, we have defined an action of \( G_w = GL(n, K) \) on the left-hand side. It remains to define an action of \( GL(n, K) \) on the right hand side such that (a) \( \Theta_\infty \) is \( GL(n, K) \) equivariant and (b) the action extends the usual action on the (smooth) generic fiber.
The action is defined by analogy with the previous action. Let \((g_0, g) \in \mathbb{Q}_p^\times \times GL(n, K)\). We let \(G\) denote the one-dimensional height \(h\) divisible \(\mathcal{O}\)-module attached to one of our abelian schemes \(A\). First suppose that we have the following integrality conditions:

(i) \(g^{-1} \in M(n, \mathcal{O})\),
(ii) \(g_0^{-1} g \in M(n, \mathcal{O})\),
(iii) \(\varpi^{m - m'} g \in M(n, \mathcal{O})\).

(It is understood that \((g_0, g)\) is any pair in \(\mathbb{Q}_p^\times \times GL(n, K)\), there exists \(a \in \mathbb{Z}\) such that \((p^{-2a} g_0, p^{-a} g)\) satisfies the above inequalities for \(m - m' \geq 0\). Under these assumptions we will define a morphism

\[
(g_0, g) : \mathcal{A}_{U(m)}(B, *) \to \mathcal{A}_{U(m)}(B, *).
\]

It will send \((A, \lambda, i, \eta^w, p)\) over \(T\) to \((A/(C \oplus C^\perp), p^{\text{val}_{\varpi}(g_0)} \lambda, i, \eta^w, p \circ g)\), where

(3.5.1) \(C_1 \subset G[\varpi^m]\) is the unique closed subgroup scheme for which the set of \(p(x)\) with \(x \in g \cdot (O^\prime)/O^\prime\) is a full set of sections;

(3.5.2) \(C = (O^\prime_{F,F} \otimes O_{F,F}, C_1) \subset A[\varpi - \text{val}_{\varpi}(g_0)]\);

(3.5.3) \(C^\perp\) is the annihilator of \(C \subset A[\varpi - \text{val}_{\varpi}(g_0)]\) inside \(A[(\varpi^\infty) - \text{val}_{\varpi}(g_0)]\) under the \(\lambda\)-Weil pairing;

(3.5.4) \(p^{\text{val}_{\varpi}(g_0)} \lambda\) is the polarisation \(A/(C \oplus C^\perp) \to (A/(C \oplus C^\perp))^\vee\) which makes the following diagram commute

\[
\begin{array}{ccc}
A & \xrightarrow{p^{\text{val}_{\varpi}(g_0)} \lambda} & A^\vee \\
\downarrow & & \downarrow \\
A/(C \oplus C^\perp) & \xrightarrow{p^{\text{val}_{\varpi}(g_0)} \lambda} & (A/(C \oplus C^\perp))^\vee
\end{array}
\]

(3.5.5) \(p \circ g : \varpi^{-m'}(O^\prime)/O^\prime \to (G[\varpi^m]/C_1)(T)\) is the homomorphism making the following diagram commute

\[
\begin{array}{ccc}
\varpi^{-m'}(O^\prime)/O^\prime & \xrightarrow{p \circ g} & G[\varpi^m]/C_1(T) \\
\downarrow & & \downarrow \\
\varpi^{-m'} g(O^\prime) / g(O^\prime) & \to (G[\varpi^\infty]/C_1)[\varpi^m'](T) \\
\downarrow & & \downarrow \\
\varpi^{-m'}(O^\prime) / g(O^\prime) & \to (G[\varpi^m]/C_1)(T) \\
\downarrow & & \downarrow \\
\varpi^{-m'}(O^\prime) / O^\prime & \xrightarrow{p} & G[\varpi^m](T);
\end{array}
\]
This definition makes use of a number of properties of Drinfel’d bases that we have not made explicit here. For instance, the existence of a subgroup scheme $C_1$ as in (3.5.1) is Lemma III.2.2 of [HT].

Over the generic fiber (i.e., over $K$) one checks that this coincides with the usual action. Thus $(p^{-2}, p^{-1})$ acts in the same way as $p \in G(A_w) –$ over the generic fiber, which is Zariski dense in the integral model, hence over the whole scheme. Indeed, the diagonal element $p \in Z_G(Q)$ acts trivially on the Shimura variety, but it is the product of $p \in G(A_w)$ and $(p^2, p) \in \mathbb{Q}_p^{\times} \times GL(n, K)$. Thus $(p^{-2}, p^{-1})$ acts invertibly on the inverse system. In this way we see that this defines an action of the whole of $G_w$. We state this formally as follows

\[ (3.5.6) \) Proposition. — The formulas (3.5.1)-(3.5.5) extend to an action of $G(A_f)$ on the tower of moduli schemes $A_{U(m)}(B, \ast)$ over $O$, in such a way that the uniformization map (3.4.9) is $W_K \times G(A_f)$-equivariant.

Remark. — One can also avoid worrying about $g_0$; the action of $\mathbb{Q}_p^{\times}$ can be defined easily for general $g_0$, just by changing the polarization. Moreover, one can define an action of $g$ that fixes the polarization, but then the polarization becomes a quasi-isogeny rather than an actual homomorphism. This strategy was followed in [HT2].

4. Stratification and vanishing cycles

The present lecture continues the study of the stratification of the special fiber $S_{U}^{(h)}$ of our Shimura variety by isomorphism type of isocrystal, which in the present simple situation correspond to stratification by $p$-rank of the universal family of abelian varieties with PEL structure. The cohomology of the generic fiber can be written, in the Grothendieck group, as the sum of cohomologies of strata of the special fiber with coefficients in the vanishing cycle sheaves. This is the First Basic Identity (4.4.4), which summarizes the contribution of vanishing cycles to the determination of the cohomology of the generic fiber.

4.1. Strata in level prime to $p$: Proof of smoothness. — Let $U = U_w \times U^w$, with $U_w = GL(n, O)$, and $U^w$ sufficiently small, so that $A_{U}(B, \ast)$ has a smooth model over $O$. We return to the stratum $S_{U}^{(h)}$ defined last time; this is the subset where $G_{\ast}^{Z,G}$ is of height $h$. We prove that each $S_{U}^{(h)}$ is smooth of dimension $h$.

In fact, we can replace $A_{U}(B, \ast)_{\bar{F}}$ by any smooth locally noetherian scheme $S$ over $\bar{F}$, and consider a one-dimensional divisible $O$-module $H/S$ of height $n$. We know that, when $S = A_{U}(B, \ast)_{\bar{F}}$, then, for every $s \in S(\bar{F})$, the formal completion $S_{s}$ is isomorphic to the universal formal deformation space (over $\bar{F}$) of $H_{s}$ (we apply the Serre-Tate isomorphism in reverse). We assume $S$ has this universal property as well; it is used only in (c) of Theorem 4.1.1. Let $S^{(h)}(\bar{F}) \subset S(\bar{F})$ be the subset where the height of $H^{et}$ is $\leq h$, $S^{(h)}(\bar{F}) = S^{(h)}(\bar{F}) - S^{(h-1)}(\bar{F})$. 

ASTÉRISQUE 298
4.1.1 Theorem

(a) Under the above hypotheses, $S^{[h]}(F)$ is the set of $F$-valued points of a reduced closed subscheme $S^{[h]}$.

(b) Over $S^{(h)}$, there is a short exact sequence

$$0 \longrightarrow H^0 \longrightarrow H \longrightarrow H^{\text{ét}} \longrightarrow 0$$

where $H^0$ is a one-dimensional formal $\mathcal{O}$-module of height $n - h$ and $H^{\text{ét}}$ is étale of height $h$.

(c) For $h = 0, \ldots, n - 1$, $S^{(h)} = S^{[h]} - S^{[h-1]}$ is either empty or smooth of dimension $h$.

Proof of (a) and (b)

Step 1. — The proof is in several steps. We first note that (a) implies (b). Indeed, Messing observed in his thesis ([Me], Ch. II, Prop. 4.9) that if $S$ is a connected noetherian scheme of characteristic $p$ (or even with $p$ locally nilpotent) and $H$ is a $p$-divisible group over $S$ with $|H[p](k(s))|$ constant, then $H$ is globally an extension of a formal group by an étale group. (More generally, if $X/S$ is a finite flat scheme with constant separable rank, then it factors uniquely $X \xrightarrow{f} X' \xrightarrow{g} S$ with $f$ radicial and $g$ étale. This is first proved for fields, where it is obvious, then for complete local rings by Hensel’s lemma, then for general local rings by faithfully flat descent, using the uniqueness over the completion, and then the uniqueness implies that these local morphisms patch together globally.) So if we have (a), then over $S^{(h)}$ we have a short exact sequence with $H^{\text{ét}}$ étale, and since both $H^0$ and $H^{\text{ét}}$ are still $\mathcal{O}$-modules, the height follows by counting the order of $H[p]$ at any point.

Step 2. — Now we prove (a). The argument is due to Oort ([Oo]). The problem is local, so we may assume $S = \text{Spec}(R)$ where $R$ is a noetherian ring and $\text{Lie}(H)$ is free over $R$. By induction, we drop the assumption that $S$ be smooth (but it remains reduced) and we also drop the assumption that the complete local ring is isomorphic to the deformation ring at each point. We assume that generically, $H_s[p](k(s))$ is of order $p^g$ for some $g$; at this point the $\mathcal{O}$-action is irrelevant. First, we establish notation for Frobenius and Verschiebung maps. Let $\text{Fr}_S : S \rightarrow S$ denote the absolute Frobenius morphism. The superscript $(p)$, for schemes over $S$, denotes pullback with respect to $\text{Fr}_S$. Let

$$V : H^{\vee,(p)} \longrightarrow H^{\vee}$$

denote the $V$-operator on $\mathcal{H}$; i.e., the Cartier dual of the Frobenius homomorphism $F_H : H \rightarrow H^{(p)}$. Let $\mathcal{H} = \text{Lie} H[p]^{\vee} = \text{Lie} H^{\vee}$ (they are equal because $p = 0$ on $S$), and let $\underline{V}$ denote the differential of $V$:

$$\underline{V} : \mathcal{H}^{(p)} \longrightarrow \mathcal{H}.$$
In this version, $V_*$ is an $O$-linear map. We may also identify $H^p$ with $Fr^*_S(H)$; then composing $V_*$ with the Frobenius map

$$F_H : H \rightarrow Fr^*_S(H)$$

we obtain

$$(4.1.2) \quad V_* = V_0 \circ F_H : H \rightarrow H,$$

a Frob-linear version $V_*$ of $V_0$:

$$V_*(ay) = a^p V_0(y), \ a \in \mathbb{F}, \ y \in \Gamma(s, H).$$

$$(4.1.3) \text{Lemma.} \quad \text{For any geometric point } s \in S \text{ there is a canonical perfect pairing}$$

$$H^{V_*=1}_s \otimes H_s[p](k(s)) \rightarrow \mathbb{F}_p.$$ 

$$(4.1.4) \text{Proof.} \quad \text{This is apparently well-known, but we were unable to find a reference. Here}$$

is the proof. It is standard (cf. [Mu, p. 138]) that there is a canonical isomorphism

$$H_s \sim \rightarrow \text{Hom}(H_s[p], \mathbb{G}_a).$$

With respect to (4.1.4), $V_*$ is identified with the map $\phi \mapsto \phi \circ F_h = F_{G_a} \circ \phi$. Applied to $k(s)$-valued points, (4.1.4) yields a pairing

$$H_s[p](k(s)) \times H^{V_*=1}_s \rightarrow G_a(k(s)) = k(s),$$

which restricts to a pairing

$$(4.1.5) \quad H_s[p](k(s)) \times H^{V_*=1}_s \rightarrow k(s)^{F=1} = \mathbb{F}_p.$$ 

If $\phi \in H^{V_*=1}_s$ and $\phi(x) = 0$ for all $x \in H_s[p](k(s))$ then $\phi$ factors through the formal group of $G_a$, hence by (4.1.5) $\phi = 0$. Thus we have an injection

$$(4.1.6) \quad H^{V_*=1}_s \rightarrow \text{Hom}(H_s[p](k(s)), \mathbb{F}_p).$$

To complete the proof of the lemma, it suffices to show the order of the left-hand side of (4.1.6) is at bounded below by that of the right-hand side. Suppose $H_s[p](k(s))$ has order $p^h$; equivalently, that there is an embedding $\mu^h_p \hookrightarrow H_s[p]^{V_*}$. Then there is an embedding

$$\text{Lie}(\mu^h_p) \hookrightarrow H_s,$$

compatible with $V_*$. But the $p$-linear map $V_*$ has slope 0 on $\mu_p$, hence

$$\dim_{\mathbb{F}_p} H^{V_*=1}_s \geq \dim_{\mathbb{F}_p} \text{Lie}(\mu^h_p) = h,$$

which yields the desired bound. \qed

Let $S_n = \{ s \in S(\mathbb{F}) \mid |H_s[p](k(s))| \leq p^n \}$. It suffices to show that each $S_n$ is the set of points of a reduced closed subscheme. By the Lemma, $S_n = \{ s \in S(\mathbb{F}) \mid |H^{V_*=1}_s| \leq p^n \}$. Let $e_1, \ldots, e_m$ be a basis for $H$ over $R$, and write $V_*$ as a matrix:

$$V_*(e_i) = \sum_j v_{ij} e_j.$$
Then $H_{\nu_0} = 1$ is identified with the subscheme of $A_m^R$ defined by the equations

$$x_j = \sum_{i,j} v_{ij} x_i^p$$

via $(x_j) \mapsto \sum_j x_j e_j$; indeed

$$V_*(\sum x_i e_i) = \sum_{i,j} x_i^p v_{ij} e_j.$$ 

These equations define a quasi-finite étale covering of $S$, since the Jacobian is the identity. Generically, the degree is $p^g$; i.e., $S = S_g$. Then $S_{g-1}$ is closed. 

Proof of (c)

Step 3. — Next, we prove that the codimension of $S_{g-1}$ is at most 1. Let $T$ be the normalization of $S = \text{Spec}(R)$ in a finite separable extension of $\text{Frac}(R)$ where the étale covering is trivialized, so $H_{k(T)}$ has $p^g$ points, say $x_1, \ldots, x_{p^g}$. Since $T/S$ is finite, it suffices to prove the result with $S$ replaced by $T$. Then $T_{g-1}$ is the union of the loci $Z_i$ where the $x_i$ are not regular. Since $T$ is normal, each $Z_i$ is of codimension $\leq 1$.

Step 4. — It remains to prove smoothness. This is more subtle, and requires Drinfeld’s theory. First, it follows from Step 3 that $S^{(h)}$ is of dimension at least $h$ for all $h$. On the other hand, the separable rank of $H$ is constant over $S^{(h)}$; so over $S^{(h)}$ the connected part $H^0$ is a (smooth) formal group of height $n-h$. If $S^{(h)}$ is empty there is nothing to prove. So let $s \in S^{(h)}$ be a closed point, and consider the maps

$$\text{Spf}(R_{K,n-h,h}) \xrightarrow{\phi} S_s^{(h)} \xrightarrow{f} S^{(h)} \xrightarrow{cl} \text{Spf}(R_{K,n-h}).$$

The map $f$ is the natural immersion and $cl$ is the classifying map attached to the deformation of $H^0_s$ over $S^{(h)}_s$ given by pullback of $H^0$ to $S^{(h)}_s$.

Let $P$ be a minimal prime of $O_{S_s^{(h)}}$ and let $cl_P$ denote the restriction of $cl$ to the corresponding irreducible component. Then the map $cl_P$ corresponds to a homomorphism of rings $R_{K,n-h} \to O_{S_s^{(h)}}/P$.

Denote by $t_1, \ldots, t_{n-h-1}$ the parameters of $R_{K,n-h}$ (parametrizing deformations of $H^0_s$) and $u_1, \ldots, u_h$ the remaining parameters in $R_{K,n-h,h}$ (parametrizing extensions by $(K/O)^h$). We will show that the parameters $t_1, \ldots, t_{n-h-1}$ of Drinfeld’s map to zero in $O_{S_s^{(h)}}/P$. Assuming this, we conclude as follows. It follows that the canonical classifying map $cl_h : (S^{(h)}_s \xrightarrow{cl} \text{Spf}(R_{K,n-h,h})),$ corresponding to the deformation of $H_s$ over $S^{(h)}_s$, corresponds to a homomorphism

$$R_{K,n-h,h}/(t_1, \ldots, t_{n-h-1}) = O[[t_1, \ldots, t_{n-h-1}, u_1, \ldots, u_h]]/(t_1, \ldots, t_{n-h-1}) \to O_{S^{(h)}_s}.$$
In the above diagram, $cl_h = \phi \circ f$; in particular it is a closed immersion. It follows that $S^{(h)}$ is of dimension $\leq h$ at $s$. But we know that it is of dimension at least $h$, hence the map above is an isomorphism, and $S^{(h)}$ is smooth at $s$.

It remains to show that the deformation of $H^0$ along $S^{(h)}$ is trivial. Let $k$ be the field of fractions of the image of $g_p$; since $S^{(h)}$ is reduced, it suffices to show that the $t_i$ map to zero in $k$. Suppose one of the $t_i$ does not map to zero, with $i$ minimal. Then the $q^i$-coefficient of $f_\infty$ (= multiplication by $\omega$ on $H^0_k$) is non-zero, and this is the first non-zero coefficient. Thus $H^0_k$ is of height $i < n - h$. This contradicts the hypothesis that $H^0$ is of height $n - h$ on $S^{(h)}$. 

We will see later, when counting points (Lecture 6), that the strata are non-empty.

4.2. Generalities on vanishing cycles. — Let $T = \text{Spec}(R)$, $R$ a Henselian discrete valuation ring, with generic point $\eta$ and special point $t$ of characteristic $p$, and assume for simplicity $k(t)$ algebraically closed. Let $f : S \to T$ be a proper morphism of finite type, with fibers $S_t$ and $S_\eta$, and geometric generic fiber $S_\eta$. Let $\mathcal{F}$ be a constructible sheaf on $S_\eta$ in $\mathbb{Q}_\ell$-vector spaces, with $\ell \neq p$. The point of vanishing cycles is to calculate $H^\bullet(S_\eta, \mathcal{F})$ as the hypercohomology of a complex $R\Psi(\mathcal{F})$ on $S_t$. There is an action of $\text{Gal}(k(\overline{\eta})/k(\eta))$ on $H^\bullet(S_\eta, \mathcal{F})$, hence one wants $R\Psi(\mathcal{F})$ to be endowed with an action of $\text{Gal}(k(\overline{\eta})/k(\eta))$ (= inertia). The recipe is formal. One considers the canonical morphisms $\overline{j} : S_\eta \to S$ and $i : S_t \to S$; then 

$$R\Psi(\mathcal{F}) = \overline{j}^* R\overline{j}_*(\mathcal{F})$$

(nearby cycles). Since $f$ is proper, one knows by proper base change that 

$$H^p(S, R^i j_*(\mathcal{F})) \xrightarrow{\sim} H^p(S_t, i^* R^q j_*(\mathcal{F})) = H^p(S_t, R^q \Psi(\mathcal{F})).$$

Then the Leray spectral sequence becomes 

$$E_2^{p,q} = H^p(S_t, R^q \Psi(\mathcal{F})) \Rightarrow H^{p+q}(S_\eta, \mathcal{F}).$$

More generally, one starts with $k(s)$ perfect (e.g. finite) and takes base change over $T$ by the Witt vectors $W(k(\overline{s}))$; then the spectral sequence becomes equivariant for $\text{Gal}(k(\overline{\eta})/k(\eta))$ covering the action of $\text{Gal}(k(\overline{\eta})/k(t))$. It is known that

**(4.2.1) Fact.** — If $\mathcal{F}$ is constructible and $f : S \to T$ is a proper morphism of finite type then the nearby cycle sheaf $R\Psi(\mathcal{F})$ is constructible.

The standard reference for vanishing cycles is [SGA7]; however, Illusie’s article [II] provides an efficient introduction.

This definition has the disadvantage that one is no wiser than before unless one can compute $R^q \Psi(\mathcal{F})$. In our setting, $T = \text{Spec}(\mathcal{O})$, $S = \mathcal{A}_{U(m)}(B, \ast)$, and we restrict attention for simplicity to $\mathcal{F} = \mathbb{Q}_\ell$. Write $\overline{S}_{U(m)} = \overline{\mathcal{A}}_{U(m)}(B, \ast)$, the special fiber of $S$. 

ASTÉRISQUE 298
and write $R^q\Psi$ for $R^q\Psi(\mathbb{Q}_\ell)$, and occasionally $R^q\Psi(m)$ when the level is indicated. Then there is a spectral sequence

$$E_2^{p,q} = H^p(\mathcal{S}_{U(m)}, R^q\Psi) \implies H^{p+q}(\mathcal{A}_{U(m)}(B,*), \mathbb{Q}_\ell).$$

Passing to the limit over $U^w$, and $m$, we find

$$\lim_{U^w,m} H^p(\mathcal{S}_{U(m)}, R^q\Psi) \implies \lim_{U^w,m} H^{p+q}(\mathcal{A}_{U(m)}(B,*), \mathbb{Q}_\ell).$$

Now the right-hand side is an admissible $G(A_f)$ module (admissible means just that at every finite level the cohomology is finite-dimensional). We consider a modified Grothendieck group of $G(A_f) \times W_K$-modules: the objects are formal sums $\sum n_{I,\sigma} \Pi \otimes \sigma$ with $\Pi$ an irreducible $\mathcal{O}_K$-valued representation of $G(A_f)$ and $\sigma$ an irreducible continuous $\mathbb{Q}_\ell$-valued representation of $W_K$; the $n_{I,\sigma} \in \mathbb{Z}$. An admissible $G(A_f) \times W_K$-module is a $G(A_f) \times W_K$-module that is admissible over $G(A_f)$ and continuous over $W_K$. To an admissible $G(A_f) \times W_K$-module $\pi$ we associate $\sum n_{I,\sigma} \Pi \otimes \sigma$ as follows. If $\Pi^U \neq 0$, then $n_{I,\sigma} \pi$ is the multiplicity of $\Pi^U \otimes \sigma$ in the semisimplification of $\pi^U$ as module over the Hecke algebra $\mathcal{H}(G(A_f)/U)$ tensored with $W_K$. One checks that this is independent of $U$. Note that $\ell$-adic monodromy in $\sigma$ has been eliminated.

Write $|\pi| = \sum n_{I,\sigma} \pi \Pi \otimes \sigma$, and define

$$[H(A(B,*))] = \sum (-1)^j \left[ \lim_{U^w,m} H^j(A_{U(m)}(B,*), \mathbb{Q}_\ell) \right]$$

Recall (from §2.4) that

**Assumption (4.2.4). —** The level subgroup is always assumed to contain $\mathbb{Z}_p^\times \subset \mathbb{Q}_p^\times$.

Then the above spectral sequence yields

$$[H(A(B,*))] = \sum_{p,q} (-1)^{p+q} \left[ \lim_{U^w,m} H^p(\mathcal{S}_{U(m)}, R^q\Psi) \right].$$

Here we are making use of the fact $[F]$ that the action of $G(A_f)$ extends canonically to an action on $R\Psi$ by cohomological correspondences, covering the action on $\lim_{U^w} \mathcal{S}_{U(m)}$.

Now recall the stratification of $\mathcal{S}_{U(m)}$ by the $\mathcal{S}_{U(m)}^{(h)}$. These have been defined when $m = 0$, and for general $m$ one takes inverse images. For any constructible sheaf $\Phi$ (on any base) there is always a long exact sequence:

$$\cdots \rightarrow H^p_c(\mathcal{S}_{U(m)}^{(h)}, \Phi) \rightarrow H^p(\mathcal{S}_{U(m)}^{[h]}, \Phi) \rightarrow H^p(\mathcal{S}_{U(m)}^{[h-1]}, i_{h-1}^* \Phi) \rightarrow \cdots$$

where $i_{h-1}$ is the obvious closed immersion. By induction, we obtain a further decomposition in Groth($G(A_f) \times W_K$):

$$[H(A(B,*))] = \sum_{p,q,h} (-1)^{p+q} \left[ \lim_{U^w,m} H^p_c(\mathcal{S}_{U(m)}^{(h)}, i_h^* R^q\Psi) \right].$$
We drop the $U$ and $m$ and just write the right-hand side

$$\sum_{p,q,h} (-1)^{p+q}[H^p_c(S^{(h)}, R^q\Psi)].$$

The stability of each $S^{(h)}$ under $G(A_f)$ follows from the fact that $G(A_f)$ preserves isogeny classes, and the height of the connected formal group is an isogeny invariant.

(4.2.8) Remark. — The above decomposition presupposes that each term is an admissible $G(A_f)$-module. The condition away from $w$ is clear, so we may as well fix $U^w$ and let $m$ vary. Then admissibility comes down to the assertion that

$$\lim_m \left[ H^p_c(S^{(h)}_{U(m)}, i_\ast^h R^q\Psi) \right] = H^p_c(S^{(h)}_{U(m)}, i_\ast^h R^q\Psi)$$

for any $h$, where $\Gamma_m \subset GL(n, \mathcal{O})$ is the principal congruence subgroup. For this we can replace the limit on the left by $H^p_c(S^{(h)}_{U(m)}, i_\ast^h R^q\Psi)^{\Gamma_m}$ for all $m' > m$. More generally, if $f : Z' \to Z$ is a quotient by a finite group $\Gamma$, and if $L'$ is a constructible sheaf on $Z'$ with compatible $\Gamma$-action, we have $H^p_c(Z, (f_\ast L')^\Gamma) \longrightarrow H^p_c(Z', (L')^\Gamma).$ So it suffices to prove the

(4.2.9) Continuity lemma. — $R^q\Psi(m) \longrightarrow f_{m', m, *} R^q\Psi(m')^{\Gamma_m}.$

Here the notation is obvious. This follows formally from the definition of vanishing cycles, because $f_{m', m}$ is the special fiber of a proper flat morphism whose generic fiber is an étale covering with Galois group $\Gamma_m/\Gamma_{m'}$.

4.3. Vanishing cycles and the fundamental local representation

Now we return to the formal situation. If $X$ is a “special” formal scheme over $\text{Spf}(\mathcal{O})$, Berkovich has constructed a vanishing cycles functor $R\Psi^{\text{form}}$ from étale sheaves over the generic fiber to constructible complexes on the special fiber, which is a scheme over $\mathbb{F}_q$. The hypothesis “special” is best expressed in terms of rigid geometry, but a finite flat covering of the formal spectrum of a formal power series ring over $\mathcal{O}$ is of that type. Again, the formal completion of a proper scheme of finite type over $\mathcal{O}$ along a subscheme of the special fiber is special. Thus the formal schemes $\tilde{M}_{n-h,h;m}$ of Drinfeld’s level structures are special in Berkovich’s sense; we have seen that their connected components are isomorphic to the formal completion of $A_{U(m)}(B, \ast)$ along points in $S^{(h)}$. When $h = 0$, the special fiber is just a point, or rather a union of points, indexed by $\mathbb{Z}$ (a connected component is of the form $\text{Spf}(R_{n-h;m})$). More generally, the special fiber is a union of connected components of the form $\text{Spf}(R_{h;h;m})$ indexed by $\mathbb{Z} \times U(h;m) \backslash GL(h,K)$, where $U(h,m) \subset GL(h,\mathcal{O})$ is the principal congruence subgroup of level $m$. In any case the vanishing cycles sheaves are just unions over the connected components of vector spaces with $W_K$ action.

We define

$$\Psi^{i}_{K,\ell,n-h,h,m} = H^0(\tilde{M}_{n-h,h;m,\text{red}}, R^i \Psi^{\text{form}} \mathbb{Q}(\ell))$$
for the formal scheme $\tilde{M}_{n-h,h;m}$; here $\tilde{M}_{n-h,h;m,\text{red}}$ is an ind-profinite scheme over $k$ with $GL(n,\mathcal{O}) \times J_{n-h,h}$-action. (The “ind” comes from the fact that $G_h \subset J_{n-h,h}$ is non-compact; in fact, $\tilde{M}_{n-h,h;m,\text{red}}$ is just a countable disjoint union of profinite schemes.) We let

\begin{equation}
\Psi^{i}_{K,\ell,n-h,h} = \Psi^{i}_{n-h,h} = \lim_{m} \Psi^{i}_{K,\ell,n-h,h,m} = \lim_{m} H^{0}(\tilde{M}_{n-h,h;m,\text{red}}, R^{i} \Psi^{\text{form}}_{\ell} \mathbb{Q}_{\ell}).
\end{equation}

Then each $\Psi^{i}_{n-h,h}$ has an action of $G \times J \times \tilde{W}_{K}$. When $h = 0$, we have $G = GL(n,K)$, $J = D_{1/n}^{\times}$, and then $\Psi^{i}_{n} = \Psi^{i}_{n,0}$ is called the fundamental local representation of $G \times J \times \tilde{W}_{K}$. More precisely, the virtual representation

$$[\Psi_{n}] = \sum (-1)^{i} \Psi^{i}_{n}$$

will be treated as the fundamental local representation. All information regarding supercuspidal representations of $GL(n,K)$ is contained in the representation on $\Psi^{n-1}_{n}$.

Let $h = 0$, and identify $\tilde{M}_{n-h,0;\infty,\text{red}}$ with $\mathbb{Z}$ as in (3.3.4); let $x_{0} \in \tilde{M}_{n-h,0;\infty,\text{red}}$ correspond to $0 \in \mathbb{Z}$ (quasi-isogenies of height 0). The stalk $\Psi^{i}_{n-h,0,x_{0}}$ of $R^{i} \Psi^{\text{form}}_{\ell}$ at $x_{0}$ inherits a representation of the isotropy subgroup at $x_{0}$

$$A_{K,n-h} \subset GL(n-h,K) \times J_{n-h} \times \tilde{W}_{K}. \quad \text{(4.3.2)}$$

Writing $g$ instead of $n-h$, the group $A_{K,g}$ can be characterized as the kernel of the map

$$\delta : GL(g,K) \times J_{g} \times W_{K} \longrightarrow \mathbb{Z}$$

defined by

\begin{equation}
\delta(\gamma,j,\sigma) = w_{K}(\det(\gamma)) - w_{K}(N(j)) - w(\sigma)
\end{equation}

where $w_{K}$ is the valuation on $K$, $N : J \rightarrow K^{\times}$ is the reduced norm, and $w(\sigma)$ is the valuation on $\tilde{W}_{K}$ induced by $w_{K}$ via the reciprocity isomorphism $W^{ab}_{K} \sim K^{\times}$. It is then clear that

\begin{equation}
\Psi^{i}_{n-h,0} = c\text{-Ind}_{A_{g,K}}^{GL(g,K) \times J_{g} \times W_{K}} \Psi^{i}_{n-h,0,x_{0}},
\end{equation}

where $c\text{-Ind}$ denotes induction with compact support.

For the sake of honesty, we will also need the version including polarization; this is $\Psi^{i}_{n-h,h,\pm}$ starting from $\tilde{M}_{n-h,h;m}^{\pm}$. The action is now complicated by an extra factor of $\mathbb{Q}_{\ell}^{\times}$ in each of $G$ and $J$, and we define

$$J_{n-h,h,\pm} = D_{1/(n-h)}^{\times} \times GL(h,K) \times \mathbb{Q}_{\ell}^{\times}. \quad \text{(4.3.5)}$$

The vanishing cycles of Berkovich satisfy the same spectral sequence as in the algebraic setting:

\begin{equation}
E^{p,q}_{2} = H^{p}(Z_{s},R^{q}\Psi_{\ell}(\mathbb{Q}_{\ell})) \Longrightarrow H^{p+q}(Z_{\mathfrak{r}},\mathbb{Q}_{\ell}),
\end{equation}

where now $Z_{\mathfrak{r}}$ is the generic fiber of $Z$ in the sense of Raynaud-Berthelot – i.e. a rigid analytic space – and the cohomology on the right is Berkovich’s étale cohomology of
analytic spaces. But we don’t need this. What we do need is Berkovich’s comparison theorem, which we state in the case when the special fiber of $X$ is a single point $x \in Z_x$.

Thus let $f : Z \to \text{Spec } O$ as before, $x \in Z_x$ a geometric point, and let $X = Z_x$. Then

(4.3.6) Theorem (Berkovich, [BII, Theorem 3.1]). — There is a canonical isomorphism

$$R\Psi^\text{form}_{\mathbb{Q}_\ell} \sim (R\Psi_{\mathbb{Q}_\ell})_x.$$ 

In other words, the vanishing cycles in the algebraic category depend only on the formal completion.

The canonicity of the isomorphism implies that it commutes with all correspondences on $Z$ in a natural sense. Thus, fix an isogeny class

$$S(x) = \lim_{\leftarrow U,m} S_{U,m}(x) \subset \lim_{\leftarrow U,m} A_{U}(B, \star).$$

This is a profinite set, and its cohomology is defined as the direct limit of cohomology of finite quotients. Via the local uniformization maps (3.4.10), as $U$ and $m$ vary, Berkovich’s comparison theorem defines an isomorphism of $G(A_f) \times W_K$-equivariant sheaves on $S(x)$:

(4.3.7) $$(\Psi_{n-h,h,+}^{X} \times (I_x(\mathbb{Q})/G^{(h)}(A_f))/J_{n-h,h,+} \sim \sim R^i\Psi_{\mathbb{Q}_\ell} |_{S(x)}.$$

(4.3.8) Remarks

(i) When $h = 0$, the set $S(x)$ maps bijectively to $\lim_{\leftarrow U} S_{U,0}(x)$. This is because the group $G_x$ is connected and because, over a reduced base, $G_x[\varpi^m]$ has a unique Drinfel’d basis, namely the trivial one (exercise). Hence we may view $R^i\Psi_{\mathbb{Q}_\ell} |_{S(x)}$ as a sheaf on the $h = 0$ stratum in $\lim_{\leftarrow U} A_{U}(0)(B, \star)$, though the vanishing cycles themselves require $m \to \infty$.

(ii) For $h > 0$, this is no longer true. On the one hand, the set $\tilde{M}_{n-h,h;M}$ maps to a product, as we will see, of $\tilde{M}_{n-h,0}$ and $\tilde{M}_{0,h}$. The second factor is just $GL(h, K)$, with $G = J = GL(h, K)$ acting on right and left. This is the $GL(h)$-factor of $J_{n-h,h,+}$. In the quotient, we have therefore an extra $GL(h, K)$-term in the limit.

(4.3.9) Proposition. — Suppose $h = 0$. Then the fundamental local representation of $G \times J \times W_K$ on $\Psi_{n}^{i}$ is admissible as a $G \times J$-module (or rather $Z$-admissible: see Remark (4.3.9.1), below) and satisfies the analogue of the continuity lemma:

$$\Psi_{K,\ell,n,0,m} = (\Psi_{n}^{i})^\Gamma_{m}$$

where $\Gamma_{m} \subset G$ is the principal congruence subgroup.

Proof. — The admissibility is a consequence of

(1) Each $\Psi_{K,\ell,n,0,m}$ is constructible (i.e., the stalks are finite-dimensional).

(2) The continuity lemma.
Consider (1) first. This follows from uniformization and the constructibility of vanishing cycles in the algebraic setting, provided we know the supersingular locus \( S(0) \) is non-empty. This we have already promised to prove later (by explicitly exhibiting points). As for (2), it follows again from the continuity lemma in the algebraic setting.

\( \Box \)

**Remark.** — In fact, the above proposition is not quite true as stated, for elementary reasons: the center of \( G \times J \) translates the connected components of \( \tilde{M}_n \) and hence has no finite-dimensional invariant subspaces. The correct statement would be that, for any character \( \xi \) of the center \( Z \) of \( G \), the maximal quotient of \( \Psi^i_n \) on which \( Z \) acts as \( \xi \) is an admissible \( G \times J \)-module. Perhaps this should be called \( Z \)-admissible. In any case, this is all we need for the applications. An analogous (but more serious) correction needs to be effected for general \( h \) below.

Let \( g \) be a non-negative integer. Before continuing, we need to introduce a “compactly supported” version of \( \Psi^i_{g,0} \). Let

\[
\Psi^i_{c,g} = c \text{-Ind}^{GL(g,K) \times J_g \times W_K}_{A_g,K} (\Psi^i_{g,0,x_0})^\vee.
\]

Thus

\[
\Psi^i_{c,g} = \lim_{\longrightarrow m} H^0(M_{g,0;m,\text{red}}, (R^i \Psi_{\text{form}}^g Q_{\ell})^\vee)
\]

is just the cohomology of the dual of \( R^i \Psi_{\text{form}}^g Q_{\ell} \). The subscript \( c \) is included to reflect the fact that, for more general Shimura varieties, one obtains the analogous construction as the compactly supported cohomology, in Berkovich’s sense, of the tube over a connected component of an isogeny class; cf. [H3] and [Fa]. In general, this dual construction behaves better in general with respect to the action of the center; here the difference is slight.

I can now state one of the main theorems of [HT]:

**Theorem ([HT]).** — Let \( g \) be a non-negative integer. Let \( \pi \in A_0(g,K) \), and let \( JL(\pi) \) denote the corresponding representation of \( J = D^\times_{1,g} \) under the Jacquet-Langlands correspondence (A.1.13).

(i) We have

\[
[\Psi_g(JL(\pi))] \overset{\text{def}}{=} \sum_i (-1)^i [\text{Hom}_J(\Psi^i_{c,g}, JL(\pi))] \overset{\sim}{\longrightarrow} (-1)^{g-1} [\pi \otimes r_{\ell}(\pi)]^\vee
\]

in Groth\((G \times W_K)\), where \( r_{\ell}(\pi) \) is a \( g \)-dimensional irreducible representation of \( W_K \).

(ii) (Cf. Proposition 5.2.18 below.) Let \( \pi' \neq \pi \in A(g,K) \) be a discrete series representation. Then for all \( i \), \( \text{Hom}_J(\Psi^i_{c,g}, JL(\pi')) \) contains no \( G \)-subquotients isomorphic to \( \pi \).

(iii) Finally, \( \sigma_{\ell}(\pi) \), defined by

\[
\sigma_{\ell}(\pi) = r_{\ell}(\pi \otimes |\cdot|^{(g-1)/2}),
\]
satisfies all the conditions of the local Langlands correspondence (except possibly compatibility with \( \varepsilon \) factors).

This was conjectured by Carayol [Ca3], following earlier conjectures of Deligne. The proof of Theorem 4.3.11 is given in \( \S \) 5, assuming some consequences of the point-counting argument that will be completed in the subsequent sections.

Now recall from Lecture 3 the notion of Drinfel’d basis of type \( M \). Here \( M \subset \varpi^{-m}O^n/O^n \) is a direct summand isomorphic to \( \varpi^{-m}O^{n-h}/O^{n-h} \), and \( p \) is a Drinfel’d level structure of type \( M \) on \( G \) if \( p | M \) is a Drinfel’d structure on \( G_0 \) and \( p \mod M \) induces a Drinfel’d level structure on the étale quotient. We have the decomposition

\[
\tilde{M}_{g,h;m} = \coprod M \tilde{M}_{g,h,M}
\]

We write

\[
\Psi_{K,t,n-h,h,m} = \oplus M \Psi_{K,t,n-h,h,M}
\]

Fix one \( M = M_0(m) \) (in standard position) and let \( P_{h,0} \subset GL(n) \) the standard maximal parabolic of type \( (n-h,h) \). Let \( O_m = O/\varpi^mO \). There is an isomorphism

\[
(4.3.12) \quad \text{Ind}_{P_{h,0}(O_m)}^{GL(n,O_m)} \Psi_{K,t,n-h,h,M_0(m)} \cong \Psi_{K,t,n-h,h,m}
\]

that sends a function \( f : GL(n,O_m) \to \Psi_{K,t,n-h,h,M_0(m)} \) satisfying \( f(pg) = pf(g) \) for \( p \in P_{h,0}(O_m) \) to

\[
[GL(n,O_m):P_{h,0}(O_m)]^{-1} \sum_{g \in P_{h,0}(O_m) \setminus GL(n,O_m)} g^{-1}f(g).
\]

It is easy to check that this is an isomorphism of \( GL(n,O_m) \)-modules (on the induced representation, the action of \( h \) takes \( f \) to \( f^h(g) = f(hg) \), and

\[
\sum_g g^{-1}f^h(g) = \sum_g g^{-1}f(gh) = \sum_g hg^{-1}f(g).
\]

Let \( V_m = \Psi_{K,t,n-h,h,M_0(m)} \), \( H_m = \Psi_{K,t,n-h,h,m} \), \( I_m = \text{Ind}_{P_{h,0}(O_m)}^{GL(n,O_m)} \). The denominator (which doesn’t work integrally!!) makes the following diagram commute:

\[
\begin{array}{ccc}
I_m V_m & \longrightarrow & H_m \\
\downarrow & & \downarrow \\
I_{m'} V_{m'} & \longrightarrow & H_{m'}
\end{array}
\]

for \( m' > m \), where the right-hand side is just pullback and the left-hand side identifies \( I_m V_m \) with functions on \( O_{m'} \) that pullback from functions on \( O_m \) and take values in the image of \( V_m \) in \( V_{m'} \) under the natural pullback. Thus in the limit this defines an isomorphism

\[
(4.3.13) \quad \lim_m I_m V_m \sim \lim H_m = \Psi_{n-h,h}
\]
Let
\[ \Psi_{K,\ell,n-h,h,M_0}^i = \lim_{m \to \infty} \Psi_{K,\ell,n-h,h,M_0(m)}^i. \]

**Proposition.** — There is a canonical isomorphism of \( G \times J \times W_K \)-modules
\[ \text{Ind}_{P_{h,0}}^{GL(n,K)} \Psi_{K,\ell,n-h,h,M_0}^i \sim \Psi_{n-h,h}^i. \]

**Proof.** — Given the above constructions, it remains to identify the left-hand side via
\[ \lim_{m \to \infty} I_m V_m \sim \text{Ind}_{P_{h,0}(K)}^{GL(n,K)} \lim_{m \to \infty} V_m. \]
But this follows from the Iwasawa decomposition
\[ GL(n, K) = P_{h,0}(K) \cdot GL(n, \mathcal{O}) \]
which identifies the right-hand side with the locally constant functions in
\[ \text{Ind}_{P_{h,0}(\mathcal{O})}^{GL(n,\mathcal{O})} \lim_{m \to \infty} V_m, \]
and the fact that locally constant functions come from the left-hand side. \( \square \)

For convenience we have ignored the polarization datum (the \( + \)); now we put it back. We consider an individual \( \Psi_{K,\ell,n-h,h,M_0}^i \), as \( P_{h,0}(\mathcal{O}) \)-module. First note that it is *finite-dimensional*. This follows from Berkovich’s theorem, once we can exhibit it as the stalk (at a point of the stratum \( S^{(h)} \)) of the global vanishing cycles on the Shimura variety. On the other hand it follows as before, from the continuity lemma, that
\[ \Psi_{K,\ell,n-h,h,M_0}^i = (\Psi_{K,\ell,n-h,h,M_0}^i)^{K_F(m)}. \]
Thus \( \Psi_{K,\ell,n-h,h,M_0}^i \) is an admissible \( P_{h,0}(K) \)-module. But by a standard lemma this implies that

**Lemma.** — The unipotent radical of \( P_{h,0}(K) \) acts trivially on \( \Psi_{K,\ell,n-h,h,M_0}^i \).

For the reader’s convenience, I include the proof, taken from Lemma 13.2.3 of Boyer’s thesis [Bo], where it is attributed to Henniart.

**Proof.** — We write \( P = P_{h,0}(K), N = R_u P, L \) a Levi subgroup of \( P \). We will show that, if \( V \) is any admissible representation of \( P \), then \( N \) acts trivially on \( V \). The proof has nothing to do with \( GL(n) \). Let \( v \in V, \) and let \( U \subset P \) be an open compact subgroup such that \( v \in V^U \). By shrinking \( U \) if necessary, we may assume \( U = U_L \cdot U_N \) where \( U_L = U \cap L, U_N = U \cap N \). Choose an element \( z \) in the center of \( L \) such that \( ad(z) \) is expanding on \( N \); i.e., such that
\[ \cdots z^{-n} U z^n \subset \cdots \subset z^{-1} U z \subset U \subset z U z^{-1} \subset \cdots \]
and such that
\[ \bigcup_{n \geq 0} z^n U z^{-n} = N. \]
For $n \in \mathbb{Z}$, let $V_n$ denote the subspace of $V$ fixed by $z^{-n}Uz^n$. Thus
\begin{equation}
(4.3.15.3) \quad V_n \subset V_{n+1}
\end{equation}
for all $n$, and $v \in V_0$. On the other hand, the action of $z$ on $V$ defines an isomorphism $V_n \sim V_{n-1}$. In particular, all the $V_n$ have the same (finite) dimension. Thus the inclusions (4.3.15.3) are isomorphisms. Hence

$$V_0 = \bigcap_{n \leq 0} V_n \subset V^N$$

by (4.3.15.2). Since $v$ was arbitrary, we find that $V = V^N$, as claimed. \hfill \Box

(4.3.16) Remark. — In [HT] we used the weaker fact that, if $V$ is a smooth $P$-module which is admissible as an $L$-module, then the action of $N$ on $V$ is trivial. Proving that $\Psi_{K, l, n-h, M_\ell}$, as defined above, is an admissible $P_{h,0}(K)$-module is straightforward, as we have seen; whereas proving admissibility as $GL(n-h, K) \times GL(h, K)$-module is rather more complicated. The strategy followed in [HT] involves replacing the strata $\mathfrak{S}^{(h)}_{U(m)}$ of the Shimura variety by the “Igusa varieties of the first kind,” moduli spaces defined abstractly in such a way as to eliminate the action of the unipotent radical of $P_{h,0}(K)$. As ringed spaces, the Igusa varieties of the first kind are isomorphic to the reduced strata $(\mathfrak{S}^{(h)}_{U(m)})_{\text{red}}$; however, the structural maps to the strata of level zero differ by a power of Frobenius (precisely the power needed to annihilate the connected part of the Drinfel’d level structure of level $m$). The advantage of the present approach is that, once the adelic group action has been defined on the full integral model $\lim_{\longrightarrow m} \mathcal{A}_{U(m)}(B, \ast)$, as in §3.5, it is not necessary to define a separate adelic group action on the inverse limit of the strata $\mathfrak{S}^{(h)}_{U(m)}$. By contrast, in the approach followed in [HT], the action on the Igusa varieties of the first kind had first to be defined separately, then shown to be consistent with the action on the strata.

Now recall the $P_{h,0}(\mathcal{O}_m) \times J \times W_K$-equivariant morphism

$$\pi_1 = \pi_{1, M_{0,0}} : \tilde{M}^+_{n-h,h,M_{0,0}} \longrightarrow \tilde{M}^+_{n-h,h,M_{0,0}} \times \tilde{M}_{0,h,m}.$$ 

This is the quotient by the unipotent radical of $P_{h,0}(\mathcal{O}_m)$ (recall that the subscript $m,0$ designates a Drinfel’d structure on the connected part only). By Proposition 3.3.5 this morphism induces an isomorphism on reduced $k$-sub schemes. We write $R^i\Psi_{K, l, n-h,h,M_{0,0}}$ (resp. $R^i\Psi_{K, l, n-h,h,M_{0,0}}$) for Berkovich’s vanishing cycles sheaf $R^i\Psi_{\mathbb{Q}_l}$ over $\tilde{M}^+_{n-h,h,M_{0,0}}$, resp. over $\tilde{M}^+_{n-h,h,M_{0,0}}$. We drop $m$ from the notation for the limit over $m$. The above lemma implies, as in the proof of the Continuity Lemma 4.2.19, that $R^i\Psi_{K, l, n-h,h,M_{0,0}}$ is the pullback via $\pi_1$ of the formal vanishing cycles of $\tilde{M}^+_{n-h,h,M_{0,0}} \times \tilde{M}_{0,h,m}$. But $\tilde{M}_{0,h,m}$ is étale (even discrete) and it follows from Proposition 3.3.5 that $\tilde{M}^+_{n-h,h,M_{0,0}}$ is smooth over $\tilde{M}^+_{n-h,h,M_{0,0}}$. But smooth morphisms preserve vanishing cycles. Let $(\mathbb{Q}_l)_{0,h,m}$ denote the constant sheaf $\mathbb{Q}_l$ over the discrete scheme $\tilde{M}_{0,h,m}$. We write $P = P_{h,0}(K)$, $N = R_uP$, $L$ a Levi subgroup of $P$, which we
identify with $GL(n-h, K) \times GL(h, K)$. Recall that $J = J_{n-h} \times GL(h, K)$. It follows immediately that

**Proposition (4.3.17)**

(i) For any $m$ and any $i$, there is a canonical isomorphism of $P_{h,0} \left( \mathcal{O}_m \right) \times J \times W_K$-equivariant sheaves over $\hat{M}_{n-h, h, M_0(m)}$

$$R^i\Psi_{K,\ell,n-h,h,M_0(m),+} \sim \left[ (\pi_{1,M_0(m)})^* R^i\Psi_{K,\ell,n-h,0; m} \boxtimes \left( \mathbb{Q}_l \right)_{0,h;m} \right]_+.$$ 

Here the action of the unipotent radical is trivial on the right-hand side, and $\boxtimes$ is the external tensor product over the product $\hat{M}_{n-h, m} \times M_0, h, m$.

(ii) In the limit, the isomorphisms above patch together to an isomorphism

$$R^i\Psi_{n-h, h, +} \sim \left[ (\pi_{1,M_0})^* R^i\Psi_{n-h,0} \boxtimes \left( \mathbb{Q}_l \right)_{0,h} \right]_+$$

of $P \times J \times W_K$-equivariant sheaves.

(iii) Define

$$A_{K, n-h} \subset GL(n-h, K) \times J_{n-h} \times W_K \subset P_{h,0}(K) \times J \times W_K$$

as in (4.3.2). Regard $(A_{K, n-h} \times GL(h, K)) \cdot N$ as a subgroup of $P \times J \times W_K$ by extending the natural inclusion of $N$ in $P$ by the natural inclusion of $A_{K, n-h}$ in $GL(n-h, K) \times J_{n-h} \times W_K$ and the diagonal map $GL(h, K) \rightarrow GL(h, K) \times GL(h, K) \subset L \times J$.

Then there is a canonical isomorphism of $P \times J \times W_K$-modules

$$\Psi_{K,\ell,n-h,h,+} \sim \text{c-Ind}_{(A_{K, n-h} \times GL(h, K)) \cdot N}^{G \times J \times W_K} \Psi_{n-h,0,x_0} \otimes 1.$$ 

Here c-Ind denotes compact induction, $\Psi_{n-h,0,x_0}$ is as in (4.3.4), and $1$ is the trivial representation of $GL(h, K)$; $N$ acts trivially on the tensor product.

Here the first two assertions follow from the previous discussion, and (iii) follows from (ii) and (4.3.4) by taking cohomology. Note that compact induction of $1$ from the diagonal in $GL(h, K) \times GL(h, K)$ just gives rise to the two-sided regular representation on $C_c^\infty(GL(h, K))$.

**Corollary (4.3.18).** — The $G \times J \times W_K$ module $\Psi_{n-h, h, +}$ is admissible and continuous and parabolically induced from an admissible (continuous) $GL(n-h, K) \times GL(h, K) \times J \times W_K \times \mathbb{Q}_p^\times$-module (add the extra factor of $\mathbb{Q}_p^\times$ for the $+$).

**Remark (4.3.19).** — As in Remark 4.3.9.1, this is not quite literally true, and in this case the problem is more serious because of the presence of the $GL(h, K) \times GL(h, K)$-action on $C_c^\infty(GL(h, K))$; one has to replace the assertion by one about the maximal quotient on which $Z_G \times GL(h, K)$ acts via any fixed finite sum of irreducible representations. But this is again all we need for the applications. In the future, we will incorporate $GL(h, K)$ with the adeles away from $w$ in order to avoid this issue.
We pause to note what this implies for an isogeny class (lying above) $\overline{S}^{(h)}$:

\[(4.3.20) \quad \text{Ind}_{P_{h,0}(K)}^{GL(n,K)}(R^i\Psi_{n-h,0} \boxtimes (\mathbb{Q}_\ell)_{0,h}) \times (I_x(\mathbb{Q}) \setminus G^{(h)}(A_f))/J_{n-h,h,+} \sim R^i\Psi_{\mathbb{Q}_\ell}|_{S(x)}.
\]

4.4. The first basic identity in the Grothendieck group. — We now want to apply this to the global cohomology. Just as in the formal setting, the stratum $\overline{S}^{(h)}_{U(m)}$ is the disjoint union:

$$\overline{S}^{(h)}_{U(m)} = \bigsqcup_M \overline{S}_{U,M}^{(h)}$$

(the $h$ and $m$ are determined by $M$). Then the above decomposition becomes

\[\overline{S}^{(h)}_{U(m)} = \bigsqcup_{\gamma \in GL(n,\mathcal{O}_m)/P_{h,0}(\mathcal{O}_m)} \gamma(\overline{S}_{U,M_0(m)}).
\]

(4.4.2) Lemma. — For each fixed $U^w$, the $P_{h,0}(K)$-module

$$H = \lim_{m} H_{\ell}^p(\overline{S}_{U,M_0(m)}, R^q\Psi)$$

is admissible.

Proof. — Let $V_m = H_{\ell}^p(\overline{S}_{U,M_0(m)}, R^q\Psi)$, $\Gamma^h_m$ the principal congruence subgroup $(1 + \varpi^m M(n,\mathcal{O})) \cap P_{h,0}(\mathcal{O})$. It suffices to prove:

$$V_m = H^1_{\ell}.$$  

As in the previous discussion, this follows from the appropriate continuity lemma:

$$R^q\Psi(m) \sim f_{m',m,\varphi}\Psi(m')^{\Gamma^h_m}.$$  

This is a stalkwise calculation, hence we are reduced by uniformization to the corresponding continuity lemma for $\Psi_{n-h,h,+}^{(h)}$. Proposition 4.3.17 reduces this to the case $h=0$, which we have already proved. 

(4.4.3) Proposition (cf. [Bo]). — There is a natural isomorphism

$$H_{\ell}^p(\overline{S}^{(h)}, R^q\Psi) \sim \text{Ind}_{P_{h,0}(K)}^{GL(n,K)} \lim_{U^w,m} H_{\ell}^p(\overline{S}_{U,M_0(m)}, R^q\Psi).$$

(Note: this is non-normalized induction.) In particular, $H_{\ell}^p(\overline{S}^{(h)}, R^q\Psi)$ is an admissible $G(A_f)$-module.

Proof. — We first observe that the action of $P_{h,0}(K)$ on $\lim_{m} \overline{S}^{(h)}_{U,m}$ stabilizes $\lim_{m} \overline{S}^{(h)}_{U,M_0(m)}$. Indeed, recall that the action of $GL(n,K)$ on $\tilde{M}_{n-h,h}$ was defined by inducing from that of $P_{h,0}(K)$ (which was denoted $P_{\tilde{M}}$) on $\{\tilde{M}_{g,h,p}^{-1} \tilde{M}/\tilde{M}\}$. The same argument works globally. On the other hand, for each level $m$, the stabilizer in $GL(n,\mathcal{O})$ of $\overline{S}^{(h)}_{U,M_0(m)}$ is $P_{h,0}(\mathcal{O}_m)$ modulo $1 + \varpi^m M(n,\mathcal{O})$. In the limit, the stabilizer in $GL(n,\mathcal{O})$ of $\lim_{m} \overline{S}^{(h)}_{U,M_0(m)}$ is $P_{h,0}(\mathcal{O}_m)$. By the Iwasawa decomposition,
it follows that \( P_{h,0}(K) \) is the stabilizer in \( GL(n, K) \) of \( \varprojlim_m \overline{S}_{U,M_0(m)} \). From this it follows formally that \( H^p_c(\overline{S}^{(h)}_{M_0}, R^q\Psi) \) is the representation compactly induced from the action of \( P_{h,0}(K) \) on \( \varprojlim_m H^p_c(\overline{S}_{U,M_0(m)}, R^q\Psi) \). Since the quotient is compact, this is the full induced representation.

In more detail, the argument is just the same as in the proof of Corollary 4.3.18. ☐

Write

\[
H^p_c(\overline{S}^{(h)}_{M_0}, R^q\Psi) = \lim_{U^w,m} H^p_c(\overline{S}_{U,M_0(m)}, R^q\Psi)
\]

We thus obtain the following formula for the cohomology as \( G(A_f) \times W_K \)-module:

**(4.4.4) First Basic Identity.** — The following identity holds in \( \text{Groth}(G(A_f) \times W_K) \):

\[
[H(A(B, \ast))] = \sum_{p,q,h} (-1)^{p+q} \left[ \text{Ind}_{P_{h,0}(K)}^{GL(n,K)} H^p_c(\overline{S}^{(h)}_{M_0}, R^q\Psi) \right].
\]

We may consider an isogeny class in \( \overline{S}^{(h)}_{M_0} \) (with base point \( x \), say); this means that the level structure is arbitrary away from \( w \), but of type \( M_0 \) at \( w \). Let \( \overline{S}(x)_{M_0} \) denote this isogeny class. Then (4.3.20) yields

\[
[(R^j\Psi_{n-h,0} \boxtimes (\mathbb{Q}_x)_{0,h,m})] \times (I_x(\mathbb{Q}) \setminus G^{(h)}(A_f))] / J_{n-h,h,+} \xrightarrow{\sim} R^j\Psi Q_{\ell} |_{\overline{S}(x)_{M_0}}.
\]

Here \( G^h \) contains a factor \( GL(h, K) \), and \( M_{0,h} \xrightarrow{\sim} GL(h, K) \), so \( \overline{S}(x)_{M_0} \) can also be written (cf. (3.3.4)

\[
I_x(\mathbb{Q}) \setminus (\mathbb{Z} \times GL(h, K) \times (\mathbb{Q}_p^\times / \mathbb{Z}_p^\times) \times G(A_f^p))
\]

where the action of \( I_x(\mathbb{Q}) \) on \( \mathbb{Z} \times (\mathbb{Q}_p^\times / \mathbb{Z}_p^\times) \times GL(h, K) \) is given by composing the inclusion of \( I_x(\mathbb{Q}) \) in \( J_{n-h,h,+} = J_{n-h} \times GL(h, K) \times \mathbb{Q}_p^\times \) with the projection of the latter on \( \mathbb{Z} \times GL(h, K) \times (\mathbb{Q}_p^\times / \mathbb{Z}_p^\times) \) whose first factor is \( j \mapsto w(N(j)) \).

**Remark.** — To obtain admissibility, one has to work with \( U^w \times U_h \)-fixed vectors, for \( U_h \) open compact in \( GL(h, K) \); then the finiteness condition holds for the action of \( Z_{GL(n-h,K)} \) as discussed above.

For applications to point counting, it will be necessary to consider the stalks of \( R^j\Psi Q_{\ell} |_{\overline{S}(x)_{M_0}} \) at a point, say \( x \), in \( \overline{S}(x)_{M_0} \). Let \( G_{n-h,h,+} = GL(n-h, K) \times GL(h, K) \times \mathbb{Q}_p^\times \). It follows from (4.4.5) that

\[
R^j\Psi Q_{\ell} |_{x} \xrightarrow{\sim} \Psi_{n-h,0,x_0}^1 \psi_{n-h,0,x_0}
\]

in the notation of (4.3.4). This is a module for the group \( A_{K,n-h} \) introduced in (4.3.4).

Let

\[
J_{n-h}^0 = \ker w \circ N : D_1^{\times} / (n-h) \to \mathbb{Z}.
\]
Then \( J^0_{n-h} \) is naturally a subgroup of \( A_{K,n-h} \); moreover, \( J^0_{n-h} \times \{1\} \subset J_{n-h,h,+} \) is the isotropy group of a point \( \bar{x} \in \bar{M}_{n-h,n,+,\text{red}} \times I_2(\mathbb{Q}) \backslash G(h)(\mathbb{A}_f) \) above \( x \) for the uniformization (3.4.10). It follows easily that

\[ (4.4.8) \text{Lemma.} - R^i\Psi_{Q_\ell} \text{ is the sheaf on } \overline{S}(x)_{M_0} \text{ associated to the representation of the isotropy group } J^0_{n-h} \text{ on } \Psi^i_{n-h,0,x_0}. \]

### 5. Construction of a local correspondence

The present lecture contains most of a proof of Theorem 4.3.11, stated in the previous lecture: the fundamental local representation realizes the Jacquet-Langlands and local Langlands correspondence for supercuspidal representations, except (for the time being) for the compatibility with local epsilon factors of pairs. The proof roughly follows the lines of Boyer’s thesis [Bo], but at some points, notably in the treatment of harmonic analysis, the point of view is closer to that of [H1]. An idea discovered by P. Boyer and exploited in his thesis shows that the local supercuspidal representations are concentrated in the zero-dimensional stratum. The construction of a local correspondence then follows by a comparison of trace formulas, as in [Bo] or [H1] (the latter in the case of \( p \)-adic uniformization). This construction is also the basis of the induction that permits us to determine the (virtual) contributions of all strata to the cohomology of the generic fiber.

The proof of Theorem 4.3.11 depends on a weak qualitative consequence (Lemma 6.13.1) of the point-counting argument that will be completed in §§6-7 (the Second Basic Identity, §6.1). Theorem 4.3.11 in turn is used to provide the strong version of the point-counting argument required to prove the Main Theorem (1.3.6).

#### 5.1. Applications to supercuspidal representations.

The following argument was first developed by Boyer, in the setting of Drinfel’d modular varieties, and is the starting point for our induction.

We let \([H(A(B,*))]_0\) denote the formal sum of all \( G(A_f) \times W_K \) modules in \([H(A(B,*))]\) that are supercuspidal as \( GL(n,K) \)-modules. By definition, there is no intertwining with induced representations. Hence the First Basic Identity (4.4.4) has a supercuspidal version:

\[ (5.1.1) \quad [H(A(B,*))]_0 = \sum_{p,q} (-1)^{p+q}[H^p_c(S(\overline{S}^{(0)}), (R^q\Psi)_0)], \]

where the subscript 0 on the right also means supercuspidal, in this case under the action of \( GL(n,K)^0 = \ker w \circ \det \subset A_{K,n} \). Here we are using the fact that any supercuspidal representation of \( GL(n,K) \) restricts to a finite sum of irreducible representations of \( GL(n,K)^0 \) that intertwine with no non-supercuspidal representation.
of $GL(n, K)$. Since $\overline{\mathcal{S}}^{(0)}$ is of dimension zero, we just find
\[(5.1.2) \quad \sum_i (-1)^i [H^i(\mathcal{A}(B, *))]_0 = \sum_q (-1)^q [H^0(\overline{\mathcal{S}}^{(0)}), (R^0 \Psi)_0]. \]
Indeed, there is a stronger assertion. The spectral sequence for vanishing cycles, applied to the supercuspidal part, has the form
\[(E^2_{p,q})_0 = \lim_{U^\infty,m} H^p(\overline{\mathcal{S}}_{U(m)}, R^q \Psi)_0. \]
But the same dévissage shows that
\[(5.1.3) \quad H^p(\overline{\mathcal{S}}_{U(m)}, R^q \Psi)_0 = H^p(\overline{\mathcal{S}}^{(0)}, (R^q \Psi)_0) = 0 \text{ unless } p = 0. \]
Thus the spectral sequence degenerates at $E_2$ and we have
\[(5.1.4) \quad H^i(\mathcal{A}(B, *), \mathbb{Q}_\ell)_0 \sim H^0(\overline{\mathcal{S}}^{(0)}, (R^i \Psi)_0), \quad i = 1, \ldots, 2n - 2. \]
Now Matsushima’s formula (1.1.3), plus the complex-analytic uniformization (2.1.2) of $\mathcal{A}(B, *)$, writes the left-hand side as
\[(5.1.5) \quad |\ker^i(\mathbb{Q}, G)| \cdot \bigoplus_{\pi} H^i(\mathfrak{g}, Z_{G}(\mathbb{R}) \cdot K_{\infty}; \pi_{\infty}) \otimes \pi_f. \]
Here $\pi$ runs through automorphic representations of $G$ that are supercuspidal at $w$. Recall the base change map from the first lecture. From $\pi$, one can find a pair $(\Pi, \psi)$, with $\Pi \subset \mathcal{A}_0(\mathbb{G}(n)_{\mathbb{F}})$, which is a base change at all unramified places (for $\pi$) and all places that split in $E$. In particular, $\Pi_w$ is supercuspidal, hence $\Pi$ is cuspidal.

Clozel’s purity lemma then implies that $\pi_{\infty}$ is in the discrete series, hence only has cohomology in the middle degree $n - 1$. Indeed, suppose this were not the case. Then by Lefschetz theory, there would be an integer $0 < i \leq n - 1$ such that
\[H^n(\mathfrak{g}, Z_{G}(\mathbb{R}) \cdot K_{\infty}; \pi_{\infty}) \neq 0 \iff a \in \{n - 1 - i, n - 1 - i + 2, \ldots, n - 1 + i\}. \]
Thus $H^n(\mathcal{A}(B, *), \mathbb{Q}_\ell)$ contains $\pi_f$ for at least two distinct $a$ of the same parity. By Deligne’s purity theorem (recall that $\mathcal{A}(B, *)$ is smooth and projective), the Frobenius eigenvalues on $H^n(\mathcal{A}(B, *), \mathbb{Q}_\ell)$ at unramified places $v$ have complex absolute values $q_v^{a/2}$; thus at unramified places $v$ that split in $E$, say, the Satake parameters of $\Pi_v$ have several distinct complex absolute values of the form $q_v^{a/2}$. But $\Pi$ is cuspidal, hence every $\Pi_v$ is generic by Shalika’s theorem. Moreover, $\Pi_v$ is unitary, up to twist by a character of the determinant. The classification of generic unitary representations of $GL(n, F_v)$ (in fact, the Jacquet-Shalika estimates) shows that all the Satake parameters have the same complex absolute value (the ratio is always $\leq q_v^{1/2}$). This completes the argument.

Thus we have
\[(5.1.6) \quad H^{n-1}(\mathcal{A}(B, *)_0 \sim H^0(\overline{\mathcal{S}}^{(0)}), (R^{n-1} \Psi)_0); \]
\[(5.1.7) \quad (R^i \Psi)_0 = 0, \quad i \neq n - 1. \]
Looking more closely at $\mathcal{S}^{(0)}$ and using a comparison of trace formulas, we can use this identity to construct a candidate for the local Langlands correspondence, for supercuspidal representations, on $R^{n-1}\Psi_{n,0}$. This is how Boyer proved Carayol’s conjecture in the equal characteristic case. The present lecture carries out the analogous constructions in the mixed characteristic situation.

5.2. The basic locus and construction of a local correspondence. — We return to the basic, or supersingular, locus $\mathcal{S}^{(0)}$, for two reasons. First, this will allow us to prove Theorem (4.3.11)(i) and (ii): we construct the local correspondence, as conjectured by Carayol, on the vanishing cycles in the basic case ($h = 0$). We have seen that this determines the stalks of the vanishing cycles at for all $h$, and we use this to study the remaining strata. The other reason is that it provides a gentle introduction to the problem of counting points. The arguments generalize those of Carayol’s thesis (in the case $n = 2$) and of Boyer’s thesis (in equal characteristic). However, we have to contend with problems related to the failure of the Hasse principle, which complicates the argument slightly.

Let $x \in \mathcal{S}^{(0)}$. Recall the uniformization (3.4.10) of the isogeny class, in the case $h = 0$ (in the limit over $U^w$):

$$\Theta : \left[ I_x(\mathbb{Q}) \backslash \tilde{N}_{n,0,0}(\mathbb{F}) \times G^{(0)}(A_f) \right] / J_{n,0} \xrightarrow{\sim} \mathcal{S}(x).$$

Here $J_{n,0} = J_n \times \mathbb{Q}_p^\times$ and $\tilde{N}_{n,0,0}(\mathbb{F})$ is just $\mathbb{Z} \times \mathbb{Q}_p^\times / \mathbb{Z}_p^\times$, the first factor for the height of the quasi-isogeny, the second for the degree of the polarization. For $h = 0$, $I_x$ turns out to be an inner form of $G$. This is clearly explained in [RZ], from which we take the following Lemma:

(5.2.2) Lemma. — Let $(A, \lambda)$ be a polarized abelian variety over $\mathbb{F}_q$, with $F \subset \text{End}^{\text{d}\text{im}}(A)$, such that the Rosati involution induces complex conjugation on $F$. Let $(N, F)$ be the rational Dieudonné module of $A$ (over $\mathbb{F}$). Consider the decomposition $F \otimes \mathbb{Q}_p = \prod F_{w_i}$, [note change in notation: no more $w^*$!!] and suppose that in the corresponding decomposition of $N = \oplus_i N_i$, each $N_i$ is isoclinic. Then some power of the Frobenius endomorphism $\text{Frob}_A$ over $\mathbb{F}_q$ belongs to $F$.

Remark. — The hypotheses of the lemma are verified for $A_x$ precisely when $x \in \mathcal{S}^{(0)}$.

Proof. — For each $i$ there is a $W(\mathbb{F})$-lattice $M_i \subset N_i$, stable under $F$ and $V$, such that $F^{s_i}M_i = p^{s_i}M_i$. We may assume that $M_i$ is fixed by $O_{F,w_i}$ (this is obvious in our case, since we are starting from an $O$-module at $w$ and elsewhere it is étale, up to Cartier duality). Up to isogeny, we may also assume $\oplus M_i$ is the Dieudonné module of $A$ and $O_K \subset \text{End}(A)$. Without loss of generality we may assume all $s_i = s$ and $q = p^s$. So then $F_q M_i = p^{s_i}M_i$. Let $\text{ord}_i$ be the valuation on $K_i$ with $\text{ord}_i(p) = 1$. 
Consider the following problem in algebraic number theory: Find an element \( u \in K \) that is a unit away from \( p \) and such that
\[
\text{ord}_i(u) = r_i; \quad uu^c = q.
\]
We are allowed to replace \( q \) by \( q^m \), which replaces \( r_i \) by \( mr_i \). For \( m \) sufficiently large, the first equation can be solved. Now the existence of the polarization fixed by \( F_q \) implies \( r_i + r_i^c = s \) for all \( i \) (and this is again obvious in our situation, by duality).
Let \( u' = qu/u^c \). Then
\[
\text{ord}_i(u') = s + r_i - r_i^c = 2r_i; \quad u'(u')^c = q^2.
\]
So up to replacing \( q \) by \( q^2 \), we have solved the equation. Now \( \varepsilon = u^{-1} \text{Frob}_A \) is an automorphism of \( A \) (because it fixes \( \bigoplus_i M_i \), by the first equation) that fixes the polarization (by the second equation). Hence by Serre’s lemma we conclude that some power of \( \varepsilon \) equals 1.

(5.2.3) Corollary. — Let \( (A, \lambda) \) and \( (A', \lambda') \) be two abelian varieties over \( F \) satisfying the same assumptions. Then
\[
\text{Hom}_K^0(A, A') \otimes \mathbb{Q}_\ell \cong \text{Hom}_K(V_\ell(A), V_\ell(A')).
\]
Proof. — This follows from the proposition and Tate’s theorem
\[
\text{Hom}_K^0(A, A') \otimes \mathbb{Q}_\ell \cong \text{Hom}_{\text{Q}_\ell(\text{Frob})}(V_\ell(A), V_\ell(A')).
\]
Now return to \( (A_x, \lambda_x, i_x) \in \mathcal{S}^{(0)} \). Recall the data \( (B, *, V) \) of our original moduli problem. Let \( C = \text{End}_B^0(V) = B^{\text{op}}, \quad C_x = \text{End}_B^0(A_x) \). Recall that we have the involution \( # \) on \( C \), induced by the symplectic embedding of \( G \); let \( #_x \) be the involution on \( C_x \) induced by the polarization \( \lambda_x \). By the Corollary, we have that
\[
C_x \otimes \mathbb{Q}_\ell \cong \text{End}_B(V_\ell(A_x)).
\]
Since \( \ell \neq p \), there is a level structure, i.e. a \( B \)-invariant symplectic similitude \( V_\ell(A_x) \cong V \otimes \mathbb{Q}_\ell \), well defined \( \text{(mod } U_\ell) \). Thus
\[
(C_x \otimes \mathbb{Q}_\ell, #_x) \cong (C \otimes \mathbb{Q}_\ell, #)
\]
as \( F \otimes \mathbb{Q}_\ell \)-algebras with involution. Therefore there is an isomorphism
\[
(C_x \otimes \overline{\mathbb{Q}}, #_x) \cong (C \otimes \overline{\mathbb{Q}}, #)
\]
which induces a \( \overline{\mathbb{Q}} \) isomorphism between \( G \) and
\[
I_x = \{ \gamma \in C_x^\times | \gamma \cdot \gamma^{#x} \in \mathbb{Q}_x^\times \}.
\]
Since \( I_x \) is compact at infinity (mod center), this can only be an inner twist. (An outer twist would be of the form \( GL(a, D) \) for some division algebra \( D \) of dimension \( b^2 \) with \( ab = n \).

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
Note that \( I_{x,p} \overset{\sim}{\longrightarrow} \prod_i I_{x,w_i} \times \mathbb{Q}_p^{\times} \), as usual. Each \( I_{x,w_i} \) is an inner form of \( GL(n, K_{w_i}) \), and \( I_{x,w_i} \subset G_{w_i} \) for \( i \geq 1 \), \( I_{x,w} \subset J_\mathcal{N} \). It follows (by dimension considerations) that these inclusions are isomorphisms. The group \( G^{(0)}(A_f) \) is then just \( I_x(A_f) \), and (5.2.1) becomes

\[
(5.2.5) \quad \Theta : [I_x(Q) \backslash \tilde{M}_{n,0,+}(F) \times I_x(A_f)] / J_{n,+} \overset{\sim}{\longrightarrow} \mathcal{S}(x).
\]

How many basic isogeny classes are there? By the above corollary, we see that, if \( x, x' \in \mathcal{S}^{(0)} \), then \( A_x \sim A_{x'} \) as abelian varieties with \( B \)-action. We may assume \( A_x = A_{x'} = A \). But not necessarily as polarized abelian varieties with \( B \)-action! In any case, \( I_x \) and \( I_{x'} \) are inner forms, isomorphic at all places (at \( p \) this is because the \( p \)-divisible groups are isomorphic as polarized \( B_p \)-modules). Hence

\[
(5.2.6) \text{Lemma.} \quad \text{Up to isomorphism, the group } I_x \text{ is independent of the point } x \in \mathcal{S}^{(0)}.
\]

\textbf{Proof.} — The proof of Lemma (2.3.1) applies to the group \( I_x \). (See also Lemma (6.6.8), below.)

Two polarizations \( \lambda \) and \( \lambda' \) from \( A \) to \( A^\vee \) are equivalent (as \( B \)-morphisms inducing \( * \)) if and only if there exists \( d \in C_x = C_{x'} = B^{\text{op}} \) and \( a \in \mathbb{Q}_p^{\times} \) such that \( \lambda' = \text{ad}^\vee \lambda d \) (\( d^\vee \) being the endomorphism of \( A^\vee \) induced by \( d \)). But any two polarizations differ by an element \( \delta \in C \) via \( \lambda' = \lambda \circ \delta \), and the symmetry of \( \lambda' \) and \( \lambda \) implies that \( \delta = \delta^* \) (\( * \) = Rosati involution); since \( \lambda' \) is a polarization, \( \delta \) must be totally positive. Then \( \lambda \circ \delta = \text{ad}^\vee \lambda d \) if and only if

\[
\delta = a(\lambda^{-1}d^\vee \lambda)d = \text{ad}^* d
\]

has a solution \((a,d)\). The set of solutions of this equation is a \textit{torsor} for the group \( I_x \) (acting on \( d \) on the left), and it has a solution if and only if the torsor is trivial. The set of torsors is parametrized by \( H^1(Q, I_x) \). But there are solutions locally for all primes \( \ell \neq p \), by the existence of the level structure; at \( \infty \) because \( \delta \) is totally positive; and at \( p \) because \( I_{x,p} \) is a product of inner twists of general linear groups, hence has no \( H^1 \) by Hilbert’s theorem 90.

So the set \( \Phi_\delta \) of basic isogeny classes is mapped by this construction to a subset of \( \text{ker}^1(Q, G) \). We will see in Lecture 6, using Honda-Tate theory, that this map is \textit{surjective}. (We still haven’t shown that \( \mathcal{S}^{(0)} \) is non-empty!) Assume this for now.

\[
(5.2.7) \text{Fact.} \quad \text{The cardinality of } \text{ker}^1(Q, G) \text{ is unchanged under inner twist.}
\]

This is proved by Kottwitz [K1, § 4].

Now recall the isomorphism (4.3.7) of vanishing cycles sheaves. In the present setting, this can be rewritten

\[
(5.2.8) \quad [\Psi_{n,+} \times (I_x(Q) \backslash I_x(A_f))] / J_{n,+} \overset{\sim}{\longrightarrow} R^i\Psi_{Q,\ell} \mathcal{S}(x).
\]
It follows formally that
\begin{equation}
R^i\Psi^0 \big|_{\mathcal{S}(x)} \xrightarrow{\sim} \text{Hom}_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I_x/I_x(\mathbb{R}), \mathbb{Q}_\ell)).
\end{equation}

Here $\mathcal{A}$ denotes automorphic forms on the group $I_x$ that are trivial on $I_x(\mathbb{R})$; again this has to be modified if we use twisted coefficients. Moreover, $\Psi^i_{c,n,+}$ is the compact version of $\Psi^i_{n,+}$ (one adds a $+1$ to the definition (4.3.10)).

In what follows, we let $S(B,*) = S(I_x,*)$ be the Shimura variety itself. For any admissible virtual $G(\mathcal{A}_f)$-module $M$, we let $M[\pi^w] = \text{Hom}_{G(\mathcal{A}_f)}(\pi^w, M)$; this is a virtual module over $G_n$. Similarly, we let $M[\pi] = \text{Hom}_{G(\mathcal{A}_f)}(\pi, M)$.

\textbf{(5.2.10) Proposition.} — Let $I = I_x$ for any $x \in \Phi_h$. Let $(\rho, \psi)$ be a representation of $J_{n,+}$ (with $\rho \in \widehat{J}_n$, $\psi$ an unramified character of $Q^\times_p$). Assume $\text{JL}(\rho)$ is a supercuspidal representation of $G_n$. Consider representations $\pi^w$ of $G(\mathcal{A}_f^\times) \xrightarrow{\sim} I(\mathcal{A}_f^\times)$. Then there is an isomorphism
\begin{equation}
H^{n-1}(S(B,*), \mathbb{Q}_\ell)[\pi^w \otimes \text{JL}(\rho) \otimes \psi] \xrightarrow{\sim} \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)[\pi^w \otimes \rho \otimes \psi]^n.
\end{equation}
Moreover, for $i \neq n-1$,
\begin{equation}
H^i(S(B,*), \mathbb{Q}_\ell)[\pi^w \otimes \text{JL}(\rho) \otimes \psi] = 0.
\end{equation}

This will be proved a bit later, by comparison of trace formulas. I remark that $\pi^w$ always determines $\pi_w$ by base change (to $GL(n, F)$) and strong multiplicity one.

Now it follows from the remarks preceding the proposition that
\begin{equation}
H^0(S(0), R^i\Psi^0) \xrightarrow{\sim} \sum_{\Phi_h} \text{Hom}_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I_x/I_x(\mathbb{R}), \mathbb{Q}_\ell))
\end{equation}
\begin{equation}
\xrightarrow{\sim} \text{Hom}_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell))[\ker^1(Q,G)]^1.
\end{equation}
Here $I$ is any $I_x$ for $x \in \Phi_h$.

Recalling that $H^{n-1}(\mathcal{A}(B,*), \mathbb{Q}_\ell) = H^{n-1}(S(B,*), \mathbb{Q}_\ell)[\ker^1(Q,G)]$, it follows from the First Basic Identity (4.4.4) and (5.1.6) that, up to semi-simplification,
\begin{equation}
H^{n-1}(S(B,*), \mathbb{Q}_\ell)_0 \xrightarrow{\sim} \text{Hom}_{J_{n,+}}((\Psi^i_{c,n,+})_0, \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)).
\end{equation}

Here, as above, the subscript $0$ on the left-hand side denotes the $GL(n, K)$-supercuspidal subspace. On the right-hand side it’s essentially the same thing, but one has to be a bit careful because the center does not act semi-simply. However, any $J_{n,+}$-homomorphism from $(\Psi^i_{c,n,+})_0$ to the space of automorphic forms factors through a quotient on which the center does act semi-simply, so (5.2.12) makes sense as written. Alternatively, one can define the supercuspidal subspace of any smooth $GL(n, K)$-module using the Bernstein center; in this way one sees it is always a direct summand.

By Matsushima’s formula (1.1.3), the left-hand side of (5.2.12) is
\begin{equation}
\bigoplus_{\pi \in \mathcal{A}_0} H^{n-1}(g, Z_G(\mathbb{R})K; \pi_\infty) \otimes \pi_f.
\end{equation}
Fix $\rho$ as above and $\pi_f$ with component $JL(\rho)$ (supercuspidal) at $w$. For given $\rho$, this is always possible (see (5.2.15), below). Let $R_{\ell}(\pi_f)$ denote the semisimplified representation of $Gal(\overline{Q}/F)$ on $Hom_{GL(A_{\ell})}(\pi_f, H^{n-1}(S(B,\ast), \mathbb{Q}_{\ell})_0)$. As we saw in Lecture 1, $R_{\ell}(\pi_f)$ is the sum of some copies of an $n$-dimensional semisimple representation $R_{\ell,0}(\pi_f)$. Let $r_{\ell}(\pi_f)$ be the contragredient of $R_{\ell,0}(\pi_f)$, twisted by $\psi \circ N_{K/Q_p}$ as in § (1.3) to remove the contribution of $\psi$. Combining the above identities, we find

(5.2.13) Theorem. — Let $\rho$ be a representation of $J_n$ such that $JL(\rho)$ is supercuspidal. Then as $G_n \times W_K$-modules, we have

$$JL(\rho) \otimes r_{\ell}(\pi_f) |_{Gal(\overline{K}/K)} \sim \sum_i (-1)^{n-1+i} [Hom_{J_{n,+}}((\Psi_{c,n,+}^{-1})_0, \rho \otimes \psi) \otimes r_{\ell}(\psi \circ N_{K/Q_p}^{-1})]$$

Proof. — The first isomorphism is a summary of the preceding discussion; we simply apply $[\pi^w]$ to both sides of (5.2.12). Similarly, the isomorphism

$$[Hom_{J_{n,+}}((\Psi_{c,n,+}^{-1})_0, \rho \otimes \psi) \otimes r_{\ell}(\psi \circ N_{K/Q_p}^{-1})] \sim \sum_i (-1)^{n-1+i} [Hom_{J_{n,+}}((\Psi_{c,n,+}^{-1})_0, \rho \otimes \psi) \otimes r_{\ell}(\psi \circ N_{K/Q_p}^{-1})]$$

follows from the vanishing of $(\Psi_{c,n,+}^{-1})_0$ for $i \neq n - 1$. The final isomorphism, showing that we can ignore the $+$, is a consequence of a simple calculation of the local Galois action on the polarization, already mentioned after (3.2.3).

To complete the proof, we thus have to show that, for any $\psi$, the virtual $G_n$-module

$$[\Psi_n(\rho)] = \sum_i (-1)^i [Hom_{J_{n,+}}((\Psi_{c,n,+}^{-1})_0, \rho \otimes \psi)]$$

defined as in the statement of Theorem 4.3.11, is purely supercuspidal as a representation of $G_n$. Write

$$[M^h] = \sum_{p,q} (-1)^{p+q} [Ind_{P_{h}(K)}^{GL(n,K)} H^p(S_{M_h}, R^q\Psi)]$$

To prove that $[\Psi_n(\rho)]$ is purely supercuspidal, we will make use of the following weak version of the Second Basic Identity (Theorem 6.1.2):

(5.2.13.1) Lemma. — Let $\pi^w$ be an admissible irreducible representation of $G(\mathcal{A}^w_f)$. Let $h > 0$, and suppose $[M^h][\pi^w] \neq 0$. Then there exists a unique irreducible representation $\pi_w$ of $G_n$ such that $[H(A(B,\ast))[\pi_w \otimes \pi^w]] \neq 0$, and such that the Jacquet module $(\pi_w)_{R_n, P_h} \pi_w$ of $\pi_w$ relative to the unipotent radical of $P_h$ is non-trivial.

In other words, only “automorphic” $\pi^w$ can contribute to the virtual module $[M^h]$. However, neither this lemma nor the Second Basic Identity determines the individual
spaces $[\text{Ind}_{P_h(K)}^{GL(n,K)} H^k_c(S_{M_0}, R^q\Psi)]$. Note that the uniqueness of $\pi_w$ in the statement of the Lemma follows from the fact that, if $[H(\mathcal{A}(B, *))][\pi_w \otimes \pi^w] \neq 0$, then $\pi_w \otimes \pi^w$ admits a base change to the finite part of a cohomological automorphic representation of $GL(n,F)$; then as remarked above, strong multiplicity one for $GL(n,K)$ implies that $\pi_w$ is determined uniquely.

We admit Lemma 5.2.13.1 for the moment. For our given $\pi_w$, we thus have $\pi_w = \text{JL}(\rho)$. Now the First Basic Identity yields

\begin{equation}
[\text{Ind}_{P_h(K)}^{GL(n,K)} H^k_c(S_{M_0}, R^q\Psi)] = \sum_h [M^h][\pi^w]
\end{equation}

in Groth($G_n$). Since $\pi_w = \text{JL}(\rho)$ is supercuspidal, all the Jacquet modules $(\pi_w)_{R_h P_h}$ vanish for $h > 0$, thus (5.2.13.2) simplifies to yield

\begin{equation}
[H(\mathcal{A}(B, *))][\pi^w] = [M^0][\pi^w].
\end{equation}

Strong multiplicity one again implies that $[M^0][\pi^w]$ is isotypic for $G_h$ of type JL($\rho$).

Next, (5.2.11) implies that

\begin{equation}
[M^0][\pi^w] = \sum_i (-1)^{n-i+i} [\text{Hom}_{J_n,1}(\mathcal{F}_{c,n,1}, \mathcal{A}(I/I(\mathcal{R}), \mathbb{Q}_c)|[\pi^w]|^{ker^1(Q,G)})
= \sum_i (-1)^{n-i+i} [\text{Hom}_{J_n,1}(\mathcal{F}_{c,n,1}, \rho \otimes \psi) \otimes \mathcal{A}(I/I(\mathcal{R}), \mathbb{Q}_c)|[\pi^w |^{ker^1(Q,G)})]
\end{equation}

where the second isomorphism is a consequence of strong multiplicity one for base change, this time from $I$ to $GL(n)$. Combining (5.2.13.3) with (5.2.13.4), we see that $[\Psi_n(\rho)]$ is purely supercuspidal, as required.

Theorem 5.2.13 implies that $[r_\ell(\pi)]_{\text{Gal}(\mathbb{F}/K)}$ is purely local at $w$; i.e., it depends only on $\pi_w = \text{JL}(\rho)$. It also calculates the supercuspidal part of $\Psi_n^{-1} \sigma(\pi)$ (ignore the $+$) and proves statement (i) of the local theorem (4.3.11). It remains to justify Lemma 5.2.13.1. This will be obtained (see §6.1 and Remark 6.1.3) as a consequence of the Second Basic Identity (6.1.2)(i), whose proof occupies sections 6 and 7.

For any $\pi \in \mathcal{A}_0(n,K)$, we write $\sigma_\ell(\pi) \in \mathcal{G}(n,K)$ for the representation $r_\ell(\pi) \otimes |\det|^{(n-1)/2}$ of $W_K$ defined in this way. Not every $\pi$ can be realized as a local component of a cohomological automorphic representation of $G$. Our hypotheses imply that the central character of $\pi$ is of finite order. Conversely, assume the central character of $\pi$ to be of finite order. Then

\begin{equation}
(5.2.14) \quad \text{An approximation argument shows there is no restriction on $K$; one can always realize $K$ as some $F_w$ for a CM field $F$ of the appropriate type, and } GL(n,K) \text{ as the local component of the right kind of } G;
\end{equation}
Given such $G$ and $\pi \in A_0(n, K)$, it follows from a theorem of Clozel [C1] that one can always find a cohomological representation $\Pi$ of $G$ with local component $\pi$ at $w$, unramified outside some fixed (non-empty) set.

To extend the correspondence to general $\pi$, one notes that any $\pi$ is of the form $\pi_0 \otimes \psi \circ \det$, where $\pi_0$ has central character of finite order and $\psi$ is some character of $K^\times$. So one defines $\sigma_I(\pi) = \sigma_I(\pi_0) \otimes \psi$ viewing $\psi$ as a character of $W_K$ via local class field theory.

To show that the latter construction is well-defined, one ought to verify that

$$\sigma_I(\pi \otimes \psi \circ \det) = \sigma_I(\pi) \otimes \psi$$

when $\psi$ is a character of finite order. This follows by applying Kottwitz' theorem to the representation $r_\ell(\pi_f)$ of $\text{Gal}(\overline{\mathbb{Q}}/F)$. Indeed, Kottwitz shows that $r_\ell(\pi_f \otimes \chi) = r_\ell(\pi_f) \otimes \chi$ whenever $\chi$ is a global Hecke character of finite order. More precisely, Kottwitz shows this is true at almost all unramified places. By Chebotarev density, it is true at $w$. This argument then shows that (5.2.17) is valid for any $\psi$, not necessarily of finite order. We have thus verified that the correspondence $\sigma_I$ satisfies property (0.2) expected of the local Langlands correspondence. More such properties are verified, in a similar way, in the following section.

Meanwhile, we have already reduced part (ii) of Theorem 4.3.11, which we restate here for convenience of reference:

**Proposition.** — Let $\pi' \in A(g, K)$ be a discrete series representation which is not supercuspidal. Then for all $i$, $\text{Hom}_J(\Psi^i_{c,g}, JL(\pi'))$ contains no $G$-subquotients isomorphic to a supercuspidal representation $\pi$.

**Proof.** — Notation is as in Theorem 4.3.11. In view of (5.1.7), it suffices to prove the assertion for $i = n - 1$. The argument used in (5.2.15) applies to show that one can find $\pi^w$, as in the statement of Lemma 5.2.13.1, such that $\pi^w \otimes \pi'$ occurs with non-zero multiplicity in $A(I/I(R), \mathbb{Q}_\ell)$. Suppose $\pi$ does occur as a $G$-subquotient of $\text{Hom}_J(\Psi^i_{c,g}, JL(\pi'))$. It then follows from (5.2.12) that $\pi^w \otimes \pi$ occurs with non-zero multiplicity in $H^{n-1}(S(B, *), \mathbb{Q}_\ell)$. Then as in the paragraph following Lemma 5.2.13.1, $\pi^w \otimes \pi$, resp. $\pi^w \otimes \pi'$, admits a base change to the finite part of a cohomological automorphic representation $\Pi$, resp. $\Pi'$, of $GL(n, F)$. By strong multiplicity one $\Pi = \Pi'$, hence $\pi = \pi'$, contradiction.

**Remark.** — Since it may not be evident from the order of the arguments above, I stress that this proof does not depend on the truth of Lemma 5.2.13.1. Although it is not strictly necessary, we will be using Theorem 4.3.11 (ii) in §7 as a step in the proof of Lemma 5.2.13.1.
5.3. Compatibility with cyclic base change and automorphic induction

We have shown in (5.2.17) that $\sigma_t : 0(n, K) \rightarrow G(n, K)$ is compatible with character twists. One shows similarly it is compatible with contragredients. Moreover, because the construction is purely local, $\sigma_t$ commutes with automorphisms of $K$. These are three properties required of a local Langlands correspondence (cf. (0.6)).

We also need to know that $\sigma_t$ commutes with cyclic base change and local automorphic induction. Having established these properties, it follows by an argument due to Henniart [BHK](7) that $\sigma_t$ is a bijection $A_0(n, K) \rightarrow G_0(n, K)$ for all $K$, and that it preserves conductors. In other words, it satisfies all the requirements of the local Langlands correspondence except preservation of $\varepsilon$ factors of pairs. Thus, as explained in Lecture 1, in order to obtain the local Langlands conjecture, it suffices to establish a form of compatibility of the local correspondence with the global correspondence.

To prove compatibility with cyclic base change and local automorphic induction, we need to use a global argument again. The following discussion is based on my article [H1], in which I treated the analogous situation for Drinfel’d uniformization.

Now global base change and automorphic induction are defined for automorphic representations of $GL(n, F)$, not of $G$. So we need to use quadratic base change (from $\mathbb{Q}$ to $E$, as in Lecture 1) and descent. This works as follows: starting from a (global) $\pi \subset A(G)$, with fixed $\pi_w$, let $\Pi$ denote its base change to $GL(n, F')$ (ignoring the extra Hecke character of $E$). Let $F'/F$ be a global cyclic extension of CM fields with only one prime dividing $w$, $K' = F'_w$. The representation $\Pi \in CU(n, F)$, and $BC_{F'/F}(\Pi) \in CU(n, F')$, hence descends to a cohomological representation (or rather $L$-packet) in $A(G')$. Here we have to be careful: $G'$ is attached to a division algebra with involution ($B', \#'$) and in general $B' \neq B \otimes_F F'$. We have to choose $F'$ so that $BC_{F'/F}(\Pi)$ still has a local discrete series component at a place other than $w$ (so that it descends to a twisted unitary group). We have to verify that the parity condition is satisfied, so that we can construct $G'$ with the right signatures at $\infty$. These are easy to verify [H1, §4]. Applying Kottwitz’ theorem and Chebotarev density, we see that

\[
\sigma_t(BC_{K'/K}(\pi_w)) = \sigma_t(\pi_w)|_{W_{K'}}
\]

provided $BC_{K'/K}(\pi_w)$ is supercuspidal (so that the left-hand side is defined). This is sufficient for Henniart’s axioms.

Automorphic induction is a bit more complicated. If we start with $\Pi \in CU(n, F')$, it is not true that $AI_{F'/F}(\Pi) \in CU(n[F' : F], F)$; in fact, $AI_{F'/F}(\Pi)$ is no longer cohomological at $\infty$. This can be remedied by twisting $\Pi$ by an appropriate Hecke character $\chi$ of $F'$. If the infinity types of $\chi$ are chosen appropriately, and if $\chi \circ c = \chi^{-1}$,

(7)This is where Henniart’s numerical version of the local Langlands conjecture [He2] is invoked, in the form of the following “splitting property” [He3]: given any supercuspidal representation $\pi$ of $GL(n, K)$, there is a finite sequence of extensions $K = K_0 \subset K_1 \subset \cdots \subset K_n$, each step cyclic of prime degree, such that the image $\pi_{K_n}$ of $\pi$ under successive cyclic base change is a principal series representation. The analogous property for $G_0(n, K)$ is obvious.
then $A_{F'/F}(\Pi \otimes \chi) \in CU(n[F' : F], F)$. This has the inevitable effect of replacing the initial $\pi_{w}$ by an unramified twist, which is not a problem. Again, the details can be found in [H1] (proof of Lemma 5).

5.4. Comparison of trace formulas. — The proof of Proposition 5.2.10 is much easier than the comparison used to study the other strata, but it provides an excuse to introduce the trace formula that will provide one side of the comparison in the general case. We only need to work with anisotropic groups.

For simplicity, we assume henceforth we are not in the case $F^{+} = \mathbb{Q}$, $n = 2$, where the above comparison is a special case of Carayol’s thesis. This special case complicates the formulas because the maximal compact subgroup is not connected.

The trace in question is that of $[H(S(B, *)) = \sum_{i}(-1)^{i}[H^{i}(S(B, *))$, the representations $[H^{i}(S(B, *))$ being admissible $G(A_f)$-modules. One could also work with a fixed central character.

If $\gamma \in G(\mathbb{Q})$ and $\phi \in C_{c}^{\infty}(G(A_f))$, we define the orbital integral

\[(5.4.1) O_{\gamma}(\phi) = \int_{G(A_f)/Z_{G}(\gamma)(A_f)} \phi(\gamma^{-1}g) dg.\]

This integral depends on a choice of Haar measures (on $G$ and on $Z_{G}(\gamma)$) that will be specified.

The global trace formula for the action of Hecke correspondences on cohomology was worked out by Arthur even in the non-compact case; he studied $L^{2}$-cohomology and had to allow for boundary terms. The compact case is of course much easier to explain. Needless to say, it is equivalent to the topological Lefschetz trace formula. However, we prefer to use Arthur’s formulation, which allows a uniform treatment of isolated and non-isolated fixed points. Here is a version of Arthur’s formula adapted to our groups $G$:

(5.4.2) Cohomological trace formula ([A]). — Let $\phi \in C_{c}^{\infty}(G(A_f))$. Then

$$\text{Tr}(\phi \mid [H(A(B, *))]) = nk_{B} \sum_{\gamma} e(\gamma)[F(\gamma) : F]^{-1} \text{vol}(Z_{G}(\gamma)(\mathbb{R}))^{1-1}O_{\gamma}(\phi).$$

Note that we have written the formula for $A(B, *)$, the union of $|\ker^{1}(\mathbb{Q}, G)|$ copies of $S(B, *)$, to simplify the formulas. Here the notation needs to be explained:

(5.4.2.1) The measure on $G(A_f)$ is arbitrary (it appears on both sides).

(5.4.2.2) $e(\gamma) = (-1)^{n/[F(\gamma) : F]^{-1}}$ is the Kottwitz sign; $\gamma$ is regular if and only if $[F(\gamma) : F] = n$, and then $e(\gamma) = 1$.

(5.4.2.3) $k_{B} = 1$ if $4 \mid [B : \mathbb{Q}]$ and equals 2 otherwise.

(5.4.2.4) $\gamma$ runs over a set of representatives of $G(A)$-conjugacy classes in $G(\mathbb{Q})$ which are elliptic in $G(\mathbb{R})$. In particular (N.B.!!), even though we are working with the union of $|\ker^{1}(\mathbb{Q}, G)|$ copies of a Shimura variety, the factor $|\ker^{1}(\mathbb{Q}, G)|$ has been
incorporated into the expression as a sum over adelic (rather than rational) conjugacy classes. See Lemma 7.1.3 for an explanation.

(5.4.2.5) \( F(\gamma) \) is the subfield of \( B \) generated over \( F \) by \( \gamma \).

N.B. The fact that \( F(\gamma) \) is always a field, because \( B \) is a division algebra, is extremely important! From the standpoint of the trace formula, this is one of the special features of the twisted unitary groups we are using; it guarantees that the stabilized trace formula contains no endoscopic terms.

(5.4.2.6) \( ZG(\gamma)(\mathbb{R})_0 \) denotes the compact mod center inner form of \( ZG(\gamma)(\mathbb{R}) \) and

\[
ZG(\gamma)(\mathbb{R})_0^1 = \ker |\nu| : ZG(\gamma)(\mathbb{R})_0 \rightarrow \mathbb{R}_>^\times.
\]

(5.4.2.7) Let \( dz(\gamma) \) be the measure used to define the orbital integral, \( (dz(\gamma)_\infty)_0^1 \) the measure used to define \( \text{vol}(ZG(\gamma)(\mathbb{R})_0^1) \), and \( (dz(\gamma)_\infty)_0 \) the measure on \( ZG(\gamma)(\mathbb{R})_0 \) defined by \( (dz(\gamma)_\infty)_0^1 \) and \( dt/t \) on \( \mathbb{R}_>^\times \). Let \( dz(\gamma)_\infty \) be the measure on \( ZG(\gamma)(\mathbb{R})_0 \) compatible with \( (dz(\gamma)_\infty)_0 \) (this is well defined). Then \( dz(\gamma)_f \times dz(\gamma)_\infty \) is Tamagawa measure.

Of course I won’t prove this. The usual trace formula in the anisotropic case is a sum

\[
\sum_{\gamma} \nu(ZG(\gamma)(\mathbb{Q}) \backslash ZG(\gamma)(\mathbb{A}))O_\gamma(\phi).
\]

Here \( \gamma \) runs \( G(\mathbb{Q}) \)-conjugacy classes. For the volume term one can take Tamagawa measure. To get cohomology, one takes \( \phi_{\infty} \) to be a sum of discrete series pseudocoeficients (over the set of discrete series with cohomology in the trivial representation); this restricts attention to \( \gamma \) elliptic at \( \infty \), and the orbital integral of \( \phi_{\infty} \) is constant on elliptic conjugacy classes. Arthur’s formulation of the cohomological trace formula in \([A]\) takes roughly this form. The present version, adapted from \([HT]\), involves a partial stabilization of Arthur’s expression: we rewrite the sum over \( G(\mathbb{Q}) \)-conjugacy classes as a sum over \( G(\mathbb{A}) \)-conjugacy classes in \( G(\mathbb{Q}) \) by counting the number of the former in the latter.\(^{(8)}\) This number turns out to be related to \( \kappa_B / \ker^1(ZG(\gamma)) \), and Kottwitz’ theorem on Tamagawa numbers gives the measure term in the stated formula. The remaining terms — \( n(-1)^{n/[F(\gamma) : F]}[F(\gamma) : F]^{-1} \) — all arise by rewriting the expressions in \([A]\) coming from the archimedean place. They would be more complicated if \( F(\gamma) \) were not a field (e.g., for untwisted unitary groups). In the next two lectures I’ll use similar arguments in counting points.

\(^{(8)}\) Full stabilization goes one step further, replacing \( G(\mathbb{A}) \)-conjugacy by \( G(\overline{\mathbb{Q}}) \)-conjugacy. Apparently this is not really necessary for the point-counting argument. However it is necessary in the general situation for comparison with trace formulas for endoscopic groups, or the twisted trace formula in the setting of base change, as in (1.2.6).
One gets a completely analogous formula when $G$ is replaced by $I$:

\[(5.4.3) \quad |ker^1(\mathbb{Q}, G)| \text{Tr}(\phi^I | \mathcal{A}(I/I(\mathbb{R}))) = \kappa_B \sum_{\gamma} |F(\gamma) : F|^{-1} \text{vol}(Z_G(\gamma)(\mathbb{R})_\delta^1)^{-1}O_\gamma(\phi^I).\]

Indeed, this is the formula for cohomology of the 0-dimensional Shimura variety attached to $I$, where the sum is again over adelic conjugacy classes in $I(\mathbb{Q})$, and we have used the fact (5.2.7) that $|ker^1(\mathbb{Q}, G)| = |ker^1(\mathbb{Q}, I)|$. The only differences with (5.4.2) are that the $n$ in front has disappeared (because the discrete series $L$-packet has only one element) and the signs have disappeared (because all centralizers are compact at $\infty$). There is no restriction on $\gamma$ (all elements are elliptic in $I(\mathbb{R})$).

\[\begin{align*}
(5.4.4) \textbf{Lemma.} & \quad \text{The set of } I(A)\text{-conjugacy classes in } I(\mathbb{Q}) \text{ is in bijection with the set of } G(A)\text{-conjugacy classes in } G(\mathbb{Q}) \text{ elliptic at } \infty \text{ and at } w. \text{ This bijection preserves orbital integrals away from } w, \text{ and takes } \gamma^I_w \in J_n \text{ to the conjugacy class in } G_n \text{ with the same characteristic polynomial.} \\
\text{Proof.} & \quad \text{Let } \gamma \in I(\mathbb{Q}). \text{ It is elliptic at } \infty \text{ and at } w, \text{ hence transfers to a conjugacy class in } G(A), \text{ and the question is whether it has a representative in } G(\mathbb{Q}). \text{ This is a consequence of a general principle: if } G \text{ and } G' \text{ are inner forms and } T \subset G \text{ is a torus that transfers locally to } G' \text{ everywhere, then } T \text{ transfers globally to } G \text{ provided a certain cohomological invariant, defined by Langlands and Kottwitz (cf. [K1, §9]), vanishes (in } H^2). \text{ But this invariant vanishes if } T \text{ is elliptic (cf. Lemma (2.3.3))}. \quad \square
\end{align*}\]

To make effective use of pseudocoefficients, we fix a compact open subgroup $U^w \subset G(A_f^w)$ and consider the representations on $[H(S_{U^w}(B, *))]$ and on $\mathcal{A}(I/I(\mathbb{R}) \cdot U^w)$. This means we have to restrict attention to $U^w$-biinvariant functions. This is not a problem, since we can take $U^w$ arbitrarily small, but it has the advantage that $\Gamma_U = U^w \cap Z_G(Q) = U^w \cap Z_f(Q)$ is a cocompact subgroup of $Z_G(K) = Z_G$. Hence we can take $\phi_w$ to be a pseudocoefficient of a chosen supercuspidal $\pi_w$, relative to the set $\mathcal{A}_{d, \text{fin}}(A.1.3)$ of representations with central character trivial on $\Gamma_U$. This has the effect on the trace side of isolating representations $\pi_f$ with component $\pi_w$ at $w$. As we have seen in (5.1.6), these occur only in $H^{n-1}$, hence for such $\phi$, we have

\[(5.4.5) \quad \text{Tr}(\phi | [H^{n-1}(A(B, *))]) = (-1)^{n-1} n \kappa_B \sum_{\gamma} (-1)^{n(\gamma)} |F(\gamma) : F|^{-1} \text{vol}(Z_G(\gamma)(\mathbb{R})_\delta^1)^{-1}O_\gamma(\phi^I).\]

We take $\phi^I = \phi^w \otimes \phi_w^I$, where $\phi_w^I$ is a pseudocoefficient for $JL(\pi_w)$. The Jacquet-Langlands correspondence (A.1.13) has the following property (cf. [Ro], §3):

\[(5.4.6) \quad O_\gamma(\phi^w) = (-1)^{n-1} e(\gamma) O_{\gamma, c}(\phi_w).\]
When $\gamma$ is regular, $e(\gamma) = 1$, and in that case this relation is the defining property of the Jacquet-Langlands correspondence. (One defines matching functions to have matching orbital integrals; then the sign appears in the trace).

The other terms are the same. Dividing by $|\ker^1(\mathbb{Q}, G)| = |\ker^1(\mathbb{Q}, I)|$, this implies the trace on $[H^{n-1}(S_{U^w}(B, *))]|_{\pi_w}$ equals $n$ times the trace on $A(I/I(\mathbb{R}) \cdot U^w)|_{JL(\pi_w)}$. It then follows from linear independence of characters (A.1.2) that the representations are as indicated in the proposition.

5.5. Properties of the fundamental local representation. — In the applications to strata of positive dimension, the fundamental local representation appears as the stalk of the vanishing cycles at a point $x$ in an isogeny class $\overline{S}(x)_{M_0}$ (cf. Lemma 4.4.8). We replace $n$ by $g$ and work with the version $\Psi_{g+1}$. We write $G = G_g = GL(g, K)$, $J = J_g$ as before. Recall from (4.3.10) that

$$\Psi_{g-1} = c\text{-Ind}_{A_g, \kappa}^{G \times J \times W_K} \Psi_{g-1}^{A_g, x_0}.$$ 

Let $A'_{g, K}$ be the subgroup of $G \times J \times W_K$ generated by $A_{g, K}$ and the center $Z$ of $G$. It is the kernel of the composite of the map $\delta : G \times J \times W_K \rightarrow G$ with the map $Z \rightarrow Z/gZ$, and also contains the center $Z_J$ of $J_g$. In particular, $A'_{g, K}$ is of index $n$ in $G \times J \times W_K$. We let

$$(5.5.1) \quad T_0' = A'_{g, K} \cap (G \times J); \quad T_0 = A_{g, K} \cap (G \times J)$$

Recall from §3.3 that the subgroup

$$Z_0 = \{(x, x) \in K^x \times K^x \simeq Z \times Z_J\} \subset T_0$$

acts trivially on the moduli space, hence on the stalks $\Psi_{i, g, x_0}$ of the vanishing cycle sheaves for any $i$. The only representations $\tau$ of $T_0$ such that $\text{Hom}_{T_0}(\Psi_{i, g, x_0}, \tau) \neq 0$ are thus those on which $Z_0$ acts trivially. Let $G_{g, 0}$, (resp. $J_{g, 0}$), denote the kernel of $\delta_G : G_g \rightarrow Z$ (resp. ker $\delta_J$) and let $T_{00} = G_{g, 0} \times J_{g, 0} \subset T_0$. We define an inertial equivalence class of representations of $T_0$ to be a set of the form $\{\tau \otimes \psi\}$ where $\tau$ is an irreducible representation of $T_0$, trivial on $Z_0$, and $\psi$ runs through the set of characters of $T_0/T_{00} \cdot Z_0 \simeq Z/nZ$. The set of inertial equivalence classes of $T_0$ is denoted $\mathcal{I}_g$. If $\tau$ is a representation of $T_0/T_{00}$, we let $[\tau]$ denote its inertial equivalence class. Then we have a discrete decomposition for each $i$

$$(5.5.2) \quad \Psi_{i, g, x_0} = \bigoplus_{[\tau] \in T_0} \Psi_{i, g, x_0}^{[\tau]}$$

where $\Psi_{i, g, x_0}^{[\tau]}$ is the sum of the $\tau_j$-isotypic components for $\tau_j \in [\tau]$.

We also define inertial equivalence for representations of $J$ and $G$. Let $\rho \in A(J)$, with central character $\psi_\rho$. The inertial equivalence class of $\rho$, denoted $[\rho]$, is the set of representations $\rho \otimes \psi \circ \text{det}$, where $\psi$ runs over unramified characters. The strong inertial equivalence class of $\rho$ is the set of $\rho \otimes \psi \circ \text{det}$ where $\psi$ runs over unramified characters of finite order dividing $g$; this is the set of representations of $J_g$ inertially equivalent to $\rho$ and with central character $\psi_\rho$. The same terminology is
used for discrete series representations of \( G_g \). The cardinality of the strong inertial equivalence class of \( \rho \) is an important invariant of \( \rho \): it equals \( g/c(\rho) \), where \( c(\rho) \) is the number of distinct unramified characters \( \psi \) of order dividing \( g \) such that \( \rho \otimes \psi \simeq \rho \).

The strong inertial equivalence class of \( JL(\rho) \in \mathcal{A}(G) \) has the same cardinality as that of \( \rho \). The set of inertial equivalence classes of representations of \( J_g \) (resp. of discrete series representations of \( G_g \)) is denoted \([\mathcal{A}](J_g)\) (resp. \([\mathcal{A}](g, K)\)).

We write \( \rho \simeq \rho' \) if \( \rho \) and \( \rho' \) are inertially equivalent. Two inertially equivalent representations of \( J \), (resp. of \( G \), resp. of \( T_0 \)), have the same restriction to \( J_g,0 \), (resp. of \( G_g,0 \), resp. of \( T_{00} \)). These restrictions are not generally irreducible. It is known, and follows easily from Clifford’s theorem, that the restriction of \( \rho \) to \( J_g,0 \) is the sum of \( c(\rho) \) irreducible components, each with multiplicity one; the same holds for \( JL(\rho) \), when \( J \) is replaced by \( G \). For want of better terminology, the irreducible components of the restriction to \( J_g,0 \) (resp. \( G_g,0 \)) of a fixed \( \rho \) will be called nearly equivalent, and we say they belong to the near equivalence class \( N(\rho) \) of \( \rho \).

**Lemma.** — Let \( \tau \) be an irreducible representation of \( T_0/Z_0 \), and suppose its restriction to \( T_{00}/T_{00} \cap Z_0 \) decomposes as the (necessarily finite) direct sum

\[
\tau|_{T_{00}} = \oplus (\alpha_i) \otimes \beta_i
\]

where each \( \alpha_i \) (resp. \( \beta_i \)) is an irreducible representation of \( G_g,0 \) (resp. \( J_g,0 \)). Then

(i) The various \( \alpha_i \) (resp. \( \beta_i \)) are nearly equivalent.

(ii) Suppose \( \Psi^{g-1}_{c,g,x_0}[\tau] \neq 0 \), and \( \pi \in \mathcal{A}_0(g, K) \) is such that the \( \beta_i \) belong to the near equivalence class of \( JL(\pi) \). Then the \( \alpha_i \) belong to the near equivalence class of \( \pi \).

**Proof.** — Part (i) is obvious, and part (ii) follows from (4.3.4) and Theorem 4.3.11. \( \square \)

Let \([\tau]\) and \( \pi \) be as in Lemma 5.5.3 (ii). Such a \([\tau]\) will be called supercuspidal. It follows from 5.5.3 that \( \Psi^{g-1}_{c,g,x_0}[\tau] \) can be described alternatively as the sum of the \( \beta_i \)-isotypic components of \( \Psi^{g-1}_{c,g,x_0} \), for \( \beta_i \in N(JL(\pi)) \), or as the sum of its \( \alpha_i \)-isotypic components, for \( \alpha_i \in N(\pi') \). This justifies writing

\[
\Psi^{g-1}_{c,g,x_0}[\tau] = \Psi^{g-1}_{c,g,x_0}[JL(\pi)] = \Psi^{g-1}_{c,g,x_0}[\pi]
\]

(note the dualization implicit in the notation relevant to G). One could just as well decompose with respect to general discrete series representations, or equivalently of general representations of \( J_g \), but in that case it is better to work with the alternating sum of \( \Psi^{g-1}_{c,g,x_0}[\tau] \). Recall however (5.1) that the supercuspidal part of \( \Psi^{g-1}_{c,g,x_0}[\tau] \) vanishes for \( i \neq g-1 \). We thus put

\[
[\Psi]_{c,g,x_0} = \sum_i (-1)^k [\Psi^i_{c,g,x_0}] = \bigoplus_{[\rho] \in [\mathcal{A}](J_g)} [\Psi]_{c,g,x_0} \rho,
\]

\textbf{ASTÉRISQUE 298}
where $\rho$ runs alternatively over $[A](J_0)$, as indicated, or over $[A]_d(g,K)$. Strictly speaking, we have only proved this here for supercuspidal inertial equivalence classes; the complete result can be found in [HT].

Now fix a character $\xi$ of $K^\times$, with restriction $\xi_0$ to $O^\times$. The maximal compact subgroup $O^\times \times O^\times \subset Z \times Z_J$ is contained in $T_0$. Let $\Psi_{c,g,x_0}^{-1}(\xi_0) \subset \Psi_{c,g,x_0}^{-1}$ denote the subspace on which $(u,u') \in O^\times \times O^\times$ acts as $\xi_0(u)^{-1}\xi_0(u')$. This is an invariant subspace for the action of $A_{g,K}$, and the action of $A_{g,K}$ on $\Psi_{c,g,x_0}^{-1}(\xi_0)$ extends uniquely to a representation, denoted $\Psi_{c,g,K}^{-1}(\xi)$, of $A'_{g,K}$, such that $(x,x') \in Z \times Z_J$ acts as $\xi(x)^{-1}\xi(x')$. Let $\Psi_{c,g,\xi}^{-1}$ denote the maximal quotient of $\Psi_{c,g,K}^{-1}(\xi_0)$, on which $(x,x') \in Z \times Z_J$ acts as $\xi(x)^{-1}\xi(x')$. Then there is a canonical isomorphism
\begin{equation}
\Psi_{c,g,\xi}^{-1} \sim \text{c-Ind}_{A'_{g,K}}^{GL(g,K) \times J_g \times W_K} \Psi_{c,g,x_0,\xi}^{-1}.
\end{equation}

Combined with (5.5.2) and (5.5.4), we thus obtain a canonical decomposition
\begin{equation}
\Psi_{c,g,\xi,\text{scusp}}^{-1} \sim \bigoplus_{\rho} \Psi_{c,g,\xi}^{-1}[\rho],
\end{equation}
\begin{equation}
\Psi_{c,g,\xi}^{-1}[\rho] \sim \text{c-Ind}_{N'_{g,K}}^{G \times J \times W_K} \Psi_{c,g,x_0,\xi}^{-1}[\rho],
\end{equation}
where the subscript “scusp” designates the $G$-supercuspidal part, and $\rho$ runs over $A_0(g,K)$. By Proposition 5.2.18, the sum can also be taken over $\rho \in A(J)$ with $\text{JL}(\rho)$ supercuspidal. The component $\Psi_{c,g,\xi}^{-1}[\rho]$ is non-trivial if and only if the central character $\xi_0$ of $\rho$ equals $\xi$. There is a similar decomposition for the alternating sum $[\Psi_{c,g,\xi}]$.

**Lemma.** — Write $G = GL(g,K)$, $J = J_g$. Let $T'_0 = A'_{g,K} \cap (G \times J)$. Let $\tau_{0,0}$ (resp. $\tau_0$) denote the restriction to $T'_0$ of the representation of $A'_{g,K}$ on $\Psi_{c,g,x_0,\xi}^{-1}$ (resp. of the representation of $G \times J$ on $\Psi_{c,g,K}^{-1}(\xi_0)$). Then
(i) The representations $\tau_0$ and $\tau_{0,0}$ are admissible.
(ii) For any $a \in G \times J$, the representation $\tau_{a,0}$ of $T'_0$, defined by $\tau_{a,0}(x) = \tau_{0,0}(axa^{-1})$, has the same character as $\tau_{0,0}$.
(iii) In the Grothendieck group of $T'_0$, we have
\[ \tau_0 = g \cdot \tau_{0,0} \]
Moreover, the character of $\Psi_{c,g,\xi}^{-1}$ restricted to $G \times J$, equals zero off $T_0$.

**Remark.** — The relation (iii) requires an explanation. The characters in the formula are the actual characters of the group $T'_0$, defined as the traces of the operators defined by (A.1.1). This yields a relation of the form
\begin{equation}
\text{trace}(\tau_0)(\phi) = g \cdot \text{trace}_{Z,G,\xi}^{-1}(\tau_{0,0})(\phi). \tag{5.5.9.1}
\end{equation}
Here $\phi$ is a compactly supported function on $T_0$, transforming with respect to $\xi_0 \otimes \xi_0^{-1}$ under $O^\times \times O^\times \subset Z \times Z_J$, $\phi_{\xi}$ the extension of $\phi$ to a function on $T'_0$ transforming under $\xi \otimes \xi_0^{-1}$ under $Z \times O^\times \subset Z \times Z_J$. The left-hand side is as in (A.1.1), whereas...
trace$_Z\xi$ on the right is as in the discussion preceding (A.1.9). Note that the function $\phi_0$ is non-compactly supported only on $G$, not $J$, and the modified trace only takes account of the center of $G$; but one can just as well replace the index $Z$ by $Z_J$ in (5.5.9.1).

Proof. — Since the central character has been fixed, (i) follows from Proposition 4.3.9 (cf. Remark 4.3.9.1). Assertion (iii) is a simple consequence of (ii) and Clifford’s theorem on induced representations; the factor $g$ is the index of $T_0$ in $G \times J$. So it remains to prove (ii). The group $T_0$ is generated by its subgroup $T_0 = A_{g,K} \cap (G \times J)$ and the central subgroup $Z$ (or $Z_J$) of $G \times J$; hence it suffices to prove (ii) for the restriction of the character $\tau_{\xi,0}$ to $T_0$. Choose an element $\varphi \in W_K$ such that $\delta(\varphi) = 1 \in \mathbb{Z}$; i.e., $\varphi$ is an extension of Frobenius to the algebraic closure of $K$. There is a homomorphism $h_\varphi : G \times J \rightarrow A_{g,K}$ given by

$$h_\varphi(\gamma, j) = (\gamma, j, \varphi^d)$$

where $d = w_K(\det(\gamma)) - w_K(N(j))$.

The restriction of $h_\varphi$ to $T_0$ is the natural inclusion. It follows that the restriction to $T_0$ of $\tau_{\xi,0}$ extends to a representation of $G \times J$, hence that its character is invariant under conjugation by $G \times J$. $\square$

(5.5.10) Remarks. — Fix a representation $\rho \in \mathcal{A}_0(g, K)$ as in (5.5.8), with $\xi_\rho = \rho$. Using (iii) and the description of the near equivalence class $N(\rho)$ given above, one verifies easily that

$$\Psi_{c,g,x_0,\xi}^{g-1}[\rho] |_{T_\infty} = \frac{g}{c(\rho)} \bigoplus_{i,j} \alpha_i \otimes \beta_j$$

where $\alpha_i$ (resp. $\beta_j$) runs through $N(\rho^\vee)$ (resp. $N(JL(\rho))$). In other words, each irreducible component of the restriction of $JL(\rho)$ to $J_{g,0}$ occurs with each component of the restriction of $\rho^\vee$ to $G_{g,0}$ with the same multiplicity. The same holds when $J_{g,0}$ and $G_{g,0}$ are replaced, respectively, by the subgroups $Z_J \cdot J_{g,0} \subset J$ and $Z \cdot G_{g,0} \subset G$, each of index $g$. However, each $\alpha_i \otimes \beta_j$-isotypic component carries a representation $r_\ell(\rho, i, j)$ of the subgroup $W_{K_g} \subset W_K$, the Weil group of the unramified extension $K_g$ of degree $g$ of $K$. One can also verify that the $g$-dimensional representation $r_\ell(\rho)$, defined as in Theorem 4.3.11, decomposes as the sum of $c(\pi)$ distinct irreducible components, each of dimension $g/c(\rho)$, and that each one occurs as an $r_\ell(\rho, i, j)$ with the same multiplicity. Thus the representation of $Z_J \cdot J_{g,0} \times Z \cdot G_{g,0} \times W_{K_g}$ on $\Psi_{c,g,x_0,\xi}^{g-1}[\rho]$ refines the correspondence of Theorem 4.3.11, though it is not sufficiently fine to characterize the local Langlands correspondence, including a description of $L$-packets, for $SL(g)$.

The full compactly induced (virtual) representation $[\Psi_{c,g}]$ also has a decomposition according to inertial equivalence:

$$[\Psi_{c,g}] \overset{\sim}{\rightarrow} \bigoplus_{\rho} [\Psi_{c,g}][\rho],$$

ASTÉRISQUE 298
where
\[ [\Psi_{c,g}]_\rho = \text{c-Ind}_{A_{g,K}}^{G \times J \times W_K} [\Psi_{c,g,x_0}]_\rho, \]
with the components on the right-hand side as in (5.5.5); note that induction is now from \( A_{g,K} \).

(5.5.12). — Finally, the same analysis can be applied to the subgroup
\[ \Xi_0 = (G_g \times W_K) \cap A_{g,K} \subset G_g \times W_K. \]

Let \( I_K = W_K \cap A_{g,K} \) (the inertia group), and define a near equivalence class of (continuous \( \ell \)-adic) representations of \( I_K \) to be the set of irreducible components of the restriction to \( I_K \) of a finite-dimensional continuous irreducible \( \ell \)-adic representation of \( W_K \). The decomposition (5.5.2) represents the \( G_g,0 \)-supercuspidal part \( (\Psi^i_{c,g,x_0})_0[\sigma] \) as a direct sum of components \( (\Psi^i_{c,g,x_0})_0[\sigma] \) where \( [\sigma] \) can stand for a near equivalence class of representations of \( I_K \). More precisely, if \( (\Psi^i_{c,g,x_0})_0[\sigma] = \Psi^i_{c,g,x_0}[\pi] \) for \( \pi \in A_0(g,K) \) (so \( i = g - 1 \)), then it follows from Theorem 4.3.11 that the action of \( I_K \) is given by the near equivalence class of representations of \( I_K \) contained in a fixed irreducible \( g \)-dimensional irreducible representation of \( W_K \), namely \( r_I(\pi) \). Note that \( \sigma \) need not be of dimension \( g \) as a representation of \( I_K \), but it will necessarily occur in the restriction to \( I_K \) of an irreducible \( g \)-dimensional representation of \( W_K \); thus the notation \( [\sigma] \) can designate an irreducible \( g \)-dimensional representation of \( W_K \) up to inertial equivalence.

For general discrete series \( \pi \), it is shown in [HT] that the corresponding representations of \( I_K \) are contained rather in an indecomposable representation of \( WD_K \) whose irreducible constituents are of degree strictly less than \( g \). It then follows that, for \( \pi \) supercuspidal, \( \Psi^{g-1}_{c,g,x_0}[\pi] \) is the sum of the isotypic subspaces of \( \Psi^{g-1}_{c,g,x_0} \) for the representations in the corresponding near equivalence class \( [\sigma] \) of representations of \( I_K \) and the same holds for \( \Psi^i_{c,g,x_0}[\pi] \) when \( i \neq g - 1 \), in which case the corresponding \( [\sigma] \)-isotypic part is trivial.

The identification of \( \Psi^{g-1}_{c,g,x_0}[\pi] \) with a specific \( \Psi^i_{c,g,x_0}[\sigma] \), even for \( \pi \) supercuspidal, is only possible after determination of the Galois representations occurring in \( \sum_i (-1)^i \Psi^i_{c,g,x_0}[\rho] \) for all \( \rho \in \mathcal{A}(J_g) \). However, as in (5.5.8), we have
\[ \Psi^{g-1}_{c,g,x_0}[\pi] = \Psi^{g-1}_{c,g,x_0}[\text{JL}(\pi)] = \sum_i (-1)^{g-1+i} \Psi^i_{c,g,x_0}[\text{JL}(\pi)] \]
for supercuspidal \( \pi \).

Note that if \( U \subset GL(g,K) \) is a compact open subgroup such that \( \pi^U \neq \{0\} \), and if \( I(\pi) \subset I_K \) is an open subgroup that acts trivially on \( r_I(\pi) \), then \( \Psi^{g-1}_{c,g,x_0}[\pi]^U = \Psi^{g-1}_{c,g,x_0}[\pi]^{U \times I(\pi)} \) is a finite-dimensional semisimple module for the Hecke algebra \( \mathcal{H}(\Xi_0) \).
6. The second basic identity and Isogeny classes in the special fiber

This is where we begin to “count points,” as in the Kottwitz’ article [K5], following earlier work of Langlands and Ihara. More precisely, we use a refined version of Honda-Tate theory to describe them in purely group-theoretic terms, as point sets with adelic group (and Frobenius) actions, as disjoint unions of certain double coset spaces. The stage is thus set for the calculation, in Lecture 7, of the Lefschetz traces of sufficiently regular Hecke operators, acting on the cohomology of each $S_{X_{M_0}}$. The principal application of this calculation is the Second Basic Identity (Theorem (6.1.2)), which compares these traces to the traces of the same Hecke operators acting on the cohomology of the generic fiber. The Second Basic Identity, proved in Lecture 7, is applied in section (6.2) to derive the main compatibility theorem (1.3.6).

6.1. General strata: statement of second basic identity

(6.1.1) Notation. — For any $g$, let $G_g = GL(g, K)$; thus $G_n = G_w$. The center of $G_g$ (resp. $J_g$), is denoted $Z_g$ (resp. $Z_{J_g}$); both are canonically isomorphic to $K^\times$, for any $g$. Henceforward, we write $N_h = R_u P_h$, $N_h^{op}$ the unipotent radical of the opposite parabolic. The modulus character for $P_h$ is denoted $\delta_h = \delta_{P_h}$; it is the absolute value of the determinant of the adjoint action on $N_h$. We let $J_{n-h}$ be the twisted inner form $D_x^\times 1 / (n-h)$ of $G_{n-h}$.

Let $L_h = G_{n-h} \times G_h$ be the standard Levi subgroup of $P_h$. We let $r_{G_n, L_h} : \text{Groth}(G_n) \to \text{Groth}(L_h)$ denote the standard (normalized) Jacquet functor

$$r_{G_n, L_h} \pi = \pi_{N_h} \otimes \delta_h^{-1/2}.$$  

The Jacquet functor for the opposite parabolic is denoted $r_{G_n, L_h}^{op}$. We define a renormalized Jacquet functor

$$\text{re} r_{G_n, L_h}^{op} = r_{G_n, L_h}^{op} \otimes \delta_{P_h}^{1/2},$$  

since $\delta_{P_h}^{1/2} = \delta_{P_h}^{1/2}$; this means it has been normalized twice.

In what follows, we fix $h$ and let $\rho \in A(J_{n-h})$. First assume $JL(\rho)$ is supercuspidal. We define a map

$$\text{red}^{(h)}_{\rho} : \text{Groth}(G_n) \to \text{Groth}(G_h)$$

as the composition of

$$\text{re} r_{G_n, L_h}^{op} : \text{Groth}(G_n) \to \text{Groth}(G_{n-h} \times G_h)$$

and the map $c_\rho$ that sends $[\alpha \otimes \beta]$, with $\alpha \in A(n-h, K)$ and $\beta \in A(h, K)$, to 0 if $\alpha \neq JL(\rho)$, and to $[\beta]$ otherwise. In other words, $\text{red}^{(h)}_{\rho}$ is the renormalized Jacquet functor followed by projection on the $JL(\rho)$-component in the first variable.

For general $\rho \in A(J_{n-h})$, we replace $c_\rho$ in the preceding definition by the map $c'_\rho$ that sends $[\alpha \otimes \beta]$ to

$$\text{Tr}(\alpha)(\phi_{JL(\rho)}, \omega) \cdot [\beta].$$
Here $\phi_{JL(\rho),\omega}$ is a normalized truncated pseudocoefficient for $JL(\rho)$ given by formula (A.1.11), relative to a sufficiently large interval $\omega$.

In what follows, we fix a level subgroup $U^w$ away from $w$, and write $H^p_c(S_{M_0}^{(h)}, R^q\Psi)$ for $H^p_c(S_{M_0}^{(h)}, R^q\Psi)^{U^w}$, $[H_c(S_{M_0}^{(h)}, R\Psi)] = \sum_{p,q} (-1)^{p+q}[H^p_c(S_{M_0}^{(h)}, R^q\Psi)]$. For $\rho \in \mathcal{A}(J_{n-h})$ we define $[\Psi_{n-h}(\rho)]$ by the alternating sum in Theorem 4.3.11, with $g = n - h$. The following identity is proved by an elaborate comparison of trace formulas:

(6.1.2) Theorem

(i) (Second Basic Identity, first version): There is a countable subset $\mathcal{A}(J_{n-h})_{\text{fin}} \subset \mathcal{A}(J_{n-h})$ such that

\[
(6.1.2.1) \quad n \cdot [H_c(S_{M_0}^{(h)}, R\Psi)] = \bigoplus_{\rho \in \mathcal{A}(J_{n-h})_{\text{fin}}} \text{red}^{(h)}_\rho [H(A(B, *))] \otimes [\Psi_{n-h}(\rho)]
\]

in Groth($\mathcal{A}_f^w$) $\times L_h \times W_K$.

Remark. — Here and below, the action of $G_{n-h} \times W_K$ on the right-hand side is concentrated on the factor $[\Psi_{n-h}(\rho)]$; the action of $W_K$ on $[H(A(B, *))]$ is ignored.

(ii) The set $\mathcal{A}(J_{n-h})_{\text{fin}}$ can be chosen so that, for any $\rho \in \mathcal{A}(J_{n-h})_{\text{fin}}$, the intersection

$\mathcal{A}(J_{n-h})_{\text{fin}}[\rho] = [\rho] \cap \mathcal{A}(J_{n-h})_{\text{fin}}$

of $\mathcal{A}(J_{n-h})_{\text{fin}}$ with the inertial equivalence class $[\rho]$ of $\rho$, defined as in § 5.5, is a finite set.

Write $\mathcal{A}(J_{n-h})_{\text{fin}} = \mathcal{A}^0_{n-h} \coprod \mathcal{A}'_{n-h}$ where $\mathcal{A}^0_{n-h}$ is the subset of $\rho$ such that $JL(\rho)$ is supercuspidal, and $\mathcal{A}'_{n-h}$ are the others. Write

\[
(6.1.2.2) \quad \mathcal{R}^{h,0} = \bigoplus_{\rho \in \mathcal{A}^0_{n-h}} \text{red}^{(h)}_\rho [H(A(B, *))] \otimes [\Psi_{n-h}(\rho)];
\]

\[
(6.1.2.3) \quad \mathcal{R}^{h,\nu} = \bigoplus_{\rho \in \mathcal{A}'_{n-h}} \text{red}^{(h)}_\rho [H(A(B, *))] \otimes [\Psi_{n-h}(\rho)]
\]

(iii) (Second Basic Identity, second version): For any $\rho \in \mathcal{A}(J_{n-h})_{\text{fin}}$, with $JL(\rho)$ supercuspidal, we have the following identity in Groth($\mathcal{A}_f^w$) $\times L_h \times W_K$:

\[
(6.1.2.3) \quad \mathcal{R}^{h,0} = \bigoplus_{\rho \in \mathcal{A}^0_{n-h}} \text{red}^{(h)}_\rho [H(A(B, *))] \otimes [JL(\rho) \otimes r_\ell(\rho)^{\nu,+}].
\]

Here $r_\ell(\rho)^{\nu,+}$ is $r_\ell(\rho)^{\nu}$ twisted by the contribution $\psi \circ N_{K/Q}_p$ of $Q_{p}^\times$, which we will simply ignore.

Since we have fixed the level subgroup $U^w$ away from $w$, the countability assertion in part (i) is just a reformulation of the admissibility of $H^p_c(S_{M_0}^{(h)}, R\Psi)$ for all $p$, which
in turn follows from Lemma 4.4.2. The assertion (ii) is also a consequence of admissibility, since any unramified twist of $JL(\rho)$ has fixed vectors for the same compact open subgroups as $JL(\rho)$.

Given the definitions, (6.1.2.3) is a direct consequence of (6.1.2.1) and Theorem 4.3.11. Now (6.1.2.1) is a more precise version of Lemma 5.2.13.1. But in (5.2) we have seen a proof of Theorem 4.3.11, assuming Lemma 5.2.13.1. Thus it only remains to prove (6.1.2.1).

(6.1.3) Remark. — Actually, to prove Lemma 5.2.13.1 it suffices to prove the identity (6.1.2.1) in $\text{Groth}(G(A_f^w) \times L_h)$, i.e. ignoring the Galois action. In §6.3 it will be shown that (6.1.2.1) in $\text{Groth}(G(A_f^w) \times L_h)$, in conjunction with Theorem 4.3.11 (i) for $g < n$, actually suffices to prove (6.1.2.3).

Combining the first and second basic identities, we find:

(6.1.4)
\[
n[H(A(B,*))] = \sum \bigg[ \bigoplus_{\rho \in A_n^{0,-h}} \text{Ind}_{P_h(\mathcal{K})}^G \{\text{red}_p^{(h)} H(A(B,*)) \otimes JL(\rho) \otimes r_\ell(\rho)^{\vee,+}\} \bigg] \\
\bigoplus \sum_{h} \bigg[ \bigoplus_{\rho \in A_n^{0,-h}} \text{Ind}_{P_h(\mathcal{K})}^G \{\text{red}_p^{(h)} [H(A(B,*))] \otimes [\Psi_{n-h}(\rho)]\} \bigg].
\]

Write
\[
[H(A(B,*))] = \sum_{\pi_f} [\pi_f] \otimes [R_\ell(\pi_f)]
\]
where $[R_\ell(\pi_f)] \in \text{Groth}(\text{Gal}((\mathcal{F}/F))$. Now recall that two cohomological automorphic representations of $G$ that agree away from $w$ agree also at $w$, by strong multiplicity one for the base change to $GL(n)$. Thus we can factor out the $G(A_f^w)$ representations and we are left with the following assertion.

(6.1.5) Theorem. — Suppose $\pi = \pi^w \otimes \pi_w$ is an irreducible admissible representation of $G(A_f)$. Then in $\text{Groth}(GL(n,K) \times W_K)$ we have

\[
n[\pi_w] \otimes [R_\ell(\pi_f) |_{W_K}] = (\dim R_\ell(\pi_f)) \sum_{h,p \in A_n^{0,-h}} n-\text{Ind}_{P_h(\mathcal{K})}^{GL(n.K)} n-\text{red}_p^{(h)}[\pi_w] \otimes JL(\rho) \otimes r_\ell(\rho)^{\vee,+}(-h/2) \\
\bigoplus \sum_{h,p \in A_n^{0,-h}} \text{Ind}_{P_h(\mathcal{K})}^G [\text{red}_p^{(h)} \text{Hom}_{G(A_f^w)}(\pi^w, [H(\mathcal{B}(B,*))] \otimes [\Psi_{n-h}(\rho)])].
\]

Here we define $n-\text{red}_p^{(h)}$ by replacing $\text{re-r}^{op}$ by the normalized Jaquet functor $r^{op}$, which is just a twist by $\delta_{r_p}^{1/2}$. Similarly, $n-\text{Ind}$ is normalized induction. The twist of $\text{re-r}^{op}$ by $\delta_{r_p}^{-1/2}$ cancels the opposite twist in $n-\text{Ind}$ but introduces a new twist in the second step of the definition of $\text{red}_p$, which accounts for the twist by the unramified character $|\bullet|^{-h/2}$, which is the meaning of the final symbol (an easy calculation).
We apply Theorem 6.1.2 here, and in the subsequent applications, with a level subgroup $U^w$ such that $\pi^{U^w} \neq 0$. The proof of Theorem 6.1.5 is then very simple. The point is that
\[
\text{red}_\rho (h) \rho [H(A(B, *)) |_{W_K}] = \sum_{\pi_f} \pi^{U^w}_f \otimes \text{red}_\rho (h) [\pi_f] \otimes [R_\ell (\pi_f)],
\]
where here the term $[R_\ell (\pi_f)]$ is just a vector space without structure: all the Galois action is on the $r_\ell (\rho)^{\vee,+] (-h/2)!$. This explains the dimension factor.

(6.1.6). — As stated in [HT, Theorem V.5.4] the Second Basic Identity is an explicit expression for the $\rho$-contribution to $n \cdot [H_c (S(h, M_0, R_\Psi))]$ for any $\rho$, including $\rho \in \mathcal{A}'_{n-h}$.

The simple form asserted in (6.1.2.3) is only valid for supercuspidal JL$(\rho)$. The second summand on the right-hand side of (6.1.5), as in (6.1.4), is made more explicit in [HT, VII.1.5]. For the cases treated in the present account the crude form presented above is sufficient.

6.2. Proof of the main theorem, assuming second basic identity. — We expect that $R_\ell (\pi_f)$ equals the sum of $|\ker^1 (Q, G)|$ copies (for the different Shimura varieties in $A(B, *)$ of a fixed representation $R_0 (\pi_f)$ of dimension $n$; this is equivalent to the conjecture that the representations have multiplicity $a(\pi) = 1$. Then

(6.2.1) $[\pi_w] \otimes [R_0 (\pi_f) |_{W_K}] = \sum_{h, \rho} \text{n-Ind}_{P_h (K)}^G n-\text{red}_\rho (h) [\pi_w] \otimes JL(\rho) \otimes r_\ell (\rho)^{\vee,+] (-h/2)}$.

To simplify notation, we make the assumption that $a(\pi) = 1$. The reader can verify that, in general, the same $a(\pi)$ appears on both sides of the formula. The proof of the main theorem is now just a calculation of $\text{n-Ind}_{P_h (K)}^G n-\text{red}_\rho (h) [\pi_w]$, as $h$ varies.

We know $\pi_w$ is generic and unitary. Thus there is a parabolic subgroup $P = P_\nu$, with $\nu = (n_1, \ldots, n_r)$, and an $r$-tuple of discrete series representations $\tau_1, \ldots, \tau_r$ such that

(Split case) $\pi_w = \text{n-Ind}_{P_\nu}^G \tau_1 \otimes \cdots \otimes \tau_r$.

As explained in (1.4.5), we restrict our attention to the case where each $\tau_i$ is supercuspidal. The general discrete series is treated in § VII.1 of [HT], and requires an explicit version of the Second Basic Identity in general, as mentioned in (6.1.6). The proof in the general case makes use of non-tempered cohomology classes as well as more precise information, due to Zelevinsky, on the decomposition of induced representations.

Warning. — The notation $\text{n-Ind}$ and $\text{n-red}$ designate normalized induction and restriction, respectively, whereas $\text{re-r}^{\text{op}}$ denotes RE-normalized restriction!!!

We first recall the following theorem due to Bernstein and Zelevinski:
(6.2.2) Geometric Lemma (Bernstein-Zelevinski)

\[ \left[ \text{Ind}_{G_n(K)}^{G_n} \left( \tau_1 \otimes \cdots \otimes \tau_r \right) \right] = \sum_{\nu_i \in \nu II} \left[ \text{n-Ind}_{P_{\nu_i}}^{G_n} \tau_1 \otimes \cdots \otimes \tau_r \right] \]

Here \( P_I = P_{\nu_I}(K) \), likewise for \( P_{II} \). Note that this sum is in the Grothendieck group; a priori the Jacquet module is not semisimple. This is not a problem for us.

Next, observe that the second summand on the right-hand side of (6.1.5) contributes trivially to (6.2.1). Indeed, by the strong multiplicity one argument already used, every irreducible constituent in that summand is of the form

\[ \text{Ind}_{P_{\nu_i}(K)}^{G_n} \left[ \text{red}^{(h)}_\rho \pi_w \otimes [\Psi_{n-h}(\rho)] \right] \]

where \( JL(\rho) \) is not supercuspidal. By (A.1.4), (A.1.5), the definition of \( \text{red}^{(h)}_\rho \), and Lemma 6.2.2, such terms necessarily vanish (more details of the calculation can be found in the next two paragraphs).

It remains to consider the first (explicit) summand. We have to compute

\[ \left[ \text{n-Ind}_{P_{\nu_I}(K)}^{G_n} \text{n-Ind}_{P_{\nu_I}}^{G_n} \tau_1 \otimes \cdots \otimes \tau_r \right] \]

First, apply the Jacquet functor relative to \( N^{op} \) to \( \text{n-Ind}_{P_{\nu_I}}^{G_n} \tau_1 \otimes \cdots \otimes \tau_r \). The result is described by the Geometric Lemma.

The next step is to project this result on the \( JL(\rho) \)-isotypic component for \( G_{n-h} \). Our hypothesis that the induced representation \( \pi_w \) is irreducible (and unitary) implies, by the Bernstein-Zelevinski classification of the discrete series [BZ,Z], that the \( JL(\rho)^{\nu'} \)-isotypic component of the term corresponding to \( \nu = \nu_I \bigcup \nu II \) is trivial unless

(i) \( JL(\rho) \) is supercuspidal, and

(ii) \( \nu_I \) is a single element \( i \).

In other words, projection on \( JL(\rho) \) picks out those \( n_i = n - h \) and those \( \tau_i = JL(\rho)^{\nu'} \). Thus, letting \( \nu^i = (n_1, \ldots, n_i, \ldots, n_r) \), we have

\[ [n-\text{red}^{(h)}_\rho] \text{n-Ind}_{P_{\nu_I}}^{G_n} \tau_1 \otimes \cdots \otimes \tau_r = \sum_{n_i = n-h, \tau_i = JL(\rho)} \text{n-Ind}_{P_{\nu_I}}^{G_n} \tau_1 \otimes \cdots \otimes \tau_r \]

Now comparing this with our original formula, and using transitivity of induction (first from \( G_{n-h} \times P_{\nu'} \) to \( L_{\nu} \), then from \( P_{\nu} \) to \( G_n \)) we have

\[ [\pi_w] \otimes [R_0(\pi_f)]_{W_{\kappa}} = \sum_{h, \rho} \sum_{n_i = n-h} \text{n-Ind}_{P_{\nu_I}}^{G_n} \text{JL}(\rho) \otimes \bigotimes_{j \in \nu'} \tau_j \otimes r_\ell(\rho)^{\nu',(-h/2)} \]

But each term on the right hand side of (6.2.4) is of the form \( [\pi_w] \otimes r_\ell(\rho)^{\nu',(-h/2)} \) where \( [\pi_w] \) is fixed, \( \rho \) runs through the \( \tau_i \) and each \( \tau_i \) occurs once (for \( n - h = n_i \)). Thus we can cancel the \( [\pi_w] \) from both sides and obtain

\[ [R_0(\pi_f)]_{W_{\kappa}} = \sum_i r_\ell(\tau_i)^{\nu', \left(\frac{n_i - n}{2}\right)} \]

ASTÉRISQUE 298
If we define
\[(6.2.6) \quad r_\ell(\pi_w) = \bigoplus_i r_\ell(\pi_i) \vee \left( \frac{n_i - N}{2} \right) , \]
then we conclude
\[(6.2.7) \quad [R_0(\pi_f) |_{W_K}] = [r_\ell(\pi_w)] \otimes (\psi \circ N^{-1}_{K/Q_p}) \]
where for a change I put back the contribution of \(Q^K_p\). This is the main theorem I announced in my first lecture, under the hypotheses of (1.4.5).

The remainder of the course will therefore be devoted to proving the Second Basic Identity. The proof is a comparison of the Lefschetz trace formula, in Fujiwara’s version, for the action of Hecke operators on the vanishing cycle cohomology \([H_c(S_h^1, \mathbf{R})]\), with Arthur’s version (5.4.2) of the cohomological trace formula.

To make these notes more readable, we will often proceed as if we already knew the local Theorem 4.3.11. The reader will check that this hypothesis is only used in the counting argument in determining the local terms in the trace formula in (7.5). At that point, as well as at other crucial points along the way, the calculation will be presented in two forms, labeled “pre (4.3.11)” and “post (4.3.11)”.

**6.3. Overview of the point counting argument.** — Point counting, which in our situation is really representation counting, has two components. The first is the partition of points among isogeny classes. This can be done to various degrees of refinement. We have already seen that an isogeny class, as point set with group action, looks like
\[
[\tilde{M}^+_{n-h,h} \times (I_x(\mathbf{Q}) \setminus G(\mathbf{A}))]/J_{n-h,h} .
\]
For general Shimura varieties, the term \(\tilde{M}^+_{n-h,h}\) is replaced by something much more complicated coming from Dieudonné theory, and we are fortunate that in our special case the Dieudonné theory gives something of dimension zero, which is in fact a homogeneous space for \(J_{n-h,h}\). We will factor off the \(G_h\) term for simplicity. The first problem is to determine how many times the same set comes up. As we have seen for supersingular isogeny classes, this turns out to be a problem in Galois cohomology, and the answer, obtained by Kottwitz for general PEL type Shimura varieties at unramified places, is completely analogous to the problem of counting the number of Shimura varieties in the overall moduli problem: it is \(|\ker^1(\mathbf{Q}, I_x)|\). This takes rather a long time to establish, and the proof is expressed in terms of hermitian forms on \(V\) regarded as a module over \(B \otimes_F M\), where \(M\) is morally the extension of \(F\) generated by Frobenius acting on \(A_x\). Obviously such arguments cannot be extended to general Shimura varieties, and the solution was found by Langlands and Rapoport: instead of isogeny classes, they work with isomorphism classes of motives with additional structure. Since the theory of motives is mostly conjectural, their conjectures require further conjectures (Tate conjecture, standard conjectures for \(\ell\)-adic cohomology) to make sense; however, for PEL types, they seem to be largely established (Milne).
Milne’s article [Mi2] is a clear introduction to the Langlands-Rapoport conjectures, and formulates their extension to the case where the derived group is non-simply connected. His article includes statements of the main results of Langlands-Rapoport but not complete proofs in all cases.

Once the isogeny classes have been determined, the Lefschetz formula, insofar as it is valid, calculates the trace of a Hecke operator on cohomology (with compact support) as a sum of local terms; this is the meaning of “counting points”. It was first observed by Ihara, in the case of GL(2), that the local term corresponding to an isogeny class can be expressed in terms of orbital integrals. In Kottwitz’ formulation, the goal is to compute the zeta function, and for this he needs to count points over individual finite fields $\mathbb{F}_{q^r}$, where $q = |k(w)|$. The $p$-adic contribution is then a twisted orbital integral of a certain explicit function on $G(K_r)$, where $K_r / K$ is the unramified extension of degree $r$. In our approach, the Galois representation is entirely contained in the vanishing cycles, and the number of times a specific Galois representation occurs is determined as a sum of local multiplicities over all fixed points of the Hecke operator over $\mathbb{F}$. The result is a sum of orbital integrals, indexed by elements of $I_x(\mathbb{Q})$ as $x$ varies. It remains to solve new problems in Galois cohomology to relate these orbital integrals to orbital integrals of elements of $G(\mathbb{Q})$. Since the orbital integrals are purely local, it is reasonable to classify these elements up to $I_x(\mathbb{A})$-conjugacy, resp. $G(\mathbb{A})$-conjugacy; this is one sort of Galois cohomology problem. The next problem is to relate the two sets, especially to determine how many $I_x(\mathbb{Q})$ can give rise to a given $\gamma \in G(\mathbb{Q})$ up to $G(\mathbb{A})$-conjugacy.

Kottwitz’ Ann Arbor article is predicated upon taking the calculation one step further, classifying the contributions up to stable conjugacy (which is $G(\mathbb{Q})$-conjugacy in our setting). His articles on the subject are designed to fit into the development of the stable trace formula, and show how the stabilization of the trace formula, combined with his point counting, would completely determine the zeta functions of Shimura varieties (at least when there is no boundary). However, this turns out not to be necessary in our situation. We make only one explicit reference to the vanishing of the cohomological groups measuring obstruction to stability (the “endoscopic character groups”) for our specific $G$; this is what leads Kottwitz to call these “simple Shimura varieties,” and what allowed Clozel to attach Galois representations to automorphic representations of $GL(n)$. We also make two indirect references to the same fact. It is not clear to me whether one can still obtain a theory of bad reduction when endoscopy is present.

(6.3.1) Lemma. — Let $\pi \in \mathcal{A}_0(n-h, K)$, and define $R^i\Psi[\pi]$ to be the subsheaf of $R^i\Psi$ on which the action of $G_{n-h}$ belongs to the inertial equivalence class of $\pi'$. Then

(i) For all $i$ $R^i\Psi[\pi]$ is a pro-constructible sheaf on $\mathbb{S}^{(h)}_{\mathcal{M}_0}$, indeed is isomorphic to $\pi \otimes R^i\Psi_u$ where $R^i\Psi_u$ is constructible. Moreover,

(ii) $R^i\Psi[\pi] = 0$ for $i \neq n-h-1$;

(iii) The stalks of $R^{n-h-1}\Psi[\pi]$ are isotypic for the inertial equivalence class of $\rho = JL(\pi) \in \mathcal{A}(J)$. 

ASTÉRISQUE 298
Proof. — Fix an open compact subgroup $U \subset G_{n-h}$ such that $\pi^U \neq \{0\}$, Now the subsheaves of $U$-invariant vanishing cycles $R^i\Psi^U \subset R^i\Psi$ are constructible, hence for any near equivalence class $[\pi]$ of representations of $G_{n-h}$, the subsheaf $R^i\Psi^U[\pi]$ defined stalkwise as in (5.5.12) by the corresponding action of the Hecke algebra $H(G^1/U)$, where $G^1$ is the kernel of the character $|\det|$, is a constructible sheaf. But then (ii) and (iii) follow from (5.1.7) and Proposition 5.2.18 (i.e., Theorem 4.3.11 (ii)).

One of the main results of [HT] is that the stalkwise decomposition (5.5.2) extends, via the identification (5.5.4), to a decomposition

$$R^i\Psi = \oplus_{[\rho] \in [\mathcal{A}(J_{n-h})]} R^i\Psi[\rho]$$

of lisse sheaves on $\overline{S}_{M_0}$. One of the purposes of the present notes was to prove the main results without reference to this global decomposition, which depends on a difficult theorem of Berkovich proved in the appendix to [HT]. Lemma 6.3.1 allows us to assert that, for any geometric point $z \in \overline{S}_{M_0}$,

$$(6.3.2) \quad R^i\Psi[\pi]_z \subset R^i\Psi_z[JL(\pi)]$$

where the left-hand side is defined above and the right hand side is as in (5.5.4). Once we know (6.1.2.1) we will be able to apply Theorem 5.2.13, which implies that the inclusion in (6.3.2) becomes a (virtual) equality upon taking the alternating sum over $i$. In the absence of a complete determination of the individual $\Psi^i_{c,n-h,x_0}[\rho]$ for any $\rho$, this is the best we can do.

However, Lemma 6.3.1 does provide an important reduction:

(6.3.3) Proposition. — Assume Theorem 4.3.11 (i) for $g < n$. Let $\pi \in \mathcal{A}_0(n-h, K)$, for $h \geq 1$, $\rho = JL(\pi)$. Then as a virtual module for $W_K$, $[H_c(\overline{S}_{M_0}, R\Psi)][\pi]$ is isotypic for $r(\rho)^{\vee, +}$.

Remark. — As in §5.2, the hypothesis concerning Theorem 4.3.11 (i) follows from (6.1.2.1) applied to smaller Shimura varieties of the same type; i.e., to $\mathcal{A}(B, \ast)$ with $\dim B < n^2$. The case $g = 1$ follows without further ado from the compatibility between local and global class field theory for CM fields.

Proof. — By Lemma 6.3.1 we can rewrite

$$[H_c(\overline{S}_{M_0}, R\Psi)][\pi] = [H_c(\overline{S}_{M_0}, R\Psi[\pi])] = [H_c(\overline{S}_{M_0}, R^{n-h-1}\Psi[\pi])].$$

It then follows from (6.3.1)(iii) and (5.5.4) (which depends on Theorem 4.3.11 (i)) that $[H_c(\overline{S}_{M_0}, R^{n-h-1}\Psi)][\pi]$ is at least isotypic for the inertial equivalence class of $r(\rho)^{\vee, +}$.

Since $R^{n-h-1}\Psi[\pi]$ is constructible for any open, $\overline{S}_{M_0}$ can be written as a disjoint union of locally closed subvarieties $X_i$, on each of which $R^{n-h-1}\Psi[\pi]$ is lisse. By dèvissage – we are working in the Grothendieck group – we may replace $\overline{S}_{M_0}$...
in the statement by any of the $X_i$, say $X$. Over a pro-étale Galois cover $Y$ of $X$, $R^{n-h-1} \Psi[\pi]$ is isomorphic to a constant sheaf with fiber at any point $x$ isotypic for the inertia equivalence class of $\pi$, and the covering group of $Y$ over $X$ commutes with the action of $G_{n-h} \times W_K$. On the other hand, by (6.3.2) this fiber is contained in the supercuspidal part of $R^{n-h-1} \Psi_x[\rho]$ which, by (5.5.12), is $\Xi_0$-equivariantly isomorphic to $(\Psi_{c,n-h,x_0})_0[\rho]$. Let $I(\rho) \subset I_K$ denote a subgroup of finite index acting trivially on $(\Psi_{c,n-h,x_0})_0[\rho]$. It then follows tautologically that the fiber of the pullback to $Y$ is isotypic for subquotients of the action of the Hecke algebra (double coset algebra) $\mathcal{H}(\Xi_0//U \times I(\rho))$ on $(\Psi_{c,n-h,x_0})_0[\rho]$. Applying the Hochschild-Serre spectral sequence for the covering $Y$ of $X$, it follows that $[H_c(X, R^{n-h-1} \Psi)[\pi]]$, and hence $[H_c(\mathcal{S}^{(h)}_{M_0}, R^{n-h-1} \Psi)[\pi]]$, is again $\mathcal{H}(\Xi_0//U \times I(\rho))$-isotypic for subquotients of $(\Psi_{c,n-h,x_0})_0[\rho]$.

Since this action is semisimple (cf. (5.5.12)), we can replace the word “subquotients” by “quotients”. Then by Frobenius reciprocity, applied to $c\text{-Ind}^{G \times W_K}(\bullet)$, the action of the Hecke algebra $\mathcal{H}(G//U \times W_K/I(\pi))$ on $[H_c(\mathcal{S}^{(h)}_{M_0}, R^{n-h-1} \Psi)[\pi]]$ is isotypic for quotients of the fundamental local representation. The proposition then follows from Theorem 4.3.11 for $g = n - h - 1 < n$.

Thus the Second Basic Identity (6.1.2.3) is equivalent to the identity

$$(6.3.4) \text{ (post 4.3.11)} \quad n \cdot [H_c(\mathcal{S}^{(h)}_{M_0}, R\Psi)]_{\rho} = (n-h) \cdot \text{red}^{(h)}_\rho H(\mathcal{A}(B, *))$$

in $\text{Groth}(G(A_f^n) \times G_h)$, for all $\rho \in \mathcal{A}(J_{n-h})_{\text{fin}}$. Here the $n-h$ on the right-hand side comes from forgetting the $n-h$-dimensional representation $r_\ell(\rho)^{i, \ast}$ of $W_K$.

Corresponding to the version (6.1.2.1), we just have

$$(6.3.4) \text{ (pre 4.3.11)} \quad n \cdot [H_c(\mathcal{S}^{(h)}_{M_0}, R\Psi)]$$

$$= \bigoplus_{\rho \in \mathcal{A}(J_{n-h})_{\text{fin}}} \text{red}^{(h)}_\rho [H(\mathcal{A}(B, *)) \otimes [\Psi_{n-h}(\rho)] \in \text{Groth}(G(A_f^n)) \times G_h);$$

this is identical to (6.1.2.1), except that we are ignoring the Galois action. The latter form is (more or less) the form in which it is proved in [HT], and in which it will be proved in §7 below.

6.4. Honda-Tate theory. — I begin by recalling the Honda-Tate classification of isogeny classes of abelian varieties with $B$-action over $F$. Proofs can be found in [Ta]. In what follows, a CM field will be either a totally real field or a totally imaginary quadratic extension of a totally real field. As usual, $c$ denotes complex conjugation.

By Tate’s theorem on isogenies of abelian varieties over finite fields, we know that, up to isogeny, an abelian variety $A$ over $\overline{F}$ is determined by its Frobenius endomorphism $\pi_A : A \to A$, where $A$ is defined over some $\mathbb{F}_q$ and $\pi_A$ is the $q$-th power of $\text{Frob} : A \to A^{(q)}$. Since $A$ is also defined over any extension of $\mathbb{F}_q$, $\pi_A$ is only well-defined up to powers; i.e., in the group $\overline{\mathbb{Q}}^\times /\mu_\infty$, where $\mu_\infty$ denotes roots of 1. But we
also know that \( \pi_A \) is a \( q \)-number: it generates a CM field (or a totally real field), it is a unit away from \( p \), all its complex absolute values equal \( q^{1/2} \). So \( \pi_A^2/q \) is a \( p \)-unit all of whose complex absolute values = 1. It is thus completely determined, up to roots of unity, by its \( p \)-adic valuations. Moreover, by Honda, every \( \pi_A \) is obtained (by reducing abelian varieties with CM). This justifies the following definition:

\[(6.4.1) \text{ Definition.} \quad \text{Let } M \text{ be a CM field, and let } \mathbb{Q}[\mathfrak{P}_M] \text{ be the } \mathbb{Q} \text{-vector space with basis the places of } M \text{ above } p. \text{ For any fractional ideal } I \subset M, \text{ we let } [I] = \sum_{v|p} v(I) \cdot v \in \mathbb{Q}[\mathfrak{P}_M]. \text{ A } p \text{-adic type for } M \text{ is an } \eta \in \mathbb{Q}[\mathfrak{P}_M] \text{ such that } \eta + c_0(\eta) = [p]. \text{ Two pairs } (M, \eta) \text{ and } (M', \eta') \text{ are isomorphic if there is an isomorphism of fields } M \overset{\sim}{\longrightarrow} M'. \text{ Taking } \eta \text{ to } \eta'. \]

A finite extension of CM fields \( i : M \rightarrow N \) induces maps in both directions
\[
i_* : \mathbb{Q}[\mathfrak{P}_M] \rightarrow \mathbb{Q}[\mathfrak{P}_N]; \quad i^* \mathbb{Q}[\mathfrak{P}_N] \rightarrow \mathbb{Q}[\mathfrak{P}_M].
\]
via \( i_*(v) = \sum_{v' | v} c_{v'/v} v' \cdot i^*(v') = f_{v'/v} v \) if \( v = v' \mid M \). Let \( \sim \) denote the equivalence relation on pairs \((M, \eta)\) generated by \((M, \eta) \sim (N, i_*\eta)\). A \( p \)-adic type is an equivalence class of \((M, \eta)\).

\[(6.4.2) \text{ Exercise.} \quad \text{Every } p \text{-adic type has a unique minimal representative, up to isomorphism.} \]

If \( q = p^r \) and \( \pi \) is a \( q \)-number, let \( \mathfrak{b}(\pi) \) be the \( p \)-adic type equivalent to \((\mathbb{Q}(\pi), \frac{1}{2} | \pi| )\). Because \( \pi \) is determined (mod roots of unity) by \([\pi]\), it is easy to see that any sufficiently divisible power of \( \pi \) generates the minimal representative of \( \mathfrak{b}(\pi) \). In particular, \( \mathfrak{b}(\pi) \) is independent of \( r \), provided \( r \) is sufficiently divisible (i.e., provided \( F_q \) is sufficiently big). The preceding discussion shows that

\[(6.4.3) \text{ Theorem (Honda-Tate, [Ta]).} \quad \text{There is an equivalence between } p \text{-adic types and isogeny classes of simple abelian varieties over } F. \]

Moreover, we can determine the invariants of \( A_b \) as follows. Let \( b \) be a \( p \)-adic type with minimal representative \((M, \eta)\). Then:

\[(6.4.3.1) \text{ End}^0(A_b) \text{ is the division algebra with center } M \text{ and invariants } \frac{1}{2} \text{ at real primes, } 0 \text{ at finite primes away from } p, \text{ and } \eta_v f_{v/p} \text{ for } v \text{ dividing } p; \]

\[(6.4.3.2) \dim A_b = \frac{1}{2} [M : \mathbb{Q}][\text{End}^0(A_b) : M]^{1/2}; \]

\[(6.4.3.3) \text{ For any } v \mid p, \text{ } A_b[v^{\infty}] \text{ is a } p \text{-divisible group of height } [M_v : \mathbb{Q}_p][\text{End}^0(A_b) : M]^{1/2} \text{ and its Dieudonné module is isoclinic \text{ with slope } } \eta_v/e_{v/p}. \]

The fact that the minimal \( M = \mathbb{Q}[\pi_A] \) is the center of \( \text{End}^0(A_b) \) – note that these are endomorphisms over \( F \) – follows from Tate’s theorem that \( \pi_A \) generates the center of \( \text{End}^0(A) \) [Ta] and our choice of \( \pi_A \) over the field with \( p^r \) elements, with \( r \) sufficiently divisible.
There is a similar theory for isogeny classes of simple abelian varieties with $F$-action, for some CM field $F$. In this case, $M$ runs through CM fields containing $F$, and equivalence is defined via equivalence of embeddings over $F$. A $p$-adic type over $F$ is an $F$-equivalence class of $p$-adic types $(M, \eta)$ for CM fields $M$ containing $F$. Again each $p$-adic type over $F$ has a unique minimal representative.

(6.4.4). — Now let $B$ be a central division algebra over $F$. We now consider the category of pairs $(A, i)$ up to isogeny, with $A$ an abelian variety over $\mathbb{F}$ and $i : B \hookrightarrow \text{End}^0(A)$. This category has simple objects, and Kottwitz has shown a version of Morita equivalence: the simple objects (not necessarily simple abelian varieties!) are in bijection with $p$-adic types over $F$. Let $b$ be a $p$-adic type over $F$ with minimal representative $(M, \eta)$, and let $(A_b, i_b)$ denote the corresponding simple object in the category of abelian varieties up to isogeny with $B$-action. Then:

(6.4.4.1) $\text{End}^0(A_b)$ is the division algebra with center $M$ and invariants $\frac{1}{2} - \text{inv}_v(B \otimes_F M)$ if $v$ is real, $- \text{inv}_v(B \otimes_F M)$ at finite primes away from $p$, and $\eta_v f_v / p - \text{inv}_v(B \otimes_F M)$ for $v$ dividing $p$;

(6.4.4.2) $\dim A_b = \sqrt{\frac{n}{2}} [M : Q] \cdot [\text{End}^0_F(A_b) : M]^{1/2}$

(6.4.4.3) For any $v | p$, $A_b[v^\infty]$ is a $p$-divisible group of height $[M_v : Q_p][B : F]^{1/2}$ and its Dieudonné module is isoclinic with slope $\eta_v / \epsilon_v / p$.

Henceforward we fix an $h \in 0, \ldots, n - 1$. The goal is to classify isogeny classes $[x] \subseteq \mathcal{S}^h / (F)$. The first step is to classify isogeny classes of pairs $(A, i)$ as above with the right divisible $\mathcal{O}_v$-modules for all $v$ dividing $p$; say $(A, i)$ is of type $h$. We may assume $(A, i) = (A_b, i_b)$ with minimal representative $(M, \eta)$ as above. Let $(A', i')$ be a simple factor, $C' = \text{End}^0_F(A')$: it is a central $M$-algebra by minimality. Recall that $B$ is chosen to be a division algebra at some finite place $v$ other than $w$. Up to replacing $v$ by $v''$ (if $v$ divides $p$), we may assume $A[v^\infty]$ is an étale $p$-divisible group, by our standing Lie algebra hypothesis. Hence for any place $v'$ of $M$ above $v$, $\eta_{v'} = 0$, hence

$$\text{inv}_{v'}(C') = - \text{inv}_{v'}(B_M) = -[M_{v'} : F_{v'}] \text{inv}_v B.$$ 

Since $C'$ is a division algebra, $[C' : M]^{1/2}$ is at least the denominator of $-[M_{v'} : F_{v'}] \text{inv}_v B$, and since $B_v$ is a division algebra, the denominator of $\text{inv}_v B$ equals $n$. So

$$[C' : M] \geq (n/[M_{v'} : F_{v'}])^2 \geq n^2/[M : F]^2;$$

$$\dim A' = \frac{1}{2} n \cdot [M : Q][C' : Q]^{1/2} \geq [F^+ : Q]n^2 = \dim A,$$

where the first equality is (6.4.4.2). Hence

(6.4.5) Lemma. — $(A, i)$ is a simple object in the category of abelian varieties with $B$-action. Moreover, if $C = \text{End}^0_F(A)$, then $n = [M : F][C : M]^{1/2}$.

The last assertion just follows from equality in the above calculation, since $C = C'$. The simplicity is very important: it implies that we only have to consider fields, not
products of fields, in classifying isogeny classes. It is a reflection of the fact that $G$ has no endoscopy.

More generally, the $p$-adic type $\eta$ is completely determined by $h$ and the Lie algebra condition. We have $\eta_v = 0$ if $v$ is a place of $M$ dividing $u$ but not $w$; and this determines $\eta_w$. Moreover, $A[w^\infty]^{0}$ is a simple object in the category of $p$-divisible groups with $B_w = M(n, K)$ action. Its endomorphism algebra is just $D_{n-h}$. Hence the action of $M_w$ on $A[w^\infty]^{0}$ comes from a unique divisor $\tilde{w}$ of $w$ in $M$. Thus $A[\tilde{w}^\infty][0] = A[w^\infty][0]$ is an isoclinic formal group equal to $n$ copies of a formal group of height $(n-h)[K : \mathbb{Q}_p]$, hence has height $n[K : \mathbb{Q}_p](n-h)$, which by Honda-Tate (6.4.4.3) equals

$$[M_{\tilde{w}} : \mathbb{Q}_p][B : F]^{1/2}[C : M]^{1/2} = n[K : \mathbb{Q}_p][M_{\tilde{w}} : K][C : M]^{1/2};$$

i.e.

(6.4.6) \[ n - h = [M_{\tilde{w}} : K][C : M]^{1/2}. \]

Combining this with the lemma, we find

(6.4.7) \[ (n - h)[M : F] = n[M_{\tilde{w}} : K]. \]

Moreover, for $v \neq \tilde{w}$ dividing $w$, $A[w^\infty]$ is again étale. Next

(6.4.8) Lemma. — $M$ embeds over $F$ in $B$ (or in $B^{\text{op}}$).

Proof. — We consider the invariants of $C$ at places $v$ of $M$. For finite $v$ not dividing $w$ or $w^c$, $A[w^\infty]$ is either étale or multiplicative, hence we have $\text{inv}_v(C) = -\text{inv}_v(B_M)$. Since $M$ embeds in $C$, $M$ embeds in $B$ at such a $v$. But $B$ splits at $w$ and $w^c$, so there is no condition. Since $M$ is a CM field, it also embeds at $\mathbb{R}$. \hfill \square

Finally, we obtain the following result:

(6.4.9) Lemma. — There is a bijection between isogeny classes of pairs $(A, i)$ of type $h$ and pairs $(M, \tilde{w})$ where $M/F$ is a CM extension that embeds over $F$ in $B$, $\tilde{w}$ is a place of $M$ above $w$ such that $(n-h)[M : F] = n[M_{\tilde{w}} : K]$, and $(M, \tilde{w})$ is minimal in the sense that there is no intermediate field $M \supset N \supset F$ such that $\tilde{w}$ is inert over $N$.

Two comments are necessary. First, the minimality of the pair $(M, \eta)$ translates into minimality of $\tilde{w}$, since $\eta$ is nonzero only for $\tilde{w}$ and $\tilde{w}^c$. Next, the construction of $(A, i)$ from $(M, \tilde{w})$ follows the obvious recipe. We define the $p$-adic type $(M, \eta)$ over $F$ with

(6.4.10) \[ \eta_{\tilde{w}} = e_{\tilde{w}/w}/((n-h)f_{w/p}); \quad \eta_v = 0 \text{ if } v \mid u, v \neq \tilde{w}. \]

This determines $\eta$ uniquely, and one checks that the corresponding $(A, i)$ is of type $h$. 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
6.5. Polarized Honda-Tate theory, following Kottwitz. — That was the easy part. The hard part is counting polarizations.

(6.5.1) Proposition. — Suppose $(A, i)$ corresponds to $(M, \tilde{w})$. Then there exists a polarization $\lambda_0 : A \to A^\vee$ whose Rosati involution stabilizes $B \otimes M$ and induces $\ast \otimes c$, and a finitely generated $B \otimes M$ module $W_0$ with $\ast \times c$-hermitian alternating pairing $(\cdot, \cdot)_0 : W_0 \otimes W_0 \to \mathbb{Q}$, such that there are

1. An isomorphism of $B \otimes M \otimes_F A_f^w$-modules

$$W_0 \otimes A_f^w \xrightarrow{\sim} V^w(A)$$

taking $(\cdot, \cdot)_0$ to an $A_f^w$-multiple of the Weil pairing induced by $\lambda_0$, and

2. An isomorphism $W_{0, R} \xrightarrow{\sim} V_R$ of $B_R$-modules taking $(\cdot, \cdot)_0$ to an $\mathbb{R}^\times$ multiple of the standard pairing $(\cdot, \cdot)$ on $V_R$.

Recall that this means in particular that the signatures of $(\cdot, \cdot)_0$ are $(1, n-1)$ at $\tau_0$ and so on.

The existence of such an embedding (a $\#$-embedding of $M$ in $B^{\text{opp}}$) is proved following Kottwitz [K5, Lemma 14.1] (originally Zink [Zi, §4.4]). The main step is to show that $(A, i)$ lifts to a CM point of $A(B, \ast)$. This follows from compatibility of the polarization with the $F$ action, and the condition on dimension of eigenspaces for different $p$-adic embeddings of $F$.

Let $\#_0$ be the involution on $B^{\text{opp}} = \text{End}_B(W_0)$ induced by the pairing $(\cdot, \cdot)_0$, $G_0 = GU(W_0, \#_0)$. Here and elsewhere, $GU$ denotes the $\mathbb{Q}$-similitude group. Then we have seen $G_0$ is isomorphic locally to $G$ at all places except possibly $p$; but since $p$ splits in $E$, one sees $G_0$ is locally isomorphic to $G$ everywhere. So we may as well replace $G_0$ by $G$ (or vice versa), since our starting point is $A(B, \ast)$ rather than $S(G, X)$. Let $\phi_0 \in H^1(\mathbb{Q}, G)$ denote the class of the difference between the polarized modules $W_0$ and $V$.

We only consider pairs $(A, i)$ admitting prime-to-$w$ level structures, i.e. isomorphisms

$$V \otimes A_f^w \xrightarrow{\sim} V^w(A)$$

as $B \otimes_F A_f^w$-modules, compatible with the polarizations as before. In particular, the $\ast$-hermitian $B$-modules $W_0$ and $V$ are isomorphic at all primes except possibly $w$; but since $w$ is split they are isomorphic as well. We have seen that our points lift to CM points on one of the Shimura varieties $S^i(B, \ast)$, hence, after changing the polarization (in characteristic zero) we can assume $(W_0, \#_0) = (V, \#)$. This hypothesis simplifies the following discussion. In particular, $\phi_0 = 0$.

Let $D = \text{End}_{B_{\text{op}}}(V)$, so $D = \text{Cent}_{B^{\text{op}}}(M)$, and let $G_{[x]} \subset D$ be the unitary similitude group of $(D, \#)$. Thus $G_{[x]} \subset G$. Let $\ast_{[x]}$ be the Rosati involution on $C = \text{Aut}(A_x, i_x)$, and let $I_{[x]} = \text{Aut}((A_x, i_x, \lambda_x)) = GU(C, \ast_{[x]})$. Then $I_{[x]}$ and $G_{[x]}$ are inner forms of each other; indeed they are locally isomorphic everywhere.
except $p$ and $\infty$, because $W_{A_f^p}$ and $V^p(A_x)$ are isomorphic as $\ast \otimes \cdot$-hermitian $B_M(A_f^p)$-modules, by the proposition. However, they are not isomorphic; in particular, $I_{[x],\mathbb{R}}$ is anisotropic.

What are the equivalence classes of pairs $(V', (, )')$ where $V'$ is a $B_M$-module and $(, )'$ is a $\ast \otimes \cdot$-hermitian $\mathbb{Q}$-alternating form such that
\[(6.5.2) \quad (V', (, )') \text{ is equivalent to } (V, (, ))\]
as $\ast$-hermitian $B$-modules? On the one hand, the equivalence classes of pairs $(V', (, )')$ without condition corresponds to $\varphi_x \in H^1(\mathbb{Q}, G_{[x]}^{\epsilon})$; the condition $(6.5.2)$ means $\varphi_x$ maps to $0 \in H^1(\mathbb{Q}, G)$. So the set is in bijection with the kernel of the map $H^1(\mathbb{Q}, G_{[x]}) \to H^1(\mathbb{Q}, G)$. We call this set $H^1(\mathbb{Q}, G_{[x]}^{\epsilon})(0)$. On the other hand, this set is also in bijection with the set of $F$-embeddings $j : M \to B^{op}$ such that $\# \circ j = j \circ c$, up to $G(\mathbb{Q})$-conjugation, where $j$ goes to the $B \otimes B^{op}$-module $(V, (, ))$ considered as $B_M$-module via $j$. This is where Galois cohomology enters the picture. We call $j$ a $\#$-embedding.

6.6. Adelic partial stabilization. — Ideally one would like to consider $\#$-embeddings $j : M \to B^{op}$ up to $G(\mathbb{Q})$-conjugacy; this would lead to a stable formula in the point count. This is even possible in the present situation, but it is not necessary; it’s enough to consider $\#$-embeddings up to $G(A)$-conjugacy.

(6.6.1) Lemma. — The bijection above induces a bijection between
\begin{enumerate}
\item $G(A_f^p)$-conjugacy classes of $\#$-embeddings $j : M \to B^{op}$;
\item The kernel $H^1(\mathbb{Q}, G_{[x]}(\overline{A}_f^p))(0)$ of the map $H^1(\mathbb{Q}, G_{[x]}(\overline{A}_f^p)) \to H^1(\mathbb{Q}, G(\overline{A}_f^p))$.
\end{enumerate}

Proof. — It is clear from the preceding that elements of (1) correspond one-to-one to the images $x \in H^1(\mathbb{Q}, G_{[x]}(\overline{A}_f^p))$ of elements $y \in H^1(\mathbb{Q}, G_{[x]})(0)$. So we must show that the restriction to kernels of the localization map $H^1(\mathbb{Q}, G_{[x]})(0) \to H^1(\mathbb{Q}, G_{[x]}(\overline{A}_f^p))(0)$ is surjective.

Now it follows from [K2, Prop. 2.6] that there is a commutative diagram with exact rows:

\[(6.6.2) \quad \begin{array}{cccccc}
0 & \to & \ker^1(\mathbb{Q}, G_{[x]}) & \to & H^1(\mathbb{Q}, G_{[x]}) & \to & H^1(\mathbb{Q}, G_{[x]}(\overline{A})) \\
& & \downarrow f & & \downarrow & & \\
0 & \to & \ker^1(\mathbb{Q}, G) & \to & H^1(\mathbb{Q}, G) & \to & H^1(\mathbb{Q}, G(\overline{A}))
\end{array} \to A(G_{[x]}) \to A(G)
\]

The group $A(G)$ is what Labesse, in [L], calls $H^1_{ab}(A/\mathbb{Q}, G)$, this at least makes the sequence plausible. We need to show that $\ker(f)$ maps onto $\ker(g)$. This follows by simple diagram chase, once we show that (a) $\ker^1(\mathbb{Q}, G_{[x]}) \to \ker^1(\mathbb{Q}, G)$ is surjective,
which follows from surjectivity of $\ker^1(\mathbb{Q}, \mathbb{Q}_G) \to \ker^1(\mathbb{Q}, G)$ since the center $Z_G \subset G_{[x]}$, and (b) $A(G_{[x]}) \to A(G)$ is injective, which follows from a computation. In fact both equal 0 if $n[F^+: \mathbb{Q}]$ is odd and both are $\mathbb{Z}/2\mathbb{Z}$ if $n[F^+: \mathbb{Q}]$ is even and the natural map is the identity. (To compute $A(G)$: it is the Pontryagin dual of $\pi_0(\tilde{Z}(\mathbb{G})^\text{Gal(\mathbb{Q}/\mathbb{Q}))}$.)

That is the calculation on the $G_{[x]}$ side. There is an analogous computation on the $I_{[x]}$ side. First,

(6.6.3) Lemma. — The following five sets are equivalent:

1. Equivalence classes of polarizations $\lambda$ of $A_x$ whose Rosati involution stabilizes $B_M$ and acts as $* \otimes c$, where $\lambda \sim \lambda'$ if there are $\delta \in C^\times$ and $\mu \in \mathbb{Q}_\geq 0$ such that $\lambda' = \mu \delta^- \lambda$.

2. Equivalence classes of non-zero $\#_{[x]}$-fixed totally positive elements $\gamma \in C$ (thus $\gamma = \delta_{[x]} \delta$ over $C_{\mathbb{Q}}$, where equivalence is given by the equation $\gamma' = \mu \delta_{[x]} \gamma \delta$. (We have already seen this equivalence in the supersingular case, via $\gamma \sim \gamma_0 \circ \gamma$.)

3. Same as (2), but where $\gamma$ is either totally positive or totally negative, and where $\mu \in \mathbb{Q}^\times$.

4. $\ker[H^1(E/\mathbb{Q}, I_{[x]}(E)) \to H^1(\mathbb{C}/\mathbb{R}, I_{[x]}(\mathbb{C}))].$

5. $\ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]})]$.

Sketch of proof. — The last two are equivalent because, over $E$, $I_{[x]}$ is a product of inner twists of $GL(r)$’s, hence has no cohomology. The map from (3) to (4) takes $\gamma$ to the value of a cocycle on $c$, bearing in mind that $I_{[x]}(E) = C^\times \times \mathbb{Q}^\times$ and that $c$ acts on $I_{[x]}(E)$ by sending $(\gamma, \mu)$ to $(\mu(\gamma)^{\#_{[x]}}, \mu)$.

Thus there is a bijection between

1. Equivalence classes of polarizations $\lambda$ on $A_x$ whose Rosati involution stabilizes $B_M$ and acts as $* \otimes c$, and for which there is a prime-to-$p$-level structure compatible with the polarizations, and

2. $\ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]}))(0)] \cap H^1(\mathbb{Q}, I_{[x]}(\mathbb{A}_f^0))(0)$ where $H^1(\mathbb{Q}, I_{[x]}(\mathbb{A}_f^0))(0) = H^1(\mathbb{Q}, G_{[x]}(\mathbb{A}_f^0))(0)$ via the isomorphism $I_{[x]}(\mathbb{A}_f^0) \simeq G_{[x]}(\mathbb{A}_f^0)$.

Say $\lambda$ and $\lambda'$ are nearly equivalent if they are equivalent over $\mathbb{A}_f^0$.

(6.6.4) Lemma. — There are bijections between the following sets

1. Near equivalence classes of polarizations $\lambda$ on $A_x$ whose Rosati involution stabilizes $B_M$ and acts as $* \otimes c$, and for which there is a prime-to-$p$-level structure compatible with the polarizations

2. $H^1(\mathbb{Q}, I_{[x]}(\mathbb{A}_f^0))(0)$

3. $G(\mathbb{A}_f^0)$-conjugacy classes of $\#$-embeddings $j : M \to B^{\text{op}}$

Proof. — We have already seen bijections between (2) and (3), and (1) is in bijection with the intersection of (2) with $\ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]})].$ So we have to show
that (2) is contained in \( \ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]})] \). This is another calculation with the exact sequence
\[
H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{Q}, I_{[x]}(\mathcal{A})) = H^1(\mathbb{Q}, I_{[x]}(\mathcal{A})) \oplus H^1(\mathbb{R}, I_{[x]}) \to A(I_{[x]});
\]
indeed, the local term at \( p \) is trivial, because \( p \) splits in \( E \), hence \( I_{[x]} \) is locally at \( p \) isomorphic to a product of \( GL(r) \)'s, as in the proof of Lemma 6.6.3.

(6.6.5) Definition. — A polarized Hodge type of type \( h \) (for \( B, * \)) is a triple \((M, \bar{w}, [j])\) where \((M, \bar{w})\) is a Hodge type of type \( h \) and \([j]\) is a \( G(\mathcal{A}_f^p) \)-conjugacy class of \#-embeddings \( j : M \to B^\alpha \).

The above lemmas show that there is a surjective map from the set of isogeny classes \([x] = [(A_x, \iota_x, \lambda_x)]\) to the set \( PHT^{(h)} \) of polarized Hodge types of type \( h \).

(6.6.6) Lemma. — Let \([x]\) be an isogeny class. The fiber of this map over the image of \([x]\) consists of \( \ker[H^1(\mathbb{Q}, I_{[x]})] \) isogeny classes.

Proof. — The fiber is the set of isomorphism classes in the near equivalence class. Recall that equivalence classes are identified with the set of elements of \( \ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]})] \) whose localization in \( \mathcal{A}_f^p \) lies in
\[
H^1(\mathbb{Q}, I_{[x]}(\mathcal{A}_f^p))(0) = \ker[H^1(\mathbb{Q}, G_{[x]}(\mathcal{A}_f^p))) \to H^1(\mathbb{Q}, G(\mathcal{A}_f^p)))].
\]
Two equivalence classes are nearly equivalent if they map to the same element of \( H^1(\mathbb{Q}, I_{[x]}(\mathcal{A}_f^p)) \). But since they are already map to zero in \( H^1(\mathbb{R}, I_{[x]}) \), and since (as in the proof of Lemma 6.6.4) there is no cohomology at \( p \), we can say they differ by an element of \( \ker[H^1(\mathbb{Q}, I_{[x]})] \). On the other hand, \( \ker[H^1(\mathbb{Q}, I_{[x]})] \) is a finite group (it can be identified with the image of \( \ker[H^1(\mathbb{Q}, Z_{I_{[x]}})] \), as before, for instance) that acts faithfully on \( H^1(\mathbb{Q}, I_{[x]}) \), so the cardinality is as indicated. \( \square \)

Recall the fixed complex embedding \( \tau_0 \) of \( F \). If \( z = (M, \bar{w}, [j]) \in PHT^{(h)} \) and \( j \in [j] \), there is a unique distinguished \( \bar{\tau}_0 \) of \( \tau_0 \) to \( M \) — except in the case of \( GU(2) \) over an imaginary quadratic field, which we have deliberately excluded — defined as follows: the embedding \( j \) endows \( V \) with a structure of \( * \otimes c \)-hermitian \( B_M(\mathbb{R}) \)-module, denoted \( V_j \). This gives a set of signatures \((a_{\sigma}, b_{\sigma})\) for every real embedding \( \sigma \) of \( M^+ \). For only one such \( \bar{\tau}_0 \) is this signature indefinite, \( \bar{\tau}_0 \) restricts to our chosen \( \tau_0 \) on \( F^+ \), and we let \( \bar{\tau}_0 = \bar{\tau}_0(j) \) be its extension to \( M \) lifting \( \tau_0 \). (Think of breaking up \( V_{\tau_0} \) under the action of \( M \otimes F, \tau_0 \mathbb{C} \). It has \([M : F]\) constituents and only one of them can be indefinite.)

Now if \( j, j' \in [j] \) with \( \bar{\tau}_0(j) = \bar{\tau}_0(j') \), then the \( * \otimes c \)-hermitian \( B_M(\mathbb{R}) \)-modules \( V_j(\mathbb{R}), V_{j'}(\mathbb{R}) \) are isomorphic. Since these are the same real vector space, the isomorphism can be realized by conjugation in \( G(\mathbb{R}) \). On the other hand, the fact that \( j \) and \( j' \) are in the same \( G(\mathcal{A}_f^p) \)-conjugacy class means that \( V_j(\mathcal{A}_f^p) \overset{\sim}{\longrightarrow} V_{j'}(\mathcal{A}_f^p) \) (with their hermitian forms), and at \( p \) there is no possible difference. In the end, we have isomorphisms \( G_{j, \mathcal{A}} \overset{\sim}{\longrightarrow} G_{j', \mathcal{A}} \) where \( G_j \) is the unitary similitude group of \( V_j \), and
this isomorphism is canonical up to $G_{j,A}$-conjugation. Note that $G_j$ is realized as a subgroup of $G$ (the commutant of $j(M^\times) \cap G$).

(6.6.7) Lemma

(1) The map $j \mapsto \tau_0(j)$ is surjective.
(2) If $\tau_0(j) = \tau_0(j')$ then the isomorphism $G_{j,A} \simto G_{j',A}$ comes from an isomorphism over $\mathbb{Q}$.

Proof. — Part (1) comes from the same diagram chase as before; we find that the possible spaces $V_j \otimes \mathbb{R}$ are in bijection with the set ker$[H^1(\mathbb{R}, G) \to H^1(\mathbb{R}, G)]$ which correspond precisely to the extensions of $\tau_0$ to $M$. As for (2), the point is that the map ker$^1(\mathbb{Q}, Z_{G_j}) \to$ ker$^1(\mathbb{Q}, G_j)$ is an isomorphism. This proves in a standard way that the cocycle measuring difference between the hermitian vector spaces $V_j$ and $V_{j'}$ defines a trivial inner twist of the groups, and a similar cohomology calculation shows that the given isomorphism over $A$ can be modified to give an isomorphism over $\mathbb{Q}$. 

It remains to carry out a similar analysis for the groups $I_{[x]}$. The result is

(6.6.8) Lemma. — Suppose $[x]$ and $[x']$ are two isogeny classes with the same image $z \in PHT^{(h)}$. Then the groups $I_{[x]}$ and $I_{[x']}$ are $\mathbb{Q}$-isomorphic, and the isomorphism can be chosen compatible with the isomorphisms $I_{[x]}(A_f^w) \simto G_j(A_f^w) \simto I_{[x]}(A_f^w)$.

The proof is completely analogous to that of Lemma 6.6.7, but simpler in that there is no possible difference at the real places: both $I_{[x]}$ and $I_{[x']}$ are $\mathbb{R}$-anisotropic (modulo the center).

Let $I_z = I_{[x]}$ for any isogeny class $[x]$ lying over $z \in PHT^{(h)}$. Fix a level subgroup $U_{w,h} \subset G(A_f^w) \times G_h \times \mathbb{Q}_p^\times$ (as always, the factor in $\mathbb{Q}_p^\times$ is $\mathbb{Z}_p^\times$). We thus have a complete description of the $\mathbb{F}$-points of $\overline{S}^{(h)}$:

(6.6.9) $\overline{S}_{M_{n-h,+},U_{w,h}, A_f^w}(\mathbb{F}) = \coprod_{z \in PHT^{(h)}} \left( [\tilde{M}_{n-h,} \times (I_z(\mathbb{Q})\backslash G^{(h)}/U_{w,h})]/J_{n-h,+} \right) |\ker^1(Q,I_z)|$.

This decomposition is compatible with the action of Frobenius (on the pro-discrete set $\tilde{M}_{n-h,+}$, and it factors through a finite Galois group), and of the Hecke algebra of $G(A_f^w) \times L_{n-h,h} \times \mathbb{Q}_p^\times$. Here $L_{n-h,h} = G_{n-h} \times G_h$ acts as follows: $G_{n-h}$ acts on $\tilde{M}_{n-h,+}$ on the first factor, whereas $G_h$ acts on the $G^{(h)}$-factor. This action of $L_{n-h,h}$ commutes with $J_{n-h,+}$.

We are now almost ready to count points.

7. Comparison of trace formulas

7.1. Counting transfers from $I_z(\mathbb{Q})$ to $G(\mathbb{Q})$, following Kottwitz. — We want to determine the trace of the representation of $G(A_f^w) \times L_{n-h,h,+}$ on the cohomology
Recall the description from the last time:

\[ \left[ H_c(\overline{S}_{M_0}^{(h)}, R\Psi) \right]. \]

We will treat the étale \((G_h)\) part of the level structure on the \(p\)-divisible group together with the prime-to-\(w\)-level structures.

As in Clozel’s course, one uses a version of the Grothendieck-Lefschetz trace formula to calculate the trace as a sum over contributions of fixed points. Because we are dealing with cohomology with compact support, we need the formula proved by Fujiwara; in particular, we can only use Hecke operators that incorporate Frobenius. Because we have already determined the stalks of \(R\Psi\), the Galois representation will come along for free.

First, we work out the cohomological formalism for transferring conjugacy classes from \(I_z(\mathbb{Q})\) to \(G(\mathbb{Q})\), up to adelic conjugation. Recall that we are always excluding the case \(F^+ = \mathbb{Q}, n = 2\). We begin with some definitions.

\textbf{(7.1.1) Definition.} — An element \(j = (j_{n-h}, j_h) \in J_{n-h,h}\) is \(h\)-regular if the \(p\)-adic valuation of every eigenvalue of \(j_{n-h}\) is strictly less than the \(p\)-adic valuation of every eigenvalue of \(j_h\) (i.e. \(|j_{n-h}| > |j_h|\)). An element \(\gamma \in I_z(\mathbb{Q})\) is \(h\)-regular if its image in \(J_{n-h,h}\) is \(h\)-regular.

Note that \(h\)-regularity is a property of conjugacy classes. The same definition can be made for \(g = (g_{n-h}, g_h) \in L_{n-h,h} \subset G_n\). In that case, the parabolic associated to \(g\) (the expanding parabolic) is contained in \(P_h^{\text{par}}\). We return to this later. An element is \textit{very} \(h\)-regular if the difference in \(p\)-adic valuations is \(\gg \) \(N\) for some large integer \(N\) determined by the problem.

\textbf{(7.1.2) Lemma.} — Let \(\gamma \in I_z(\mathbb{Q})\) be \(h\)-regular, with \(z = (M, \tilde{w}, [j])\). Then \(F(\gamma) \supset M\).

This is a simple argument with ramification groups of primes of \(F(\gamma)\) above \(w\), and uses the minimality of \(M\), and is related to (7.3.4) below. See Lemma V.2.2 of \([HT]\) for details.

Since \(I_z(\mathbb{R})\) is anisotropic modulo center, every element of \(I_z(\mathbb{Q})\) is elliptic; in particular, is semisimple. However, they are not necessarily regular. One could restrict attention to regular elements by using a trick due to Labesse, but this trick only works for forms of \(GL(n)\). Thus we work out the general case. The following analysis is based on Kottwitz’ article \([K2]\).

\textbf{(7.1.3) Lemma.} — Let \(\gamma \in I_z(\mathbb{Q})\). The number of \(I_z(\mathbb{Q})\)-conjugacy classes in the \(I_z(\mathbb{A})\)-conjugacy class of \(\gamma\) equals \(|\ker^1(\mathbb{Q}, Z_{I_z}(\gamma))/\ker^1(\mathbb{Q}, I_z)|\).

\textbf{Proof.} — If \(\gamma\) and \(\gamma' \in I_z(\mathbb{Q})\) are conjugate over \(\mathbb{A}\), their centralizers in \(I_z\) are inner forms of another that become isomorphic over \(\mathbb{A}\). In this way one sees that the
number is the cardinality of
\[ \ker[\ker^1(\mathbb{Q}, Z_{I_z}(\gamma))] \rightarrow \ker^1(\mathbb{Q}, I_z)]. \]
The Lemma follows from the surjectivity of this map, which follows from the fact (already used in §6.6) that
\[ \ker^1(\mathbb{Q}, Z_{I_z}) \rightarrow \ker^1(\mathbb{Q}, I_z) \]
is an isomorphism, and likewise for \( Z_{I_z}(\gamma). \)

Recall that the group \( I_z \) depends only on \( z \) (up to isomorphism and \( G(h)(A) \)-conjugacy), whereas the inner forms \( G_{[a]} \subset G \) depend also on the choice of an extension \( \tau_0 \) of \( \tau_0 \) to \( M \) (up to isomorphism and \( G(A) \)-conjugacy). We let \( G_z, \tau_0 \) denote this \( \mathbb{Q} \)-group.

(7.1.4) — Now we discuss transfer from \( I_z \) to \( G_z, \tau_0 \) to \( G \). Note that we can discuss \( h \)-regular elements in \( G_z, \tau_0(\mathbb{Q}) \), since it comes with an embedding in \( L_{h,n-h} \) at \( w \).

Consider the following three sets:

(7.1.4.1) The set \( T(h) \) of pairs \((z, [a])\) where \( z \in \text{PHT}^{(h)} \) and \([a]\) is an \( h \)-regular \( I_z(A) \)-conjugacy class in \( I_z(\mathbb{Q}) \).

(7.1.4.2) The set \( G(h) \) of triples \((z, \tau_0, [\gamma])\) where \( z = (M, \bar{w}, [j]) \in \text{PHT}^{(h)} \), \( \tau_0 \) is as above, and \([\gamma]\) is an \( h \)-regular \( G_z, \tau_0(A) \)-conjugacy class in \( G_z, \tau_0(\mathbb{Q}) \) that is \( \mathbb{R} \)-elliptic and has elliptic image in \( G_{n-h} \subset L_{n-h} \).

(7.1.4.3) The set \( FP^{(h)} \) of equivalence classes of pairs \((\gamma, \bar{w})\) where \( \gamma \in G(\mathbb{Q}) \) is an \( h \)-regular \( \mathbb{R} \)-elliptic element and where \( \bar{w} \) is a place of \( F(\gamma) \) above \( w \) such that
\[ (n-h)[F(\gamma) : F] = n[F(\gamma)_{\bar{w}} : F_w]. \]
The pairs \((\gamma, \bar{w})\) and \((\gamma', \bar{w}')\) are equivalent if \( \gamma \) and \( \gamma' \) are conjugate by an element of \( G(A) \) inducing an isomorphism \( F(\gamma)_{\bar{w}} \sim F(\gamma')_{\bar{w}} \) identifying \( \bar{w} \) with \( \bar{w}' \).

Note that \( M \) has disappeared from (7.1.4.3). The decomposition into isogeny classes gives us elements as in (7.1.4.1), and we want to get to (7.1.4.3). Note also that elements of (7.1.4.1) can embed in \( G^{(h)}(A) \), as follows: The embedding of \( G(\mathbb{Q}) \) in \( G(A_f^p) \) is obvious. To get embeddings at primes other than \( w \) dividing \( p \), embed \( G(\mathbb{Q}) \) in \( G(\mathbb{Q}_p) \), then project as in (2.4.1) on the factors other than \( w \). Finally, to obtain an embedding in \( J_{n-h} \times G_h \), it suffices to show that the field \( F(\gamma)_{\bar{w}} \) embeds in \( D_{n-h} \), and this follows from the equality of degrees (7.1.4.4).

Recall that the group \( G_z, \tau_0 \) comes with an embedding in \( G \).

(7.1.5) Lemma. — The map \( G(h) \rightarrow FP^{(h)} \), sending \((z, \tau_0, [\gamma])\) with \( z = (M, \bar{w}, [j]) \) to \((\gamma, \bar{w}')\), where \( \bar{w}' \) is the unique place of \( M(\gamma) = F(\gamma) \) above the place \( \bar{w} \) of \( M \), is a bijection.

Proof. — First note that \( \gamma \) being \( h \)-regular, \( M(\gamma) = F(\gamma) \). The \( G_{n-h} \)-ellipticity implies that \( F(\gamma) \otimes_M M_{\bar{w}} \) is a field, hence that \( \bar{w}' \) exists. Let \((\gamma, \bar{w}) \in FP^{(h)} \), and let
$M \subset F(\gamma)$ be the minimal subfield containing $F$ for which $\bar{w}$ is inert from $M$ to $F(\gamma)$. (The existence of such an $M$ is left as an exercise). Then $(M, \bar{w})$ is a Honda-Tate parameter. To obtain the polarization, let $j : M \hookrightarrow B^{op}$ be the tautological embedding. This endows $V_{\mathbb{R}}$ with the structure of $B_M \otimes_{\mathbb{Q}} \mathbb{R}$-module, and since $j$ comes from an element $\gamma$ already in $G$, this module has an $\ast \otimes c$-hermitian $\mathbb{R}$-alternating pairing. The invariants $(a_{\tau}, b_{\tau})$ of this pairing, for $\tau : M^+ \to \mathbb{R}$, pick out a unique complex place $\bar{\tau}_0$ except in the excluded case $n = 2$. This defines an element $((M, \bar{w}, [j]), \bar{\tau}_0, [\gamma]) \in G^{(h)}$ above $(\gamma, \bar{w})$, and it is clearly unique.

The other comparison is deeper. There is a map $\phi : G^{(h)} \to I^{(h)}$ sending $(z, \bar{\tau}_0, [\gamma])$ to $(z, [a])$ where $a \in I_z(\mathbb{Q})$ is conjugate to $\gamma$ in $I_z(A) \xrightarrow{\sim} G_{z, \bar{\tau}_0}(A)$. The existence of such an $a$ is the most difficult step in the counting argument. The following lemma asserts that $a$ exists and is unique up to $I_z(A)$-conjugacy.

(7.1.6) Lemma. — The map $\phi$ is well-defined; i.e., the $A^{\tau}_n$-conjugacy class of $\gamma$ has a representative in $I_z(\mathbb{Q})$. Moreover, the map $\phi$ is surjective, and the fiber above $(z, [a])$ has cardinality $[F(a) : F] = [F(\gamma) : F].$

Proof. — We first associate a well-defined adelic conjugacy class $a_A \subset I_z(A)$ to $(z, \bar{\tau}_0, [\gamma])$. Away from $\infty$ and $w$ there is nothing to say. Since $\gamma$ is $\mathbb{R}$-elliptic, it transfers to any inner form over $\mathbb{R}$. More precisely, its transfer is well-defined as a stable conjugacy class (up to conjugacy over $\mathbb{C}$). But $I_z(\mathbb{R})$ is compact modulo center, so $\mathbb{C}$-conjugacy and $\mathbb{R}$-conjugacy coincide. Finally, at $w$, we need to show that the image of $\gamma$ in $I_{n-h}$ transfers to a well-defined conjugacy class in $D_{n-h}^c \times G_h$. But this follows from the hypothesis that the image of $\gamma$ in $G_{n-h}$ is elliptic.

We view $a_A$ as an $I_z(A)$ conjugacy class that contains a representative in $I_z(\overline{\mathbb{Q}}) = G_{z, \bar{\tau}_0}(\overline{\mathbb{Q}})$, namely $\gamma$. The problem is now to determine whether or not it has a representative in $I_z(\mathbb{Q})$. In [K2], for any connected reductive group $H$ with simply-connected derived group, Kottwitz constructed an obstruction $\text{obs}(\gamma, [a_A])$ where the first term is an $H(\overline{\mathbb{Q}})$-conjugacy class and the second an $H(A)$-conjugacy class, whose vanishing is equivalent to the existence of a representative in $H(\mathbb{Q})$. (The hypothesis that it be simply connected is removed by Labesse, and the connectedness is likewise replaced by the hypothesis that the group of components is cyclic.) This obstruction class belongs to the group Kottwitz denotes $\mathfrak{H}(I^0/\mathbb{Q})$, the group of endoscopic characters; here $I^0$ is the centralizer of the transfer to the quasi-split inner form of $I_z$ (or of $G_{z, \bar{\tau}_0}$) of $\gamma$. (By a theorem of Kottwitz, $\gamma$ always transfers to the quasi-split inner form.) But this is precisely the group that vanishes for every possible $I^0_z$, as Clozel showed in his course. (In [HT] the argument is given on p. 180.) If this were not the case, we would have to restrict $G^{(h)}$ to the set of $(z, \bar{\tau}_0, [\gamma])$ for which the Kottwitz obstruction vanishes. This would lead to a different formula in the end, but still presumably in the direction of the stable trace formula.
Remark. — More generally, the Kottwitz invariant for a triple coming from a polarized abelian variety should be related in a simple way to this obstruction invariant.

In any case, we have shown the existence of \((z, [a]) \in T^h\). Now we have to determine the cardinality of its inverse image under \(\phi\). In the first place, its inverse image is non-empty. Indeed, the argument above applies just as well in the opposite direction, showing that \([a]\) transfers to a rational conjugacy class in \(G_z, \tilde{\tau}_0\). This already decomposes the inverse image into \([M : F]\) subsets, one for each choice of \(\tilde{\tau}_0\). It remains to show that each subset has \([F(a) : M] = [M(a) : M]\) distinct elements (except in the excluded case). Remember that we are counting \(G_{z, \tilde{\tau}_0}(A)\)-conjugacy classes, not \(G_z, \tilde{\tau}_0(\mathbb{Q})\)-conjugacy classes! But the \(I_z(A)\)-conjugacy class of \([a]\) determines the \(G_{z, \tilde{\tau}_0}(A_f)\)-conjugacy class uniquely. Indeed, the groups only differ at \(w\), but there the transfer from \(D_{n-h}^0\) to \(G_{n-h}\) is injective. So the only ambiguity is at \(\infty\), and indeed at \(\tilde{\tau}_0\), since elsewhere \(G_{z, \tilde{\tau}_0}\) is compact mod center. The question is then to count conjugacy classes in \(G_{z, \tilde{\tau}_0}(\mathbb{R})\) stably conjugate to \(a\), and as before these are in bijection with extensions of \(\tilde{\tau}_0\) to \(M(a)\). Indeed, they are parametrized by

\[
\ker[H^1(\mathbb{R}, Z_{G_{z, \tilde{\tau}_0}(a)}) \rightarrow H^1(\mathbb{R}, G_{z, \tilde{\tau}_0}(a))],
\]

(kernel as map of pointed sets), and this set also parametrizes equivalence classes of \(*\otimes\cdot\)-hermitian \(B \otimes M(a)(\mathbb{R})\)-modules that are equivalent to the given \(B_M(\mathbb{R})\)-module.

So the calculation is as before.}

7.2. Acceptable functions and Fujiwara’s trace formula. — For the next step we need to work at finite level. Let \(U_h^w = U_h^w \times U_h^w\) for some compact open subgroup \(U_h \subset G_h = L_{0, h}\). We introduce a class of acceptable functions

\[\phi \in C_c^\infty(G(A_f^w) \times L_{n-h, h}/U_h^w)\]

where the symbol \(/\) designates bi-invariance. These functions act as correspondences on \(\mathcal{S}_{U_h^w}\) and on the complex \(R\Psi\), hence define operators on \([H_c(\mathcal{S}_{M_0}^h, R\Psi)]|_{U_h^w}\). We assume \(\phi\) factors as \(\phi^\wedge \otimes \phi_w\), with \(\phi_w = \phi_{w, n-h} \otimes \phi_{w, h}\). Say \(\phi\) is \(h\)-regular (resp. very \(h\)-regular) if \(\phi_w\) is supported in the set of \(h\)-regular (resp. very \(h\)-regular) elements of \(L_{n-h, h}\).

The goal is to determine the trace of \(\phi\) on \([H_c(\mathcal{S}_{M_0}^h, R\Psi)]|_{U_h^w}\) for all \(\phi\). This would suffice to prove the Second Basic Identity in the form (6.3.2), but this is both impossible and unnecessary. Here is one way of stating Fujiwara’s trace formula [F] in our present situation:

(7.2.1) Theorem. — Let \(\phi \in C_c^\infty(G(A_f^w) \times L_{n-h, h}/U_h^w)\), and suppose \(\phi\) is very \(h\)-regular (depending on \(U_w^w\) and \(U_h^w\)). Then

\[
\text{Tr}(\phi | [H_c(\mathcal{S}_{M_0}^h, R\Psi)]|_{U_h^w}) = \sum_{z \in PHT^h} |\ker^1(\mathbb{Q}, I_z)| \text{Tr} \left( \phi \mid H^0([M_{n-h, +} \times (I_z(\mathbb{Q})\backslash G_f^h)/U_h^w] \cap J_{n-h, +}, [R\Psi]_z) \right).
\]
The above formula requires a few comments. The left-hand side being a trace on cohomology, the right-hand side must be a sum over fixed points. But the fixed points can be regrouped among \((G(A_w) \times L_{n-h,h}/U_h) \times W_K\)-invariant subsets, and we choose to regroup them according to isogeny classes, which are zero-dimensional. Then it is purely formal that the sum over fixed points in an isogeny class can be rewritten as a trace on cohomology: the Lefschetz formula is also valid for zero-dimensional varieties. The groups

\[
H^0(\tilde{M}_{n-h,+} \times (I_z(Q)\backslash G^{(h)}/U_h)/J_{n-h,+}, [R\Psi]_z)
\]

are smooth, but not generally admissible, representations of \(G^{(h)}\). However, under the hypothesis that \(\phi\) is very \(h\)-regular, the set of fixed points of \(\phi\) on \(\tilde{M}_{n-h,+} \times (I_z(Q)\backslash G^{(h)}/U_h)/J_{n-h,+}\) is finite, so we formally define the trace to be the sum over the fixed points of local terms, whose definition is recalled below. This is a bit ad hoc but has the right formal properties for our present purposes; moreover, it is the form in which the trace formula will be used.

Next, because the strata are not proper, Fujiwara’s theorem requires that a correspondence be twisted by a high power of Frobenius in order to eliminate wild local terms at the boundary. This is the reason for the condition that \(\phi\) be very \(h\)-regular. Fujiwara’s theorem is proved for varieties, i.e. noetherian schemes, hence we need to work at finite level; in principle, the degree of \(h\)-regularity depends on the choice of level subgroup. One could have worked with a general \(\phi\), twisted by a sufficiently high power of Frobenius, but in fact the twist by Frobenius is built into the \(h\)-regularity condition. This is a consequence of what Carayol calls the congruence formula for strata, which basically comes down to the formula (3.1.4). For details, see [HT, Lemma V.1.3].

Recall from (4.3.4) that the stalk of \(R^i\Psi\) at a point in the \(h\)-stratum is isomorphic to the representation of \(A_{K,n-h}\) on \(\Psi^{i}_{c,n-h,0,x_0}\). Recall also the decomposition (5.5.2), (5.5.4), (5.5.5) of the alternating sum \([\Psi_{c,n-h,0,x_0}]\) as a sum over inertial equivalence classes \([\rho]\in [A](J_{n-h})\), and the corresponding decomposition (5.5.11) for the cohomology. There is also a version \([\rho, +]\) incorporating the action of the extra factor \(Q_p^\times /\mathbb{Z}_p^\times\), whose definition is left to the reader. This gives a decomposition of the virtual sheaf of vanishing cycles \([R\Psi]_z\) over the zero-dimensional pro-variety \(S(z)\), and hence an expression for the cohomology space on the right-hand side of (7.2.1):

\[
H^0(S(z), [R\Psi]_z) = \bigoplus_{[\rho]\in [A](J_{n-h})} H^0(S(z), [\Psi]|_{[\rho]}).
\]

We rewrite Fujiwara’s trace formula accordingly:

(7.2.3) Corollary. — Under the hypotheses of Theorem 7.2.1,

\[
\text{Tr}(\phi | [H_c(S_{M_0}, R\Psi)](U_h)) = \sum_{z\in \text{PHT}^{(h)}} \sum_{[\rho]\in [A](J_{n-h})} |\ker^1(Q,I_z)|_{[z,\rho]} \phi
\]
where
\[ t_\zeta[\rho](\phi) = \text{Tr} \left( \phi \mid H^0([L_{n-h} \times (I_\zeta(\mathbb{Q}) \setminus G^{(n)}/U_h^w)]/J_{n-h+}, [R\Psi]_{\zeta[\rho]}) \right). \]

The meaning of the trace on the right-hand side is as above.

(7.2.4) Remark. — As remarked following (6.3.1), [HT] obtains the corresponding decomposition globally over the h-stratum, as a sum of lisse sheaves indexed by inertial equivalence classes of representations of \( J_{n-h} \).

7.3. Expression for trace of acceptable functions, and transfer to \( G \)

Our ultimate goal is to prove the formula (6.3.4), in its “pre (4.3.11) version, namely

(7.3.1) \[ \eta[H_c(S^{(h)}_{M_0}, R\Psi)] = \bigoplus_{[\rho] \in [A((J_{n-h})_{\text{fin}}[\rho])} \text{red}^{(h)}(H(A(B, \ast)) \otimes [\Psi_{n-h}(\rho')] \in \text{Groth}(G(A^w) \times G_h). \]

Formula (7.3.1) is understood as an equality in Groth\((G(A^w) \times L_{n-h})\). Notation is as in Theorem 6.1.2(ii); in particular, the sums on both sides are finite.

To prove (7.3.1), we prove the traces on the two sides are equal for a sufficiently large family of test functions \( \phi = \phi^w \otimes \phi_w \), with \( \phi_w = \phi_{w,n-h} \otimes \phi_{w,h} \) as above. The functions \( \phi^w \) and \( \phi_{w,h} \) are chosen arbitrarily, whereas \( \phi_{w,n-h} \) has to be chosen so that the resulting \( \phi \) is very \( h \)-regular. One verifies without difficulty that such a set of functions suffices to separate characters, the point being admissibility of the two sides; here the finiteness of the sets \( A((J_{n-h})_{\text{fin}}[\rho]) \) is crucial. For example, by Theorem A.1.5 of the appendix, one can choose \( \phi_{w,n-h} \) to be a pseudocoefficient for any fixed \( JL(\rho') \), with \( \rho' \in A((J_{n-h})_{\text{fin}}[\rho]) \), relative to the set \( JL(A((J_{n-h})_{\text{fin}}) \subset A_d(n-h, K) \) (cf. (A.1.3)). Moreover, the condition (A.1.11) guarantees that, for any pair \( (\phi^w, \phi_{w,h}) \) the choice of \( \phi_{w,n-h} \) can be made consistently with the condition that \( \phi \) be very \( h \)-regular.

Fix \( \rho' \in A(J_{n-h})_{\text{fin}}[\rho] \) and let \( \pi' = JL(\rho') \). To fix ideas, and to simplify the formulas the first time around, we assume

(7.3.2) Hypothesis. — \( \pi' \) is supercuspidal and \( \phi_{w,n-h} \) is a pseudocoefficient for \( \pi' \), denoted \( \phi_{\pi',\omega} \) in the notation of (A.1).

Here \( \omega \) is an interval \( [a, b] \subset \mathbb{Z} \) chosen to guarantee the \( h \)-regularity condition, and long enough (i.e. \( m = \frac{b-a+1}{n-h} \in \mathbb{Z} \) is sufficiently large) to guarantee that \( \phi_{\pi', \omega} \) picks out \( \pi' \) among its unramified twists occurring in \( JL(A(J_{n-h})_{\text{fin}}) \). Set \( \phi^w_\omega = \phi^w \otimes \phi_h \). A test function of the form \( \phi = \phi^w_\omega \otimes \phi_{\pi', \omega} \) as above – in particular, satisfying the \( h \)-regularity condition – will be called acceptable for \( \rho' \). We verify (7.3.1) by proving equality of traces for all test functions acceptable for \( \rho' \), for all \( \rho' \in A(J_{n-h})_{\text{fin}}[\rho] \). In the final paragraphs of §7.6 we explain what needs to be modified when Hypothesis
(7.3.2) is relaxed; i.e., when $\phi_{w,n-h}$ is taken to be an arbitrary test function and $\rho$ is an arbitrary representation of $J_{n-h}$.

For $a \in I_z(\mathbb{Q})$, define the orbital integral

\[(7.3.3) \quad O^h_{[a]}(\phi^w_h) = \int_{Z(a) \backslash G(A^h_\gamma) \times G_h} \phi^w_h(gag^{-1})dg.\]

Here $Z(a)$ is the centralizer of $a$ in $G^{(h)}(A_f)$. In the applications, only $h$-regular $a$ contribute non-vanishing orbital integrals. We may thus assume $a$ to be $h$-regular. It is not too difficult\(^{(9)}\) to see that this implies

\[(7.3.4) \quad Z_G(a) = Z_{I_z}(a)\]

[HT, Lemma V.2.2], so $Z(a)$ is the adelicization of the $\mathbb{Q}$-group $Z_{I_z}(a)$ (though $G^{(h)}(A_f)$ is not adelic). Via the embedding of $I_z(\mathbb{Q})$ in $J_{n-h}$, $a$ defines a local conjugacy class $[a] \subset J_{n-h}$, necessarily elliptic. We let $[\gamma(a)]$ denote the transfer of $[a]$ to a conjugacy class in $G_n-h$; i.e., an element $\gamma \in [\gamma(a)]$ becomes conjugate to an element $a \in [a]$ under an isomorphism $J_{n-h} \simeq G_{n-h}$ over $\mathbb{K}$. (All conjugacy classes in $J_{n-h}$ transfer to the quasi-split inner form $G_{n-h}$).

To save space, volumes are denoted $v$ rather than $\text{vol}$. Here is an expression for the contribution of $z \in PHT^{(h)}$ to the trace formula:

\[(7.3.5) \text{Theorem.} \quad \text{Fix } \rho' \in \mathcal{A}(J_{n-h})_{\text{fin}}[\rho], \text{ and let } \phi = \phi^w_h \otimes \phi_{\pi',\omega} \text{ be a test function acceptable for } \rho'. \text{ Then}
\]

\[
\text{Tr}(\phi | H^n([\tilde{M}_{n-h,+} \times S^h_{U_h}(z)]/J_{n-h,+}, [R\Psi]_z[\rho])) = (n-h) \sum_{[a]} e(\gamma(a))O^h_{[a]}(\phi^w_h) \cdot G^{(h)}_{\gamma(a)}(\phi_{\pi',\omega})v(Z_{I_z}(a)(\mathbb{Q}) \backslash Z_{I_z}(a)(A_f)).
\]

Here $[a]$ runs through $h$-regular $I_z(\mathbb{Q})$-conjugacy classes in $I_z(\mathbb{Q})$, and the volume $v(Z_{I_z}(a)(\mathbb{Q}) \backslash Z_{I_z}(a)(A_f))$ is normalized as for $h = 0$. Moreover, $[\gamma(a)] \subset G_{n-h}$ is the transfer of the conjugacy class $[a] \in J_{n-h}$, as above. Finally $e(\gamma(a))$ is the Kottwitz sign $(A.1.12 \text{ bis})$.

The proof of this formula is based on a standard argument for translating point counting problems on double coset spaces into sums of orbital integrals, and will be our last order of business. We note here that this calculation presupposes Theorem 4.3.11, as well as Hypothesis 7.3.2, and hence suffices to prove the strong version of the Second Basic Identity. In (7.6) we will first obtain the weaker version.

\(^{(9)}\)The point is subtle, however, and deserves to be stressed, as it lies at the heart of the difference between the approach to point counting in [HT] and that in [K5]. The proof in [HT, Lemma V.2.2], which simultaneously establishes Lemma 7.1.2, is elementary, but we have not yet seen how it generalizes to other Shimura varieties.
The first subtlety involves rewriting the volume factor, using Kottwitz’ results on Tamagawa numbers [K3]. The formula is
\[ v(Z_{I_z}(a)(\mathbb{Q}) \setminus Z_{I_z}(a)(A_f)) = \kappa_B |\ker^1(\mathbb{Q}, Z_{I_z}(a))|^{-1} v(Z_{I_z}(a)(\mathbb{R})^1)^{-1} \]
where \( \kappa_B \) and the measures are as in our discussion of Arthur’s formula; in particular, \( \kappa_B = |A(Z_{I_z}(a))| = 2 \) if \( |B : \mathbb{Q}| \) is divisible by 4 and 1 otherwise. This is an explicit computation (cf. p. 167 of [HT]).

In particular, we can rewrite the expression in Theorem 7.3.5 as
\[ (7.3.6) \quad (n - h)\kappa_B \sum_{[a]} \epsilon(\gamma(a)) O_{[a]}^h(\phi_w) \cdot O_{[\gamma(a)]}^{G_{n-h}}(\phi_{n-h}) \ker^1(\mathbb{Q}, Z_{I_z}(a))|^{-1} v(Z_{I_z}(a)(\mathbb{R})^1)^{-1} \]
Next, to rewrite Theorem 7.3.5 as a sum over \( I_z(A) \)-conjugacy classes, we note that if \( a, a' \in I_z(\mathbb{Q}) \) are \( I_z(A) \)-conjugate, then their centralizers are inner forms of each other that become isomorphic over \( A \), and their Tamagawa measures agree under this isomorphism. Thus
\[ (7.3.7) \quad O_{[a]}^h(\phi_w) \vol(Z_{I_z}(a)(\mathbb{R})^1)^{-1} = O_{[a']^1}^h(\phi_w) \vol(Z_{I_z}(a')(\mathbb{R})^1)^{-1}. \]
So it suffices to count the number of \( I_z(\mathbb{Q}) \)-conjugacy classes in an \( I_z(A) \)-conjugacy class, and this is
\[ (7.3.8) \quad \ker(\ker^1(\mathbb{Q}, Z_{I_z}(a)) \to \ker^1(\mathbb{Q}, I_z)) = \ker^1(\mathbb{Q}, Z_{I_z}(a))/\ker^1(\mathbb{Q}, I_z) \]
because the map on \( \ker^1 \)'s is surjective, a fact we have already used several times. Write \([a]/\mathbb{Q}\) for \( I_z(\mathbb{Q})\)-conjugacy classes, \([a]/A\) for \( I_z(A)\)-conjugacy classes, and write \( v(a) = \vol(Z_{I_z}(a)(\mathbb{R})^1) \). Then
\[ (7.3.9) \quad \Tr(\phi) [H_c(S_{\mathfrak{m}_0}^h [R\Psi])_{l}^{l_{\gamma}(a)}] = \sum_{z \in PHT(h)} \sum_{[a] \in [A](J_{n-h})} |\ker^1(\mathbb{Q}, I_z)|_{l_{z,[a]}(\phi)} \]
\[ \quad = \sum_{z \in PHT(h)} |\ker^1(\mathbb{Q}, I_z)|_{l_{z,[\rho]}(\phi)} \]
\[ \quad = (n - h)\kappa_B \sum_{z,[a]} v(a)^{-1} \epsilon(\gamma(a)) O_{[a]}^h(\phi_w) \cdot O_{[\gamma(a)]}^{G_{n-h}}(\phi_{n-h}) \]
\[ \quad = (n - h)\kappa_B \sum_{z,[a]} v(a)^{-1} \epsilon(\gamma(a)) O_{[a]}^h(\phi_w) \cdot O_{[\gamma(a)]}^{G_{n-h}}(\phi_{n-h}). \]
The first equality is (7.2.3), and the third is (7.3.5). The second follows from our choice of \( \phi_{n,-h} \) to be a pseudocoeficient for \( \pi' \), which by (5.2.18) eliminates all \([\alpha] \neq [\rho] \). This is the step that will have to be treated in greater generality at the end of \( \S \) 7.6. The final line summarizes the discussion following (7.3.5). Note that the passage from \([a]/\mathbb{Q}\) to \([a]/A\) is just what it takes to eliminate the \( \ker^1 \)'s, thanks to Kottwitz’ theorem on Tamagawa numbers [K3]. This is a central step in the point.
counting argument, and more generally of the theory of the stable trace formula (this point was also made in Clozel’s course).

We now use the comparison with $G^{(h)}$, and then with $FP^{(h)}$, to rewrite this as

$$
(n - h)\kappa_B \sum_{(\gamma, \hat{\omega}) \in FP^{(h)}} [F(\gamma) : F]^{-1}e(\gamma)O^h_{[\gamma]}(\phi^w) \cdot O^{G_n - h}_{[\gamma]}(\phi_{n-h})v(a)^{-1}
$$

(7.3.10)

This expression is a bit schizophrenic, because it involves a sum over $\gamma \in G(\mathbb{Q})$, but two of the terms are still expressed in terms of the $\gamma$. To remove all trace of $a$, we consider these terms in turn. First, $v(a) = v(\gamma) = \text{vol}(Z_G(\gamma)(\mathbb{R})_0)$ where $Z_G(\gamma)$ is of course the centralizer of $\gamma$ in $G$, $Z_G(\gamma)(\mathbb{R})_0$ is the compact mod center inner form of $Z_G(\gamma)(\mathbb{R})$,

$$
Z_G(\gamma)(\mathbb{R})_0^1 = \ker |\nu| : Z_G(\gamma)_0 \rightarrow \mathbb{R}_{>0}^*.
$$

Moreover, $Z_G(\gamma)$ is given Tamagawa measure as before. Next, we can obviously replace the orbital integral over $[a]$ in $G(A_F^\infty) \times G_h$ by the orbital integral over $[\gamma]$, since the two give rise to the same conjugacy class. Thus the product simplifies, and the final formula is

$$
\text{Tr}(\phi \mid [H_{c}(S_{M_0}^{(h)}, R\Psi)]) = (n - h)\kappa_B \sum_{(\gamma, \hat{\omega}) \in FP^{(h)}} [F(\gamma) : F]^{-1}e(\gamma)O^{G(A_F^\infty)}_{[\gamma]}(\phi^w) \cdot O^{L_{n-h}}_{[\gamma]}(\phi_w)
$$

(7.3.11)

We have removed the superscript $(U^{\infty}_h)$ because it is built into our choice of functions $\phi$. By definition of $FP^{(h)}$, the $\gamma$‘s that enter into the above sum have the property that their $G_{n-h}$ components transfer to $J_{n-h}$, hence are elliptic.

7.4. Descent, comparison with global trace formula, and second basic identity. — Recall the cohomological version of the trace formula we used to obtain the comparison for the supersingular locus.

$$
\text{Tr}(\Phi \mid [H_{c}(A(B, *))]) = n\kappa_B \sum_\gamma e(\gamma)[F(\gamma) : F]^{-1}e(\gamma)^{-1}O_{[\gamma]}(\Phi)
$$

(7.4.1)

Here $\Phi = \Phi_w \otimes \phi^w \in C_c^\infty(G_n \times G(A_F^\infty))$ and we have written $v(\gamma)$ for $\text{vol}(Z_G(\gamma)(\mathbb{R})_0^1)$, as in (7.3.11). To compare this with our final version (7.3.11) of the trace formula for the stratum $S_{M_0}^{(h)}$, we need a way to compare orbital integrals on $G_n$ with orbital integrals on $L_{n-h}$. This is provided by the following proposition.

(7.4.2) Proposition (Descent of orbital integrals). — Let $\phi_w = \phi_{n-h} \otimes \phi_h \in C_c^\infty(L_{n-h,h})$ be an $h$-regular test function, and suppose the orbital integrals of $\phi_{n-h}$ are supported on the elliptic set. Then there exists a test function $\Phi_w \in C_c^\infty(G_n)$ that satisfies the following three properties:

(7.4.2.1) If $\gamma \in G_n$ is a semi-simple element not conjugate to an element of $L_{n-h,h}$, then $O^{G_n}_{[\gamma]}(\Phi_w) = 0$;
(7.4.2) For any $\gamma \in L_{n-h,h}$,
\[
O_{\gamma}^{G_n}(\Phi_w) = \sum_{s(\gamma)} O_{s(\gamma)}^{L_{n-h,h}}(\phi_{n-h} \otimes \phi_h)
\]
where $s(\gamma)$ runs through the set of $L_{n-h,h}$-conjugacy classes in the $G_n$-conjugacy class of $\gamma$ (i.e., $s \in G_n$ takes $\gamma$ to $L_{n-h,h}$) such that the $G_n-h$-factor of $s(\gamma)$ is elliptic;

(7.4.3) Let $\pi$ be an irreducible admissible representation of $G_n$, with
\[
[r_{G_n,L_h}^{op}(\pi)] = \sum m_{\alpha,\beta}[\alpha \otimes \beta]
\]
for $\alpha \in A(n-h,K)$ and $\beta \in A(h,K)$. Then
\[
\text{Tr}(\pi)(\Phi_w) = \sum m_{\alpha,\beta} \text{Tr}(\alpha)(\phi_{n-h}) \text{Tr}(\beta)(\phi_h).
\]

The ellipticity hypothesis in the above proposition is superfluous, but is satisfied in our present situation. The existence of $\Phi_w$ satisfying simultaneously the orbital integral conditions (7.4.2.1-2) and the trace condition (7.4.2.3) is a special case of descent of orbital integrals. Actually, the map in the other direction is called descent; the $h$-regularity condition is required in order to prove existence of a map in the indicated direction. The proof of Proposition 7.4.2 is sketched in [DKV, Appendice 1, 4.d] and (in more detail) in [HT, Lemma VI.3.2].

Applying (7.4.1) with this choice of $\Phi_w$, for $\phi_{n-h} = \phi_{\pi',w}$ a test function acceptable for $\rho'$, we find
\[
(7.4.3) \quad (n-h) \text{Tr}(\Phi | [H(A(B,*)] = n(n-h)k_B \sum e(\gamma) |F(\gamma) : F|^{-1} v(\gamma)^{-1} \sum s O_{s(\gamma)}^{G(A_f^\psi)}(\phi_w) O_{s(\gamma)}^{L_{n-h,h}}(\phi_{n-h} \otimes \phi_h)
\]

Now note that there is a one-to-one correspondence between $\tilde{\omega}$ as in (7.3.11) and $s$ as in (7.4.2): each $s$ defines the subfield $F_w(\gamma) \subset M_{n-h}(K)$ – a subfield because of the ellipticity condition – hence a completion $F(\gamma)_{\tilde{\omega}}$ above $w$ that satisfies the degree condition.

Formula (7.4.3) does not require $\pi'$ to be supercuspidal. If we now return to Hypothesis 7.3.2 – in particular, $\pi'$ is supercuspidal – we can combine (7.4.3) with (7.3.11), and obtain
\[
(7.4.4) \quad (n-h) \text{Tr}(\Phi | [H(A(B,*)]) = n \cdot \text{Tr}(\phi | [H_{c}(\tilde{S}^{\psi}_{M_0}, R\Psi)^U [\rho]]).
\]

The absence of $[\rho]$ on the left-hand side should cause no alarm; $\phi_{n-h,w}$ has been chosen in (7.3.2) to cut out only the part of $[H(A(B,*)])$ coming from $[\rho]$. Indeed, if
\[
[H(A(B,*))] = \sum a(\tau_f) \tau_f \psi \otimes \tau_w \in \text{Groth}(G(A_f))
\]
where \( \tau \) runs through a set of cohomological automorphic representations, then the trace relation (7.4.2.3) implies that

\[
\text{Tr}(\Phi | [H(A(B, *))]) = \sum_{\tau} \text{Tr}(\phi^w | \tau^w) \text{Tr}(\Phi_w | \tau_w).
\]

Here the symbol \( \sum_{\tau'} \) indicates that the sum is taken over those \( \tau \) such that

\[
[r_{G_n, L_h}^{op}(\tau_w)] = \sum m_{\alpha, \beta} [\alpha \otimes \beta] \text{ such that, for some } \alpha \text{ that occurs,}
\]

(7.4.5)

\[
\text{Tr}(\alpha)(\phi_{\pi'; \omega}) \neq 0;
\]

in other words, such that \( \alpha \) is inertially equivalent to \( \pi' \). But since we are working at finite level, the set of all \( \alpha \)'s that arise in this way is finite. Hence, by expanding \( A(J_{n-h})_{\text{fin}} \) if necessary, we can arrange that (7.4.5) only holds for \( \alpha = \pi' \), and in that case, as we know, \( \text{Tr}(\pi'')(\phi_{\pi', \omega}) = 1 \). It then easily follows that

(7.4.6) \[
\text{Tr}(\Phi | [H(A(B, *))]) = \sum_{\rho' \in A(J_{n-h})_{\text{fin}}[\rho]} \text{Tr}(\phi^w \otimes \phi_w | JL(\rho') \otimes \text{red}_{\rho}^{(h)} H(A(B, *)))
\]

Indeed, only \( JL(\rho') = \pi' \), for our chosen \( \pi' \), gives a non-zero contribution to the right-hand side of (7.4.6). By varying \( \rho' \), we see that the identity (7.4.6) is valid for every \( \rho' \in A_{n-h}^0 \). Now the Second Basic Identity, or rather the identity of traces (7.3.1) for test functions satisfying (7.3.2), follows by combining (7.4.6) and (7.4.4).

To obtain the pre-(4.3.11) version, we let \( \psi_{n-h} \) be arbitrary subject to the \( h \)-regularity condition. As noted, Proposition 7.4.2 holds without the ellipticity hypothesis, and we let \( \Phi_w \) be the function constructed there. On the other hand, let \( \Phi_w(\pi', \omega) \) be the function of Proposition 7.4.2 associated to \( \phi_{\pi', \omega} \otimes \phi_h \). Then

\[
\text{Tr}(\phi^w \otimes \phi_w | \text{red}_{\rho'}^{(h)}[H(A(B, *))] \otimes [\Psi_{n-h}(\rho')])
\]

\[
= \text{Tr}(\phi_{n-h} | [\Psi_{n-h}(\rho')]) \cdot \text{Tr}(\phi^w \otimes \phi_h | \text{red}_{\rho'}^{(h)}[H(A(B, *))])
\]

\[
= \text{Tr}(\phi_{n-h} | [\Psi_{n-h}(\rho')]) \cdot \text{Tr}(\phi^w \otimes \phi_{\pi', \omega} \otimes \phi_h | r_{G_n, L_h}^{op}(H(A(B, *)))
\]

\[
= \text{Tr}(\phi_{n-h} | [\Psi_{n-h}(\rho')]) \cdot \text{Tr}(\Phi_w(\pi', \omega) \otimes \phi^w | [H(A(B, *))])
\]

Returning to (7.4.3), we thus have

(7.4.3 (pre 4.3.11)) \[
\sum_{\rho' \in A(J_{n-h})_{\text{fin}}[\rho]} \text{Tr}(\phi^w \otimes \phi_w | \text{red}_{\rho'}^{(h)}[H(A(B, *))] \otimes [\Psi_{n-h}(\rho')])
\]

\[
= n k_B \sum_{\gamma} e(\gamma)[F(\gamma) : F]^{-1} u(\gamma)^{-1} \sum_{s} O_{s(\gamma)}^{(A_{s(\gamma)}^{(\gamma)})}(\phi^w)O_{s(\gamma)}^{L_{n-h, h}(\phi_{\pi', \omega} \otimes \phi_h)} \cdot \text{Tr}(\phi_{n-h} | [\Psi_{n-h}(\rho')])
\]

To prove the Second Basic Identity under Hypothesis 7.3.2, once Theorem 4.3.11 has been established, it thus remains to justify Theorem 7.3.5. For the general case, we need to show that the analogue of Theorem 7.3.5 holds, for any test function \( \phi \).
with the term \((n-h)O_{[\gamma(a)]}^{G_{\ast-h}}(\phi_{\pi_{\ast}}, \omega)\) replaced by
\[
(7.4.7) \sum_{\rho' \in A(J_{n-h})} O_{[\gamma(a)]}^{G_{\ast-h}}(\phi_{\pi_{\ast}}, \omega) \cdot \text{Tr}(\phi_{n-h} | [\Psi_{n-h}(\rho')]).
\]
For a general function \(\phi_{n-h}\) one also has to sum over all \([\rho]\), as in Corollary 7.2.3.

Calculation of the fixed point contribution is the subject of the two remaining sections.

7.5. Fixed point formalism in double coset spaces. — We consider the following abstract situation. We have three totally disconnected groups \(Y, G, J,\) and a discrete group \(I\) that embeds (discretely) in \(Y \times J\). There is also a discrete abelian group \(\Delta,\) and surjective maps \(\delta_G : G \to \Delta, \delta_J : J \to \Delta;\) the composite \(I \to J \to \Delta\) is surjective.

The group \(J\) is assumed to act continuously on a locally noetherian scheme \(M\) over the field \(\mathbb{F}_p,\) compatibly with a surjective map \(\delta_M : M \to \Delta.\) We assume \(M\) is given with a \(J\)-equivariant (open or closed) locally finite covering, and let \(M_i, i = 1, \ldots, N,\) denote the disjoint union of the \(i + 1\)-fold intersections of this covering; the restriction of \(\delta_M\) to each \(M_i\) is also surjective. We assume the set of connected components of \(M_i/J\) is finite for all \(i\) (equivalently, for \(i = 0),\) that the stabilizer \(J_\alpha\) of any connected component \(M_\alpha\) of any \(M_i\) is an open compact subgroup of \(J,\) and that the action of \(J_\alpha\) on \(M_\alpha\) factors through a finite quotient. We also assume there is an action of \(G\) on \(M\) that factors through \(\delta_G,\) compatible with \(\delta_M.\) It follows that the stabilizer in \(G\) of any \(M_\alpha\) is exactly \(G(0) = \ker \delta_G.\)

Finally, we assume the action of \(G \times J\) on \(M\) lifts to a \(G \times J\)-equivariant constructible-admissible \(\ell\)-adic complex \(\Psi^*\) on \(M.\) This means that, for any open compact subgroup \(\mathcal{U} \subset G(0),\) the sheaf \(\mathcal{H}^j(\Psi^*)\) is a constructible \(J\)-equivariant sheaf on \(M.\) Then the action of \(J_\alpha\) on \(\mathbb{H}^\ast(M_\alpha, \Psi^*)\) factors through a finite quotient.

To simplify notation, and because this is the only case we need, we assume \(M = M_1,\) a principal homogeneous space for \(\Delta,\) with fixed component \(M_\alpha\) denoted \(M_0;\) we write \(J(0)\) for \(J_\alpha.\) In our applications, \(J\) is the compact mod center group \(D^\times_{(n-h)} J(0)\) its unique maximal compact subgroup, \(M\) is zero-dimensional, and \(J/J(0) \cong \mathbb{Z}\) acts transitively on \(M.\) However, the arguments presented below can be applied simplicially to the \(\check{C}\)ech complex of the more general \(M\) discussed above. Similarly, we replace the complex \(\Psi^*\) by one of the cohomology sheaves \(\mathcal{H}^j(\Psi^*),\) which we denote simply \(\Psi,\) or by the alternating sum \([\Psi] = \sum_j (-1)^j [\mathcal{H}^j(\Psi^*)],\) a virtual representation of \(G \times J.\) Additional properties of \(\Psi,\) satisfied in the applications, will be specified below.

For any open compact subgroup \(U \subset Y,\) let \(S_U = [M \times (Y \times J(0))]/J,\) with the profinite topology. The group \(G\) acts on \(S_U\) via the action on \(M.\) For simplicity, we write \(\delta\) instead of \(\delta_G, \delta_J.\) Let \(y \in Y, f\) the characteristic function of the double coset \(UYU.\) Let \(\phi \in C_c^\infty(G).\)
First, assume $\phi$ supported on $G(d) = \delta^{-1}(d)$ for some fixed $d \in \Delta$. The pair $(f, \phi)$ defines a Hecke correspondence on $S_U \times S_U$: it is the set of pairs of classes of points modulo the groups acting on the right and left. Note that $x$ is determined modulo $U(y) := U \cap yUy^{-1}$, so the correspondence is in bijection with the set of points $s \in S_{U(y)}$. We may as well take $j = 1$. A fixed point of the correspondence is a class $[\delta, x, 1]$ such that $[d \cdot \delta, xy, j] = [\delta, x, 1]$; i.e., such that there are $u \in U$, $a \in I$, and $j \in J$ such that

$$(d \cdot \delta, xy, 1) = (\delta(j)\delta, axu, aj).$$

Thus

$$(7.5.2) \quad a = j^{-1} = x(yu^{-1})x^{-1}, \quad \delta(a)d = 1.$$ 

Assume $U$ is sufficiently small, in a sense to be determined momentarily; then the first condition in (7.5.2) determines $a$ uniquely. Indeed, if $\beta$ is another element of $I$ satisfying the same condition, then $x^{-1}\beta x$ and $x^{-1}ax$ are both in $y \cdot U$, so

$$\beta^{-1}a \in x^{-1}yUy^{-1}x \cap I = \{1\},$$

where the last equality is what we mean by “sufficiently small”; a standard argument shows that any open compact $U$ contains a subgroup of finite index that is sufficiently small in this sense. On the other hand, we can replace $[\delta, x, 1]$ by $[\delta(j)\delta, \beta xu, \beta j]$ for some $v \in U \cap yUy^{-1}$, $\beta \in I$, $j = \beta^{-1} \in J$. Then $a$ is replaced by $\beta a \beta^{-1}$. So the conjugacy class $[a]$ of $a$ in $I$ is a well-defined invariant of the fixed point $s$, and we denote this invariant $[a(s)]$.

Now given $a \in I$, let $\text{Fix}(f \otimes \phi, a)$ denote the set of fixed points $s$ with $[a(s)] = [a]$. If $\delta(a)d \neq 1$, the second condition of (7.5.2) shows that $\text{Fix}(f \otimes \phi, a)$ is empty. If $\delta(a)d = 1$, one checks easily that

$$(7.5.3) \quad |\text{Fix}(f \otimes \phi, a)| = |M/(Z_I(a) \cap U(y)) \times X_I(g, a)/U(y)|,$$

where $X_I(g, a) = \{x \in Y \mid x^{-1}ax \in yU\}$ and $U(y) = U \cap yUy^{-1}$.

Remark. — Suppose we are in the setting of §86, (7.3); i.e., $Y = G(A_f^u) \times G_h$, $U = U_h^w$, $G = G_{n-h}$, $J = J_{n-h}$, $\Delta = \mathbb{Z} = J/J(0)$ acts simply transitively on the set $M = M_{n-h,+}$. The set on the right-hand side of (7.5.3) is then the same as

$$(7.5.4) \quad Z_I(a) \backslash \{x \in Y \times J \mid x^{-1}ax \in yU \times J(d)\}/U(y) \times J(0),$$

where $J(d) = \delta^{-1}(d)$ for $d = \delta(g)$ as above, and $J(0)$ acts on $J(d)$ by right translation. Note that the condition that $x^{-1}ax \in J(d)$ is equivalent to the condition $\delta(a) = d$, and imposes no restriction on $x$.

Henceforward, we assume that we are in the situation (7.5.4), i.e., in the situation of the Second Basic Identity. In particular, notation is as in (7.5.4).

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
Now, since $U$ is small, we find that any double coset $Z_I(a) \cdot x \cdot (U(y) \times J(0))$ is the disjoint union over $b \in Z_I(a)$ of $b \cdot x \cdot (U(y) \times J(0))$. It follows that the cardinality in (7.5.4) equals (7.5.5) \[ \text{vol}(U(y)J(0))^{-1} \cdot \text{vol}(\{ x \in Z_I(a) \setminus (Y \times J) \mid x^{-1}ax \in yU \times J(0) \}). \]

The Haar measures are arbitrary but have to be used consistently, and of course the discrete groups are given the counting measure. This is where the orbital integrals arise: the cardinality in (7.5.5) equals (7.5.6) \[ \text{vol}(U(y)J(0))^{-1} \cdot \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O^{Y \times J}_{[a]}(\chi_{y,U} \cdot \chi_d) \]
where $\chi_d$ is the characteristic function of $J(0)$ and $\chi_{y,U}$ is the characteristic function of $y \cdot U$. This is obviously non-canonical, since it depends on the choice of $y$. One makes it canonical by summing over representatives of $UyU/U$ and dividing by $|(UyU)/U|$, and we obtain finally that (7.5.7) Proposition. — Under the hypothesis that $\phi$ is supported on $G(d)$ and $f$ is the characteristic function of $UyU$, the number of fixed points $s \in S_U$ with $[a(s)] = [a]$ equals \[ [\text{vol}(U) \text{vol}(J(0))]^{-1} \cdot \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O^{Y \times J}_{[a]}(f \cdot \chi_d). \]

More generally, let $\omega = [a, b] \subset \mathbb{Z}$ be an interval as in (A.1.10), and assume $\phi$ has support in $G(\omega) = \delta^{-1}(\omega)$, and $f \in C^{\infty}_c(Y)$. Then the number of fixed points $s$ with $[a(s)] = [a]$ equals \[ [\text{vol}(U) \text{vol}(J(0))]^{-1} \cdot \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O^{Y}_{[a]}(f)O^{J}_{[a]}(\chi_\omega). \]

Here $\chi_\omega$ is the characteristic function of $J(\omega) = \delta^{-1}_J(\omega)$.

The formula in the final paragraph follows by linearity.

For fixed $a$ and $\omega$, the orbital integral $O^{J}_{[a]}(\chi_\omega)$ is given as follows: (7.5.8) \[ O^{J}_{[a]}(\chi_\omega) = \text{vol}(J/Z_J(a)), \delta(a) \in \omega; \quad O^{J}_{[a]}(\chi_\omega) = 0, \delta(a) \notin \omega. \]

The measure on $Z_J = K^\times$ is normalized by (A.1.8), and one sees readily that (7.5.9) \[ \text{vol}(J(0))^{-1} \text{vol}(J/Z_J(a)) = [J : Z_J \cdot J(0)] \cdot \text{vol}(Z_J(a)/Z_J)^{-1} = g \cdot \text{vol}(Z_J(a)/Z_J)^{-1}. \]

Thus the cardinality in (7.5.7) can be rewritten: (7.5.10) \[ |\text{Fix}(f \otimes \phi, a)| = g \cdot \text{vol}(U)^{-1} \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O^{Y}_{[a]}(f) \cdot \text{vol}(Z_J(a)/Z_J)^{-1}. \]

Now the $G \times J$-equivariant constructible $\ell$-adic sheaf $\Psi$ on $M$ descends to a constructible $\ell$-adic sheaf, still denoted $\Phi$, on $S_U$. The function $f \otimes \phi$ acts as a Hecke operator on $\Psi$ over $S_U$. The normalization of $f \otimes \phi$ as Hecke operator is given by
integrating over $Y \times G$; one verifies easily that this amounts to multiplying the Hecke correspondence defined above by $\text{vol}(U)$. Let

\[ [H(S_U, \Psi)] = \sum_{i,j} (-1)^{i+j} H^{i}(S_U, \mathcal{H}^{j}(\Psi^*)) \]

In the application to (7.3.5), and indeed under our assumption that $\Delta$ acts transitively on $M$, only $i = 0$ contributes to the above sum. Assuming both sides are finite, the Lefschetz fixed point formula yields the following formula for the trace of $f \otimes \phi$, acting on the cohomology of $\Psi$:

\begin{equation}
(7.5.11) \quad \text{Tr}(f \otimes \phi \mid [H(S_U, \Psi)]) = \text{vol}(U) \sum_{[a] \in I(\mathbb{Q})} \sum_{s \in \text{Fix}(f \otimes \phi, a)} \text{Loc}_s(f \otimes \phi, [\Psi]).
\end{equation}

Here as above, the sum is over conjugacy classes $[a]$ in $I(\mathbb{Q})$.

This is the framework in which we have stated Fujiwara’s trace formula (Theorem 7.2.1). Here $\text{Loc}_s(f \otimes \phi, [\Psi])$ is a local term that is in general quite complicated. In the situation discussed in the lectures, however, a non-trivial local term is just the alternating sum of local traces at an isolated fixed point of a (transversal) correspondence on a smooth variety, hence is just given by the trace of $f \otimes \phi$ acting on the (virtual) stalk of $[\Psi]$ at $s$. One checks that this is independent of $s \in \text{Fix}(f \otimes \phi, a)$, and indeed is independent of $f$ (since $Y$ acts trivially on $[\Psi]$). For fixed $a$, the local term is given by

\begin{equation}
(7.5.12) \quad \text{Loc}_s(f \otimes \phi, [\Psi]) = \text{trace}(\phi \otimes a \mid [\Psi]_0)
\end{equation}

where $[\Psi]_0$, the stalk of $[\Psi]$ at $M_0$, is a virtual representation space for $T_0 = (\delta_G \times \delta_J)^{-1}\{0\} \subset G \times J$. Note however that $\phi$ is acting via an integral, hence the trace depends on a measure on $G$, whereas $a$ is acting as an element of a translate of the compact open subgroup $J(0)$. In this sense, the expression (7.5.12) is not symmetric in the two variables.

Combining (7.5.12) with (7.5.9) and (7.5.10), we obtain (when both sides are finite)

\begin{equation}
(7.5.13) \quad \text{Tr}(f \otimes \phi \mid [H(S_U, \Psi)]) = g \cdot \sum_{[a] \in I(\mathbb{Q})} \text{vol}(Z_I(a) \setminus Z_Y \cdot J(a)) \cdot O_{\mathbb{Q}}(f) \cdot \text{vol}(Z_J(a)/Z_J)^{-1} \cdot \text{trace}(\phi \otimes a \mid [\Psi]_0).
\end{equation}

**Remark.** — Nowhere in the present section have we made use of Hypothesis 7.3.2 or Theorem 4.3.11. In particular, the formula (7.5.13) holds unconditionally.

7.6. Completion of the calculation. — Now we specialize to the situation of (7.3), with $g = n - h$, taking $\phi = \phi_{\pi''}$, $\pi = \Phi_{\pi''}$, as in (A.1), and taking the alternating sum $[R\Psi][\rho]$ for $[\Psi]$. We continue to write $I$ for $I_z$ and drop the subscript $z$ elsewhere. Here $[\rho]$ is an inertial equivalence class in $\mathcal{A}(J_{n-h})$ such that $JL(\rho)$ is supercuspidal, and $\pi'$ is inertially equivalent to $JL(\rho)$. Once we have established (6.1.2.1) – hence
Theorem 4.3.11 – (5.1.6) implies we can replace $[\Psi][\rho]$ by $(-1)^{n-h-1}\Psi^{n-h-1}[\rho]$, in the notation of (5.5). Thus it follows from (A.1.12) and (5.5.9) that

\[(n-h)\cdot \text{trace}(\phi \otimes a \mid \Psi^{n-h-1}[\rho]_0) = (-1)^{n-h-1}\text{trace}_{Z_{\varphi,\xi}}(\phi_{\varphi} \otimes a \mid \Psi^{n-h-1}[\rho]_\xi)\]

for any appropriate central character $\xi$. Here the extension of the compactly supported function $\phi\otimes a$ to the function $\phi_{\xi} \otimes a$, compactly supported modulo $Z_0$, is as in (5.5.9.1). Indeed, if $\xi \neq \psi_{w}$, then

\[
\text{trace}(\phi \otimes a \mid \Psi[\rho]_\xi) = 0
\]

because $\phi$ is a pseudocoefficient relative to $\mathcal{A}(n-h,K)_{\text{fin}}$. On the other hand, if $\xi = \psi_{w}$, then the formula above holds (cf. (5.5.9.1)) and the right hand side can be simplified:

\[(n-h)/m \text{ arises as follows. The denominator comes from the normalization (A.1.11), and arises from the distinction between the modified trace of (A.1.9) and the unmodified trace; replacing $\phi_{\psi;\omega}$ by $(\phi_{\psi;\omega})_\xi$ amounts to undoing the truncation without compensating for the denominator. On the other hand, the numerator $(n-h)$ is the coefficient on the right-hand side of the formula}

\[(\Psi^{n-h-1}[\rho]_\xi = (n-h) \bigoplus_{\rho' \in [\rho]_\xi} \rho' \otimes \text{JL}(\rho'),\]

\[\text{as representation of } J_{n-h} \times G_{n-h}; \text{this is just Theorem 4.3.11 with the action of } W_K \text{ forgotten.}

Comparing (7.6.1) and (7.6.2), the specialization of (7.5.13) to the situation of Theorem 7.3.5 becomes

\[\text{Tr}(f \otimes \phi \mid H^0([M_{n-h-\varnothing} \times S_{\psi,\omega}^h(z)]/J_{n-h-\varnothing}, [R\Psi]_\xi[\rho]) = \text{Tr}(f \otimes \phi \mid [H(S_{n-h-\varnothing}, [\Psi])]) =
\]

\[(-1)^{n-h-1}(n-h) \sum_{[a] \in I(\mathbb{Q})} \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O_{[a]}(f) \cdot \text{vol}(Z_J(a)/Z_J)^{-1} \cdot R_{(\psi;\omega)}(\gamma)(a)
\]

\[= (n-h) \sum_{[a] \in I(\mathbb{Q})} e(\gamma(a)) \text{vol}(Z_I(a) \setminus Z_{Y \times J}(a)) \cdot O_{[a]}(f) \cdot O_{\gamma(a)}(\phi)
\]

where $[\gamma] \in G_{n-h}$ transfers to $[a] \in J$ and $e(\gamma)$ is the Kottwitz sign. The last equality follows from Proposition (A.1.12 bis); as indicated above, the truncation is no longer pertinent.

Recalling our notation, we rewrite the last expression in (7.6.4):

\[(n-h) \sum_{[a] \in I_h(\mathbb{Q})} e(\gamma(a))O_{[a]}(\phi_{\psi;\omega}) \cdot \text{vol}(Z_{I_h}(a)(\mathbb{Q}) \setminus Z_{I}(a)(\mathbb{Q}))
\]

By the choice of $\phi_{\psi;\omega}$ the sum runs over $h$-regular conjugacy classes $[a]$.
This completes the proof of Theorem 7.3.5, assuming Theorem 4.3.11, i.e. (6.1.2.1). To complete the proof of the Second Basic Identity, we need to eliminate this assumption and relax hypothesis (7.3.2). The calculation in §7.5 is valid without these assumptions, the only change coming in the determination of the local term \( \text{trace}(\phi \otimes a \mid [\Psi]_0) \).

In (7.5.13) we take \( [R\Psi]_\xi [\rho] \) for \( [\Psi]_0 \). It suffices to show that, for general \( \phi = \phi_{n-h} \), assumed to have zero trace on any \( \pi \in \mathcal{A}(n-h, K)_\text{fin} \) with \( \psi_\pi \neq \xi \), we have

\[
(n-h) \text{trace}(\phi \otimes a \mid [\Psi]_0) = \sum_{[\rho]} \sum_{\rho' \in \mathcal{A}(J_{n-h})_{\text{fin}},[\rho]} O^{G_{n-h}}(\phi_{\pi',\omega}) \cdot \text{Tr}(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]),
\]

where the right-hand side is the expression appearing in (7.4.7). But the special hypotheses have only been used in (7.6.3). In the general case we have

\[
(7.6.3 \text{ (pre 4.3.11)}) \quad [\Psi][\rho]_\xi = \bigoplus_{\rho' \in [\rho]_\xi} \rho'^{\chi_J} \otimes [\Psi_{n-h}(\rho')].
\]

Using Proposition (A.1.12 bis), we now find that

\[
(7.6.2 \text{ (pre 4.3.11)}) \quad \text{trace}(\phi_{n-h} \otimes a \mid [\Psi]^U[\rho]_\xi) = \frac{1}{m} \sum_{\rho' \in [\rho]_\xi} \chi_{\text{JL}(\pi',\nu)}(a) \text{Tr}((\phi_{n-h} \mid [\Psi_{n-h}(\rho')]).
\]

We conclude as above. This completes the proof of the Second Basic Identity.

8. Strata in Shimura varieties of PEL type

This final section, which does not correspond to any of the lectures given during the special semester at the IHP, describes possible extensions to general Shimura varieties of the geometric techniques presented in the previous lectures. The first two subsections elaborate on material contained in [H3], and prove some of the claims made there. The final subsection explains recent results of L. Fargues, who has proved a number of the results predicted in [H3] for Shimura varieties of PEL type.

The reader is expected to be familiar with the basic properties of Shimura varieties over number fields (existence of canonical models and the like). A good general reference for Shimura varieties is the article [Mi1].

8.1. Presentation of the problem. — As in (3.1), we denote by \( K \) the fraction field of the Witt vectors of the algebraic closure of \( \mathbb{F}_p \), and let \( \sigma \) denote the Frobenius acting on \( K \). If \( G \) is a reductive group over \( \mathbb{Q}_p \), let \( B(G) \) denote the set of \( \sigma \)-conjugacy classes in \( G(K) \), i.e., equivalence classes for the relation

\[
b \sim h \cdot b \cdot \sigma(h)^{-1}, \quad h \in G(K)
\]
For any \( Q_p \)-rational representation \((\tau, V)\) of \( G \), an element \( b \in B(G) \) defines a structure of isocrystal on \( N_\tau = V \otimes_{Q_p} K \) by defining
\[
\phi = \tau(b) \otimes \sigma : V \otimes_{Q_p} K \rightarrow V \otimes_{Q_p} K.
\]
If \( G = GL(V) \), then any isocrystal with underlying vector space \( V \otimes_{Q_p} K \) arises this way; \( b \) is the matrix of \( \phi \) with respect to some basis of \( V \otimes_{Q_p} K \), and the \( \sigma \)-linearity of \( \phi \) implies that changing the basis replaces \( b \) by a \( \sigma \)-conjugate matrix. For general \( G \) and \( \tau \), the isocrystal \( N_\tau \) has “additional structure” in the sense that invariants of \( G \) in tensor powers of \( V \) give rise to \( \phi \)-fixed vectors (“crystalline Tate classes”) in the corresponding tensor powers of \( N_\tau \). When \( G \) is the similitude group of a non-degenerate bilinear form on \( V \), then \((N_\tau, \phi)\) has a polarization of the corresponding type in the category of isocrystals; when \( V \) is a \( C \)-module for some \( Q_p \)-algebra \( C \), and \( G \subset GL_C(V) \), then one obtains a map \( C \rightarrow \text{End}(N_\tau, \phi) \). Combining these two kinds of structure, one obtains the sort of isocrystals arising from the Dieudonné modules of abelian varieties of PEL type. The moduli spaces of such abelian varieties are Shimura varieties. The present lecture describes the stratification of the special fibers of such Shimura varieties at primes of \( E(G, X) \) dividing \( p \), and the conjectural stratification of the (conjectural) special fiber of general Shimura varieties, in terms of isocrystals.

We briefly recall the formalism of Shimura varieties. Suppose \( G \) is a reductive group over \( \mathbb{Q} \). Let \( X \) be a \( G(\mathbb{R}) \)-conjugacy class of homomorphisms \( h : R_{C/\mathbb{R}}^* G_m \rightarrow G_{\mathbb{R}} \) so that the pair \((G, X)\) satisfies the axioms defining a Shimura variety. Thus \( X \) is naturally a finite union of isomorphic hermitian symmetric spaces, and for every open compact subgroup \( K \in G(\mathbb{A}_f) \), where \( \mathbb{A}_f \) defines the ring of finite adeles,
\[
K \text{Sh}(G, X)(\mathbb{C}) = G(\mathbb{Q}) \backslash X \times G(\mathbb{A}_f)/K.
\]
is the set of complex points of a quasi-projective algebraic variety, with canonical model over a certain number field \( E = E(G, X) \) (the reflex field). The reflex field does not depend on \( K \), and the natural continuous action of \( G(\mathbb{A}_f) \) on
\[
\text{Sh}(G, X)(\mathbb{C}) = \lim_K K \text{Sh}(G, X)(\mathbb{C})
\]
is rational over \( E \). In particular, for any irreducible admissible representation \( \pi \) of \( G(\mathbb{A}_f) \), the \( \pi \)-isotypic component of the \( \text{Hom}_{G(A_f)}(\pi, H^i(\text{Sh}(G, X), \mathbb{Q}_l)) \) (étale cohomology) is naturally a representation space \( H^i[\pi] \) for \( \text{Gal}(\overline{E}/E) \), easily seen to be finite-dimensional.

For any point \( h \in X \), we let \( \mu_h : G_{m, \mathbb{C}} \rightarrow G_{\mathbb{C}} \) denote the (complex) cocharacter associated to \( h \): identifying the complexification of \( R_{C/\mathbb{R}} G_m \) with \( \mathbb{C}^* \times \mathbb{C}^* \), we have
\[
\mu_h(z) = h_{\mathbb{C}}(z, 1).
\]
The conjugacy class of \( \mu_h \) depends only on \( X \), and its field of definition is precisely \( E(G, X) \). We may regard \( \mu_h \), or simply \( \mu \), as a character of a maximal torus of the
complex dual group $\widehat{G}$ of $G$, hence as an extreme weight, necessarily minuscule, of a certain irreducible representation of $\widehat{G}$. Let $r_\mu$ denote the representation of the $L$-group $L^G_{\nu}$ (relative to $E(G, X)$) constructed by Langlands [La]; its restriction to $\widehat{G}$ is just the minuscule representations with extreme weight $\mu$. In [La] Langlands expressed the expectation that most of the middle-dimensional $\ell$-adic cohomology of $\text{Sh}(G, X)$ would break up as a sum in $\text{Groth}(G(A_f) \times \text{Gal}(\overline{E}/E))$:

$$H^{\dim X}(\text{Sh}(G, X), \mathbb{Q}_\ell) = \oplus \pi_f \otimes r_\ell(\pi_f) \oplus \text{endoscopic contributions} \quad (8.1.2)$$

where the sum on the right is taken over those admissible irreducible representations of $G(A_f)$ occurring in stable cohomological $L$-packets (the meaning of “most” above) and $r_\ell(\pi_f)$ is a $\mathbb{Q}_\ell$-valued representation of $\text{Gal}(\overline{E}/E)$) of dimension $\dim r_\mu$. Moreover, at a place $v$ where $\pi_f$ is unramified, the local component $\pi_v$ of $\pi_f$ is classified, via the Satake isomorphism, by a semi-simple conjugacy class $s(\pi_v) \in L^G(\mathbb{Q}_v)$, and up to conjugacy, geometric Frobenius is given by the formula

$$r_\ell(\pi_f)(\text{Frob}_v) = r_\mu(s(\pi_v)). \quad (8.1.3)$$

For the Shimura varieties considered in the present article, and for those attached to twisted unitary groups with general signatures, this identity is established for almost all unramified places, up to multiplicities, by Kottwitz in [K4]. The article [K5] also contains results on general PEL-type Shimura varieties that strongly support the predictions of [La].

Assuming one has a $\pi_f$ that contributes to the non-endoscopic part of the right-hand side of (8.1.2), how can (8.1.3) be extended to ramified places? Naturally, one assumes the Satake parameter will be replaced more generally by a parameter given by the (in general still conjectural) local Langlands correspondence for $G$, but this begs the question of how ramified local representations arise in the cohomology of $\text{Sh}(G, X)$. If $v$ is a place of $E$ dividing a rational prime $p$ at which the group $G$ is unramified (briefly: $v$ is an unramified place for $\text{Sh}(G, X)$), and if $K_p$ is a hyperspecial maximal compact subgroup of $G(\mathbb{Q}_p)$, then for sufficiently small compact open subgroups $K' \subset G(A_f^p)$, one expects $K_p' \backslash \text{Sh}(G, X)$ to have good reduction at $v$ (cf. [K5] for the PEL case). Let $K_p \mathcal{S}$ denote the special fiber. Guided by our experience with the Shimura varieties treated in [HT], one would then expect $K_p \mathcal{S}$ to have a stratification in terms of isocrystals. Moreover, assuming the $K_p' \backslash \text{Sh}(G, X)$ have reasonable integral models for open subgroups $K_p' \subset K_p$, one would expect the stratification to lift to the corresponding special fibers $K_p' \mathcal{S}$ in such a way that the vanishing cycles are well-behaved along the strata. This latter hope is certainly too optimistic – no one knows how to generalize the theory of Drinfel’d level structures – but it is reasonable to assume that different kinds of ramified contributions to $r_\ell(\pi_f)$ correspond to the different strata, just as one saw in (6.1) that the $n-h$-dimensional irreducible representations of the local Galois group arise from the stratum $\mathcal{S}^{(h)}$. 

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
What can we mean by “different kinds” of ramification? We need a concept playing the role for a general $G$ that $n - h$-dimensional irreducible representations play for $GL(n)$, as $h$ varies from 0 to $n - 1$. In the preceding lectures, the $n - h$-dimensional irreducible representation was attached to a supercuspidal representation of the factor $GL(n-h)$ of a Levi subgroup of the maximal standard parabolic of $GL(n)$ of partition type $(n-h, h)$. Closer examination reveals that the same $n-h$-dimensional irreducible Galois representation occurs for irreducible admissible representations of $GL(n)$ induced from standard parabolics of partition type $(n_1, \ldots, n_r)$, where at least one of the $n_j$ equals $n - h$. The following section describes a relation between stratifications – in most cases conjectural – of general Shimura varieties, and irreducible components of restrictions of the Langlands representation $r_\mu$ to Levi factors of parabolic subgroups of $^LG$. This relation serves in [H3] to motivate conjectures on the cohomology of Rapoport-Zink $p$-adic period domains, and their relation to the cohomology of Shimura varieties. Partial results in this direction, due to L. Fargues, are described in the final section.

8.2. Classification of strata. — For the moment, we set aside the global arithmetic motivation and concentrate on the formal properties of isocrystals with additional structure, as analyzed by Kottwitz in [K6, K7]. Let $F$ be a finite extension of $\mathbb{Q}_p$, and let $\Gamma = \text{Gal}(\overline{F}/F)$. Let $G$ be a quasi-split reductive group over $F$, and fix an $F$-rational Borel subgroup $B_0$ of $G$, with Levi factor $T_0$ and unipotent radical $N_0$; this determines an order on the root lattice of $G$ and, dually, on that of the complex dual group $\overline{G}$. Let $A \subset T_0$ be a maximal $F$-split torus, with cocharacter group $X_*(A)$, and let $\Phi_0 \subset \text{Hom}(X_*(A), \mathbb{Z})$ denote the set of roots of $A$ in $N_0$. Define $\mathfrak{A}$, $\mathfrak{C}_\mathbb{Q} \subset \mathfrak{A}$, as in [K7, pp. 267-268]:

$$\mathfrak{A}_\mathbb{Q} = X_*(A) \otimes \mathbb{Q}; \quad \mathfrak{A} = X_*(A) \otimes \mathbb{Z};$$

$$\mathfrak{C}_\mathbb{Q} = \{ x \in \mathfrak{A}_\mathbb{Q} \mid \langle \alpha, x \rangle \geq 0 \forall \alpha \in \Phi_0 \}.$$

Let

$$\varpi : B(G) \rightarrow \mathfrak{C}_\mathbb{Q}$$

be the Newton map, defined as in [RR] and [K6, loc. cit.]. When $G = GL(n)$, $\varpi$ is the map that associates to an isocrystal its set of slopes with multiplicities, ordered in accordance with the choice of $P_0$; for general $G$, one can obtain $\varpi$ by embedding $G$ faithfully in an appropriate $GL(n)$ and using Tannakian arguments.

If $P$ is a standard parabolic subgroup, let $AP \subset A$ be a split component, and define $\mathfrak{A}_P = X_*(AP) \otimes \mathbb{Z}$ and $\mathfrak{A}_{P, \mathbb{Q}}$ as above. Then $\mathfrak{A}_P$ is naturally a subset of $\mathfrak{A}$, and indeed the chamber $\mathfrak{C}_\mathbb{Q}$ is a disjoint union over standard parabolics of the corresponding walls $\mathfrak{A}_P^\pm$ (see [K7, 5.1] for this notation; we omit the subscript $\mathbb{Q}$ for the walls). Let $\mathfrak{A}_P^+ \supset \mathfrak{A}_P^-$ denote the corresponding closed chamber.

Following Kottwitz [K7, §6], we let $B(G, \mu) = B(G_F, \mu)$ be the set of $\delta \in B(G)$ satisfying the following condition:
(8.2.1) Under the natural map
\[ B(G) \rightarrow X^*(Z(G)\Gamma) = H_0(\Gamma, \pi_1(G)) \]
(see [K6, §3] for the first version of the maps, [Mi2, Prop. B.27] for the second) the image of \( \delta \) is the negative of the class of \(-\mu_x\) (see [Mi2, 6.1.4] for an explanation), and such that
\[ (8.2.2) \quad \overline{\nu}(\delta) \leq \mu_{\mathfrak{A}}. \]
Here \( \mu_{\mathfrak{A}} \in \overline{\mathfrak{A}}_Q \) is what Kottwitz denotes \( \mu_2 \), and the order \( \leq \) is the usual lexicographic order. The \( \delta \in B(G, \mu) \) are precisely those such that, up to replacing \( \mu \) in its conjugacy class, the pair \((\delta, \mu)\) is weakly admissible in the sense of [RZ]. Equivalently, the filtered isocrystal induced by \((\delta, \mu)\) on any \( p \)-adic representation of \( G \) is weakly admissible in Fontaine’s sense.

Recall that \( \delta \in B(G) \) is basic if \( M(\delta) = G \). Let \( B(G)_b \) denote the set of basic classes. Condition (8.2.1) determines a unique element \( \delta(\mu) \in B(G)_b \) (cf. [K7, 6.4]). For \( \delta \in B(G) \), let \( P_\delta \subset G \) be the unique standard parabolic subgroup such that \( \overline{\nu}(\delta) \in \mathfrak{A}^+_{P_\delta} \). If \( P = LU \) is a standard parabolic, let \( B(G)_{L,r} = \{ \delta \in B(G) \mid P_\delta = P \} \). Then \( B(G) = \bigsqcup_L B(G)_{L,r} \), where \( L \) runs through standard Levi subgroups of \( G \) (i.e., containing the chosen \( T_0 \)). Here we are referring to [K7, (5.1.1)], but we have replaced his notation \( B(G)_P \) by \( B(G)_{L,r} \). If \( L \) is a standard Levi subgroup, then there is a natural map \( i_{LG} : B(L) \rightarrow B(G) \) [K6, §6]. Note that \( \mathfrak{A}^+_P \) is a chamber in \( \mathfrak{A}_P \), the \( \mathfrak{A} \) associated to \( M \). Thus there is a Newton map \( \overline{\nu} : B(L) \rightarrow \mathfrak{A}_P \); let \( B(L)_b^+ \) (resp. \( \overline{\mathfrak{B}(L)}^+_b \)) denote the subset whose image under this Newton map is contained in \( \mathfrak{A}^+_P \) (resp. in \( \overline{\mathfrak{A}^+_P} \)). Then \( i_{LG} \) is injective on \( \overline{\mathfrak{B}(L)}^+_b \), and \( B(G)_{L,r} = i_{LG}(B(L)_b^+) \).

We now assume \( F = \mathbb{Q}_p \), and let \( E \) be the field of definition of the conjugacy class of \( \mu \). Let \( \Gamma_E = \text{Gal}(\overline{E}/E) \). Consider the Langlands representation \( r_\mu \) of \( \hat{L} \), taken relative to \( E \). Let \( P = LU \subset G \) be a standard \( \mathbb{Q}_p \)-rational parabolic. The representation \( r_\mu \) decomposes, upon restriction to \( \hat{L} \), as a sum of irreducible components \( \mathcal{C}_0(L, \mu) \), each intervening with multiplicity one. Indeed, \( \mu \) is a minuscule weight, with stabilizer \( W_\mu = W_\mathfrak{q} \) for a certain parabolic subalgebra \( \mathfrak{q} \subset \mathfrak{g} \) defined over \( \overline{\mathbb{Q}} \). The irreducible components of \( r_\mu \) are indexed by the set of \( \Gamma_E \)-orbits in \((W_P \setminus W_G/W_\mathfrak{q})\) where \( W_P \) is the absolute Weyl group of \( L \), or equivalently of its Langlands dual \( \hat{L} \). The highest weight of the component corresponding to \( w \), relative to the standard ordering induced by \( P_0 \), is the one in the orbit containing \( w \mu \). Let \( W_G(L) \subset W(G) \) denote the subgroup of elements normalizing \( L \). Since \( L \) is \( \mathbb{Q}_p \)-rational, the action of \( \Gamma \) on \( W_G \) stabilizes \( W_G(L) \). We identify two elements \( \lambda, \lambda' \in \mathcal{C}_0(L, \mu) \) if they are associate; i.e., if there is an element of \( W_G(L) \) that takes \( \lambda \) to \( \lambda' \). Let \( \mathcal{C}(L, \mu) \) be the set of equivalence classes for this relation.
Remark 8.2.3. — By definition,

\[ \Gamma_E = \{ \sigma \in \Gamma \mid \sigma(\mu) \in W_G(\mu) \}. \]

It follows that there is a bijection between the set of \( \Gamma_E \) orbits in \( W_G(\mu) \) and the set of \( \Gamma \)-orbits in the \( W_G \rtimes \Gamma \)-orbits of \( \mu \) in \( X^*(T^0) \). Thus \( \mathcal{C}(L, \mu) \) can be identified with the set of \( W_P \rtimes (W_G(L) \rtimes \Gamma) \) orbits in the \( W_G \rtimes \Gamma \)-orbit of \( \mu \) in \( X^*(T_0) \). In particular, we can replace \( \Gamma_E \)-orbits by \( \Gamma \)-orbits in the following discussion.

We index the elements of \( \mathcal{C}(L, \mu) \) by their highest weights; if \([\lambda] \in \mathcal{C}(L, \mu)\) consists of several elements of \( \mathcal{C}_0(L, \mu) \), we take the one with the highest weight relative to the standard ordering on \( X^*(\hat{T}_0) \) defined by \( P_0 \) (for which \( \mu \) is a highest weight). Each component in \( \mathcal{C}_0(L, \mu) \) is obviously minuscule: its weights form an orbit under \( W_P \). Now restriction to the center \( Z(\hat{L}) \) defines a one-to-one correspondence

\[ (8.2.4) \quad \{ \text{minuscule highest weights of } \hat{L} \} \leftrightarrow X^*(Z(\hat{L})). \]

Indeed, for semisimple groups, this follows from Proposition 8 of [Bu, Ch. VIII, §7], and the generalization to arbitrary reductive groups is immediate. The bijection of (8.2.4) is \( W_G(L) \rtimes \Gamma \)-equivariant \((W_P \text{ acts trivially on both sides})\) and induces a bijection, which we denote \( \beta_L \), between the set \( \mathcal{M}(\hat{L}) \) of \( \Gamma \)-orbits in the set of minuscule weights of \( \hat{L} \) and \( X^*(Z(\hat{L})^\Gamma) \). We may identify \( \mathcal{C}(L, \mu) \) with a subset of \( \mathcal{M}(\hat{L}) \).

Lemma 8.2.5. — Let \( \mathcal{S}(L, \mu) \) be the set of \( W_G(L) \)-orbits of elements \( \chi \in X^*(Z(\hat{L})^\Gamma) \) such that

1. \( \chi \mid_{X^*(Z(\hat{G})^\Gamma)} = \beta_G(\mu) \);
2. \( \mu \mid_{X^*(Z(\hat{L})^\Gamma)} \geq \chi \);
3. \( \chi(H_\alpha) \in \{0, -1, 1\} \) for all roots \( \alpha \) of \( (G, T_0) \); here \( H_\alpha \) is the standard vector in \( \text{Lie}(T_0) \).

Then the map \( \beta_L \) restricts to a bijection

\[ \beta_L : \mathcal{C}(L, \mu) \leftrightarrow \mathcal{S}(L, \mu). \]

Here the order in the inequality is that defined by \( P_0 \) on \( X^*(\hat{T}_0)^\Gamma \).

Proof. — It is clear that \( \beta_L \) takes values in \( \mathcal{S}(L, \mu) \). Thus we need to show that every element of \( \mathcal{S}(L, \mu) \) comes from the \( W_G \) orbit of \( \mu \). In other words, we need to show that, if \( \lambda \) is a minuscule weight of \( \hat{L} \) satisfying (1), (2), and (3), then \( \lambda = w\mu \) for some \( w \in W_G \). But it follows from (3) and Proposition 6 of [Bu, loc. cit.] that \( \lambda = w\mu' \) for some dominant minuscule weight \( \mu' \) of \( (G, T_0) \). Then (1) and (8.2.4) imply that \( \mu = \mu' \). Condition (2) is in fact redundant.

On the other hand, let

\[ B(G, \mu)_L = B(G, \mu) \cap i_{LG}(B(L)_b) = B(G, \mu) \cap i_{LG}(\overline{B(L)_b}^+) \].
Note that $B(G,\mu)_L$ is not generally contained in $B(G)_{L,\Gamma}$. Let
\[ \alpha_L : B(L)_b \longrightarrow X^*(Z(L)^\Gamma) \]
denote the bijection of [K6, Proposition 5.6]. To any element $\delta \in B(L)_b$ we can associate its Kottwitz invariant $\alpha(\delta) = \alpha_G(i_{L,G}(\delta)) \in X^*(Z(\hat{G})^\Gamma)$. Then $\alpha(\delta)$ is the restriction of $\alpha_L(\delta)$ to $X^*(Z(\hat{G})^\Gamma)$.

**Lemma 8.2.6.** — There is a natural bijection $\text{Strat}_L : C(L,\mu) \rightarrow B(G,\mu)_L$ uniquely determined by the property that, if $\text{Strat}_L(w\mu) = i_{L,G}(\delta_L)$, then the pair $(\delta_L, w\mu)$ is weakly admissible for $L$.

**Proof.** — The condition of weak admissibility is precisely the analogue of (8.2.1), namely that
\[ \alpha_L(\delta_L) = \beta_L(w\mu). \]
Since $\alpha_L$ is a bijection on basic classes, this condition certainly determines $\text{Strat}_L$ uniquely. It thus remains to be shown that $\alpha_L$ defines a bijection between $\overline{B}(L)_b^+$, which we identify with $B(G,\mu)_L$ via $i_{L,G}$, and $S(L,\mu)$. It follows from (2) of Lemma 8.2.5 that $S(L,\mu) \subset \alpha_L(B(G,\mu)_L)$. Moreover, every element of $B(G,\mu)$ satisfies (1) of Lemma 8.2.5. On the other hand, the order on $\overline{B}(L)_b^+$ defined by the Newton map is compatible with that on $X^*(Z(\hat{L})^\Gamma)$, i.e., by pairings with the vectors $H_\alpha$ for simple roots $\alpha$. Since $\beta_G(\mu) \geq \alpha_L(\delta_L) \geq 0$, for $\delta_L \in \overline{B}(L)_b^+$, with $\mu$ minuscule, it follows that $\alpha_L(\delta)$ satisfies (3) as well. This completes the proof. \[ \square \]

We now let $C(\mu) = \coprod_L C(L,\mu)$, where $L$ runs through the classes of standard Levi subgroups of $G$.

**Corollary 8.2.7.** — There is a natural surjective map
\[ \text{Strat} : C(\mu) \longrightarrow B(G,\mu) \]
given on $C(L,\mu)$ by $\text{Strat}_L$.

Indeed, the map is surjective because
\[ B(G) = \cup_L i_{L,G}(\overline{B}(L)_b^+) \]
as $L$ runs over the set of standard Levi subgroups of $G$. Note, however, that the map is not generally injective. Indeed, the union in (8.2.8) is not disjoint in general. However, this is the only source of ambiguity. To $b \in B(G,\mu)$, we let $\text{Rep}(b) = \text{Strat}^{-1}(b)$; it is a set of pairs $(L,\lambda)$, with $\lambda \in C(L,\mu)$, partially ordered by inclusion in the obvious sense. It contains a maximal element $(M = M(b),\lambda_0)$ with the property that $b \in B(G)_{M(b)}$; here $M(b)$ is defined as above.

**Lemma 8.2.9.** — With the above notation, there is a bijection between $\text{Rep}(b)$ and the set of $P(b)$ of standard parabolics $P \subset M = M(b)$ that transfer to the inner form $J(b)$ of $M$ defined by the basic $\sigma$-conjugacy class $b$. 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
Proof. — We have seen that $\text{Rep}(b)$ is in bijection with the set of pairs $(P = LU, b_L)$ where $b_L \in B(L)_b$ is such that $i_{LM}(b_L) = b$. Thus the lemma comes down to the assertion that $b$ is $\sigma$-conjugate to an element of $B(L)_b$ if and only if $P$ transfers to $J(b)$. We let $M_{ad}$ be the adjoint group of $M$ and $b_{ad}$ the image of $b$ in $B(M_{ad})_b$. There is a bijection

$$j: H^1(\Gamma, M_{ad}) \xrightarrow{\sim} B(M_{ad})_b$$

(cf. [K7, 3.2]) and $j^{-1}(b_{ad})$ is the cohomology class defining the inner form $J(b)$.

One direction is simple. Suppose $b = i_{LM}(b_L)$ as above. It follows that the inner form $J(b_L)$ of $L$ defined by $b_L$ transfers to $J(b)$, hence is necessarily a Levi subgroup of a rational parabolic $P_b \subset J(b)$.

To construct a map in the other direction, we may as well assume $M = M_{ad}$, since both sides of the purported bijection are unchanged when $M$ is replaced by $M_{ad}$. Thus $b = b_{ad}$. Let $Q(b)_0$ be a standard minimal parabolic subgroup of $J(b)$, with Levi subgroup $L(b)_0$ and anisotropic kernel $I(b)_0$. Let $Q_0$ be a standard parabolic subgroup of $M$ that transfers to $Q(b)_0$, and let $I_0 \subset Q_0$ be the reductive subgroup corresponding to $I(b)_0$. It is standard that the cohomology class in $H^1(\Gamma, M_{ad})$ defining the inner form $J(b)$ is represented by a class in $H^1(\Gamma, I_0)$; i.e., $b$ is $\sigma$-conjugate to $b_I \in B(I_0) \cap j_{Ib}(H^1(\Gamma, I_0))$, where for any reductive group $H$, there is a natural bijection

$$j_H : H^1(\Gamma, H) \longrightarrow B(H)_b$$

as in (8.2.9.1). Let $P = LU$ be a standard parabolic subgroup of $M_{ad}$ that transfers to $P_b \subset J(b)$, and let $L_b \subset P_b$ be a Levi subgroup, necessarily an inner form of $L$. Then $I_0 \subset L$. The obvious commutative diagram then shows that $b = i_{LM}(b_I) \in \text{Im}[H^1(\Gamma, L) \to B(L)_b \to B(M)]$, hence a fortiori belongs to the image under $i_{LM}$ of the image of $b_I$ in $B(L)_b$. \hfill $\Box$

8.2.10 Example. — We work out the stratification in the case of a Shimura variety $\text{Sh}(G, X)$ uniformized by the symmetric space associated to a unitary similitude group of signature $(k, n - k)$, for some integer $0 \leq k \leq n$. For simplicity, we assume $G$ to be the unitary similitude group, as in (1.2), relative to a central simple algebra over an imaginary quadratic field $E$; however, we now assume $G(\mathbb{R}) \cong GU(k, n - k)$. For an appropriate choice of Shimura datum $(G, X)$, the corresponding representation $\rho_p$ of the dual group $\hat{G} \cong GL(n, \mathbb{C}) \times \mathbb{C}^\times$ of $G$ is of the form $\wedge^k \text{St} \otimes \nu$, where $\text{St}$ is the standard representation of $GL(n)$ and $\nu$ is a character which we simply ignore. We consider a prime $p$ that splits in $E$, so that $G(\mathbb{Q}_p) \cong GL(n, \mathbb{Q}_p) \times \mathbb{Q}_p^\times$. Let $K = K_p \times K^p \subset G(\mathbf{A}_f)$ be a level subgroup, with $K_p$ hyperspecial. The special fiber of $K \text{Sh}(G, X)$ then naturally carries a family $\mathcal{H}$ of $p$-divisible groups of height $n$ and dimension $k$, generalizing the family considered in [HT]. (There is an “additional structure” coming from the character of $\mathbb{Q}_p^\times$, but this plays no role in the following discussion.)
The strata correspond to the Dieudonné-Manin classification of isogeny classes of $p$-divisible groups in terms of the slope decomposition. The set $B(G, \mu)$ can then easily be identified with the set of partitions

$$(k, n) = \sum_{i=1}^{m} (r_i, s_i)$$

where $r_i, s_i$ are non-negative integers satisfying $r_i \leq s_i$ for all $i$, and the rational numbers $r_i/s_i$ are all distinct. The order in the sum is immaterial. The geometric point $x \in S$ belongs to the stratum $S((\{r_i, s_i\}))$ if and only if the $p$-divisible group $\mathcal{H}_x$ is isogenous to a $p$-divisible group of the form $\prod_i(\mathcal{H}_{r_i/s_i})^{d_i}$, where $\mathcal{H}_{r_i/s_i}$ is a simple $p$-divisible group of slope $r_i/s_i$ and $d_i$ is the greatest common divisor of $r_i$ and $s_i$. The centralizer in $G(K)$ of the corresponding slope morphism is then $M((\{r_i, s_i\})) = \prod_{i=1}^{m} GL(s_i, K) \times K^\times$, and the associated twisted form is

$$J((\{r_i, s_i\})) = \prod_{i=1}^{m} GL(d_i, D_{r_i/s_i}) \times Q_p^\times,$$

where $D_{r_i/s_i}$ is the division algebra of dimension $(s_i/d_i)^2$ with invariant $r_i/s_i$. The set of standard parabolics of $M((\{r_i, s_i\})$ that transfer to $J((\{r_i, s_i\}))$ is in one-to-one correspondence with the set of $m$-tuples $(\delta_i)$, where each $\delta_i$ is a divisor of $d_i$.

On the other hand, to each partition $n = \sum_{j=1}^{t} n_j$ corresponds a standard Levi factor $L = L(\{n_j\}) \equiv \prod_j GL(n_j, \mathbb{Q}_p) \times \mathbb{Q}_p^\times$ of $G(\mathbb{Q}_p)$, and the Langlands dual of $L$ has the same form. If we write $\hat{G} \equiv GL(V) \times \mathbb{C}^\times$, for some $n$-dimensional complex vector space $V$, then $\hat{\mathcal{L}}$ is the stabilizer of a decomposition $V = \sum V_j$, with $\dim V_j = n_j$. The restriction of $\wedge^k V$ breaks up as the sum of the irreducible $\hat{\mathcal{L}}$-invariant subspaces

$$\oplus_{k = k_1 + \cdots + k_t} \wedge^{k_1} V_1 \otimes \cdots \otimes \wedge^{k_t} V_t,$$

where $k = k_1 + \cdots + k_t$ runs through partitions of $k$. Thus $\mathcal{C}(\mu)$ is the set of partitions $(k, n) = \sum_{i=1}^{t} (k_i, n_i)$, and the map $\text{Strat} : \mathcal{C}(\mu) \to B(G, \mu)$ consists in replacing the partition $(k, n) = \sum_{i=1}^{t} (k_i, n_i)$ by the one obtained by adding together all pairs $(k_i, n_i)$ with fixed $k_i/n_i$. It is easy to check that the above description of parabolics transferring to $J((\{r_i, s_i\}))$ is compatible with Proposition 8.2.9.

The book [HT] and the previous lectures are concerned with the specific case $k = 1$, and the classification is valid whether or not the base field $E$ is imaginary quadratic. The partition (8.2.10.1) then has at most two terms:

$$(1, n) = (1, n - h) + (0, h)$$

where the second term is present if and only if $h \neq 0$. The first term corresponds to the connected part of the $p$-divisible group, the second to the étale part. Then $\text{Strat}^{-1}(1, n - h)$ consists of a single element, whereas $\text{Strat}^{-1}(0, h)$ consists of all partitions of $h$. In other words, $\text{Strat}^{-1}(1, n - h)$ corresponds to of standard parabolic subgroups of $GL(n)$ contained in $P_h$ and containing the $GL(n - h)$-component of its Levi factor.
8.3. Results of Fargues. — As at the end of (8.1), we consider the special fiber $K_p\mathcal{S}$ at a place $v$ of the reflex field $E$ of the Shimura variety $K_p\cdot K_r \Sh(G, X)$ with good reduction at $v$. Let $G_p = G_{Q_p}$, and let $\mu$ be the cocharacter of $G$ associated to the Shimura datum $(G, X)$, viewed as a $\overline{Q}_p$-cocharacter. Then one expects $K_p\mathcal{S}$ to decompose as a disjoint union of locally closed reduced subvarieties
\begin{equation}
K_p\mathcal{S} = \bigsqcup_{b \in B(G_p, \mu)} K_p\mathcal{S}(b).
\end{equation}
When $(G, X)$ is a PEL type, $K_p\mathcal{S}$ is a moduli space for abelian varieties with additional structure (at least in the unramified cases considered in [K5]). Then the stratification (8.3.1) is known to exist: $K_p\mathcal{S}(b)$ is the reduced subscheme whose geometric points classify abelian varieties of the given PEL type and with isocrystal (with additional structure) of type $b$. That (8.3.1) defines a stratification is a consequence of Grothendieck’s theorem on specialization of isocrystals, as generalized by Rapoport and Richartz [RR]. Let $K_p\mathcal{S}(\geq b)$ denote the closure of $K_p\mathcal{S}(b)$ in the special fiber. It then follows from the results of [RR] that $K_p\mathcal{S}(\geq b)$ is a finite union of strata $K_p\mathcal{S}(b')$, for $b' \in B(G_p, \mu)$ such that $\mathfrak{p}(b') \leq \mathfrak{p}(b)$ for a natural partial ordering (the Newton polygon associated to $b'$ lies above that associated to $b$).

For the rest of this discussion $K^p$ will be fixed. We assume for simplicity that $G$ is anisotropic (modulo center). For any open subgroup $K'_p \subset K_p$, we consider the rigid-analytic space $\Sh_{K'_p}^\rig$, associated to the Shimura variety $K'_p \cdot K_r \Sh(G, X)$ (Fargues considers various versions of rigid-analytic spaces, including Huber’s adic spaces and Berkovich’s analytic spaces; here we will not be precise). Let $\Sh_{K_p}^{\rig, \geq b} \subset \Sh_{K_p}^{\rig}$ denote the (open) tube over the closed subvariety $K_p\mathcal{S}(\geq b)$ of the special fiber: $\Sh_{K_p}^{\rig, \geq b}$ is the set of points of $\Sh_{K_p}^{\rig}$ whose specialization lies in $K_p\mathcal{S}(\geq b)$. For any open subgroup $K'_p \subset K_p$, we define $\Sh_{K'_p}^{\rig, \geq b}$ to be the fiber product of $\Sh_{K_p}^{\rig, \geq b}$ with $\Sh_{K_p}^{\rig}$ over $\Sh_{K'_p}^{\rig}$; note that this can be defined without reference to an integral model of $K'_p \cdot K_r \Sh(G, X)$. We let $\Sh_{K'_p}^{\rig, b}$ denote the complement of $\Sh_{K'_p}^{\rig, \geq b'}$, for $\mathfrak{p}(b') < \mathfrak{p}(b)$, in $\Sh_{K'_p}^{\rig, \geq b}$.

Let $\pi_f$ be a representation of $G(A_f)$ contributing to non-endoscopic cohomology in (8.1.2). We will soon assume $\pi_p$ to be supercuspidal, but for the moment we let $P \subset G_p$ be the parabolic subgroup, with Levi subgroup $L$, and assume that $\pi_p$ is isomorphic to the representation induced from a discrete series representation $\tau_p$ of $L$. Then the Langlands parameter attached to $\pi_p$ is (conjecturally) given by a homomorphism $\sigma(\pi_p) : WD_{E_v} \to L(L(\overline{\mathbb{Q}_l}))$. Compatibility of local and global correspondences, generalizing Theorem 1.3.6, amounts to the hypothesis that the restriction to $WD_{E_v}$ of $r_f(\pi_f)$ to $WD_{E_v}$ is equivalent to $r_\mu \circ \sigma(\pi_p)$. In particular, by the discussion preceding Remark 8.2.3,
\begin{equation}
(r_f(\pi_f)|_{WD_{E_v}} = \bigoplus_{\lambda \in C(L, \mu)} r_f(\pi_f)_\lambda,
\end{equation}
where we have grouped together irreducible summands that are associate.
Let \([H^\bullet(\Sh(G, X), \overline{q}_\ell)]\) denote the direct limit, over \(K'_p \cdot K_p\), of the alternating sum of the \(\ell\)-adic cohomology groups of \(K'_p \cdot K_p \Sh(G, X)\). We define \([H^\bullet(\Sh^{\rig,b}, \overline{q}_\ell)]\) analogously, using this time the \(\ell\)-adic cohomology of the indicated rigid space. Roughly speaking, the stratification gives rise to an identity in the Grothendieck group of \(G(\mathbb{A}_f) \times WD_{E_v}\), analogous to the First Basic Identity (4.4.4):

\[
(8.3.3) \quad [H^\bullet(\Sh(G, X), \overline{q}_\ell)] = \sum_{b \in B(G, \mu)} [H^\bullet_c(\Sh^{\rig,b}, \overline{q}_\ell)].
\]

The heuristic expectation is that, if \(b = \Strat_L(\lambda)\), then the representation \(r_\ell(\pi_f)_\lambda\) is realized on the compactly supported cohomology \(\lim_{\rightarrow} K'_p H^\bullet_c(\Sh^{\rig,b}, \overline{q}_\ell)\). In [HT] the partition \((1, n) = (1, n - h) + (0, h)\) of (8.2.10.2) corresponds to the stratum here denoted \(\overline{S}^{(h)}\). The Second Basic Identity, and more precisely the calculations (6.2.3)-(6.2.7), show that the \(\pi_p\) that contribute to the cohomology of \(\overline{S}^{(h)}\) are precisely those for which \(\pi_w\) is induced from a parabolic subgroup of \(P_h\) corresponding to a partition of \(h\). As indicated at the end of the previous section, this is just the fiber of \(\Strat\) lying above the partition \((1, n) = (1, n - h) + (0, h)\).

In particular, if \(\pi_p\) is supercuspidal, there is only one \(\lambda\) in the sum (8.3.2)\(^{(10)}\), namely \((L = G, w\mu = \mu)\), and \(\Strat(\lambda)\) is the basic stratum. The heuristic expectation is then the

\[\text{(8.3.4) Conjecture. — Let } b_0 \in B(G, \mu) \text{ denote the basic class. If } \pi_p \text{ is supercuspidal, then } r_\ell(\pi_f)_\lambda \text{ comes exclusively from the contribution of } [H^\bullet_c(\Sh^{\rig,b_0}, \overline{q}_\ell)] \text{ to the right-hand side of (8.3.3)}. \]

In Groth(G(\mathbb{A}_f))).

This conjecture was verified in [HT] for the Shimura varieties considered there, and is asserted as (5.1.4) above. As we have seen, the proof of this conjecture is based on Boyer’s trick, which proves that the cohomology of the non-basic strata is induced from parabolic subgroups, because the strata themselves are induced. For more general Shimura varieties this trick fails; it is easy to see that the strata are generally not induced. However, Fargues proves:

\[\text{(8.3.5) Theorem (Fargues). — Suppose } G \text{ is the unitary similitude group of a division algebra } B \text{ of degree } n^2 \text{ over a CM field of the form } F = F^+ E, \text{ as in (1.1). Suppose } B \text{ is locally everywhere either split or a division algebra. Let } p \text{ be a prime unramified in } F. \text{ Suppose either}
\]

\[\begin{align*}
\text{(1) } & p \text{ splits in } E; \text{ or} \\
\text{(2) } & p \text{ is inert in } E \text{ and } n = 3
\end{align*}
\]

\[\text{Then Conjecture (8.3.4) is true.}\]

\(^{(10)}\)This does not mean the representation \(r_\ell(\pi_f)|_{WD_{E_v}}\) is necessarily irreducible, or even indecomposable.
In case (1) Fargues actually makes the slightly stronger assumption that \( p \) is inert in \( F \), but this is merely to simplify the exposition.

In the absence of Boyer’s trick, Fargues proves Theorem 8.3.5 by proving vanishing of the trace of a supercuspidal matrix coefficient \( \phi \) against the sum \( \sum_i (-1)^i H^*_c(\text{Sh}^\text{rig,b}, \overline{\mathbb{Q}}_\ell) \) appearing on the right hand side of the formula in (8.3.4). By (8.3.2), this is equivalent to showing that the trace of \( \phi \) on \( [H^*_c(\text{Sh}(G,X), \overline{\mathbb{Q}}_\ell)] \) equals the trace of \( \phi \) on \( [H^*_c(\text{Sh}^\text{rig,b}, \overline{\mathbb{Q}}_\ell)] \).

The trace of \( \phi \) on \( [H^*_c(\text{Sh}(G,X), \overline{\mathbb{Q}}_\ell)] \) is given by the cohomological trace formula (5.4.2). To compute the trace of \( \phi \) on \( [H^*_c(\text{Sh}^\text{rig,b}, \overline{\mathbb{Q}}_\ell)] \), Fargues carries out a preliminary analysis of fixed-point contributions of isogeny classes, as in Lectures 6 and 7. However, for a variety of reasons, this analysis, unlike the analysis in Lecture 5, does not calculate the trace of a Hecke operator on cohomology, even of the basic stratum, unless the Hecke operator has first been twisted by a high power of Frobenius, as required by Fujiwara’s trace formula. Since the cohomological trace formula (5.4.2) has no room for twisting by Frobenius, there seems to be an insurmountable obstacle.

Fargues overcomes this obstacle by making use of the Galois representation \( r_\ell(\pi_f) \), whose restriction to \( WD_{F_\tau} \) can be determined by combining the results of [K4] (at unramified primes away from \( p \)), the Main Theorem 1.3.6 (for a Shimura variety of signature \((1,n-1)\) attached to an inner form of \( G \)), and Chebotarev’s density theorem. In particular, he finds that \( r_\ell(\pi_f)|_{WD_{F_\tau}} \) depends only on \( \pi_p \), which allows him to “twist by Frobenius” for fixed \( \pi_p \) at the level of the cohomological trace formula.

For general PEL types of type \( \tilde{A} \) and \( C \), Rapoport and Zink have shown in [RZ] that the basic stratum \( \text{Sh}^\text{rig,b} \) admits a rigid-analytic uniformization by a tower of moduli spaces \( \tilde{M}(G_p,\mu)\mathcal{K}_\ell' \); a special case of this uniformization is (3.4.10). Using this uniformization, Fargues determines the trace of \( [H^*_c(\text{Sh}^\text{rig,b}, \overline{\mathbb{Q}}_\ell)] \) by constructing a Hochschild-Serre spectral sequence, simultaneously generalizing that of [H1] and the (much simpler formula) (5.2.11), in terms of the cohomology \( [H^*_c(\tilde{M}(G_p,\mu), \overline{\mathbb{Q}}_\ell)] \), as defined by Berkovich or Huber. These cohomology groups are smooth modules for \( G_p \times J_{b_0} \times WD_{F_\tau} \), where \( J_{b_0} \) is the inner form of \( G_p \) given as the group of self-quasiisogenies of the \( p \)-divisible group attached to any point in the basic stratum. In Theorem 8.3.5 (1), \( J_{b_0} \) is the multiplicative group of a division algebra with invariant \( r/n \) for some \( r \) prime to \( n \). Using Theorem (8.3.5) and the existence of the local Langlands correspondence for \( GL(n) \), Fargues then proves

(8.3.6) **Theorem (Fargues).** — Under the hypotheses of Theorem 8.3.5 (1), let \( \pi \) be an irreducible admissible representation of \( J_{b_0} \) corresponding to a supercuspidal representation \( JL(\pi) \) of \( G_p \) via the Jacquet-Langlands correspondence. Then

\[
\sum_i (-1)^i \text{Hom}_{J_{b_0}}(H^*_c(\tilde{M}(G_p,\mu), \overline{\mathbb{Q}}_\ell), \pi) = [JL(\pi)] \otimes r_\mu \circ \tilde{\sigma}_\ell(JL(\pi)).
\]

Here \( \tilde{\sigma}_\ell \) is a certain normalized twist of the local Langlands correspondence. Fargues obtains similar results for \( U(3) \), but the presence of \( L \)-packets complicates the statement.
The Local Langlands Correspondence

Appendices

A.1. Traces, pseudocoefficients, and the Jacquet-Langlands correspondence. — In the present section, $K$ is a finite extension of $\mathbb{Q}_p$. Let $g$ denote a positive integer, and let $\mathcal{A}(g, K)$ denote the set of equivalence classes of irreducible admissible representations of $G_g = GL(g, K)$. Let $\mathcal{A}_d(g, K)$ (resp. $\mathcal{A}_t(g, K)$) the subset of discrete series (resp. tempered) representations. For any representation $\pi \in \mathcal{A}(g, K)$, the trace $\text{Tr}(\pi)$ is a distribution, defined on $C^\infty_c(G_g)$ as the trace of the finite rank operator

\begin{equation}
\pi(\phi) = \int_{G_g} \phi(g)\pi(g)dg.
\end{equation}

Note that the trace depends linearly on the choice of Haar measure. It is known thanks to Harish-Chandra that $\text{Tr}(\pi)$ is represented by a locally $L^1$-function $\chi_\pi$, defined on the regular semi-simple elements $G_{g, \text{reg}}$. (Of course Harish-Chandra’s theorem is valid for any reductive algebraic group over $\mathbb{Q}_p$.) It is also known that

(A.1.2) Linear independence of characters. — Any relation $\sum_{\pi \in A} a_\pi \text{Tr}(\pi) = 0$, where $A \subset \mathcal{A}(g, K)$ is a finite subset and $a_\pi \in \mathbb{C}$, is trivial.

Let $\mathcal{A}_{d, \text{fin}} \subset \mathcal{A}_d(g, K)$ be any countable subset with the following property:

(A.1.3). — For any $\pi \in \mathcal{A}_d(g, K)$, the set of unramified characters $\psi$ of $K^\times$ such that $\pi \otimes \psi \circ \det \in \mathcal{A}_{d, \text{fin}}$ is finite.

In other words, for any $\pi \in \mathcal{A}_d(g, K)$, the intersection $\mathcal{A}_{d, \text{fin}}(\pi)$ of $\mathcal{A}_{d, \text{fin}}$ with the inertial equivalence class of $\pi$ is finite. Let If $\pi \in \mathcal{A}_{d, \text{fin}}$, a pseudocoefficient for $\pi$, relative to $\mathcal{A}_{d, \text{fin}}$, is a function $\phi_\pi \in C^\infty_c(G_g)$ such that

\begin{equation}
\text{Tr}(\pi')(\phi_\pi) = 1;
\end{equation}

\begin{equation}
\text{Tr}(\pi')(\phi_\pi) = 0 \text{ if either } \pi' \in \mathcal{A}_{d, \text{fin}}, \pi' \neq \pi \text{ or } \pi' \in \mathcal{A}_t(g, K), \pi' \notin \mathcal{A}_{d, \text{fin}}.
\end{equation}

(A.1.5) Theorem ([DKV], [HT, I.3])

(i) For any set $\mathcal{A}_{d, \text{fin}}$ satisfying (A.1.3) and any $\pi \in \mathcal{A}_{d, \text{fin}}$, a pseudocoefficient $\phi_\pi$ for $\pi$ (relative to $\mathcal{A}_{d, \text{fin}}$) exists.

(ii) If $\pi$ is supercuspidal, then $\text{Tr}(\pi')(\phi_\pi) = 0$ for any $\pi' \neq \pi$ (not necessarily tempered).

(iii) For general $\pi \in \mathcal{A}_{d, \text{fin}}$, let $\pi' \in \mathcal{A}(g, K)$ be a non-tempered representation such that $\text{Tr}(\pi')(\phi_\pi) \neq 0$. Then $\pi'$ belongs to the block of $\pi$; i.e., there is an unramified character $\psi$ of $K^\times$, a proper standard parabolic subgroup $P \subset G_g$, and an irreducible admissible representation $\tau$ of $P$ such that $\pi' \otimes \psi \circ \det$ and $\pi$ are Jordan-Hölder constituents of $\text{n-Ind}_{P}^{G_g} \tau$.

The pseudocoefficients, and the block of $\pi$, can also be defined cohomologically, as Euler-Poincaré functions; cf. [SS], Proposition III.4.1 and Corollary III.4.8.
Without the restriction to $A_{d,\text{fin}}$, the theorem is false, because a given $\pi$ can be twisted by an arbitrary unramified character, and the family of such twists is continuous. For $g = 1$ the existence of pseudocoefficients without restriction would imply that the Fourier transform is defined on a discrete space, which is false.

Pseudocoefficients are not unique. For the purposes of the present notes, we primarily need them for supercuspidal $\pi$, in which case the construction is relatively simple. Let $Z_g$ denote the center of $G_g$. Let $\phi_{\nu',\nu}$ be any matrix coefficient of the contragredient $\pi^\vee$ of $\pi$:

$$\phi_{\nu',\nu}(g) = \langle \pi^\vee(g)\nu', \nu \rangle$$

for some $\nu' \in \pi^\vee$, $\nu \in \pi$ such that $\langle \nu', \nu \rangle \neq 0$. Let $\psi_\pi$ denote the central character of $\pi$; then

$$\phi_{\nu',\nu}(zg) = \psi_\pi^{-1}(z)\phi_{\nu',\nu}(g), \quad z \in Z_g.$$  

Let $C_c^\infty(G_g, \psi_\pi^{-1})$ denote the space of functions compactly supported modulo $Z_g$ and satisfying (A.1.6). Since $\pi$ is supercuspidal, the matrix coefficient $\phi_{\nu',\nu}$ belongs to $C_c^\infty(G_g, \psi_\pi^{-1})$.

If $\pi'$ is any admissible representation of $G_g$ with central character $\psi = \psi_\pi$, then any function $f \in C_c^\infty(G_g, \psi^{-1})$ defines a trace class operator $\pi'(f)$ on $\pi'$ by the formula

$$\pi'(f) = \int_{G_g/Z_g} f(g)\pi'(g)d\hat{g}.$$  

Here $d\hat{g}$ is an invariant measure on $G_g/Z_g$. We write $\text{Tr}_{Z_g,\psi}(\pi) = \text{Tr}_{Z_g,\psi_\pi}(\pi)$ to distinguish the trace of the operator defined by (A.1.7) from that defined via (A.1.1). Once and for all, we choose our Haar measure $dz$ on $K^\times$ so that

$$\int_{O^\times} dz = 1,$$  

and define $d\hat{g}$ to be the quotient measure $dg/dz$. Then (cf. [DKV], A.3.g)

$$\text{Tr}_{Z_g,\psi}(\pi)(\phi_{\nu',\nu}) = d(\pi)^{-1} \phi_{\nu',\nu}(1)$$

where $d(\pi)$ is the formal degree (which depends on the choice of Haar measure on $G_g$). Thus by choosing $\nu'$ and $\nu$ appropriately, we may assume $\text{Tr}_{Z_g,\psi_\pi}(\pi)(\phi_{\nu',\nu}) = 1$; we then write $\phi_Z = \phi_{\nu',\nu}$.

As above, we let $\delta = w_K \circ \det : G_g \to \mathbb{Z}$, with $w_K$ the valuation on $K$. Let $a < b$ be a pair of integers, with $b - a + 1$ an integral multiple of $g$, say

$$b - a + 1 = mg.$$  

Let $\omega$ denote the interval $[a, b]$. For any locally constant function $f$ on $G_g$, we define the $\omega$-truncation $t_\omega(f) \in C_c^\infty(G_g)$ by

$$t_\omega(f)(g) = f(g), \quad \delta(g) \in \omega; \quad t_\omega(f)(g) = 0 \text{ otherwise.}$$
Then it is easy to see that, for any interval \([a, b]\) as above with \(m\) sufficiently large, relative to the set \(\mathcal{A}_{d,\text{fin}}\), the function

\[
\phi = \phi_{\pi,\omega} = \frac{1}{m} t_\omega(\phi_Z)
\]

is a pseudocoefficient for \(\pi\) relative to \(\mathcal{A}_{d,\text{fin}}\). (In any case, \(\phi_{\pi,\omega}\) has trace zero on any tempered representation not inertially equivalent to \(\pi\), and for large enough \(m\), \(\phi_{\pi,\omega}\) separate elements of \(\mathcal{A}_{d,\text{fin}}(\pi)\).) In particular, we can assume all \(\phi_\pi\) have support in elements of arbitrarily small (or arbitrarily large) determinant. This is important in the applications of Fujiwara’s theorem.

Henceforward we drop the assumption that \(\pi\) be supercuspidal. The truncation can be defined for any pseudocoefficient and has the properties indicated above. If \(\gamma \in G_g\) is a semisimple element and \(f \in C_\infty^c(G_g)\), the orbital integral \(O_\gamma(f) = O^{JL}_\gamma(f)\) is defined as in (5.4.1). The orbital integral \(O_\gamma(f)\) depends on the choice of Haar measure \(dg\) on \(G_g\), which has already been fixed (and is reflected in the choice of \(\phi_\pi\)), and on the Haar measure on the centralizer \(Z(\gamma) \subset G_g\). Let \(d\gamma\) denote the quotient measure on the quotient \(Z(\gamma)/Z_g\) (recall (A.1.9)). Then

\begin{proposition}
\[O_\gamma(\phi) = \frac{1}{m} t_\omega(\chi_{\pi^\vee})(\gamma),\]
where
\[\text{vol}(Z(\gamma)/Z_g) = \int_{Z(\gamma)\backslash G_g} 1 \, d\gamma.\]
\end{proposition}

The vanishing of the non-elliptic orbital integrals is the Selberg principle. The expression of the elliptic orbital integrals in terms of the character is well-known; cf. [DKV], A.3, and the normalization (A.1.11) introduces the factor \(1/m\) as well as the truncation.

The Jacquet-Langlands correspondence is a bijection

\[
\mathcal{A}_d(G_g) \xrightarrow{\text{JL}} A(J_g).
\]

The notation JL designates the bijection in either direction. It is characterized by the following character identity

\[
\chi(\text{JL}(\pi))(a) = (-1)^{g-1}\chi(\pi)(\gamma).
\]

if \(\gamma\) is an elliptic regular element and \(a \in J_g\) transfers to \(\gamma\). Thus the expression in Proposition (A.1.12) can be rewritten

\[
O_\gamma(\phi) = (-1)^{g-1} \frac{1}{m} \text{vol}(Z(\gamma)/Z_g)^{-1} t_\omega(\chi_{\text{JL}(\pi^\vee)})(\gamma)
\]
for \( \gamma \) elliptic regular; the truncation for \( J_g \) is defined by analogy with that for \( G_g \). Both sides of this formula are defined for general elliptic elements, and the formula extends with the addition of signs:

**(A.1.12 bis) Proposition.** — For \( \gamma \in G_g \) elliptic, \( a \in J_g \) an element whose conjugacy class transfers to the conjugacy class of \( \gamma \), the following identity holds

\[
O_\gamma(\phi) = (-1)^{g-1}e(\gamma) \text{vol}(Z(\gamma)/Z_g)^{-1} \frac{1}{m} t_\omega(\chi_{JL(\pi'_g)})(a).
\]

Here \( e(\gamma) \) is the Kottwitz sign (cf. [L, 1.7.1]).

We let \( \mathcal{A}(J_g)_{\text{fin}} \) denote the image under JL of \( \mathcal{A}_d,\text{fin} \). Then the analogue of Theorem A.1.5 holds for \( \mathcal{A}(J_g)_{\text{fin}} \); indeed, the pseudocoeficients can be constructed starting from matrix coefficients just as for \( GL(g, K) \).

If \( \pi \) is an admissible representation of \( G(A) \), or of \( G(A_f) \), then \( \text{Tr}(\pi) \), defined just as in the local case, exists as a distribution on \( C_c^\infty(G(A)) \). If \( \pi = \otimes_v \pi_v \) is irreducible and \( \phi = \otimes_v \phi_v \) is decomposed with \( \phi_v \in C_c^\infty(G(\mathbb{Q}_v)) \), almost everywhere equal to the characteristic function of a maximal compact subgroup, then

\[
\text{(A.1.14)} \quad \text{Tr}(\pi)(\phi) = \prod_v \text{Tr}(\pi_v)(\phi_v).
\]

**A.2. \( L \) and \( \varepsilon \) factors, and some results of Henniart.** — In this section \( F \) is a number field, \( v \) designates a (variable) place of \( F \), and \( K \) denotes a local field of characteristic zero, generally arising as the completion \( F_v \) of \( F \) at \( v \). The notation of (A.1) for \( K \) remains in force, except that \( K \) can now be an archimedean field, in which case the notion of “irreducible admissible representation” needs to be modified accordingly. By \( \mathcal{A}_0(n, F) \) we denote the set of cuspidal automorphic representations of \( GL(n, F) \): i.e., the irreducible constituents of the space \( \mathcal{A}_0(GL(n, F) \backslash GL(n, A_F)) \) of global cusp forms.

Let \( \mathcal{A}_{\text{gen}}(n, K) \) denote the set of generic irreducible admissible representations of \( GL(n, K) \). Let \( n \) and \( m \) denote two positive integers, \( n \geq m \), and let \( \pi, \pi' \in \mathcal{A}(n, K) \), \( \pi' \in \mathcal{A}(m, K) \). Let \( \Pi \in \mathcal{A}_0(n, F) \), \( \Pi' \in \mathcal{A}_0(m, F) \). We fix an global additive character \( \psi : ad_F / F \rightarrow \mathbb{C}^\times \), and a local additive character \( \psi_K : K \rightarrow \mathbb{C}^\times \); if \( K = F_v \) we assume \( \psi_K \) to be the restriction of \( \psi \) to \( K \).

We momentarily let \( N \) be a positive integer, and let \( \sigma_0 \in \mathcal{G}(N, K) \). Let \( L(s, \sigma_0) \) denote the local Artin L-factor of \( \sigma_0 \), which is a product of \( \Gamma \)-functions if \( K \) is archimedean. Langlands and Deligne (cf. [De2]) have defined local constants \( \varepsilon(s, \sigma_0, \psi_K) \) which are entire nowhere-vanishing functions of \( s \in \mathbb{C} \), and which are compatible with the global functional equations of Artin-Hecke \( L \)-functions in the following sense. Let \( \Sigma_0 \) be an \( N \)-dimensional representation of the global Weil group of \( F \), and let \( L(s, \Sigma_0) \) denote its global \( L \)-function. For any place \( v \) of \( F \), let \( \Sigma_{0,v} \in \mathcal{G}(N, F_v) \) denote the restriction of \( \Sigma_0 \), and let \( \psi_v \) denote the restriction of the
additive character $\psi$. Then there is a functional equation

\[(A.2.1) \quad L(s, \Sigma_0) = \varepsilon(s, \Sigma_0)L(1-s, \hat{\Sigma}_0); \quad \varepsilon(s, \Sigma_0) = \prod_v \varepsilon(s, \Sigma_{0,v}, \psi_v)\]

Note that the product of the local $\varepsilon$-factors is independent of the choice of additive character.

The local factors are characterized by a number of appealing properties, described in detail in [De2]. We simply recall that, for $N = 1$, they are defined by Gauss sums as in Tate’s thesis; they are multiplicative in the sense that

\[(A.2.2) \quad \varepsilon(s, \sigma_0 \oplus \sigma_1, \psi_K) = \varepsilon(s, \sigma_0, \psi_K) \cdot \varepsilon(s, \sigma_1, \psi_K),\]

hence define functions on the Grothendieck group of virtual representations of $WD_K$; finally, they are inductive in degree zero: if $K'/K$ is a finite extension, and $\sigma'$ is a virtual representation of dimension zero of $WD_{K'}$, then

\[(A.2.3) \quad \varepsilon(s, \sigma', \psi_K \circ \text{Tr}_{K'/K}) = \varepsilon(s, \text{Ind}_{K'/K}(\sigma'), \psi_K).\]

These properties are used in (1.4).

Now suppose $\sigma \in G(n, K)$, $\sigma' \in G(m, K)$, and let $N = nm$. Then we can define $\varepsilon(s, \sigma \otimes \sigma', \psi)$, which arises as the local factor in a functional equation of the form (A.2.1) for the tensor product of two representations of the Weil group of $F$. Motivated by the expectation of a local Langlands correspondence, one would then expect to be able to attach analogous local factors to pairs of representations $\pi, \pi'$ as above. This can be done, and with the notation introduced above, there is a global functional equation

\[(A.2.4) \quad L(s, \Pi \otimes \Pi') = \prod_v \varepsilon_v(s, \Pi_v \otimes \Pi'_v, \psi_v)L(1-s, \Pi'_{\vee} \otimes \Pi''_{\vee})\]

already encountered in (1.4.2). Moreover, the local epsilon factors of pairs satisfy the following analogue of (A.2.2):

\[(A.2.4) \quad \varepsilon(s, \pi \boxplus \pi', \psi_K) = \varepsilon(s, \pi, \psi_K) \cdot \varepsilon(s, \pi', \psi_K),\]

with notation as in (1.4).

The two constructions of these local factors, respectively in [JPSS] and [Sh], characterize them in terms of local harmonic analysis on general linear groups over $K$, or more precisely in terms of local functional equations generalizing those found in Tate’s thesis for $n = m = 1$. However, the two characterizations look quite different, and in both cases apply only when $\Pi_v$ and $\Pi'_v$ are generic, as is automatically the case when they arise as local components of cuspidal automorphic representations. In the general case, local factors can be defined ad hoc using the classification of all representations via induction from generic representations.

With these preliminaries out of the way, we can now explain some results proved by Henniart long before [HT] and [He5], which are used in a crucial way in both proofs.
Henniart’s numerical local Langlands correspondence [He2], and the splitting principle it implies [He3], have already been invoked (cf. the Introduction and the footnote to (5.3)). The following theorem was mentioned in the introduction:

**Theorem A.2.5 (He4, Théorème 1.1).** — Let $K$ be a non-archimedean local field and $n \geq 2$. Let $\pi_1, \pi_2 \in \mathcal{A}_0(n, K)$. Suppose for all integers $m < n$ and all $\pi' \in \mathcal{A}_0(m, K)$ we have the equality

$$\varepsilon(s, \pi_1 \otimes \pi', \psi_K) = \varepsilon(s, \pi_2 \otimes \pi', \psi_K).$$

Then $\pi_1$ and $\pi_2$ are equivalent.

As noted in the introduction, this theorem implies in rather straightforward fashion that there is at most one family of local correspondences satisfying properties (0.1)-(0.8). The key property (0.8) is obtained in (1.4), for representations induced from characters, from a global identity of $L$-functions with functional equation. In the setting of (1.4), this yields the equality

$$(A.2.6) \prod_{w \in S} \gamma_w(s, \Pi(\chi)_w \otimes \Pi(\chi')_w, \psi_w) = \prod_{w \in S} \gamma_w(s, \text{Ind}_{F_2,w/F_1,w} r_\ell(\chi_w) \otimes \text{Ind}_{F_2,w/F_1,w} r_\ell(\chi'_w), \psi_w).$$

Here $S$ is the finite set of primes where the data are ramified (including all places $w$ at which either of the local $\varepsilon$-factors is non-trivial and all places where one doesn’t know a priori that $L_w(s, \Pi(\chi) \otimes \Pi(\chi')) = L_w(s, \text{Ind}_{F_2/F_1} r_\ell(\chi) \otimes \text{Ind}_{F_2/F_1} r_\ell(\chi'))$), and

$$\gamma_w(?) = \frac{\varepsilon_w(s, ?, \psi_w)L_w(1 - \frac{?}{2})}{L(s, ?)}.$$

In particular, the place of interest $v$, at which $F_v = K$, belongs to $S$. Using an argument originating in [De2], and applied in the automorphic setting in [He1], one shows that one can twist by characters highly ramified at all $w \in S - \{v\}$ to simplify all the $\varepsilon$ factors on both sides except for the one at the place $v$ of interest, at which $F_v = K$. It then becomes obvious that the $\varepsilon$ factors in (A.2.6) away from $v$ match on the two sides. A weight argument serves to eliminate the local $L$-factors in (A.2.6), and all that remains is the equality (1.4.4) of $\varepsilon$ factors at $v$.

**References**


SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005


[HT2], Deformations of automorphic Galois representations, manuscript, 1998.


p-ADIC AUTOMORPHIC FORMS ON REDUCTIVE GROUPS

by

Haruzo Hida

Abstract. — In these lecture notes, we will prove vertical control theorems for ordinary p-adic automorphic forms and irreducibility of the Igusa tower over unitary and symplectic Shimura varieties.

Résumé (Formes automorphes p-adiques sur les groupes réductifs). — Nous démontrons le contrôle vertical pour les formes automorphes ordinaires p-adiques et l’irreductibilité de la tour d’Igusa pour les variétés de Shimura symplectique et unitaire.

Contents

1. Introduction ....................................................148
2. Elliptic Curves ................................................155
3. Vertical Control for Elliptic Modular Forms ....................166
4. Hecke Equivariance of the Eichler-Shimura Map ..............170
5. Moduli of Abelian Schemes .....................................182
6. Shimura Varieties .............................................197
7. Formal Theory of Automorphic Forms .........................207
8. Vertical Control for Projective Shimura Varieties ..............219
9. Hilbert Modular Forms ..........................................230
10. Igusa Towers ................................................241
References ........................................................252

Key words and phrases. — p-adic automorphic form, Hecke algebra, Shimura variety, Igusa tower.

Ten lectures at Centre Émile Borel, UMS 839, in Institut Henri Poincaré (CNRS/UPMC), Paris, France, from March 29 to June 14 in 2000. The author is partially supported by the following grant from NSF: DMS 9988043 and DMS 0244401.

© Astérisque 298, SMF 2005
1. Introduction

Let $p$ be a prime. What I would like to present in this series of lectures is the theory of families of $p$-ordinary $p$-adic (cohomological) automorphic forms on reductive groups. After going through basics of the theory of $p$-adic automorphic forms, we would like to study

1. Vertical Control Theorem (VCT: construction of $p$-adic families);
2. $p$-adic $L$-functions (in Symplectic and Unitary cases);
3. Galois representations;
4. the Iwasawa theoretic significance of $p$-adic $L$-functions.

1.1. Automorphic forms on classical groups. — Let $G/\mathbb{Z}$ be an affine group scheme whose fiber over $\mathbb{Z}_p$ is a classical Chevalley group; so, unitary groups are included (dependent on the choice of $p$). Take a Borel subgroup $B$ and its torus $T$. When $G$ is split over $\mathbb{Q}$, we may embed $G$ into $GL(n)/\mathbb{Q}$. Let $B$ be the Borel subgroup (we can take it to be the group of upper triangular matrices in $G$). Let $T$ be the group of diagonal matrices. We have a splitting $B = T \ltimes U$ for the unipotent radical $U$ of $B$. On the quotient variety $G/U$ (which is a $T$-torsor over the projective flag variety $G/B$), $T$ acts by $gUt = gtU$, and hence $T$ acts on the structure sheaf $\mathcal{O}_{G/U}$ by $t\phi(gU) = \phi(gtU)$. This action gives rise to an order on $X(T) = \text{Hom}(T, \mathbb{G}_m)$ so that the positive cone in $X(T)$ is made of $\kappa \in X(T)$ such that the $\kappa$-eigenspace $L(\kappa)$ on the global sections of $\mathcal{O}_{G/U}$ is non-trivial. We then have a representation $L(\kappa; A) = L_G(\kappa; A)$ on $L(\kappa)$ given by $\phi(gU) \mapsto \phi(h^{-1}gU)$ for $h \in G(A)$, as long as $T$ is split over a ring $A$. When $G = SL(2)$, $T \cong \mathbb{G}_m$, $X(T) \cong \mathbb{Z}$ by $\kappa \leftrightarrow n$ if $\kappa(x) = x^n$, and $L(\kappa; A)$ is the symmetric $\kappa$-th tensor representation of $SL(2)$, which can be realized on the space of homogeneous polynomials of degree $n$ so that $\alpha \in SL(2)$ acts on a polynomial $P(X, Y)$ by $P(X, Y) \mapsto P((X, Y)^{\alpha^{-1}})$.

There are two ways of associating a weight to automorphic forms on $G$: One is to consider the cohomology group $H^d(\Gamma, L(\kappa; A))$ of an appropriate degree $d$ for a given arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, and we call harmonic automorphic forms spanning $H^d(\Gamma, L(\kappa; \mathbb{C}))$ automorphic forms of (topological) weight $\kappa$. This way works well for any classical (or more general reductive) groups.

When the symmetric space of $G$ is isomorphic to a (bounded) hermitian domain $\mathcal{H}$ with origin $0$, like (the restriction of scalar to $\mathbb{Q}$ of) $F$-forms of $Sp$ or $SU(m, n)$ over totally real fields $F$, we have another way to associate a weight to holomorphic automorphic forms. In this case, we have $\mathcal{H} \cong G(\mathbb{R})/C_0$ for the stabilizer $C_0$ of $0$, which is a maximal compact subgroup of $G(\mathbb{R})$. In the simplest case of $SL(2)/\mathbb{Q} = Sp(2)/\mathbb{Q}$, $C_0 = SO_2(\mathbb{R})$ and $\mathcal{H} = H = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ with $G(\mathbb{R})/SO_2(\mathbb{R}) \cong H$ by $g \mapsto g(\sqrt{-1})$. As is well known that $H$ is holomorphically equivalent to the open unit disk in $\mathbb{C}$ by $z \mapsto \frac{z - \sqrt{-1}}{z + \sqrt{-1}}$. 

ASTÉRISQUE 298
The group $G_0$ can be regarded as a group of real points with respect to a twisted complex conjugation in the complexification $C$ of $G_0$. In the case of $SL(2)/Q$, $SO_2(R)$ can be regarded as $S^1$ in $G_m(C)$ by $(\gamma d) \mapsto c\sqrt{-1} + d \in S^1$, and $S^1$ is the set of fixed points of the twisted “complex conjugation”: $x \mapsto \overline{x}^{-1}$ in $G_m(C) = \mathbb{C}^\times$. Generalizing this example, we see that the compact group $U(n)$ is the subgroup of $GL_n(C)$ fixed by the complex conjugation: $x \mapsto x^\overline{1}$. Any holomorphic representation $\rho : C \rightarrow GL(V(\rho))$ gives rise to a holomorphic complex vector bundle $\tilde{V} = (G(R) \times V)/C_0$ by the action $(g, v) \mapsto (gu, u^{-1}v)$ for $u \in C_0$. Since $H$ is simply connected, we can split $\tilde{V} \cong H \times V$ as holomorphic vector bundles; so, we have a linear map $J_\rho(g, z) : V_z \rightarrow V_{g(z)}$ for each given $g \in G(R)$ which identifies the fibers $V_z$ and $V_{g(z)}$ of $\tilde{V}$. Thus we have a function $J_\rho : G(R) \times H \rightarrow GL(V)$ satisfying

(1) (Cocycle Relation) $J_\rho(gh, z) = J_\rho(g, h(z))J_\rho(h, z)$ for $g, h \in G(R)$ ;
(2) (Holomorphy) $J_\rho(g, z)$ is holomorphic in $z$.

When $G = SL(2)$, then $G_0 = SO_2(R) \subset C = \mathbb{C}^\times$ whose irreducible complex representation is given by

$$\begin{pmatrix} \cos(\theta) - \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mapsto \rho \begin{pmatrix} * & * \\ * & d \end{pmatrix} = (ci + d)^k = e^{i\theta}.$$

In this case, $J_\rho(g, z) = (cz + d)^k$. This goes as follows: Split $GL_2(R) = P C_0$ for $P$ made of upper triangular matrices with right lower corner 1. For $z = x + iy$, define $p_z = (\frac{y}{1})$. Then for $g \in SL_2(R)$, write $gp_z = p_{g(z)}u$ with $u \in C_0$, and we have $\rho(u) = \rho(p_{g(z)}u)$ by computation. Indeed, $J(g, z)$ sends $(v, p_z)$ to $(uv, p_{g(z)}u)$.

One can view the complexification $C$ as a real algebraic group; let $T_C$ be a maximal real torus of $C$. To any character $\kappa$ of $T_C$, we can attach a rational representation $L_C(\kappa; C) = \rho_\kappa$ of $C$. Let $V(\kappa) = L_C(\kappa; C)$. For an arithmetic discrete subgroup $\Gamma \subset G(Q)$, a holomorphic automorphic form of (coherent) weight $\kappa$ is a holomorphic function $f : H \rightarrow L_C(\kappa; C)$ satisfying $f(\gamma(z)) = J_\rho(\gamma, z)f(z)$ for all $\gamma \in \Gamma$ (with some additional growth condition if $\Gamma \backslash H$ is not compact). Again the space of holomorphic automorphic forms is trivial unless the weight $\kappa$ is positive (with respect to a fixed Borel subgroup $B$).

Often the complex manifold $\Gamma \backslash H$ is canonically algebraizable, giving rise to an algebraic variety (or a scheme) $X_\Gamma$, called canonical models or Shimura varieties, defined over a valuation ring $W$ in a number field with residual characteristic $p$. At the same time, we can also algebraize the vector bundle $\tilde{V}(\kappa)$ associated to $V(\kappa)$. Thus we often have a coherent sheaf $\omega^{\kappa}_A$ on $X_\Gamma$ giving rise to $\tilde{V}(\kappa)$ after extending scalar to $\mathbb{C}$. The global sections of $H^0(X_\Gamma, \omega^{\kappa}_A)$ for $W$-algebra $A$ are called $A$-integral automorphic forms of weight $\kappa$. Note that, $T_C$ is isomorphic to $T$, because they are maximal tori in the same group $G$. Thus we can and will identify $T$ and $T_C$ (with compatible choice of Borel subgroups $B$ and $B_C = B \cap C$). On $X_\Gamma$, we may regard the $\Gamma$-module $L_C(\kappa; A)$ as a locally constant sheaf associating to an open subset $U \subset X_\Gamma$.

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
sections over $U$ of the covering space $\tilde{L}_G(\kappa; A) = \Gamma \backslash (D \times L_G(\kappa; A))$ over $X_\Gamma$. Here the quotient $\Gamma \backslash (D \times L_G(\kappa; A))$ is taken through the diagonal action. Thus each positive weight $\kappa \in X(T)$ gives two spaces of automorphic forms:

$$H^d(X_\Gamma, L_G(\kappa; A)), \quad H^0(X_\Gamma, L_G(\kappa; A)) = G^\kappa(\Gamma; A).$$

There is (at least conjecturally) a correspondence $\kappa \mapsto \kappa^*$ such that

$$H^0(X_\Gamma, L_G(\kappa; A)) \rightarrow H^d(X_\Gamma, L_G(\kappa^*; \mathbb{C}))$$

by a “generalized Eichler-Shimura isomorphism” which is supposed to be equivariant under Hecke operators. If such equivariance holds, we say that the two modules: the source and the image are equivalent as Hecke modules. In the example of $SL(2)/\mathbb{Q}$, we have $\kappa \in X(T) = X(\mathbb{G}_m) = \mathbb{Z}$ and $\kappa^* = \kappa - 2$ with:

$$G_\kappa(\Gamma; \mathbb{C}) \rightarrow H^1(X_\Gamma, L_{SL(2)}(\kappa - 2; \mathbb{C})) \quad (\Gamma \subset SL_2(\mathbb{Z}))$$

via $f \mapsto$ the cohomology class of $[f(z)(X - zY)^{\kappa-2}dz]$. This is valid if $\kappa \geq 2$.

1.2. $p$-Adic interpolation of automorphic forms. — We would like to interpolate these two sets of spaces $\{H^0(X_\Gamma, L_G(\kappa; \mathbb{C}))\}_\kappa$ and $\{H^d(X_\Gamma, L_G(\kappa; W))\}_\kappa$ when the weights $\kappa$ vary continuously in $\text{Hom}_{\text{pro-\mathbb{p}}}(T(\mathbb{Z}_p), \mathbb{Q}_p^\times)$. On these two spaces, there is a natural action of Hecke operators; so, we want this interpolation to take into account the Hecke operators. To describe our idea of how to interpolate automorphic forms, we write $W$ for the $p$-adic completion of $W$. What we would like to do in the two cases is:

(1) (Universality) Construct a (big) space $V$ which is a compact module over $W[[T(\mathbb{Z}_p)]]$ such that the $\kappa$-eigenspace $V[\kappa]$ contains canonically the space $H^d(X_\Gamma, L_G(\kappa; W))$ in the topological case, resp. $H^0(X_\Gamma/W, \mathbb{C}^\kappa)$ in the coherent case as $W[[T(\mathbb{Z}_p)]]$-modules.

(2) (Hecke operators) Establish a natural action of Hecke operators on $V$, and show the inclusion in (1) is Hecke equivariant.

(3) (VCT) Find an appropriate $W[[T(\mathbb{Z}_p)]]$-submodule $X \subset V$ of co-finite type ($\Leftrightarrow W$-dual is of finite type) such that $X$ is stable under Hecke operators and $X[\kappa]$ is canonically isomorphic, as Hecke modules, to a well-described subspace of automorphic forms of weight $\kappa$ if $\kappa \gg 0$.

The item (3) is called a vertical control theorem of the subspace $X$. Examples of the VCT are given as Theorem 3.2 for elliptic modular forms, Theorem 3.3 for $p$-adic family of elliptic modular forms, Theorem 8.5 for automorphic forms on unitary groups, Theorem 9.1 for Hilbert modular forms and Corollary 9.3 for Hilbert modular Hecke algebras. A more general result on VCT can be found in [H02] and [PAF]. In [H02] page 37 and [GME] 3.2.3, Hecke operators $T$ are defined for a given (geometric) modular form $f$ as a sum $f(T(A_S)) = \sum_\alpha f(A_\alpha/S)$ of the values of $f$ at abelian schemes $A_\alpha$ with a specific isogeny $\alpha : A \rightarrow A_\alpha$ of a given degree. This is perfectly fine if the degree is invertible on the base scheme $S$, but otherwise if $S$ is of characteristic $p$.
and the degree is $p$, one has to replace the sum by the trace from the (possibly purely inseparable) extension of $S$ over which the isogeny is defined (as was originally done for elliptic modular forms in Katz’s definition in [K3] 3.11). Thus the argument proving the control theorem in these works has to be modified slightly. This adjustment will be described in the present lecture notes in 3.1.3, 7.1.6 and 8.2.2. The author is grateful to Eric Urban for his pointing out this error in the above cited works (except for [PAF]) of the author.

We will mainly deal with the coherent case where $G$ admits Shimura varieties which are given as moduli of abelian varieties with PEL structure. However at some point, we need to use some results obtained in the topological case; so, a couple of lectures will be devoted also to describe the situation in topological cases. In any case, I will often suppose for simplicity that $G$ to be $U(m, n)$ or its $F$-inner forms over a totally real field $F$, although we also give expositions for $GSp(2g)$ from time to time.

In the coherent case, we shall define $V$ to be the space of formal functions on an formal pro-scheme, called the Igusa tower, classifying abelian schemes with a level $p^\infty$ structure in addition to a PEL structure outside $p$. We will prove the vertical control for the space $X = V^{\text{ord}}$ of nearly $p$-ordinary automorphic forms and prove that its $W$-dual $\text{Hom}_W(V^{\text{ord}}, W)$ is $W[[T(Z_p)]]$-projective of finite type.

Actually, we have for any classical group a good definition of nearly $p$-ordinary cusp forms, that is, a cusp form is called nearly $p$-ordinary if it has the property that the Newton polygon of the Hecke polynomial at $p$ is equal to the hypothetical Hodge polygon mechanically constructed out of the weight $\kappa$ (of the motive attached to the cusp form). We can prove that the Newton polygon is always on or above the Hodge polygon (without recourse to hypothetical motives); so, a nearly $p$-ordinary form has minimal possible Newton polygon (see Section 4).

One would expect that $\text{Hom}_W(V^{\text{ord}}, W)$ should be $W[[T(Z_p)]]$-projective of finite rank if $G$ is associated to a bounded hermitian domain. Contrary to this, when we deal with the group like $GL(n)$ ($n > 2$), the module $\text{Hom}_W(V^{\text{ord}}, W)$ is of finite type over $W[[T(Z_p)]]$, but it is known to be of torsion. Natural questions are:

(Q1) When can one expect that the space $V^{\text{ord}}$ is $W[[T(Z_p)]]$-coprojective (that is, its $W$-dual is projective)? What is the (expected) minimal value of $\kappa$ at which the vertical control holds? What happens if one specializes to a very low weight? If $V^{\text{ord}}$ is co-torsion, what is the Krull dimension of the $W[[T(Z_p)]]$-module $\text{Hom}_W(V^{\text{ord}}, W)$? What is its characteristic power series if $\text{codim}(V^{\text{ord}}) = 1$ in $\text{Spec}(W[[T(Z_p)]])$?

It turns out that all these questions are quite arithmetic, as we will see it in the course. In the elliptic modular case, the lowest weight where VCT holds is 2. However, as Buzzard and Taylor studied, there is a good criterion via Galois representations to guarantee the limit at weight 1 to be a true modular form (not just $p$-adic), which
played an important role in their proof of the Artin conjecture for some icosahedral cases.

In the simplest example of $SL(2) / \mathbb{Q}$, we take an arbitrary $p$-adically complete $W$-algebra $A = \lim_{\longrightarrow n} A/p^n A$. We consider a test object $(E, \phi_p, \phi_N)/A$ made of an elliptic curve $E$, a level $p^\infty$-structure $\phi_p : \mu_{p^\infty} \hookrightarrow E$ (that is a closed immersion of ind-group schemes) and a level $N$-structure $\phi_N$, like a point of order $N$ (here, an inclusion of $\mathbb{Z}/N\mathbb{Z}$ into the set of $N$-torsion elements $E[N]$ in $E$), all these data being defined over $A$. A $p$-adic modular form $f$ is a functorial rule associating an element of $A$ to a test object $(E, \phi_p, \phi_N)/A$. Thus we have $f(E, \phi_p, \phi_N) \in A$, and for each $p$-adically continuous $W$-algebra homomorphism $A \xrightarrow{\rho} B$,

$$f((E, \phi_p, \phi_N)/A, \rho \times B) = \rho(f(E, \phi_p, \phi_N)).$$

A $p$-ordinary modular form which is an eigenform of $T(p)$ has by definition a $p$-adic unit eigenvalue for $T(p)$. In general, $p$-ordinary modular forms are linear combinations of such $p$-ordinary eigenforms (we will give a more conceptual definition in the text). The evaluation of $f$ at the Tate curve Tate($q$) at the cusp infinity yields the $q$-expansion:

$$f(q) = f(\text{Tate}(q)\infty, \phi_p^{\text{can}}, \phi_N^{\text{can}}) = \sum_{n=0}^{\infty} a(n, f)q^n.$$

We can deduce from the irreducibility of the Igusa tower that $V$ is isomorphic to the $p$-adic completion of

$$W[[q]] \cap \left( \sum_{k=0}^{\infty} \sum_{\alpha} G_k(\Gamma_1(Np^n)) \right).$$

Here we have embedded $G_k(\Gamma_1(N))$ into $\mathbb{C}[[q]]$ by the Fourier expansion, writing $q = \exp(2\pi i z)$.

In the topological case of $SL(2)$, $V$ is given by $H^1(\Gamma', \mathcal{C}(\tilde{\Gamma}'/U(\mathbb{Z}_p), W))$, where $\mathcal{C}(\tilde{\Gamma}'/U(\mathbb{Z}_p), W)$ is the space of continuous functions on $\tilde{\Gamma}'/U(\mathbb{Z}_p)$, $\Gamma' = \Gamma_1(N) \cap \Gamma_0(p)$ and $\tilde{\Gamma}$ is the closure of $\Gamma_0(p)$ in $SL_2(\mathbb{Z}_p)$. Then $L(k; W) \hookrightarrow C(\tilde{\Gamma}'/U(\mathbb{Z}_p); W)$ induces a map $H^1(\Gamma', L(k; W)) \hookrightarrow V$.

For any $W[[T(\mathbb{Z}_p)]]$-submodule $X \subset V$ satisfying (VCT), the eigenvalue $\lambda(t)$ of a Hecke operator $t$ on $X$ is algebraic over $W[[T(\mathbb{Z}_p)]]$. In fact, the Hecke algebra $\mathfrak{h}$ in $\text{End}_W[W[[T(\mathbb{Z}_p)]]](X)$ generated by (appropriate) Hecke operators is an algebra over $W[[T(\mathbb{Z}_p)]]$ of finite (generic) rank (or even of torsion). Take an irreducible component $\text{Spec}(\mathfrak{I})$ of $\text{Spec}(\mathfrak{h})$. The operator $t$ projected to $\mathfrak{I}$, written as $\lambda(t)$ (that is, $\mathfrak{h} \rightarrow \mathfrak{I}$ is the projection), can be considered to be an algebraic function (that is, global section of the structure sheaf) on $\text{Spec}(\mathfrak{I})$. In particular, if $P \in \text{Spec}(\mathfrak{I})(W) = \text{Hom}_{W-\text{alg}}(\mathfrak{I}, W)$ with $P|_{W[[T(\mathbb{Z}_p)]]} = \kappa$ for $\kappa \gg 0$, $\lambda(t)(P) = P(\lambda(t))$ is the eigenvalue of $t$ occurring in either $H^0(X_\Gamma, \omega^\kappa)$ or $H^d(X_\Gamma, L(\kappa; W))$. In the simplest case of $SL(2)$, we have $T(\mathbb{Z}_p) = \mathbb{Z}_p^\times = u^{\mathbb{Z}_p} \times \Delta$ for a finite group $\Delta$. Thus $W[[T(\mathbb{Z}_p)]] = \Lambda[\Delta]$ for
\[ \Lambda = W[[u^{\mathbb{Z}_p}]] \cong W[[X]] \] (a formal power series ring) via \( u^s \mapsto (1 + X)^s = \sum_{n=0}^{\infty} \binom{s}{n} X^n \).

Note that

\[ \kappa((1 + X)^s) = \kappa(u^s) = u^\kappa s = (1 + X)^s|_{X = u^{-1}}. \]

The algebra homomorphism \( \kappa : \Lambda \to W \) is the “evaluation” at \( X = u^\kappa - 1 \)!

Thus if \( \mathcal{I} = \Lambda, \lambda(T(n))(\kappa) = \lambda(T(n))(u^\kappa - 1) \) (viewing \( \lambda(T(n)) \) as a power series) gives a \( p \)-adic analytic interpolation of Hecke eigenvalues. In general, we get the \( p \)-adic interpolation of Hecke operators parameterized by \( \text{Spec}(\mathcal{I}) \).

**1.3. \( p \)-Adic Automorphic \( L \)-function.** — Since the specialization \( \lambda(T(n))(u^\kappa - 1) \) is the Hecke eigenvalue occurring in the space of cusp forms, it can be considered as a complex number uniquely (by fixing embeddings \( i_\infty : \overline{\mathbb{Q}} \to \mathbb{C} \) and \( i_p : \overline{\mathbb{Q}} \to \mathbb{Q}_p \)). Thus we can think of automorphic \( L \)-functions \( L(s, \lambda(\kappa)) \) made out of such eigenvalues; for example, the modular Hecke \( L \)-function of \( GL(2) \):

\[ L(s, \lambda(\kappa)) = \sum_{n=1}^{\infty} i(\lambda(T(n))(u^\kappa - 1))n^{-s}, \]

writing \( i = i_\infty \circ i_p^{-1} \). Supposing that \( L(m, \lambda(\kappa)) \) for a fixed integer \( m \) has rationality (up to a transcendental factor or a period \( \Omega(\kappa) \)), a natural question we then ask is:

(Q2) **Is it possible to interpolate \( p \)-adically the value \( \frac{L(m, \lambda(\kappa))}{\Omega(\kappa)} \)? Is it possible to find \( L_\lambda \in \mathcal{I} \) such that \( L_\lambda(\kappa) = L(m, \lambda(\kappa)) \) for \( \kappa \gg 0 \)?

This problem of course involves a subtle question of how to normalize the factor \( \{\Omega(\kappa)\}_\kappa \) in the aggregate (varying \( \kappa \in X(T) \)) to get an “optimal” integrality; so, it is more involved than proving rationality (see Section 9 for some examples and [H96] for a general theory). Once we are successful in constructing canonical \( p \)-adic \( L \)-functions, we could ask more specifically:

(Q3) **When is the \( p \)-adic \( L \)-function analytic? Where could it have singularity? If there is a singularity, what is the residue?**

See [H96] for some examples and conjectural discussions on these questions.

**1.4. Galois Representations.** — Once an irreducible component \( \mathcal{I} \) of the Hecke algebra is given, one would expect to have a Galois representation \( \rho_\lambda : \text{Gal}(\overline{\mathbb{Q}}/F) \to GL_n(\mathcal{I}) \) such that the Hecke polynomial at \( \ell \neq p \) gives rise to the characteristic polynomial of the Frobenius element. We could partially and conjecturally answer the question (Q3) that the associated \( p \)-adic \( L \)-function has singularity at \( P \) if the Galois representation specialized at \( P \), that is, \( \rho_\lambda \mod P \) contains the trivial representation (a \( p \)-adic Artin conjecture, See [H96]). We then further ask:

(Q4) **For a given representation \( \rho_\lambda \) as above, is there any good way to associate a Selmer group \( \text{Sel}(\rho_\lambda) \) so that the characteristic element in \( \mathcal{I} \) of the Selmer group should be equal to the \( p \)-adic \( L \)-function or its numerator?**
See [MFG] Chapter 5 for a general description of Selmer groups. If this is affirmative, then this would describe the zero-set of the $p$-adic $L$-function. Related to this, VCT is often useful to identify the nearly $p$-ordinary Hecke algebra with the nearly $p$-ordinary universal deformation ring if at one weight the deformation ring with the given weight is identified with the Hecke algebra of the specific weight (see for example, [HM] Section 4.3). The argument proving VCT often yields another type of control theorem: so-called the horizontal control theorem (HCT), giving a precise description of the behavior of a Hecke algebra if one add primes outside $p$ to the level of the Hecke algebra. This horizontal control (HCT) is used in the case of $GL(2)$ to construct the Taylor-Wiles systems, which in turn proves the identification of the Hecke algebra of a specific weight with the deformation ring. See [MFG] Chapter 3 and [GME] Chapter 3 for these topics.

1.5. Plan of the lectures. — I will try to answer some of these questions in the lecture in some specific cases in a concrete way and in some other cases conjecturally. Here is a plan:

(1) In a first few lectures in April, 2000 (Sections 1-3), I will recall the theory in the elliptic modular case with some proofs as a prototype of the theory and basic properties of nearly ordinary automorphic forms on general groups.

(2) Lectures in May, 2000 (Sections 4-8) will be devoted to prove the VCT for unitary groups. I will describe the proof in the cocompact case in details (and touch briefly the non-cocompact case taking Hilbert modular varieties as an example: Section 9).

(3) Lectures in June, 2000 would first discuss applications of VCT and the $q$-expansion principle in the Hilbert and the elliptic modular cases (Section 9), and in Section 10, we shall give a sketch of a proof of the $q$-expansion principle of $p$-adic automorphic forms for split symplectic groups and quasi-split unitary groups (acting on a tube domain).

Some of the papers and preprints of mine related to these subjects can be downloaded from my web site: www.math.ucla.edu/~hida.

Although we have tried to give details of the proofs of the material described above in these notes, many results have to be taken for granted here in these notes. The book [PAF] covers similar materials with more details and contains a proof (different from the one presented in Section 10 of these notes) of the irreducibility of the Igusa tower over the mod $p$ canonical models (in a more general setting).

The author wishes to thank the audiences of the lectures for their interest and patience and the organizers of the automorphic semester at l’institut Henri Poincaré for their invitation.
2. Elliptic Curves

In this lecture, I try to sketch a proof of the VCT in the elliptic modular case. There are several different approaches:

(1) Through the moduli theory of elliptic curves; this is what we do ([H86a] and [GME] Chapter 3).

(2) Through studying of topological cohomology groups and jacobians of modular curves. This way has an advantage of producing at the same time Galois representations into $GL_2(I)$, where $I$ is a big ring (which is finite and often flat over $W[[X]]; [H86b]$).

(3) Through the theory of $p$-adic Eisenstein measures and $p$-adic Rankin convolution theory. This method was found by A. Wiles and explained in the elliptic modular case in my book: [LFE] Chapter 7.

(4) As an application of the identification of Hecke algebras and universal Galois deformation rings at many different weights (done by Wiles and Taylor). This method is exposed in my book [MFG].

We follow the first method. A shorter proof than the original in 1986 can be found in my book [PAF] Chapter 3 and also in my Tata lecture notes: Control Theorems and Applications, which can be downloaded from my web site. Also Chapter 3 of the book [GME] contains a more down-to-earth description of the proof.


2.1.1. Definition of Elliptic Curves. — For a given scheme $S$, a proper smooth curve $f : E \to S$ is called an elliptic curve if it satisfies the following conditions:

(E1) $E$ has a section $0 = 0_E \in E(S)$ (thus $f \circ 0 = 1_S$);

(E2) $\dim_S E = 1$, and $E$ is geometrically connected (this means that each geometric fiber of $E$ over a geometric point is connected and of dimension 1);

(E3) $f_*\Omega_{E/S}$ (equivalently $R^1 f_* \mathcal{O}_E$ by Grothendieck-Serre duality) is locally free of rank 1 (genus = 1).

There is no harm to assume that $S$ is connected, as we do from now on. For any $S$-scheme $\phi : T \to S$, the fiber product $E_T = E \times_S T$ is again an elliptic curve with the zero section $0_T = 0_E \times 1_T$. For two elliptic curves $E$ and $E'$ over $S$, an $S$-morphism $h : E \to E'$ is always supposed to take $0_E$ to $0_{E'}$.

2.1.2. Cartier Divisors. — A closed subscheme $D \subset E$ is called an effective Cartier divisor (relative to $S$) on $E$ if $f_*\mathcal{O}_D = f_*(\mathcal{O}_E/I(D))$ given by an invertible sheaf of ideals $I(D)$ is $S$-flat (so locally free). We define $\mathcal{L}(D) = I(D)^{-1}$ and put $\deg(D) = \deg(\mathcal{L}(D)) = \text{rank}_S f_*(\mathcal{O}_D)$. In particular, the $0$ section gives rise to a divisor $[0]$ of degree 1 given by $\mathcal{O}_{[0]} \cong \mathcal{O}_S$. We then think of $I(m[0]) = I([0])^m$ and $\mathcal{L}(m[0]) = \ldots$
\[ I(m[0])^{-1} \text{ for } m \in \mathbb{Z}. \] The line bundle \( \mathcal{L}(m[0]) \) can be regarded as the sheaf of meromorphic functions on \( E \) with sole singularity at 0 having pole of order equal to or less than \( m \) at 0.

Write \( \text{Div}^r(E/S) \) for the set of all degree \( r \) effective divisors relative to \( S \). The association \( T \mapsto \text{Div}^r(E_T/T) \) is a contravariant functor by pull-back of divisors \( D/E' \mapsto D_{E'/E} = D \times_E E'_T \).

If \( S = \text{Spec}(k) \) for an algebraically closed field \( k \), \( k\)-rational effective divisors can be identified with positive linear combinations of points on \( E(k) \). We have \( \text{deg}(\sum \rho m_P[P]) = \sum \rho m_P \). We can thus consider the group \( \text{Div}(E/k) \) of all formal linear combinations (including negative coefficients) of points on \( E \). Then \( \text{deg} : \text{Div}(E/k) \to \mathbb{Z} \) is a well defined homomorphism given by the above formula.

2.1.3. Picard Schemes. — For any scheme \( X \), we define \( \text{Pic}(X) \) as the set of all isomorphism classes of invertible sheaves on \( X \). The association \( X \mapsto \text{Pic}(X) \) is a contravariant functor by the pull-back of invertible sheaves, and \( \text{Pic}(X) \) is a group by tensor product. We define, for each \( S \)-scheme \( \phi : T \to S \)

\[ \text{Pic}_{E/S}(T) = \text{Pic}(E_T)/\phi^* \text{Pic}(T). \]

We can extend the degree map to \( \text{deg} : \text{Pic}_{E/S}(T) \to \mathbb{Z}^{\pi_0(T)} \) for the set \( \pi_0(T) \) of connected components. Indeed, for any algebraically closed field \( k \) and a geometric point \( s : \text{Spec}(k) \to T \), the fiber \( E(s) = E \times_{S,\phi,s} s = E_T \times_T s \) is an elliptic curve over the field \( k \) and \( \text{deg}(\mathcal{L}) = \text{deg}(\mathcal{L}(s)) \) for the pull back \( \mathcal{L}(s) \) at \( s \), which is well defined.

By this fact, we can define

\[ \text{Pic}_{E/S}(T) = \{ \mathcal{L} \in \text{Pic}_{E/S}(T) \mid \text{deg}(\mathcal{L}) = r \text{ for all connected component of } T \}. \]

Here is Abel’s theorem (e.g. [GME] 2.2.2):

\[ (\text{Abel}) \quad \text{Pic}_{E/S}(T) \cong E(T) = \text{Hom}_S(T, E) \text{ by } \mathcal{L}(P) \otimes \mathcal{L}([0])^{r-1} \mapsto P \]

Thus an elliptic curve is a group scheme with the identity 0. If \( \phi : C \to C' \) is a non-constant \( S \)-morphism of two smooth geometrically connected curves, \( |\phi^{-1}(s)| \) is constant for geometric points \( s \) of \( C' \), that is, \( \phi_* \mathcal{O}_{C} \) is locally free of finite rank. We write this number as \( \text{deg}(\phi) \). Thus \( \bigwedge^{\text{deg}(\phi)} \phi_* \mathcal{O}_C \) is an invertible sheaf on \( C' \). If \( \phi : E \to E' \) is an \( S \)-morphism of elliptic curves, by our convention, \( \phi \) takes \( 0_E \) to \( 0_{E'} \), and hence, at the side of the Picard scheme, it is just \( \mathcal{L} \mapsto \bigwedge^{\text{deg}(\phi)} \phi_* \mathcal{L} \); so, obviously \( \phi \) is a homomorphism of group schemes.

2.1.4. Invariant Differentials. — By (E3), for a dense affine open subset \( \text{Spec}(A) \) of \( S \), \( H^0(E, \Omega_{E/A}) = \omega_0 \) for a 1-differential \( \omega \). For each point \( P \in E(S) \), \( T_P : x \mapsto x + P \) gives an automorphism on \( E \). Since we can therefore bring any given cotangent vector at 0 to \( P \) isomorphically to a cotangent vector at \( P \), each cotangent vector at 0 extends to a global section of \( \Omega_{E/S} \). Thus \( T_P \omega = \omega \) (cf. [GME] 2.2.3).
2.1.5. Classification Functors. — An important fact from functorial algebraic geometry is: we can associate to each $S$-scheme $X$, a contravariant functor $X : S\text{-SCH} \to \text{SETS}$ such that $X(T) = \text{Hom}_S(T, X)$. This association is fully faithful; in other words, writing $CTF$ for the category of contravariant functors from $S$-schemes to $\text{SETS}$, we have $\text{Hom}_S(X, Y) \cong \text{Hom}_{CTF}(X, Y)$ by $X \mapsto Y \mapsto \phi(T) : X(T) \to Y(T)$ given by $\phi(T)(T \dashrightarrow X) = \phi \circ f$ (e.g. [GME] Lemma 1.4.1). This is intuitively clear because an algebraic variety is just a function associating to each ring $R$ its $R$-integral points $X(R) = X(\text{Spec}(R))$. I leave the verification of this to the reader as an exercise (the inverse is given by $\text{Hom}_{CTF}(X, Y) \ni F \mapsto F(X)(1_X)$ where $F(X) : X(X) \to Y(X) = \text{Hom}_S(X, Y)$).

Here is an example of how to use the faithfulness: Let $N$ be a positive integer. Since $E(T)$ is a group, $x \mapsto Nx$ gives a functorial map $N(T) : E(T) \to E(T)$; so, an endomorphism of elliptic curves $N : E \to E$. We define its kernel $E[N] = E \times_{E,N,A} S$:

$$
\begin{array}{ccc}
E[N] & \longrightarrow & E \\
\downarrow & & \downarrow \\
S & \longrightarrow & N \\
0 & \longrightarrow & E.
\end{array}
$$

It is clear that $E[N](T) = \text{Ker}(N(T))$. It is known that $\deg N = N^2$ and if $N$ is invertible over $S$, $E[N](k) \cong (\mathbb{Z}/N\mathbb{Z})^2$ for all algebraically closed fields $k$ with $\text{Spec}(k) \hookrightarrow S$.

We consider the following functor:

$$
P_{\Gamma_1(N)}(A) = [(E, P, \omega)/A]
$$

from the category $\text{ALG}$ of $\mathbb{Z}$-algebras into $\text{SETS}$, where $\omega$ is a nowhere vanishing invariant differential, $P$ is a point of order exactly $N$, that is, $m \mapsto mp$ induces an isomorphism $\mathbb{Z}/N\mathbb{Z}/A \hookrightarrow E$ of group schemes defined over $A$ and $[\cdot] = \{\cdot\}/ \cong$ is the set of all isomorphism classes of the objects inside the brackets. Here $\mathbb{Z}/N\mathbb{Z}$ as a group functor associates with $T$ the constant group $(\mathbb{Z}/N\mathbb{Z})^{\tau_A(T)}$.

Therefore $\mathcal{O}_{(\mathbb{Z}/N\mathbb{Z})/S} = \bigoplus_{\mathbb{Z}/N\mathbb{Z}} \mathcal{O}_S$; so, the structure sheaf of $\mathbb{Z}/N\mathbb{Z}$ is free of finite rank $N$. Such a group scheme is called a locally free group scheme (of rank $N$). There is another example: Start with the multiplicative group scheme $G_m$ (as a functor $G_m(A) = A^\times$ and as a scheme $\text{Spec}(\mathbb{Z}[t, t^{-1}])$, we consider the kernel $\mu_N$ of $N : x \mapsto x^N$ as a functor $\mu_N(A) = \{\zeta \in A \mid \zeta^N = 1\}$ and as a scheme

$$
\mu_N = \text{Spec}(\mathbb{Z}[t]/(t^N - 1)) = \text{Spec}(\mathbb{Z}[(\mathbb{Z}/N\mathbb{Z})]).
$$

Then $\mu_N$ is a locally free group scheme of rank $N$. If $N > 1$, it is not isomorphic to $(\mathbb{Z}/N\mathbb{Z})$, since for any prime $p$, $\mu_p(\mathbb{F}_p) = \{1\}$ but $(\mathbb{Z}/p\mathbb{Z})(\mathbb{F}_p) = \mathbb{Z}/p\mathbb{Z}$ for a prime $p$.

We consider a version of the functor $P_{\Gamma_1(N)}$ defined as follows:

$$
P_{\Gamma_1(N)}(A) = [(E, \phi_N : \mu_N \hookrightarrow E[N], \omega)/A]
$$
2.1.6. Cartier Duality. — The two functors $\mathcal{P}_{\Gamma_1(N)}$ and $\mathcal{P}'_{\Gamma_1(N)}$ are isomorphic by the following theory of Cartier duality: If $G$ is a locally free group scheme of rank $N$ over $S$, there exists a group scheme $\mathcal{G}/S$ such that $\mathcal{G}(T) = \text{Hom}_T(G_T, \mathbb{G}_m/T) = \text{Hom}_T(G_T, \mathbb{G}_m/T)$, where $\mathbb{G}_m/S = \mathbb{G}_m \times S$ and $\mu_N/S = \mu_N \times S$ over $\text{Spec}(\mathbb{Z})$. We have $\hat{G} \cong G$ in an obvious manner, and $\mathbb{Z}/\mathbb{Z} = \mu_N$ by $\zeta(m) = \zeta^m$ for $\zeta \in \mu_N(A)$ and $m \in (\mathbb{Z}/\mathbb{Z}))(A)$.

Let $E/S$ be an elliptic curve. The section $0 : S \to E$ induces a section of $f^* : \text{Pic}(S) \to \text{Pic}(E)$; so, we have a splitting:

\[
\text{Pic}(E_T) = f_T^* \text{Pic}(T) \oplus \text{Ker}(0_T^*) \quad \text{and} \quad \text{Ker}(0_T^*) = \text{Pic}_{E/S}(T),
\]

regarding $\text{Pic}_{E/S}(T)$ as a set of isomorphism classes of invertible sheaves whose restriction to $0$ is trivial, that is, $0^*\mathcal{L}$ is isomorphic to $\mathcal{O}_S$.

Let $P \in \text{Ker}(\pi)$ for a non-constant $S$-morphism $\pi : E \to E'$. Then

\[
\pi^* : E' \cong \text{Pic}_{E'/S} \longrightarrow \text{Pic}_{E/S} = E
\]
is an $S$-homomorphism. We are going to show that $\text{Ker}(\pi^*) = \text{Ker}(\pi)$. Take $\mathcal{L} \in \text{Ker}(\pi^*)$, and take an open covering $E' = \bigcup_i U_i$ such that $\mathcal{L}|_{U_i} = f_i^{-1}\mathcal{O}_{U_i}$. Since $0^*\mathcal{L} = \mathcal{O}_S$, one can assume $f_i \circ 0_{E'} = f_j \circ 0_{E'}$ for all $i \neq j$ on $U_i \cap U_j$. Let $h_i = f_i \circ \pi$; we have $\pi^*\mathcal{L}|_{V_i} = h_i^{-1}\mathcal{O}_{V_i}$ for $V_i = \pi^{-1}(U_i)$. Let $P \in (\text{Ker }\pi)(T)$; then on $P^{-1}(V_i) \cap P^{-1}(V_j)$, we have

\[
h_i \circ P = f_i \circ \pi \circ P = f_i \circ 0_{E'} = f_j \circ 0_{E'} = h_j \circ P.
\]

This implies that $h_i \circ P$'s glue to give a global section $h \circ P \in \Gamma(T, \mathcal{O}_T^\times) = \mathbb{G}_m(T)$, getting a homomorphism $\text{Ker}(\pi^*) \rightarrow \text{Ker}(\pi)$, which can be easily verified to be an isomorphism (because twice this operation yields an identity map of $\text{Ker}(\pi^*)$). Since $N^* = N$ as we can see easily, we get $E^*[N] = E[N]$. Writing the pairing as $\langle \cdot, \cdot \rangle : E[N] \times E[N] \rightarrow \mu_{N/S}$, we get $\langle \phi(P), Q \rangle = \langle P, \phi^*(Q) \rangle$; so, $\phi \mapsto \phi^*$ is an involution with $\phi^* \circ \phi = \text{deg}(\phi) \geq 0$ (a positive involution).

For a given additive level $N$-structure $\phi_N : \mathbb{Z}/\mathbb{Z} \rightarrow E[N]$, by duality, we get $\pi_N : E[N] \rightarrow \mu_N$ which has a section $\phi_N^*$ well determined modulo $C = \phi_N(\mathbb{Z}/\mathbb{Z})$. Thus $(E/C, \phi_N) : \mu_N \hookrightarrow (E/C)[N], \omega')$ is well defined as an element of $\mathcal{P}_{\Gamma_1(N)}(A)$, where $\omega'$ coincides with $\omega$ at the identity (because the projection $E \rightarrow E/C$ is a local isomorphism; that is, an étale morphism). The inverse: $\mathcal{P} \rightarrow \mathcal{P}'$ is given by

\[
(E', \phi_N^* : \mu_N \hookrightarrow E'[N], \omega') \mapsto (E'' = E'/\text{Im}(\phi_N^*), \phi_N : \mathbb{Z}/\mathbb{Z} \rightarrow E''[N], \omega'')
\]
similarly. Since $(E/C)/\phi_N^*(\mu_N) = E/E[N] \cong E$, we have $\mathcal{P}' \cong \mathcal{P} \cong \mathcal{P}'$ and hence equivalence.

2.2. Moduli of Ordinary Elliptic Curves and the Igusa Tower. — We now study the scheme $Y_1(N)$ representing $\mathcal{P}_{\Gamma_1(N)}$ over $\mathbb{Z}[1/N]$-algebras. This eventually leads us to the vertical control theorems in the elliptic modular cases.
2.2.1. Moduli of level 1 over $\mathbb{Z}[\frac{1}{6}]$. — Hereafter, we assume until Section 3 (for simplicity) that 6 is invertible in any algebra we consider. Let $(E, \omega)_{/A}$ be a couple of an elliptic curve and a nowhere vanishing differential. We choose a parameter $T$ at 0 so that

$$\omega = (1 + \text{higher terms of } T)dt.$$ 

By the Riemann-Roch theorem, $\dim H^0(E, \mathcal{L}(m[0])) = m$ if $m > 0$. We have two morphisms $x, y : E \to \mathbb{P}^1$ such that

1. $x$ has a pole of order 2 at 0 with the leading term $-T^{-2}$ in its Taylor expansion in $T$ (removing constant term by translation);
2. $y$ has a pole of order 3 with leading term $-T^{-3}.$

Out of these functions, we can create bases of $H^0(E, \mathcal{L}(m[0]))$:

- $H^0(E, \mathcal{L}(2[0])) = A + Ax$, $H^0(E, \mathcal{L}(3[0])) = A + Ax + Ay$. This implies that $x$ has a pole of order 2 at 0 and $y$ has order 3 at 0. They are regular outside 0;

- Out of these functions $1, x, y$, we create functions with pole of order $n$ at 0 as follows:

  - $n \leq 4 : 1, x, y, x^2$ (dim = 4)
  - $n \leq 5 : 1, x, y, x^2, xy$ (dim = 5)
  - $n \leq 6 : 1, x, y, x^2, xy, x^3, y^2$ (dim = 6).

Comparing the leading term of $T^{-6}$, one sees that the seven sections $1, x, y, x^2, xy, x^3, y^2$ of $H^0(E, \mathcal{L}(6[0]))$ have to satisfy the following relation:

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6.$$

We can kill in a unique way the terms involving $xy$ and $y$ by a variable change $y \mapsto y + ax + b$. Indeed, by the variable change $y \mapsto y - \frac{a_1 x}{2} + \frac{a_3}{2}$, we get the simplified equation:

$$y^2 = x^3 + b_2 x^2 + b_4 x + b_6.$$

Again a variable change: $x \mapsto x - b_3/3$ simplifies the equation to

$$y^2 = x^3 + c_2 x + c_3.$$

Since $\mathcal{L}(3[0])$ is very ample ($\deg(\mathcal{L}(3[0])) = 3 \geq 2g + 1$), finally making a variable change $2y \mapsto y$ (so now the $T$-expansion of $y$ begins with $-2T^{-3}$), we get a unique equation out of $(E, \omega)_{/A}$:

$$y^2 = 4x^3 - g_2(E, \omega)x - g_3(E, \omega) \text{ for } g_2(E, \omega), g_3(E, \omega) \in A.$$

In other words, $E \subset \mathbb{P}^3_{/A}$ is given by

$$\text{Proj}(A[X, Y, Z]/(ZY^2 - 4X^3 + g_2(E, \omega)XZ^2 + g_3(E, \omega)Z^3)).$$
It is easy to see that this equation gives a smooth curve of genus 1 having \(0 = \infty = (0, 1, 0)\) in \(\mathbb{P}^2\) if \(\Delta = \Delta(E, \omega) = g_2^3 - 27(g_3)^2 \in A^\times\). We recover the differential \(\omega\) by \(dz/y\). This shows that, writing \(R = \mathbb{Z}[\frac{1}{6}, g_2, g_3, \frac{1}{2}]\) for variables \(g_2\) and \(g_3\),

\[
\mathcal{P}_{\Gamma_1(1)}(A) \cong \text{Hom} \mathbb{Z}[\frac{1}{6}-\text{alg}] (R, A) = \mathcal{M}_1(A),
\]

where \(\mathcal{M}_1 = \text{Spec}(R)\) for \(R = \mathbb{Z}[\frac{1}{6}, g_2, g_3, \frac{1}{2}]\). We have the universal elliptic curve and the universal differential \(\omega\) given by

\[
(E, \omega)_{/\mathcal{M}_1} = \left( \text{Proj}(R[X, Y, Z]/(ZY^2 - 4X^3 + g_2XZ^2 + g_3Z^3)), \frac{dx}{y} \right).
\]

For each couple \((E, \omega)_{/A}\), we have a unique \(\varphi \in \mathcal{M}_1(A) = \text{Hom}_S(\text{Spec}(A), \mathcal{M}_1)\) \((S = \text{Spec}(\mathbb{Z}[\frac{1}{6}])\) such that

\[
(E, \omega)_{/A} \cong \varphi^*(E, \omega) = (E, \omega) \times_{\mathcal{M}_1} \text{Spec}(A).
\]

If we change \(\omega\) by \(\lambda \omega\) for \(\lambda \in A^\times = G_m(A)\), the parameter \(T\) will be changed to \(\lambda T\) and hence \((x, y)\) is changed to \((\lambda^{-2}x, \lambda^{-3}y)\). Thus \((E, \lambda \omega)_{/A}\) will be defined by

\[
(\lambda^{-3}y)^2 = 4(\lambda^{-2}x)^3 - g_2(E, \lambda \omega)(\lambda^{-2}x) - g_3(E, \lambda \omega).
\]

This has to be equivalent to the original equation by the uniqueness of the Weierstrass equation, and we have

\[
g_j(E, \lambda \omega) = \lambda^{-2j} g_j(E, \omega).
\]

Again by the uniqueness of the Weierstrass equation, we find that

\[
\text{Aut}((E, \omega)_{/A}) = \{1_E\}
\]
as long as 6 is invertible in \(A\).

2.2.2. Moduli of \(\mathcal{P}_{\Gamma_1(N)}\). — Consider \((E, P, \omega)\) for a point \(P \in E[\ell](A)\) of order \(\ell\) for a prime \(\ell\). We have a unique \(\varphi \in \mathcal{M}_1(A)\) such that

\[
\varphi_E : (E, \omega)_{/A} \cong \varphi^*(E, \omega) = (E, \omega) \times_{\mathcal{M}_1} \text{Spec}(A).
\]

We thus have a commutative diagram:

\[
\begin{array}{ccc}
E & \xrightarrow{\varphi} & \mathcal{M}_1 \\
\varphi_E \downarrow & & \varphi \downarrow \\
\text{Spec}(A) & \xrightarrow{P} & E & \rightarrow & \text{Spec}(A).
\end{array}
\]

Then \(P\) induces a unique morphism \(\varphi_P = \varphi_E \circ P : \text{Spec}(A) \to (E[\ell] - \{0\})(A)\). This shows that, over \(\mathbb{Z}[\frac{1}{6d}]\),

\[
\mathcal{P}_{\Gamma_1(\ell)}(A) \cong \mathcal{P}_{\Gamma_1(\ell)}'(A) \cong (E[\ell] - \{0\})(A).
\]

Similarly, over \(\mathbb{Z}[\frac{1}{6N}]\)

\[
\mathcal{P}_{\Gamma_1(N)}(A) \cong \mathcal{P}_{\Gamma_1(p)}'(A) \cong \left( E[N] - \bigcup_{N > d|N} E[d] \right)(A).
\]
We put $\mathcal{M}_{\Gamma_1(N)} = E[N] - \bigcup_{d \mid N} E[d]$. Thus we have proven

**Theorem 2.1.** — There is an affine scheme $\mathcal{M}_{\Gamma_1(N)} = \text{Spec}(R_{\Gamma_1(N)})$ defined over $\mathbb{Z}[\frac{1}{N}]$ such that

$$\mathcal{P}_{\Gamma_1(N)}(A) \cong \mathcal{P}_{\Gamma_1(N)}'(A) \cong \text{Hom}_{\mathbb{Z}[\frac{1}{N}]}(R_{\Gamma_1(N)}, A) = \mathcal{M}_{\Gamma_1(N)}(A)$$

for all $\mathbb{Z}[\frac{1}{N}]$-algebras $A$. The scheme $\mathcal{M}_{\Gamma_1(N)}/\mathcal{M}_1$ is an étale covering of degree $\varphi(N)$ for the Euler function $\varphi$.

The fact that the covering is étale finite follows from the same fact for $E[N]$ over $\mathbb{Z}[\frac{1}{N}]$ since $E[N](k) \cong (\mathbb{Z}/N\mathbb{Z})^2$ for all algebraically closed fields $k$ with characteristic not dividing $N$. Since $\mathcal{M}_1$ is affine, any finite covering of $\mathcal{M}_1$ is affine.

2.2.3. **Action of $G_m$.** — The group scheme $G_m$ acts on the functor $\mathcal{P}_{\Gamma_1(N)}$ in the following way: $(E, \phi, \omega) \mapsto (E, \phi, \omega)_{/A}$ for $\lambda \in G_m(A)$. This induces an action of $G_m$ on $\mathcal{M}_{\Gamma_1(N)}$ and hence on $R_{\Gamma_1(N)}$.

Here is a general fact on the action of $G_m$. Let $X$ be an $A$-module. Regard $X$ as a functor from $A$-$\text{ALG}$ to the category of $A$-modules $A$-$\text{MOD}$ by $X(B) = X \otimes_A B$. If a group scheme $G_A$ has a functorial action: $G_A \times X \to X$, we call $X$ a schematic representation of $G$. It is known (e.g. [GME] 1.6.5) that if $X$ has a schematic action of $G_{m/A}$, then

$$X = \bigoplus_{\kappa \in \mathbb{Z}} X[\kappa]$$

such that $X[\kappa](B) = \{x \in X \mid \lambda \cdot x = \lambda^\kappa x\}$, that is, $X[\kappa]$ is the eigenspace for the character $G_m(B) \to B^\times$ taking $z \in G_m(B) = B^\times$ to $z^\kappa$.

The action of $G_{m/A}$ on $\mathcal{M}_{\Gamma_1(N)}$ gives rise to a schematic action on $R_{\Gamma_1(N)}$ (because it was defined by functorial action). Thus we can split

$$R_{\Gamma_1(N)/A} = \bigoplus_{\kappa \in \mathbb{Z}} R_\kappa(\Gamma_1(N); A),$$

where on $f \in R_\kappa(\Gamma_1(N); A)$, $G_m$ acts by the character $-\kappa$.

Since $f \in R_\kappa(\Gamma_1(N); A)$ is a functorial morphism:

$$\mathcal{M}_{\Gamma_1(N)}(B) = \mathcal{P}_{\Gamma_1(N)}(B) \to A^1(B) = B,$$

we may regard $f$ as a function of $(E, \phi, \omega)_{/B}$ with $f((E, \phi, \omega)_{/B}) \in B$ satisfying

- (G0) $f((E, \phi, \omega)_{/B}) = \lambda^{-\kappa} f((E, \phi, \omega)_{/B})$ for $\lambda \in B^\times = G_m(B)$;
- (G1) If $(E, \phi, \omega)_{/B} \cong (E', \phi', \omega')_{/B}$, then $f((E, \phi, \omega)_{/B}) = f((E', \phi', \omega')_{/B})$;
- (G2) If $\rho : B \to B'$ is a morphism of $A$-algebras, then $f((E, \phi, \omega)_{/B} \times_B B') = \rho(f((E, \phi, \omega)_{/B}))$.

If a graded ring $A = \bigoplus_j A_j$ has a unit $u$ of degree 1, $A = A_0 \otimes_{\mathbb{Z}} \mathbb{Z}[u, u^{-1}]$ and $\text{Spec}(A) = \text{Spec}(A_0) \times G_m$ by definition; so, $\text{Proj}(A) = \text{Spec}(A)/G_m = \text{Spec}(A_0)$. If $A$ has a unit of degree $n > 0$, then $\text{Proj}(A) = \text{Proj}(A^{(n)}) = \text{Spec}(A_0)$ for $A^{(n)} = \bigoplus_{j=0}^{n} A_j$. 

**SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005**
2.2.4. Compactification

Theorem 2.2

We consider the functor defined over $\mathbb{Z}[\frac{1}{N}]$-$ALG$ given by

$\mathcal{E}_{\Gamma_1(N)} (A) = \left[ (E, \phi_N : \mu_N \hookrightarrow E[N])/A \right] .$

By definition, $\mathcal{E}_{\Gamma_1(N)} = \mathbb{G}_m \backslash \mathcal{P}_{\Gamma_1(N)}$. Since $\text{Proj}(\mathcal{P}_{\Gamma_1(N)})$ gives the quotient by $\mathbb{G}_m$ of $\text{Spec}(R_{\Gamma_1(N)})$ (see [GME] Theorem 1.8.2), we conclude

**Theorem 2.2 (Shimura, Igusa).** — We have an affine curve

$$Y_1(N) = \text{Proj}(R_{\Gamma_1(N)}) = \mathbb{G}_m \backslash \mathcal{M}_{\Gamma_1(N)}$$

defined over $\mathbb{Z}[\frac{1}{N}]$, which is locally free of finite rank over $M_1 = \text{Proj}(R) = \mathbb{P}^1(J) - \{\infty\}$. For all geometric point $\text{Spec}(k)$ of $\text{Spec}(\mathbb{Z}[\frac{1}{N}])$, we have $Y_1(N)(k) = \left[ (E, \phi_N)_k \right]$. The above assertion holds for any $\mathbb{Z}[\frac{1}{N}]$-algebra $A$ in place of $k$ if $N \geq 4$.

Here a “geometric point” means that $k$ is an algebraically closed field. It is well known that $\Gamma_1(N) \backslash \mathfrak{H}$ classifies all elliptic curves with a point of order $N$ over $\mathbb{C}$ for $\mathfrak{H} = \{ z \in \mathbb{C} \mid \text{Im}(z) > 0 \}$; so, we conclude

$$Y_1(N)(\mathbb{C}) = \Gamma_1(N) \backslash \mathfrak{H}.$$  

Thus $Y_1(N)(\mathbb{C})$ is an open Riemann surface.

2.2.4. Compactification. — For any $\mathbb{Z}[\frac{1}{N}]$-algebra $A$, we put

$$G(A) = A[g_2, g_3] = \mathbb{Z}[\frac{1}{N}, g_2, g_3] \otimes A.$$  

Let $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ be the integral closure of $G(\mathbb{Z})$ in the graded ring $R_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$.

To see that $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ is a graded ring, we write $\tilde{r}$ for the non-trivial homogeneous projection of highest degree of $r \in R_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$. If $r \in R_{\Gamma_1(N)/\mathbb{Z}[\frac{1}{6N}]}$ is integral over $G(\mathbb{Z})$, $\tilde{r}$ satisfies an equation $P(X) = X^n + a_1X^{n-1} + \cdots + a_n = 0$ with $a_j \in G(\mathbb{Z})$. Then $\tilde{r}$ satisfies $\tilde{P}(X) = X^n + \tilde{a}_1X^{n-1} + \cdots + \tilde{a}_n = 0$, and $\tilde{r}$ is integral over $G(\mathbb{Z})$. Then by induction of the degree of $\tilde{r}$, we see that $G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}])$ is graded (cf. [BCM] V.1.8).

We put for any $\mathbb{Z}[\frac{1}{6N}]$-algebra $A$

$$G_{\Gamma_1(N)}(A) = G_{\Gamma_1(N)}(\mathbb{Z}[\frac{1}{6N}]) \otimes A = \bigoplus_{k=0}^{\infty} G_k(\Gamma_1(N); A).$$

We then define $X_1(N)/A = \text{Proj}(G_{\Gamma_1(N)}(A))$. By definition, $X_1(N)$ is the normalization of $\text{Proj}(G) = \text{Proj}(G^{(12)}) = \mathbb{P}^1(J) (J = (12g_2)^3/\Delta)$ for $G^{(12)} = \bigoplus_{k=0}^{\infty} G_{12k}$ in $Y_1(N)$. As classically known, $J^{-1}$ has $q$-expansion starting with $q$, that is, $J^{-1} \in q\mathbb{Z}[q]$ (see [IAT] (4.6.1)). Thus the completion of the local ring of $\mathbb{P}^1(J)$ at
the cusp $\infty$ is isomorphic to $\mathbb{Z}_p[[q]]$. Moreover we have the Tate curve (e.g. [GME] 2.5):

$$\text{Tate}(q) = \text{Proj}(\mathbb{Z}[[q]](\frac{1}{p})[X,Y,Z]/(ZY^2 - 4X^3 + g_2(q)XZ^2 + g_3(q)Z^3)),$$

which extends the universal curve over $\mathbb{P}^1(J) - \{\infty\}$ to $\mathbb{P}^1(J)$ locally at the cusp $\infty$.

Since $\text{Tate}(q)(\mathbb{A}[[q]]) \supset (\mathbb{A}[[q]])^\times / q^\mathbb{Z}$ (see [GME] Theorem 2.5.1 (2)), we may think $\text{Tate}(q)$ to be a "quotient" $\mathbb{G}_m/\mathbb{Z}[[q]]/q^\mathbb{Z}$ of $\mathbb{G}_m$; so, it has a canonical level structure $\phi_N^{\text{can}} : \mu_N \hookrightarrow \mathbb{G}_m \rightarrow \text{Tate}(q)$. The Tate couple $(\text{Tate}(q), \phi_N^{\text{can}}/\mathbb{Z}[[q]])$ is a test object over $\mathbb{Z}[[q]]|q^{-1}|$; so, by the universality of $Y_1(N)$, we have a morphism

$$\iota_\infty : \text{Spec}(\mathbb{Z}(\frac{1}{p})[[q]][q^{-1}]) \rightarrow Y_1(N).$$

Since we may regard the Tate curve as a universal formal deformation of a stable curve of genus 1 (with the level structure $\phi_N^{\text{can}}$) centered at the $\mathbb{Z}(\frac{1}{p})$-point represented by an ideal (q) of $\mathbb{Z}(\frac{1}{p})[[q]]$ ([GME] 2.5.2-3), the morphism $\iota_\infty$ is infinitesimal isomorphism centered at the cusp $\infty$ (by the universality of the $Y_1(N)$ and the universality of the Tate curve). Since $X_1(N)$ is the normalization of $\mathbb{P}^1(J)$ in $Y_1(N)$, we conclude that the formal completion along the cusp $\infty$ on $X_1(N)$ is canonically identified with $\mathbb{A}[[q]]$ by $\iota_\infty$. Replacing the level structure $\phi_N^{\text{can}}$ by $\phi_N^{\text{can}} \circ \alpha$ for $\alpha \in SL_2(\mathbb{Z}/N\mathbb{Z})$, basically by the same argument, the local ring at the cusp $\alpha(\infty)$ of $X_1(N)/\mathbb{A}$ is given by $\mathbb{A}[\mu_d][(q^{1/d})]$ for a suitable divisor $d|N$. We need to extend scalar to $\mathbb{A}[\mu_d][(q^{1/d})]$ because the Tate curve $(\text{Tate}(q), \phi_N^{\text{can}} \circ \alpha)$ is only defined over $\mathbb{A}[\mu_d][(q^{1/d})]$ for a suitable divisor $d|N$ dependent on the choice of $\alpha$. This point is a bit technical, and we refer the reader to a more detailed account, which can be found in [AME] Chapter 10 and [GME] 3.1.1. Thus $X_1(N)$ is smooth at the cusps, and moreover $f \in G_k(\Gamma_1(N); \mathbb{A})$ is a function of $(E, \phi_N, \omega)$ satisfying (G0-2) and

$$(G3) \ f(\text{Tate}(q), \phi_N, \omega) \in \mathbb{A}[\zeta_N][(q^{1/N})]$$

for any choice of $\phi_N$ and $\omega$.

Since $\Gamma_1(N) \setminus (\mathcal{F} \cup \mathbb{P}^1(\mathbb{Q}))$ is a smooth compact Riemann surface and is the normalization of $\mathbb{P}^1(J)$ in $Y_1(N)(\mathbb{C})$, we conclude

$$X_1(\mathbb{C}) = \Gamma_1(N) \setminus (\mathcal{F} \cup \mathbb{P}^1(\mathbb{Q})).$$

The space $G_k(\Gamma_1(N); \mathbb{C})$ is the classical space of modular forms on $\Gamma_1(N)$ of weight $k$. Since $\text{Tate}(q)$ is the “quotient” $\mathbb{G}_m/\mathbb{Z}[[q]]/q^\mathbb{Z}$, it has a canonical differential $\omega_{\text{can}}$ induced by $\mathbb{Z}$ identifying $\mathbb{G}_m = \text{Spec}(\mathbb{Z}[t, t^{-1}])$. In particular,

$$f(q) = f(\text{Tate}(q), \phi_N^{\text{can}}, \omega_{\text{can}}) = \sum_{n=0}^{\infty} a(n; f)q^n \; \text{with} \; q = \exp(2\pi i z)$$

coincides with the Fourier expansion of $f$ at the infinity if $f \in G_k(\Gamma_1(N); \mathbb{C})$.

2.2.5. Hasse Invariant. — Let $A$ be a ring of characteristic $p$ and $(E, \omega)$ be an elliptic curve over $S = \text{Spec}(A)$. On each affine open subset $U = \text{Spec}(\Gamma(U, \mathcal{O}_E))$ in $E$, the Frobenius endomorphism $x \mapsto x^p$ induces a morphism $F_{\text{abs}} : U \rightarrow U$. These glue each other to the absolute Frobenius endomorphism $F_{\text{abs}} : E_A \rightarrow E_A$. Note here that
$F_{\text{abs}}$ acts non-trivially on the coefficient ring $A$. We can define the relative Frobenius map: $E \to E^{(p)} = E \times_{S,F_{\text{abs}}} S$ by $F_{\text{abs}} \times_S f$ for the structure morphism $f : E \to S$. This relative Frobenius is the classical map taking homogeneous coordinates of $E$ to their $p$-powers.

Let $T_{E/S}$ be the relative tangent bundle; so, its global section $H^0(E, T_{E/S})$ is the $A$-dual of $H^0(E, \Omega_{E/S})$, and $H^0(E, T_{E/S})$ is spanned by a dual base $\eta = \eta(\omega)$. One can identify $H^0(E, T_{E/S})$ with the module of $\mathcal{O}_S$-derivations $\text{Der}_{\mathcal{O}_S}(\mathcal{O}_{E,0}, \mathcal{O}_S)$ (cf. [GME] 1.5.1). For each derivation $D$ of $\mathcal{O}_{E,0}$, by the Leibnitz formula, we have

$$D^p(xy) = \sum_{j=0}^p \binom{p}{j} D^{p-j} x D^j y = x D^p y + y D^p x.$$ 

Thus $D^p$ is again a derivation. The association: $D \mapsto D^p$ induces an $F_{\text{abs}}$-linear endomorphism $F^*$ of $T_{E/S}$. Then we define $H(E, \omega) \in A$ by $F^* \eta = H(E, \omega) \eta$. Since $\eta(\lambda \omega) = \lambda^{-1} \eta(\omega)$, we see

$$H(E, \lambda \omega) \eta(\lambda \omega) = F^* \eta(\lambda \omega) = F^*(\lambda^{-1} \eta(\omega))$$

$$= \lambda^{-p} F^* \eta(\omega) = \lambda^{-p} H(E, \omega) \eta(\omega) = \lambda^{-p} H(E, \omega) \lambda \eta(\omega) = \lambda^{1-p} H(E, \omega) \eta(\omega).$$

Thus we get

$$H(E, \lambda \omega) = \lambda^{1-p} H(E, \omega).$$

Then $H$ is a modular form of weight $p-1$ defined over $\mathbb{F}_p$:

$$H(E, \omega) \in G_{p-1}(\Gamma_1(1), \mathbb{F}_p).$$

We compute $H(E_{\infty}, \frac{dw}{w})$. The dual of $\frac{dw}{w}$ is given by $D = w \frac{dw}{w}$. The action of $F$ keeps $D$ intact, because $D(w) = w$ (so $D^p(w) = w$). On the tangent space, $F$ acts as identity, and hence $H(E_{\infty}, \omega) = 1$.

An important fact is:

$$H(E, \omega) = 0 \iff E \text{ is super singular}.$$

This is because:

1. If $E_{\mathbb{F}_p}$ is ordinary, then $E[p] \cong \mu_p \times (\mathbb{Z}/p\mathbb{Z})$ over $\mathbb{F}_p$;
2. $\mu_p = \text{Spec}(\mathbb{F}_p[t]/(t^p - 1))$ shares the tangent space with $\mathbb{G}_m$, because they are both of dimension 1 infinitesimally;
3. $F^2 = p$ up to units in the super singular case.

The zero locus of a section of a line bundle is a divisor; hence, on the moduli space, $X_1(N)$ for $p \nmid N$, the points in $X_0(N)(\mathbb{F}_p)$ corresponding to super-singular elliptic curves are finitely many.

2.2.6. Igusa Curves. — Let $W = \mathbb{Z}_p$ and $W_m = W/p^m W$. Fix $N$ with $p \nmid N$. We have a lift of Hasse invariant in $G_{p-1}(\Gamma_1(1); \mathbb{Z}_p)$, which is the Eisenstein series $E$ normalized so that $a(0, E) = 1$. By Von Staut theorem, the $q$-expansion $E(q)$ of $E$ is congruent to 1 modulo $p$; so, $E \mod p$ coincides with $H$. Let $(E, \phi_N)_{/M}$ be the
genus 1 semi-stable curve (completed by appropriate Tate curves at the cusps) over $M = X_1(N)/W$. Let $M_m = X_1(N)/W_m = X_1(N) \times_W W_m$. Define $S_m \subset M_m$ by the open subscheme of $M_m$ on which $E$ is invertible. The scheme $S_m$ does not depend on the choice of the lift $E$, since $E \equiv E' \equiv H \mod p$ guarantees $M_m[\frac{1}{E}] = M_m[\frac{1}{E'}]$ for any other lift $E'$ as long as $p$ is nilpotent in the base ring. We write $S_\infty$ for the formal completion $\lim_m S_m$ of $S$ along $S_1$.

Since we have defined $X_1(N)$ by $\text{Proj}(G_{\Gamma_1(N)})$, the invertible sheaf $\mathcal{O}^k (k > 0)$ associated to the $k$-th graded piece is ample. To see for which $k$, $\mathcal{O}^k$ becomes very ample, we recall that an invertible sheaf of degree $\geq 2g + 1$ over a curve of genus $g$ is very ample by Riemann-Roch theorem (see [GME] Proposition 2.1.4). Computing the genus of $X_1(N)$ (e.g. [GME] Theorem 3.1.2), the invertible sheaf $\mathcal{O}^k$ corresponding to $G_k(\Gamma_1(N); A)$ is very ample if $k > 2$ and $N > 4$ (or $k > 2$). Thus $S_m$ is affine, and $S_m = \text{Spec}(V_{m,0})$ for a $W_m$-flat algebra $V_{m,0}$. We consider the functors

$$E^\text{ord}_\alpha(A) = \left[\left((E, P, \phi_N) / A\right) \right. \quad \text{and} \quad E^\text{ord}_\alpha(A) = \left[\left((E, \mu_{p^q} \longrightarrow E[p^q], \phi_N) / A\right) \right.$$ where $P$ is a point of order $p^q$. Then we see that

$$E^\text{ord}(A) \cong E^\text{ord}_\alpha(A) = \left(\left(E[p^q]\right)^{\text{et}} - E[p^q-1]\right) / S_m(A)$$

for all $W_m$-algebras $A$. We write $T_{m,\alpha}/S_m = \left(\left(E[p^q]\right)^{\text{et}} - E[p^q-1]\right) / S_m$, which is an étale covering of degree $p^{q-1}(p-1)$. It is a classical result of Igusa that $T_{m,\alpha}$ is irreducible (and hence connected; see [GME] 2.9.3), although we do not need this irreducibility here. We will come back to the proof of the irreducibility of the Igusa tower over more general Shimura varieties later in Lecture 10. Since $S_m$ is affine, $T_{m,\alpha}$ is also affine. We write $T_{m,\alpha} = \text{Spec}(V_{m,\alpha})$. We have a tower of $W_m$-flat algebras:

$$V_{m,0} \subset V_{m,1} \subset \cdots \subset V_{m,\alpha} \subset \cdots .$$

These algebras are étale over $V_{m,0}$ and $\text{Gal}(V_{m,\alpha}/V_{m,0}) \cong (\mathbb{Z}/p^\alpha\mathbb{Z})^\times$. Over $V_{m,\alpha}$, we have a canonical isomorphism

$$I_{\text{can}} = \phi_{p^q} : \mathbb{Z}/p^\alpha\mathbb{Z} \cong P_{\alpha} = E[p^q]^{\text{et}}.$$  

We then define $V_{m,\infty} = \bigcup_{\alpha} V_{m,\alpha}$ and

$$V = V_{\Gamma_1(N)} = \lim_m V_{m,\infty} \quad \text{and} \quad V = V_{\Gamma_1(N)} = \lim_m V_{m,\infty}.$$ 

The space $V_{\Gamma_1(N)}$ is the space of $p$-adic modular forms on $\Gamma_1(N)$. By taking the Cartier dual of $\mathbb{Z}/p^\alpha\mathbb{Z} \hookrightarrow E[p^q]$, we may regard $f \in V_{m,\alpha}$ as a rule associating an element of $A$ to $(E, \phi_p : \mu_{p^q} \hookrightarrow E[p^q], \phi_N) / A$ satisfying the conditions similar to (G0-3). Each element $f \in V_{\Gamma_1(N)} \otimes_W A$ for a $W$-algebra $A = \lim_m A/p^mA$ is a function of $(E, \phi_p, \phi_N)$ satisfying the conditions similar to (G0-3) (see [GME] (Gp,1-3) in page 230).
3. Vertical Control for Elliptic Modular Forms

3.1. Vertical Control Theorem. — We have a $p$-divisible module $V_{T_1(N)}$ on which $\text{Gal}(V_{m,\infty}/V_{m,0}) = \mathbb{Z}_p^\times = T(p)$ acts continuously. Here $T = \mathbb{G}_m$. We shall construct a projector $e$ acting on $V$ out of the Hecke operator $U(p)$ commuting with the action of $\mathbb{Z}_p^\times = \text{Gal}(V_{m,\infty}/V_{m,0})$. The important features of $e$ are

- $e = \lim_{n \to \infty} U(p)^n$;
- $V^{\text{ord}} = eV$ has Pontryagin dual which is projective over $W[[\mathbb{G}_m(\mathbb{Z}_p)]]$;
- For any $k \geq 3$, there is a canonical isomorphism
  
  $V^{\text{ord}}[-k] \cong \varepsilon H^0(S, \omega^k \otimes T_p) = e H^0(M, \omega^k \otimes T_p)$ ($T_p = \mathbb{Q}_p/\mathbb{Z}_p$),

where $V[-k] = \{ f \in V | zf = z^{-k}f \forall z \in \mathbb{Z}_p^\times \}$. We hereafter write $H^q_{\text{ord}}$ for $eH^q$ and $G^q_{\text{ord}}$ for $eG_k$.

3.1.1. Axiomatic treatment. — Let $\omega^k = G_{\Gamma_1(N)}(k) = \mathcal{O}(k)$ for the embedding of $X_1(N) = \text{Proj}(G_{\Gamma_1(N)})$ into the projective space. Then $\omega^k = \omega^k \otimes k$. Computing the genus of $X_1(N)$, the Riemann-Roch theorem tell us that $\omega^k$ is very ample if $k \geq 3$ (see [GME] Proposition 2.1.4 and Theorem 3.1.2). Therefore $\omega^k$ is the pull back of $\mathcal{O}(k)$ of the target projective space. Let $(E, \phi_N, \omega)$ be the universal elliptic curve over $Y_1(N)$. For each triple $(E, \phi_N, \omega)$ defined over $A$ (called a test object), we have a unique $\iota : \text{Spec}(A) \to Y_1(N)$ such that $\iota^*(E, \phi_N, \omega) = (E, \phi_N, \omega)$. For each section $f \in H^0(Y_1(N), \omega^k)$, we define

$$\iota^* f = f(E, \phi_N, \omega) \omega^k.$$ 

The function $(E, \phi_N, \omega) \mapsto f(E, \phi_N, \omega)$ satisfies (G0-2). The condition (G3) assures that $f$ extends to $X_1(N)$. This shows

$$H^0(X_1(N)/A, \omega^k) = G_k(\Gamma_1(N); A)$$

for all $\mathbb{Z}[(1/\mathcal{O}_A)]$-algebra $A$.

Let $(E, \phi_p, \phi_N)$ be the universal elliptic curve over $S_m$. Pick a section $f \in H^0(S_m, \omega^k)$. Since $\mu_p \infty$ carries a canonical differential $\omega_{\text{can}} = dt/t$, writing $\mu_p \infty = \text{Spec}(\mathbb{Z}[1/\ell]/(t^{p^n} - 1))$, we may regard $f$ as a function of $(E, \phi_p, \phi_N)$ by $f(E, \phi_p, \phi_N) = f(E, \phi_p, \phi_N, \omega_{\text{can}})$. For each $(E, \phi_p, \phi_N) \in S_m^{\text{ord}}(A)$ for a $W_m$-algebra $A$, we have a unique morphism $\iota : \text{Spec}(A) \to T_m, \infty$ such that $(E, \phi_p, \phi_N) = \iota^*(E, \phi_p, \phi_N)$. Then $\iota^* f$ is just a function of $(E, \phi_p, \phi_N)$ such that $f(E, z^{-1}\phi_p, \phi_N) = z^k f(E, \phi_p, \phi_N)$ for $z \in \text{Gal}(V_{m,\infty}/V,0) = \mathbb{Z}_p^\times$. This shows that

$$V_{m,\infty}[k] = H^0(S_m, \omega^k) \quad \text{and} \quad V[k] = H^0(S_W, \omega^k \otimes T_p) = H^0(S_W, \omega^k) \otimes T_p,$$

where $T_p = \mathbb{Q}_p/\mathbb{Z}_p$. The last identity follows, since $S$ is affine. This shows that $V[k]$ is $p$-divisible, and its direct summand $eV[k]$ is also $p$-divisible.

We consider the following condition:

(F) $\text{corank}_w eV[k] = \text{rank}_w \text{Hom}(eV[k], T_p)$ is finite for an integer $k$. 

{\footnotesize
\text{ASTÊRISQUE} 298
In practice, this condition is often proven by showing
\[(3.1) \quad H^0_{\text{ord}}(S/W, \omega^k \otimes T_p) = H^0_{\text{ord}}(X_1(N)/W, \omega^k \otimes T_p) = G_k^{\text{ord}}(\Gamma_1(N), T_p).\]
The left-hand-side (LHS) of (3.1) is $p$-divisible, since $S$ is affine. The (RHS) is of finite corank since $X_1(N)$ is projective. Thus $eV[k]$ is $p$-divisible of finite corank.

Decompose $\mathbb{Z}^\times_p = \Gamma_T \times \Delta$ for a $p$-profinite group $\Gamma_T$ and a prime-to-$p$ finite group $\Delta$. For simplicity suppose that $p > 2$. Then $\Gamma_T$ is isomorphic to $\mathbb{Z}_p$ and for its generator $\gamma$, we have $W[[\Gamma_T]] \cong W[[X]] = \Lambda$ via $\gamma \mapsto 1+X$ (that is, $\gamma^s \mapsto (1+X)^s = \sum_{j=0}^{\infty} \binom{s}{j} X^j$), and $W[[\mathbb{Z}^\times_p]] = \Lambda[\Delta]$. Let $V^{\text{ord}}$ be the Pontryagin dual module of $eV$. If $eV[k]$ is of finite corank for one $k$, then by duality, we have
\[(3.2) \quad V^{\text{ord}}[\chi]/(X + 1 - \gamma^k)V^{\text{ord}}[\chi] = V^{\text{ord}} \otimes_{W[[T(\mathbb{Z}_p)]]} W \cong \text{Hom}_W(H^0_{\text{ord}}(X_1(N), \omega^k), W) \cong \text{Hom}_W(G_k^{\text{ord}}(\Gamma_1(N); W), W) \quad (\chi = k|\Delta).\]

In the middle equality, we have assumed (3.1). Here the subscript or superscript “ord” implies the image of $e$. Decompose $V^{\text{ord}}$ by the character of $\Delta$ as follows:
$$V^{\text{ord}} = \bigoplus_{\chi \in \Delta} V^{\text{ord}}[\chi].$$

If $z \mapsto z^k$ coincides with $\chi$ on $\Delta$, then $V^{\text{ord}}[\chi] \otimes_{\Lambda,k} W = V^{\text{ord}} \otimes_{W[[T(\mathbb{Z}_p)]]} W$. By Nakayama’s lemma, we have a surjective homomorphism of $\Lambda$-modules:
$$\pi : \Lambda^s(\chi) \rightarrow V^{\text{ord}}[\chi],$$
where $s = s(\chi) = \text{corank}_W eV[k]$. If (F) holds for one $k$, it holds for all $\kappa$ inducing $\chi$, and $\pi$ has to be an isomorphism by the following reason: The number $s$ is the minimum number of generators of $V^{\text{ord}}[\chi] \otimes_{\Lambda,\kappa} W$ over $\kappa$. We know that this module is $W$-free, because its dual $V[\kappa]$ is $p$-divisible; so, it is free of rank $s$. The morphism $\pi$ induces an isomorphism modulo $(1 + X) - \gamma^\kappa$ for all $\kappa$ inducing $\chi$. Then
$$\text{Ker}(\pi) \subset \bigcap_{\kappa} \text{Ker}(\pi \mod (1 + X - \gamma^\kappa)) = 0,$$
and we get

**Theorem 3.1.** Suppose that (F) holds for one $k$. Write $H^0_{\text{ord}}$ for $eH^0$ and $G_k^{\text{ord}}$ for $eG_k$. Then $V^{\text{ord}}[\chi]$ is $\Lambda$-free of finite rank $s(\chi)$, and if (3.1) holds for $k$, then
$$V^{\text{ord}} \otimes_{W[[T(\mathbb{Z}_p)]]} W \cong \text{Hom}_W(G_k^{\text{ord}}(\Gamma_1(N); W), W).$$

### 3.1.2. Bounding the $p$-ordinary rank
Since $S_1$ is affine, we have
$$H^0(S_1, \omega^k) = H^0(S/W, \omega^k) \otimes W 1.$$  
If $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_j$ is a sequence of linearly independent sections in $H^0_{\text{ord}}(S_1, \omega^k)$, we can lift them to $f_i \in H^0(S/W, \omega^k)$ so that $\overline{f}_i = (f_i \mod p)$. Since $S = M[1/\ell]$, we have
$$H^0(S/W, \omega^k) = \lim_{\longleftarrow} H^0(M/W, \omega^{k+n(p-1)})/E^n.$$  

---

**SOCIETÉ MATHÉMATIQUE DE FRANCE 2005**
Thus $E^m f_i \in H^0(M_{/W}, \omega^{k+m(p-1)})$ for all $i = 1, \ldots, j$ for sufficiently large $m$, and they are linearly independent. We now assume

(C) $e(Ef) = E(ef)$ for all $f \in H^0(S_1, \omega^k)$.

By this, $e(E^m f_i)$ are still linearly independent in $H^0_{\text{ord}}(M_{/K}, \omega^{k+m(p-1)})$; so, we have

$$\dim_K H^0_{\text{ord}}(M_{/K}, \omega^{k+m(p-1)}) \geq j.$$ 

If $\text{rank}_W H^0_{\text{ord}}(S_1, \omega^k)$ is infinite, the rank of $H^0_{\text{ord}}(M_{/K}, \omega^{k+m(p-1)})$ grows as $m \to \infty$.

The condition (F) for all $k$ follows from

(F') $\dim_K G^\text{ord}_k(\Gamma_1(N), K)$ is bounded independent of $k$ ($K = \Q_p$).

Actually, the Eichler-Shimura isomorphism combined with a calculation of group cohomology $H^1_{\text{ord}}(\Gamma_1(N), L(k; K))$ proves much stronger

(E) If $k \geq 3$, $\dim_K G^\text{ord}_k(\Gamma_1(N), K)$ depends only on $k \text{ mod } p - 1$ ([LFE] 7.2).

The projector $e$ will be constructed in the following subsection.

3.1.3. Construction of the projector. — Let $(E, \phi_p, \phi_N)_A$ be a test object. Suppose that $A$ is flat over $\Z_p$. Each subgroup $C$ of order $p$ outside the image of $\phi_p$ is étale over $A[\frac{1}{p}]$; so, we can think of the quotient $(E/C, \phi_p, \phi_N)$ defined over an étale finite extension $B$ of $A[\frac{1}{p}]$. We define

(U) $f[U(p)](E, \phi_p, \phi_N) = \frac{1}{p} \sum_{C} f(E/C, \phi_p, \phi_N)$.

Computing $q$-expansion, we know

$$a(n, f[U(p)]) = a(np, f).$$

So the operator preserves integral structure over $A$. The above construction of $U(p)$ works well for triples $(E, \phi_p, \phi_N)$ over general scheme $T$ as long as $T$ is flat over $\Z_p$. Thus we have $U(p)$ operator well defined over $S[\frac{1}{p}]$.

We shall extend the definition of $U(p)$ to $A$ with $p$-torsion following Katz [K3] 3.10. For the universal elliptic curve $E$ over $S$, we have a non-split exact sequence

$$0 \longrightarrow \mu_p \longrightarrow \mathbf{E} \longrightarrow \mathbf{E} \longrightarrow \mathbf{E} \longrightarrow 0.$$ 

To have an étale subgroups $C$ in $\mathbf{E}[p]$, we need to split the above sequence via base-change from $S$ to its finite flat covering $S'$. By the deformation theory of elliptic curves by Serre-Tate (which we will expose in Lecture 8), for each closed point $x \in S_1(\overline{\mathbb{F}_p})$, we have a canonical identification of the formal completion $\hat{S}_x$ of $S$ along $x$ with the formal multiplicative group $\hat{G}_{m/W}$ over the Witt ring $W$ of $\overline{\mathbb{F}_p}$. Then the above extension is equivalent to

$$0 \longrightarrow \mu_p \longrightarrow T_p \longrightarrow \mathbf{Z}/p\mathbf{Z} \longrightarrow 0,$$
where the group scheme \( T_p/G_m \) is defined as follows (cf. [GME] Example 1.6.5 in page 43):

\[
T_p = \text{Spec} \left( \prod_{i=0}^{p-1} \frac{\mathbb{Z}[t, t^{-1}][x]}{(x^p - t^i)} \right).
\]

Thus \( T_p \) is a finite flat group scheme over \( G_m = \text{Spec}(\mathbb{Z}[t, t^{-1}]) \), and \( E[p] \times_S \widehat{S}_x \cong T_p \times_{G_m} \widehat{S}_x \). For any commutative ring \( R \)

\[
T_p(R) = \{ (x, i/p) \mid x^p = t^i, \ x \in G_m(R), \ i/p \in p^{-1}\mathbb{Z}/\mathbb{Z} \} = \text{Ker}(G_m(R)/x^2 \xrightarrow{t \mapsto t^p} G_m(R)/x^2).
\]

This shows that \( \hat{O}_{S',x} \) has to be isomorphic to the formal completion of the ring \( \mathbb{W}[t^{1/p}, t^{-1/p}] = \mathbb{W}[t, t^{-1}][x]/(x^p - t) \) along \( x = 1 \). Thus \( S' \) is a finite flat covering of \( S \) radicel (or purely inseparable) at the special fiber over \( p \). In any case, we have the trace map \( \text{Tr}_\varphi : \mathcal{O}_{S'} \to \mathcal{O}_S \) and the inclusion \( \iota : \mathcal{O}_S \hookrightarrow \mathcal{O}_{S'} \). We also have the Frobenius map \( \varphi : \mathcal{O}_S \to \mathcal{O}_{S'} \). In other words, \( S' \) is the moduli of quadruples \((E, \phi_p, C, \phi_N)\) for an étale subgroup \( C \subset E \), and the Frobenius map \( \varphi \) for general base is induced by the correspondence:

\[
(E, \phi_p, \phi_N) \mapsto (E[p]) = E/\phi_p(\mu_p), \phi'_p : \mu_p \cong E[p^2]\varphi(\mu_p), E[p]^\text{ét}, \phi_N),
\]

where \( \phi'_p \) is induced by

\[
\mu_p \xrightarrow{\phi_p} \phi_p(\mu_p) \xrightarrow{\zeta \mapsto \zeta^{1/p}} E[p^2]^\varphi/\phi_p(\mu_p).
\]

Then it is easy to check that the \( U(p) \) operator coincides with \( \frac{1}{p} \text{Tr}_\varphi \) after inverting \( p \).

We thus use the formula (U) heuristically over general base \( A \) under the understanding that \( \sum C \) in (U) indicates \( \text{Tr}_\varphi \) if \( A \) has non-trivial \( p \)-torsion. In other words, in \( \sum C \), the étale subgroups \( C \) is counted with multiplicity \( p \) if \( A \) has \( p \)-torsion; so, \( U(p) \) is divisible by \( p \). In particular, \( p \cdot T(p) = p \cdot U(p) + \varphi \) is the \( p \)-adic lift of the congruence relation of Eichler and Shimura as given in [K3] 3.11.3.3.

Since \( E \equiv 1 \mod p \), we confirm (C). Let

\[
G_{\Gamma_1(N)}(A) = \bigoplus_{k=0}^{\infty} G_k(\Gamma_1(N); A).
\]

One can prove the \( p \)-adic density of \( G_{\Gamma_1(N)}(W)[\frac{1}{p}] \cap V \) in \( V \) ([GME] Corollary 3.2.4 and Theorem 3.2.10). Using this fact, we can show that \( \lim_{n \to \infty} U(p)^{n!} \) exists. The final result is as follows:

**Theorem 3.2 (VCT).** For all \( k \geq 3 \), we have

\[
V^\text{ord} \otimes_{W[[\mathbb{Z}_p^\text{ord}]]} W \cong \text{Hom}_W(G_k^\text{ord}(\Gamma_1(N); W), W).
\]

Similarly, if we write \( V^\text{cusp} \) for the subspace of cusp forms in \( V^\text{ord} \) and write \( V^\text{cusp} \) for its Pontryagin dual (that is the cuspidal quotient of \( V^\text{ord} \)), the above result holds for
spaces of cusp forms replacing $V^\text{ord}$ and $G^\text{ord}_k$ by $V^\text{ord}_\text{cusp}$ and the subspace $S^\text{ord}_k$ of cusp forms in $G^\text{ord}_k$.

3.1.4. Families of $p$-ordinary modular forms.— Let $a(n) : V^\text{ord} \to \mathbb{T}_p$ be the linear map associating $f$ its coefficient of $q^n$ in the $q$-expansion; so, $a(n)$ is in the dual $V^\text{ord}$.

We now consider $G(\chi; \Lambda) = \text{Hom}_\Lambda(V^\text{ord}[\chi], \Lambda)$.

With each $\phi \in G(\chi; \Lambda)$, we associate its $q$-expansion

$\phi(q) = \sum_{n=0}^{\infty} \phi(a(n)) q^n \in \Lambda[[q]]$.

**Theorem 3.3.** For each $k \geq 2$, we have

1. $G(\chi; \Lambda)$ is $\Lambda$-free of finite rank;
2. $G(\chi; \Lambda) \otimes_{\Lambda,k} W \cong G^\text{ord}_k(\Gamma_1(N) \cap \Gamma_0(p), \chi \omega^{-k}; W)$;
3. the above identification is induced by $\phi \mapsto \sum_n a(n, \phi(\gamma^k - 1)) q^n \in W[[q]]$.

**Proof.** The $\Lambda$-freeness follows from the freeness of $V^\text{ord}[\chi]$. We only prove the assertion when $k$ induces $\chi$ on $\Delta$; so, $\chi \omega^{-k}$ is trivial. We have

$G(\chi; \Lambda) \otimes_{\Lambda,k} W \cong \text{Hom}_\Lambda(V^\text{ord}[\chi], \Lambda) \otimes_{\Lambda,k} W$

$\cong \text{Hom}_W(V^\text{ord}[\chi] \otimes_{\Lambda,k} W, W) \cong \text{Hom}_W(\text{Hom}_W(G^\text{ord}_k(\Gamma_1(N); W); W))$

$= G^\text{ord}_k(\Gamma_1(N); W) = G^\text{ord}_k(\Gamma_1(N) \cap \Gamma_0(p); W)$.

We leave the verification of the specialization of $q$-expansion to the audience.

There is a version of this type of results for $\Gamma_0(N)$ and also for cusp forms, which is valid for all weights $k \geq 2$ (see [GME] Chapter 3 in particular Theorem 3.2.17).

### 4. Hecke Equivariance of the Eichler-Shimura Map

Let $G_{/\mathbb{Z}(p)} (\mathbb{Z}(p) = \mathbb{Q} \cap \mathbb{Z}_p)$ be a connected reductive group (split over $\mathbb{Z}_p$). For simplicity, we often assume that $G$ has trivial center. We shall prove semi-simplicity of the commutative Hecke algebra acting on the nearly ordinary cohomology group $H^q_{n,\text{ord}}(X(U), L) \subset H^q_1(X(U), L)$ for a modular variety $X(U)$ associated to an arbitrary $p$-power level open compact subgroup $U$ of $G(\mathbb{A}^\infty)$. Here the locally constant or coherent sheaf $L$ on $X(U)$ is associated to a rational representation of $G$ twisted by a finite order character. Although we have assumed that $G$ is split over $\mathbb{Z}_p$, the argument works equally well for a connected reductive group $G_{/\mathbb{W}}$ split over $\mathbb{W}$ (see [PAF] Chapter 5). Here $\mathbb{W}$ is a valuation ring over $\mathbb{Z}(p)$ of a number field, and $\mathbb{W} = \varprojlim_n \mathbb{W}/p^n\mathbb{W}$ is the $p$-adic completion of $\mathbb{W}$.
4.1. Semi-simplicity of Hecke Algebras. — We shall describe the semi-simplicity of the Hecke algebra acting on topological cohomology groups. Later we relate the topological and the coherent cohomology groups by the generalized Eichler-Shimura map, which shows the semi-simplicity of the Hecke algebra acting on the (degree 0) coherent cohomology.

4.1.1. Jacquet Modules. — Let \( \pi \) be an admissible semi-simple representation of \( G(\mathbb{Q}_p) \) on a vector space \( V = V(\pi) \) over a field \( K \) of characteristic 0 (in this lecture, \( K \) is just a characteristic 0 field no more no less). Contrary to the tradition, I always suppose that \( V \) is a right \( G(\mathbb{Q}_p) \)-module. Let \( B \) be a Borel subgroup with split torus \( T = B/N \) for the unipotent radical \( N \).

We have a Haar measure \( du \) of \( N(\mathbb{Q}_p) \) with \( \int_{N(\mathbb{Z}_p)} du = 1 \). We then define

\[
V(B) = V(B, \pi) = \{ v - v\pi(n) \in V(\pi) \mid v \in V(\pi), n \in N(\mathbb{Q}_p) \},
\]

and put \( V_B = V_B(\pi) = V/V(B) \), which is called the Jacquet module. We take a sufficiently large open compact subgroup \( U_w \subset N(\mathbb{Q}_p) \) for each \( w = v - v\pi(n) \in V(B) \) so that \( n \in U_w \). Then we see that \( \int_U v\pi(u)du = 0 \) for every open subgroup \( U \) with \( U_w \subset U \subset N(\mathbb{Q}_p) \). By this fact, we can conclude that the association \( V \mapsto V_B \) is an exact functor. Later we shall give a canonical splitting \( V^N = V_B \oplus V(B)^N \) as Hecke modules, where \( V^N = H^0(N(\mathbb{Z}_p), V) \) (Bernstein-Casselman).

Let \( V' \) be a \( K \)-vector space. A function \( f : G(\mathbb{Q}_p) \to V' \) is called smooth if it is locally constant (uniformly under the left translation). In other words, there exists an open compact subgroup \( C_f \subset G(\mathbb{Q}_p) \) such that \( f(kg) = f(g) \) for all \( g \in G(\mathbb{Q}_p) \) and \( k \in C_f \). For each admissible \( T(\mathbb{Q}_p) \)-module \( V' \), we define \( \text{Ind}_T^G V' \) to be the space of smooth functions on \( G(\mathbb{Q}_p) \) such that \( f(gb) = f(g) \overline{b} \) for all \( b \in B(\mathbb{Q}_p) \), where \( \overline{b} \) is the projection of \( b \) in \( T(\mathbb{Q}_p) \). Then we let \( G(\mathbb{Q}_p) \) act on \( \text{Ind}_T^G V' \) by \( f(g)g' = f(g'g) \) for \( g \in G(\mathbb{Q}_p) \). This representation \( \text{Ind}_T^G V' \) is the smooth induction of \( V' \) from \( B \) to \( G \). In this definition, we may replace \( B \) by a parabolic subgroup of \( G \) and \( T \) by the reductive part of \( P \). Hereafter all representations of \( G, B \) and \( T \) are assumed to be smooth admissible.

Since the smooth induction preserves admissibility ([BZ] 2.3), \( \text{Ind}_T^G V' \) has composition series, and hence its semi-simplification \( (\text{Ind}_T^G V')^{ss} \) is well defined. The beauty of the theory of admissible representations is its purely algebraic nature; so, we do not need to assume any analytic assumptions; in particular, our representations are often not unitary as is clear from our main result Theorem 4.2.

The following results are due to Jacquet and Bernstein-Zelevinsky and are well known [BZ]:

1. (Frobenius reciprocity) \( \text{Hom}_B(V_B, V') \cong \text{Hom}_G(V, \text{Ind}_B^G V') \);
2. If \( \pi \) is absolutely irreducible, then \( \dim_K V_B \leq |W| \), where \( W \) is the Weyl group of \( T \) in \( G \) (Bernstein-Zelevinsky);
(3) If \( \pi \) is absolutely irreducible and \( V_B \neq 0 \), then \( \text{Ind}_B^G \lambda \to V \) for a character \( \lambda : T(Q_p) \to K^\times \) (Jacquet);

\[ (\text{Ind}_B^G \lambda)^{ss} \cong (\text{Ind}_B^G \lambda_w)^{ss} \text{ for all } w \in W \] (Bernstein-Zelevinsky),
where "ss" indicates semi-simplification, \( \lambda^w(t) = \lambda(wtw^{-1}) \) and \( \lambda = \delta_B^{1/2} \lambda \) for the right module character \( \delta_B \) of \( B \): \( \int_{N(Q_p)} \phi(u)du = \delta_B(b) \int_{N(Q_p)} \phi(b^{-1}xb)du \). We have the following corollary of the above facts:

**Corollary 4.1.** — Suppose that \( \pi \) is irreducible and that \( V_B[\lambda] \neq 0 \), where \( \lambda = \delta_B^{1/2} \lambda \) for the module character \( \delta_B \) on \( B \). Then \( \pi \) is a quotient of \( \text{Ind}_B^G \lambda \). If \( \lambda^w(t) \) for \( w \in W \) are all distinct, \( V_B \subset \bigoplus_{w \in W} \lambda^w \) as \( T(Q_p) \)-modules.

**Proof.** — Since the algebra in \( \text{End}_K(V_B) \) generated by the action of \( T \) is a finite dimensional commutative algebra, the \( \lambda \)-eigenspace is non-trivial if and only if the maximal \( \lambda \)-quotient is non-trivial. Thus, we have a morphism of \( T \)-modules: \( V_B \to V(\lambda) \). Since we have \( (\text{Ind}_B^G \lambda)^{ss} \cong (\text{Ind}_B^G \lambda_w)^{ss} \), by Frobenius reciprocity, all eigenvalues \( \lambda^w \) can show up as a quotient of \( (\text{Ind}_B^G \lambda)_B \) whose dimension is bounded by \( |W| \). Thus if all characters \( \lambda^w \) are distinct, we have \( V_B \subset (\text{Ind}_B^G \lambda)_B \cong \bigoplus_{w \in W} \lambda^w \). Since \( V \to V_B \) is exact, this is enough to conclude the assertion.

### 4.1.2. Double Coset Algebras.

We consider the double coset algebra made of formal linear combinations of double cosets of a subgroup in a semi-group. This type of algebra is considered in [IAT] 3.1 and often called a Hecke ring. We shall use the terminology “double coset algebra” to avoid confusion with Hecke algebras later we shall study.

Let

\[
D = \{ x \in T(Q_p) \mid xN_B(Z_p)x^{-1} \supset N_B(Z_p) \}
\]

which is called the expanding semi-group in \( T(Q_p) \). Write \( B = B(Z_p) \) and \( N = N(Z_p) \) for simplicity. Define so-called Iwahori subgroups by

\[
U_0(r) = \{ u \in G(Z_p) \mid u \text{ mod } p^r \in B(Z/p^rZ) \}
\]

\[
U_1(r) = \{ u \in G(Z_p) \mid u \text{ mod } p^r \in N(Z/p^rZ) \}.
\]

These subgroups \( S \) have the Iwahori decomposition: \( S = N'TN \cong N' \times T' \times N \) for open compact subgroups \( T' \subset T(Q_p) \) and \( N' \) in the opposite unipotent \( 'N = 'N(Z_p) \).

Each \( x \in D \) shrinks \( 'N : x'N_x^{-1} \subset 'N \). Then we have

\[
N\xi N = \bigsqcup_{u \in \xi^{-1}N\xi N} N\xi u = \bigsqcup_{u \in N\xi N\xi^{-1}} Nu\xi,
\]

\[
B\xi B = \bigsqcup_{u \in N\xi N\xi^{-1}} Bu\xi \text{ and } S\xi S = \bigsqcup_{u \in N\xi N\xi^{-1}} Su\xi,
\]

where \( S \) is an Iwahori subgroup. By this fact, \( \Delta_N = NDN, \Delta = \Delta_B = BDB \) and \( \Delta_S = \Delta_p = SDS \) are sub-semigroups of \( G(Q_p) \), and the double coset algebras.
generated additively over \( \mathbb{Z} \) by double cosets of the group in the semigroup are all isomorphic as algebras:

\[ R = R(N, \Delta_N) \cong R(B, \Delta_B) \cong R(S, \Delta_S). \]

Further these algebras are commutative: \( T(\xi)T(\eta) = T(\xi\eta) \) for \( T(\xi) = N\xi N \) and \( \xi, \eta \in D \) (cf. [IAT] Chapter 3 and [H95] Section 2). We let \( R \) act on \( v \in V^N = H^0(N(\mathbb{Z}_p), V) \) by

\[ v|T(\xi) = v|[N\xi N] = \sum_{u \in \xi^{-1}N\xi \setminus N} v\pi(\xi u) = \int_{\xi N\xi^{-1}} v\pi(u)\pi(\xi)du, \tag{4.3} \]

and similarly for \( \pi \in V_B \) in place of \( v \in V^N \); then the projection: \( V^N \to V_B \) is \( R \)-linear. Here the Haar measure \( du \) is normalized so that \( \int_{N(\mathbb{Z}_p)} du = 1 \).

Let \( \Sigma \) be the set of maximal (proper) parabolic subgroups \( P \supset B \). Decomposing \( P = M_P N_P \) for reductive part \( M_P \supset T \) and the unipotent radical \( N_P \), we can identify the set \( \Sigma \) with the following set of co-characters:

\[ \{ \alpha_P : \mathbb{G}_m \to G \mid \alpha_P(p) \text{ generates } Z(M_P)(\mathbb{Q}_p) \cap D \text{ modulo } Z(M_P)(\mathbb{Z}_p) \}, \]

where \( Z(M_P) \) is the center of \( M_P \). Then \( \{ \xi_\alpha = \alpha(p) \}_{\alpha \in \Sigma} \) generate \( D/T(\mathbb{Z}_p) \), and \( R \cong \mathbb{Z}[T(\xi_\alpha)] \) if the center of \( G \) is trivial. If \( G = GL(n) \),

\[ \Sigma = \{ \alpha_j \mid \alpha_j(p) = \text{diag}[1_j, pI_{n-j}] \}. \]

For \( \xi = \prod_{\alpha \in \Sigma} \xi_\alpha \), we have

\[ N(\mathbb{Q}_p) = \bigcup_{j=0}^{\infty} \xi^j N \xi^{-j}. \]

We still denote by \( T(\xi) \) the action of \( N\xi N \) on \( V^N \). The formula (4.3) defines as well an action of \( T(\xi) \) on \( V_B \). We see easily from (4.3) that \( T(\xi^i) = T(\xi)^i \) and for each finite dimensional subspace \( X \subset V(\mathbb{B}) \), \( T(\xi)|_X \) is nilpotent on \( X^N \) by (4.3).

For any \( R \)-eigenvector \( v \in V^N \) with \( \pi t = \lambda(t)\pi \) (\( t \in T(\mathbb{Q}_p) \), \( \pi = v \mod V(\mathbb{B}) \)), we get

\[ v|[N_x N] = [N : x^{-1}Nx]\lambda(x)v = |\det(\text{Ad}_N(x))|_p^\lambda(x)v, \tag{4.4} \]

where \( |^\lambda \) is the standard absolute value of \( \mathbb{Q}_p \) such that \( |p|_p^{-1} = p \) and \( \text{Ad} \) is the adjoint representation of \( T \) on the Lie algebra of \( N \).

4.1.3. Rational representations of \( G \) — Let us first define a canonical splitting:

\[ V^N = V_B \oplus V(\mathbb{B})^N \text{ as } R \text{-modules.} \]

We have by definition, \( V^N = V^N(\mathbb{Z}_p) \cup V^U_1(\mathbb{R}) \). The subspace \( V_\mathbb{R} = V^U_1(\mathbb{R}) \) is finite dimensional and is stable under \( R \). By Jordan decomposition applied to \( T(\xi) \) (\( \xi = \prod_{\alpha \in \Sigma} \xi_\alpha \)), we can decompose uniquely that \( V_\mathbb{R} = V_\mathbb{R}^\circ \oplus V_\mathbb{R}^{nil} \) so that \( T(\xi) \) is an automorphism on \( V_\mathbb{R}^\circ \) and is nilpotent on \( V_\mathbb{R}^{nil} \). We may replace \( T(\xi) \) by \( T(\xi^a) = T(\xi)^a \) for any positive \( a \) in the definition of the above splitting. Since \( T(\xi) \) is nilpotent on any finite dimensional subspace of \( V(\mathbb{B}) \), \( V_\mathbb{R}^\circ \) injects into \( V_B \); so, dim \( V_\mathbb{R}^\circ \) is bounded.
by \( \dim V_B \leq |W| \). For any \( T \)-eigenvector \( \tau \in V_B \), lift it to \( v \in V \). Then for sufficiently large \( j \), \( v\tau(\xi^{-j}) \) is in \( V_N \). Since \( \tau\pi(\xi^{-j}) \) is a constant multiple of \( \tau \), we may replace \( \tau \) and \( v \) by \( \tau\pi(\xi^{-j}) \) and \( \tau\pi(\xi^{-j}) \), respectively. Then for sufficiently large \( k \), \( w = vT(\xi^k) \in V_\sigma \). Then \( \tau\pi T(\xi^{-k}) \) is equal to \( \tau \) for the image \( \bar{\tau} \) in \( V_B \). This shows the splitting: \( V_N \cong V_B \oplus V(B)^N \) as \( R \)-modules when the action of \( T \) on \( V_B \) is semisimple. In general, taking a sufficiently large \( r \) so that \( V_r \) surjects down to \( V_B \). We apply the above argument to the semi-simplification of \( V_r \) under the action of the Hecke algebra. Thus \( V_\sigma = \bigcup_r V_r \cong V_B \), and this concludes the proof.

Let \( G(\mathbb{Z}) \subset G(\mathbb{A}^\infty) \) denote a maximal compact subgroup hyperspecial everywhere (by abusing notation; see [Tt] for hyperspecial compact subgroups). We assume that the \( p \)-component of \( G(\mathbb{Z}) \) is given by \( G(\mathbb{Z}_p) \). We now assume \( K \) to be a finite extension over \( \mathbb{Q}_p \). Let \( \mathcal{O} \) be the \( p \)-adic integer ring of \( K \). We write \( U = U_0(r) \) for \( r > 0 \). Recall the Iwahori decomposition \( U = N' T(\mathbb{Z}_p) N \). We consider the space of continuous functions: \( \mathcal{C}(A) = \{ \phi : U/N(\mathbb{Z}_p) \to A \} \) for \( A = \mathcal{O} \) and \( K \). We would like to make \( \mathcal{C} \) a left \( \Delta_p^{-1} \)-module for the opposite semi-group \( \Delta_p^{-1} \) of \( \Delta_p = \Delta U \). For that, we first define a left action of \( \Delta_p \) on \( Y_\ell = U/N(\mathbb{Z}_p) \). Since \( U \) acts on \( Y_\ell = U/N(\mathbb{Z}_p) \) by left multiplication, we only need to define a left action of \( D \). Pick \( y N(\mathbb{Z}_p) \in Y_\ell \) and by the Iwahori decomposition, we may assume that \( y \in N' T(\mathbb{Z}_p) \subset U \) and consider \( y N \). Then for \( d \in D \),

\[
dydN^{-1} = dyd^{-1}N^{-1} = dyd^{-1}N(\mathbb{Q}_p) \text{ and } dyd^{-1}N(\mathbb{Q}_p)
\]

is well defined in \( G(\mathbb{Q}_p)/N(\mathbb{Q}_p) \). Since conjugation by \( d \in D \) expands \( N(\mathbb{Z}_p) \) and shrinks \( N' \), \( dyd^{-1} \in U \), and the coset \( dyd^{-1}N(\mathbb{Q}_p) \cap U = dyd^{-1}U \) is a well defined single coset of \( N \), which we designate to be the image of the action of \( d \in D \). We now let \( \Delta_p^{-1} \) act on \( \mathcal{C} \) by \( d\phi(y) = \phi(d^{-1}y) \). In this way, \( \mathcal{C} \) becomes a \( \Delta_p^{-1} \)-module.

We consider the algebro-geometric induction module:

\[
(4.5) \quad L(\kappa; K) = \left\{ \phi : G/N \to K \in H^0(G/N, \mathcal{O}_{G/N}) \mid \phi(yt) = \kappa(t)\phi(y) \quad \forall t \in T \right\},
\]

where \( \mathcal{O}_{G/N} \) is the structure sheaf of the scheme \( G/N \). We let \( G \) act on \( L(\kappa; K) \) by \( g\phi(y) = \phi(g^{-1}y) \). Then \( L(\kappa; K) = \text{ind}_{\mathcal{O}}^G \kappa^{-1} \) (following the normalization of induction as in [RAG] I.3.3), which is the induction in the category of scheme theoretic representations (that is, polynomial representations). We call \( \kappa \) dominant if \( L(\kappa; K) \neq 0 \). We write this representation as \( \rho_\kappa = \rho_\kappa^G : G \to GL(L(\kappa; K)) \).

We restrict functions in \( L(\kappa; K) \) to \( Y_\ell = U/N(\mathbb{Z}_p) \) and regard \( L(\kappa; K) \subset \mathcal{C}(K) \). Then multiply \( L(\kappa; K) \) by a character \( \varepsilon : T(\mathbb{Z}/p^r \mathbb{Z}) = U_0(r)/U_1(r) \to \mathbb{O}^\times \) (regarding it as a function on \( \mathcal{C}(\mathcal{O}) \)). Since \( \mathbb{Q}_p^\times = \mathbb{Z}_p^\times \times p^\mathbb{Z} \), we can decompose \( T(\mathbb{Q}_p) = T(\mathbb{Z}_p) \times (p^\mathbb{Z})^\times \) for the rank \( r \) of \( T \), and we can extend \( \varepsilon \) to \( T(\mathbb{Q}_p) \) requiring it to have constant value \( 1 \) on \( (p^\mathbb{Z})^\times \). In this way, we get the twisted \( \Delta_p^{-1} \)-module \( L(\kappa \varepsilon; K) = \varepsilon L(\kappa; K) \subset \mathcal{C} \). The pull-back \( \Delta_p^{-1} \)-action preserves \( L(\kappa \varepsilon; \mathcal{O}) = L(\kappa \varepsilon; K) \cap \mathcal{C}(\mathcal{O}) \) but original \( \rho_\kappa \) may not be. Then for \( \xi \in D \),

\[
(4.6) \quad \text{the action of } \xi^{-1} \in \Delta_p^{-1} \text{ is given by } \kappa(\xi)^{-1}\varepsilon(\xi)\rho_\kappa(\xi^{-1})).
\]
Since the action of $\xi$ on $Y_U$ is conjugation: $x \mapsto \xi x \xi^{-1}$, the front $\kappa(\xi)^{-1}$ comes from the definition of $L(\kappa; K)$ in (4.5): $\phi(\xi x \xi^{-1}) = \kappa(\xi^{-1}) \rho_x(\xi^{-1}) \phi(x)$. By definition, the new action is optimally integral.

**Example 4.1.** — To illustrate our integral modification of the action, let us give an example in the simplest non-trivial case: Let $L(\kappa; K)$ be the space of homogeneous polynomial of two variable $(X, Y)$ of degree $n > 0$. Then we let $G = \text{GL}(2)$ act on $\phi(X, Y) \in L(\kappa; K)$ by $(a \ b ~ c \ d)^{-1} \phi(X, Y) = (ad - bc)^v \phi(dX - bY, -cX + aY)$ for an integer $v \in \mathbb{Z}$. Then $L(\kappa; K) = \text{ind}_G^H_\kappa^{-1}$ for $\kappa : \text{diag}[a, d] \mapsto (ad)^v a^n$ for the upper triangular Borel subgroup $B \subset \text{GL}(2)$. If the integer $v$ is negative, the lattice $L(\kappa; \mathcal{O})$ is obviously not stable under the action of the diagonal matrices

$$D = \{ z \text{ diag}[1, d] \mid 0 \neq d \in \mathbb{Z}_p, z \in \mathbb{Q}_p^* \}.$$  

The modified (integral) action defined above is just

$$(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix})^{-1} \circ \phi(X, Y) = \phi(dX, Y) = d^{-v} \phi(1 0)^{-1} (1 0)^{-1} \phi(X, Y) = \kappa(1 0)^{-1} (1 0)^{-1} \phi(X, Y).$$

4.1.4. Nearly $p$-Ordinary Representations. — Hereafter we assume that $\kappa$ is an element $\kappa_0$ of $X(\mathbb{T})$ up to finite order character of $T(\mathbb{Z}_p)$. Let $U$ be an open subgroup of $G(\mathbb{Z})$. We consider the associated modular variety:

$$X(U) = G(\mathbb{Q}) \backslash G(\mathbb{A}) / UC_{\infty+} \cong \bigsqcup_{\Gamma} X_{\Gamma},$$

where $C_{\infty+}$ is the identity connected component of the maximal compact subgroup of the Lie group $G(\mathbb{R})$ and $\Gamma$ runs over the following finite set

$$\{ G(\mathbb{Q}) \cap tUG(\mathbb{R})t^{-1} \mid t \in (G(\mathbb{Q}) \backslash G(\mathbb{A}) / UG(\mathbb{R})) \}$$

where $G_+(\mathbb{R})$ is the identity connected component of $G(\mathbb{R})$. For the symmetric space $X = G_+(\mathbb{R}) / C_0$, we have written $X_{\Gamma} = \Gamma \backslash X$. For any $\mathcal{O}$-module $A$, we define a right action of $u \in UC_{\infty+}$ on $L(\kappa; A)$ by $\phi|u = \rho_u(\kappa) \phi$ if $U_p \subset U_0(r)$ for some $r > 0$.

We define the covering space $X(U)$ of $X(U)$ by

$$X(U) = G(\mathbb{Q}) \backslash (G(\mathbb{A}) \times L(\kappa; A)) / UC_{\infty+} \cong \bigsqcup_{\Gamma} X_{\Gamma},$$

where $\gamma(x, \phi)u = (\gamma x u, \phi|u)$ for $\gamma \in G(\mathbb{Q})$ and $u \in UC_{\infty+}$, and we define $X_{\Gamma} = \Gamma \backslash (X \times L(\kappa; A))$ by the diagonal action. We use the same symbol $L(\kappa; A)$ for the sheaf of locally constant sections of $X(U)$ over $X(U)$.

We consider the limit, shrinking $S$,

$$L(A) = L^\otimes(\kappa; A) = \lim_{\mathcal{S}} H^\otimes_0(X(S), L(\kappa; A)).$$

Here $H^\otimes_0(X(S), L(\kappa; A))$ $(A = K$ or $\mathcal{O})$ is the image of the compactly supported cohomology group $H^\otimes_0(X(S), L(\kappa; A))$ in $H^\otimes(X(S), L(\kappa; K))$. On the space $L(K)$, the group $G(\mathbb{A}^\infty)$ acts from the right via a smooth representation, which is completely
reducible. Thus in particular, we have an action on $H^0(U, \mathcal{L}^q (\kappa; K)) = \mathcal{L}^q (\kappa; K)^U$ of the double coset algebra

$$R_U = R(U, G(\mathbb{A}^{p\infty}) \times \Delta_p) \cong R(U^{(p)}, G(\mathbb{A}^{p\infty})) \otimes R$$

of double cosets $U \times U$ with $x \in G(\mathbb{A}^{p\infty}) \times \Delta_p$, where $U = U_p \times U^{(p)}$ and we have assumed that $U_p = U_0(r)$.

We take $\xi \in D$ such that $N(\mathbb{Q}_p) = \bigcup \xi^i N(\mathbb{Z}_p) \xi^{-i}$. We may assume that $\xi = \prod_{\alpha \in \Sigma} \xi_\alpha$. Then $T(\xi)$ acts on $\mathcal{L}^q (\kappa; \mathcal{O})^N$ $(N = N(\mathbb{Z}_p))$ through the $\Delta_p^{-1}$-module structure on $L(\kappa; \mathcal{O})$. We write this operator as $T$. On the other hand, $T(\xi)$ acts on $\mathcal{L}^q (\kappa; K)^N$ through the action of $G(\mathbb{A}^\infty)$ via the rational representation $\rho_\kappa$. The corresponding operator will be written by the same symbol $T$. Since the action through $\rho_\kappa (\xi^{-1})$ and the modified integral action of $\xi^{-1} \in \Delta_p^{-1}$ differs by the scalar $\kappa (\xi^{-1}) (4.6)$, the two operators $T$ and $T$ are related on the image of $\mathcal{L}^q (\kappa; \mathcal{O})^N$ by

$$(4.9) \quad T(\xi) = \kappa (\xi)^{-1} T(\xi).$$

When $\kappa = 0$ (the identity character), the action of the Hecke operator is (truly canonically) induced by the Hecke correspondence $T(\xi) \subset (X(U) \times X(U))$, and in this case, $T(\xi) = T(\xi)$. If $\kappa > 0$, we may relate cohomology groups of the sheaf $L(\kappa; K)$ as a part of the cohomology group with constant coefficients of a certain self-product $Z$ of copies of the universal abelian scheme over $X(U)$. Since the Hecke operator then has interpretation as an isogeny action on the universal abelian scheme, it can be regarded as the action induced by the Hecke correspondence in $Z \times Z$. The action of $\mathbb{T}(\xi)$ and $T(\xi)$ uses different action of $\Delta_p^{-1}$. This action of $\Delta_p^{-1}$ determines the part of the cohomology group over $Z$ identified with the cohomology group over $X(U)$ with locally constant (but non-constant) coefficients. Thus the motivic realization of the two operators $T(\xi)$ and $\mathbb{T}(\xi)$ could be actually different, and the operator $\mathbb{T}(\xi)$ may not even have motivic realization (as in the Hilbert modular case of non-parallel weight). For example, in Scholl’s construction [Sc] of the Grothendieck motive associated to an elliptic Hecke eigenform $f$, if one changes the action of congruence subgroup $\Gamma$ of $SL_2(\mathbb{Z})$ by a power of determinant character, the physical sheaf over $X_T(\mathbb{C})$ obtained is the same, but its rational structure (including the Galois action) different. In this way, we can construct the motive associated to the standard $p$-adic Galois representation $\rho_f$ of $f$ and its Tate twists $\rho_f(m)$ as the étale realization of motives directly realized over a self-product of the universal elliptic curve. For a Hilbert Hecke eigenform $f$, we could twist $\rho_f$ locally at each $p$-adic place by a power of the $p$-adic cyclotomic character, but this twist may not extend to a global twist because the exponent of the cyclotomic character depends on the $p$-adic place. In particular, if $f$ is of non-parallel weight, the process of defining $\mathbb{T}(\xi)$ corresponds to untwisting $\rho_f$ to reach a $p$-ordinary Galois representation at each $p$-adic place $p$, which cannot be performed globally; so, the operator $\mathbb{T}(\xi)$ may not have a motivic interpretation.
For any $U = U_p \times U^{(p)}$ with $U_p \supset N = N(\mathbb{Z}_p)$, the limit $e = \lim_{n \to \infty} T(\xi)^n$ exists as an endomorphism of $H^0(X(U), L(\kappa; A))$ for $A = \mathcal{O}$ and $K$. Thus the limit $e$ extends to an endomorphism of $L^B(\kappa; A) N$ for $A = \mathcal{O}$ and $K$. It is easy to see, if $U_p \supset N$,
\begin{equation}
H^0(U, e(L^B(\kappa; K) N)) = e(H^0(U, L(\kappa; K))).
\end{equation}
We write $L^B_{n, \text{ord}}(\kappa; A)$ for $e(L^B(\kappa; A) N)$. An irreducible representation $\pi$ of $G(A^{\infty})$, which is a subquotient of $L^B(\kappa; K)$, is called \textit{nearly ordinary} of $p$-type $\kappa$ if $e(V(\pi)^N)$ does not vanish for the representation space $V(\pi)$ of $\pi$.

4.1.5. Semi-simplicity of Interior Cohomology Groups. — Let $\pi$ be a cohomological automorphic representation of $p$-type $\kappa$. Suppose $\pi_p$ is a subquotient of $\text{Ind}_B^G \tilde{\lambda}$ (this is automatic if $\pi$ is nearly $p$-ordinary). Then for its $p$-component $\pi_p$ (acting on $V := V(\pi_p)$), we find a character $\lambda : T(\mathbb{Q}_p) \to K^\times$ with the above property such that $V_B[\tilde{\lambda}] \neq 0$ and
\[
|\text{det}(\text{Ad}_N(x))^{-1}\tilde{\lambda}(x)|_p = |\text{det}(\text{Ad}_N(x))|_p \lambda(x)|_p \leq |\kappa(x)|_p.
\]
The equality holds if and only if $\pi$ is $p$-nearly ordinary (in this case, automatically $V_B \neq 0$ and $\text{Ind}_B^G \tilde{\lambda} \to \pi_p$ because $V^N \cong V_B \oplus V(B)^N$ as $R$-modules).

For the moment, suppose that $G(\mathbb{Q}_p) = \text{GL}_n(\mathbb{Q}_p)$ and write $\lambda(\text{diag}[t_1, \ldots, t_n]) = \prod_{i=1}^n \lambda_i(t_i)$. Define the Hecke polynomial (at $p$) by
\[
H_\pi(T) = \prod_{i=1}^n (1 - \lambda_i(p)T),
\]
and write $\Delta_N$ for the Newton polygon of $H_\pi(T)$. Define the Hodge polygon $\Delta_H$ of $\pi$ to be the Newton polygon of $\prod_{i=1}^n (1 - (\kappa\rho)_i(p)T)$. Then the above inequality implies
\[
\Delta_N \geq \Delta_H
\]
and the two extreme ends of the two polygons match.

We return to a general group $G$ and assume that $\pi$ is nearly $p$-ordinary. By definition,
\[
\int_{N(\mathbb{Q}_p)} \phi(u)du = \delta_B(b) \int_{N(\mathbb{Q}_p)} \phi(b^{-1}xb)du.
\]
This shows that
\begin{equation}
\delta_B = |\text{det} \circ \text{Ad}_N|_p^{-1}.
\end{equation}
By definition, $2\rho = \text{det} \circ \text{Ad}_N$ is a sum of positive roots, and $\rho$ is a sum of fundamental weights with respect to $B$. This shows
\begin{equation}
|\lambda|_p = |\kappa\rho|_p.
\end{equation}
Note that $\kappa$ is non-negative with respect to $B$ because $\kappa$ is dominant. Since $\kappa \geq 0$, $\kappa\rho > 0$, that is, $\kappa\rho$ is in the interior of the Weyl chamber of $B$. This shows that if
Because $W$ acts simply transitively on Weyl chambers and each element in the interior of the chamber of $\lambda$ has the maximum $p$-adic absolute value in its conjugates under $W$. In particular, we get

**Theorem 4.2.** — Let $\pi$ be an irreducible nearly ordinary representation of $p$-type $\kappa$. Then there exists a character $\lambda : T(\mathbb{Q}_p) \to K^\times$ such that $\lambda \mapsto V(\pi_p)_B \mapsto \oplus_{w \in W} \tilde{\lambda}^w$ and $|\lambda|_p = |\rho\kappa|_p$, where $\rho$ is the sum of fundamental weight with respect to $B$ and $|\cdot|_p$ is the absolute value on $K$. Moreover $\epsilon H^0(N(\mathbb{Z}_p), V(\pi_p))$ is one dimensional, on which $T(\xi) = U(\xi)(U \in D)$ acts by scalar $|\rho(\xi)|_p \lambda(\xi)$.

Now suppose that $U = U_p \times G(\mathbb{Q}^p)$ with $U_p = U_0(r)$ for $r > 0$. By the above theorem, we get the following semi-simplicity of the Hecke algebra (for cohomological nearly ordinary cusp forms of $p$-type $\kappa$) from the fact that the spherical irreducible representation of $G(\mathbb{Q}_l)$ has a unique vector fixed by (any given) maximal compact subgroup:

**Corollary 4.3.** — Let the notation and the assumption be as above. Then the Hecke module $\epsilon H^0_l(X(U), L(\kappa; K))$ is semi-simple.

Note that the projector $e_\epsilon = e_p$ is actually defined over $\mathcal{L}^2(\kappa; \mathbb{Q})$. Thus the above semi-simplicity remains true on $\epsilon S H^0_l(X(U), L(\kappa; \mathbb{Q}))$ for $\epsilon_S = \pi_{x \in S} e_\epsilon$ with a finite set of primes $S$ (where $G$ is split over $\mathbb{Z}_l$) and a subgroup $U$ of level $M$, which is a product of powers of primes in $S$. For such nearly $S$-ordinary automorphic forms, semi-simplicity of the Hecke operator action is always true.

### 4.2. The Eichler-Shimura Map

Before starting detailed study of the nearly ordinary part of coherent cohomology groups, we shall make explicit a generalized Eichler-Shimura map for unitary groups and hence the association of the weight: $\kappa \mapsto \kappa^*$ so that $H^0_{\text{cusp}}(X_\Gamma, \mathbb{Q}^{\kappa^*}) \mapsto H^d(X_\Gamma, L(\kappa^*; \mathbb{C}))$. To construct the map, we briefly recall an explicit shape of the symmetric domain of unitary groups.

#### 4.2.1. Unitary groups

Define the complex unitary group $G$ by

$$G = U(m, n)(\mathbb{R}) = \{ g \in GL_{m+n}(\mathbb{C}) \mid g I_{m,n} g^* = I_{m,n} \},$$

where $I_{m,n} = \text{diag}([1_m, -1_n]) = \begin{pmatrix} 1_m & 0 \\ 0 & -1_n \end{pmatrix}$.

We want to make explicit the quotient space $X = G/C_0$ following [AAF] 3.2. We consider

$$\mathcal{Y} = \left\{ Y \in GL_{m+n}(\mathbb{C}) \mid Y^* I_{m,n} Y = \text{diag}[T, S] \text{ with } 0 < T = T^* \in M_m(\mathbb{C}), 0 > S = S^* \in M_n(\mathbb{C}) \right\}.$$
Write $Y = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. By a simple calculation, we have

$$Y^* I_{m,n} Y = \begin{pmatrix} A^* C & A^* D \\ C^* A & D^* B \end{pmatrix} = \text{diag}[T, S].$$

Since $A^* A > C^* C \geq 0$, $A$ is invertible. Similarly $D$ is invertible. Put $z = BD^{-1}$. We then see

$$A^* B = C^* D \iff (CA^{-1})^* = A^{-*} C^* = BD^{-1} = z,$$

$$B^* B - D^* D = D^* (z^* z - 1) D < 0$$

and

$$Y = \begin{pmatrix} 1_m & z \\ z^* & 1_n \end{pmatrix} \text{diag}[A, D]$$

with $z^* z < 1$ thus we get

(4.14) \[ \mathcal{D} \times GL_m(\mathbb{C}) \times GL_n(\mathbb{C}) \cong \mathcal{Y} \]

by $(z, A, D) \mapsto Y(z) \text{diag}[A, D]$ for $Y(z) = \begin{pmatrix} 1_m & z \\ z^* & 1_n \end{pmatrix}$. Here

$$\mathcal{D} = \{ z \in M_{m,n}(\mathbb{C}) \mid z^* z < 1 \}.$$ 

Since $Y \mapsto gY$ for $g \in G$ takes $Y$ into itself isomorphically, we have

$$gY(z) = Y(g(z)) \text{diag}[h(g, z), j(g, z)] \quad h(g, z) = \mathfrak{p} + \mathfrak{d} z \quad \text{and} \quad j(g, z) = cz + d$$

if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$.

By these formulae, it is clear that for the zero matrix $0 \in \mathcal{D}$,

$$G/C_0 \cong \mathcal{D} \quad \text{via} \quad g \mapsto g(0)$$

and $C_0 = U(m) \times U(n)$. Therefore the complexification $C$ of $C_0$ is $GL_m(\mathbb{C}) \times GL_n(\mathbb{C})$.

The functions $^t h^{-1}$ and $j$ correspond to the standard representation of $GL_m$ and $GL_n$, respectively. Since

$$Y(w)^* I_{m,n} Y(z) = \begin{pmatrix} 1-wz^* & z-w \\ wz^* & w^*z^* \end{pmatrix},$$

replacing $z$ and $w$ by $z + \Delta z$ and $w$, we get

$$\text{diag}[^t h(g, z), j(g, z)] \begin{pmatrix} 1-g(w)^*g(z)^* & \Delta g(z) \\ \Delta g(w)g(z)^*-1 \end{pmatrix} \text{diag}[h(g, z), j(g, z)]$$

$$= \text{diag}[^t h(g, z), j(g, z)] Y(g(z)) I_{m,n} Y(g(z) + \Delta z) \text{diag}[h(g, z), j(g, z)]$$

$$= Y(z)^* g^* I_{m,n} g Y(z + \Delta z) = \begin{pmatrix} 1-wz^* & \Delta z \\ \Delta z^* & w^*z^* \end{pmatrix}.$$ 

From this, we conclude

(4.15) \[ dg(z) = ^t h(g, z)^{-1} dz \quad j(g, z)^{-1} \]

We can show (see Shimura’s books: [EPE] (6.3.9) and [AAF] Section 3):

$$\det(h(g, z))) = \det(g)^{-1} \det(j(g, z)).$$
This can be shown also as follows: On the diagonal torus $T_\mathbb{C} \subset U(m) \times U(n)$, for $g = \text{diag}[t_1, \ldots, t_m, t_{m+1}, \ldots, t_{m+n}]$, $t_j$ satisfies $t_j^{-1}$ and
\[ j(g, z) = cz + d = \text{diag}[t_{m+1}, \ldots, t_{m+n}] \quad \text{and} \quad h(g, z) = \overline{a} + \overline{d}z = \text{diag}[t_1^{-1}, \ldots, t_m^{-1}]. \]

Then $j$ (resp. $h$) corresponds therefore to (resp. the contragredient of) the standard representation of $GL_n(\mathbb{C})$ (resp. $GL_m(\mathbb{C})$); so, the corresponding highest weight character, after applying “det”, is:
\[
\text{diag}[t_1, \ldots, t_{m+n}] \mapsto \prod_{j=m+1}^{m+n} t_j \quad \text{(resp. } \prod_{j=1}^m t_j^{-1}).
\]

This relation coincides with the above formula of Shimura. We thus embed the product $U(m) \times U(n)$ into $GL(m) \times GL(n)$ by $g \mapsto J(g) = (h(g, 0)^{-1}, j(g, 0))$. We also write $J(g, z) = (h(g, z)^{-1}, j(g, z)).$

Writing $dz = \bigwedge_{i,j} dz_{ij}$, we get
\[
d\log = \det(g)^n \det(j(g, z))^{-m-n} \, dz.
\]

Write $\mu_{m,n} \in X(T)$ for the character
\[
\mu_{m,n}(\text{diag}[t_1, t_2, \ldots, t_{m+n}]) = (t_1 \times t_2 \times \cdots \times t_m)^{-n} \times (t_{m+1} \times \cdots \times t_{m+n})^m.
\]

Suppose that $\kappa \geq \mu_{m,n}$, and write $\kappa^* = \kappa - \mu_{m,n}$. We try to find a non-zero polynomial function $p_{\kappa^*}: D \to \text{Hom}_\mathbb{C}(L_C(\kappa^*; \mathbb{C}), L_G(\kappa^*; \mathbb{C}))$ in $z$ such that
\[
p(\alpha(z)) \circ \rho_{\kappa^*}^G(J(\alpha, z)) = \alpha p(z) \quad (\rho_{\kappa^*}^G = \text{ind}_B^G \kappa^*)
\]
for all $\alpha \in G$, where $C = GL(m) \times GL(n)$. Since $D = G/C_0$, if it exists, such a function is unique. By the above (hypothetical) formula, we could define $p$ by
\[
p(\alpha(0)) \circ \rho_{\kappa^*}^C(J(\alpha, 0)) = \alpha p(0),
\]
if we find an appropriate map $p(0) \in \text{Hom}_\mathbb{C}(L_C(\kappa^*; \mathbb{C}), L_G(\kappa^*; \mathbb{C}))$. If we change $\alpha$ by $\alpha u$ for $u \in U(m) \times U(n)$, then we have
\[
p(\alpha(0)) \circ \rho_{\kappa^*}^C(J(\alpha)J(u)) = C \alpha p(0)
\]
\[
\iff p(\alpha(0)) \circ \rho_{\kappa^*}^C(J(\alpha)) = \alpha C p(0) \circ \rho_{\kappa^*}^C(J(u))^{-1}.
\]
Such a map $p(0)$ with $C \alpha p(0) \circ \rho_{\kappa^*}^C(J(u))^{-1} = p(0)$ exists because $GL(m) \times GL(n)$ is identified with a subgroup of $GL(m+n)(\mathbb{C}) = U(m+n)(\mathbb{C})$ (thus it corresponds to the identity inclusion: $L_C(\kappa^*; \mathbb{C}) = \rho_{\kappa^*}^C \hookrightarrow \text{ind}_{P}^{U(m+n)} \rho_{\kappa^*}^C = L_G(\kappa^*; \mathbb{C}) = \rho_{\kappa^*}^C$, for $P = \text{diag}[GL(m), GL(n)]B$). Take $\kappa^*|_{T_\mathbb{C}}$ to be the highest weight $\omega_n$ associated to the standard representation of $GL(n)$. Then $\kappa^*$ corresponds to the standard representation of $U(m, n)$, and we have $p_{\omega_n}(z) = (\frac{1}{d}) x$ for $x \in \mathbb{C}^n$. We verify easily that $g p_{\omega_n}(z)(x) = p_{\omega_n}(z)(j(g, z)x)$. Thus $p(z)$ is a polynomial in $z$ in this special case. Similarly to the above, if $\kappa^*|_{T_\mathbb{C}} = \omega_m$ corresponds to the contragredient of the standard representation of $GL(m)$, then $\kappa^*$ is associated to the complex conjugate of the standard representation of $U(m, n)$, and we have $p_{\omega_m}(z)(x) = (\frac{1}{d^*}) x$ for $x \in \mathbb{C}^m$.
Again we verify that $\varphi_{p^{\infty}}(z)(x) = p_{\infty}(z)(\varphi(x))$, and $p_{\infty}(z)$ is a polynomial in $z$.

For general projected image of the tensor product of copies of $p_\infty$ (resp. $L_G(p_\infty ; C)$) is a quotient of $L_G(p_\infty ; C)^{\otimes t} \otimes L_G(p_\infty ; C)^{\otimes s}$ (resp. $L_G(p_\infty ; C)^{\otimes t} \otimes L_G(p_\infty ; C)^{\otimes s}$). The general $p_{\infty}(z)$ is a constant multiple of the projected image of the tensor product of copies of $p_{\infty}(z)$ and hence is a polynomial in $z$.

We define for $f \in H^0(X_T; \omega_\infty^\kappa)$ a holomorphic differential with values in $L(\kappa^*; C)$ by

$$\omega(f) = p_{\kappa^*}(z)(f)dz.$$ 

Note that here $L_G(\kappa; C) = L_G(\kappa^*; C) \otimes L_G(\mu_m; C)$ and that $L_G(\mu_m, n; C)$ is one-dimensional; so, we can identify $L_G(\kappa; C)$ with $L_G(\kappa^*; C)$ canonically as vector space, and thus, the above definition is consistent. We can easily verify that $\alpha^* \omega(f) = \rho^G(\alpha^*) \omega(f)$.

**Theorem 4.4.** — Assume that $\kappa \geq \mu_m, n$. Then the association: $f \mapsto [\omega(f)] \in H^d(X_T, L(\kappa^*; C))$ for $d = \dim_C D$ induces the embedding:

$$H^0_{\text{cusp}}(X_T, \omega^\kappa) \hookrightarrow H^d(X_T, L_G(\kappa^*; C)),$$

where $[\omega(f)]$ is the de Rham cohomology class of $\omega(f)$.

As an exercise, compute $\kappa^*$ when $G = GSp$.

4.2.2. Hecke equivariance. — We are going to show that the Eichler-Shimura map is equivariant under Hecke operators and is compatible with normalization of Hecke operators.

We have normalized the Hecke operator on the topological cohomology group taking the action of $\xi \in \Delta_B$ normalized as $\rho^G_\kappa(\xi^{-1}) = \kappa^{-1}(\xi) \rho^G(\xi^{-1})$. Note that here, for any algebraic character $\chi : G \to \mathbb{Z}_m$, $\tilde{\rho}_\kappa = \rho_\kappa \otimes \chi = \rho_\chi$.

We normalize again in the same way the action on $\omega^\kappa$ taking the action of $\rho^G_\kappa(\xi^{-1}) = \kappa^{-1}(\xi) \rho^G(\xi^{-1})$ in addition to the division by $\mu(\xi)$ ($\mu = \mu_m, n$). Let $T = T(\xi)$ and write also coset representatives by $\xi$. Recalling $\kappa^* = \kappa \mu^{-1}$ and noting that $(\kappa^*)^{-1}(\xi) \rho_\kappa^{-1}(\xi^{-1}) = \kappa^{-1}(\xi) \rho_\kappa(\xi^{-1})$ for $\xi \in D$, we have

$$\omega(f)|_{\text{top}} = \sum_{\xi} (\kappa \mu^{-1}(\xi))^{-1} \rho^G_\kappa(\xi^{-1}) p(\xi) (f(\xi)) d(\xi)$$

$$= \sum_{\xi} p(z) ((\kappa \mu^{-1}(\xi))^{-1} \rho_\kappa(\xi^{-1}) - \mu(\xi)^{-1} f(\xi)) \mu(\xi)^{-1} dz$$

$$= p(z) \left( \mu(\xi)^{-1} \sum_{\xi} \kappa^{-1}(\xi) \rho_\kappa(\xi) J(\xi, z)^{-1} f(\xi) \right) dz = \omega(f)|_{\text{coh}}.$$ 

In short, the extra modification of the action of the Hecke operator $T(\xi)$ by the character $\mu$ on the coherent cohomology is absorbed by $d(\xi) = \mu(\xi)^{-1} dz$ on the topological cohomology. Hence the normalization of Hecke operators at $p$ is identical on the left-hand-side and the right-hand-side of the Eichler-Shimura map.
5. Moduli of Abelian Schemes

We recall the construction of moduli spaces of abelian schemes. The theory of moduli varieties of abelian varieties has been studied mainly by Shimura and Mumford in the years 1950’s to 1960’s. Shimura proved in the late 1950’s to the early 1960’s the existence of the moduli varieties over a canonically determined number field relative to a given endomorphism ring, a level $N$-structure and a polarization. This of course gives a moduli over the integer ring of the field with sufficiently large number of primes inverted.

Basically at the same time, Grothendieck studied the moduli of subschemes in a given projective scheme $X_S$ (flat over $S$) and also that of the Picard functors. The existence of a moduli scheme, the Hilbert scheme $\text{Hilb}_{X/S}$, of closed flat subschemes of $X_S$ enabled Mumford, via his theory of geometric quotients of quasi-projective schemes ([GIT]), to construct moduli of abelian schemes with level $N$-structure over $\mathbb{Z}[rac{1}{N}]$.

We recall here the construction of Grothendieck and Mumford briefly, limiting to the cases which we will need later. We will redo the construction of Shimura varieties with a canonical family of abelian varieties in the following lectures.

5.1. Hilbert Schemes. — In this subsection, we describe the theory of the Hilbert scheme which classifies all closed $S$-flat subschemes of a given projective variety $X_S$. This is a generalization of the earlier theory of Chow coordinates which classifies cycles on a projective variety. The theory is due to A. Grothendieck and main source of the exposition here is his Exposé 221 in Sémin. Bourbaki 1960/61.

5.1.1. Grassmannians. — Let $GL(n) : \text{ALG} \to \text{GP}$ for the category of groups $\text{GP}$ be the functor given by $GL(n)(A) = GL_n(A)$. This functor is representable by a group scheme $GL(n) = \text{Spec} \left( \mathbb{Z} \left[ t_{ij}, \frac{1}{\det(t_{ij})} \right] \right)$. We may extend the functor to the category of schemes $\text{SCH}$ by $GL(n)(S) = \text{Aut}_{\text{Sch}}(\mathcal{O}_S^n)$. We recall the notation $[\ast] = \{ \ast \}/\cong$ introduced in 2.1.5 which implies the set of isomorphism classes of the objects: "$\ast$" in the bracket. Then we define a contravariant functor $\text{Grass} : \text{SCH} \to \text{SETS}$ by

$$\text{Grass}_{\mathcal{O}_S^n}(S) = \left[ \pi : \mathcal{O}_S^n \to \mathcal{F} \mid \pi : \mathcal{O}_S\text{-linear surjective, } \mathcal{F} \text{ locally } \mathcal{O}_S\text{-free of rank } m \right].$$

For each morphism $f : T \to S$, the pullback $f^*\pi : \mathcal{O}_T^n \to f^*\mathcal{F}$ gives contravariant functoriality. The quotient $\pi : \mathcal{O}_S^n \to \mathcal{F}$ is isomorphic to $\pi' : \mathcal{O}_S^n \to \mathcal{F}'$ if we have the following commutative diagram:

$$
\begin{array}{cccccc}
0 & \longrightarrow & \text{Ker}(\pi) & \longrightarrow & \mathcal{O}_S^n & \xrightarrow{\pi} & \mathcal{F} & \longrightarrow & 0 \\
| & & | & & | & | & | & | & |
0 & \longrightarrow & \text{Ker}(\pi') & \longrightarrow & \mathcal{O}_S^n & \xrightarrow{\pi'} & \mathcal{F}' & \longrightarrow & 0
\end{array}
$$

ASTÉRISQUE 298
with exact rows. The stabilizer of $\pi : O^a_S \to O^m_S$ can be identified with the maximal parabolic subgroup
\[ P = \{ (a \ b) \in GL(n) \mid d \text{ is of size } m \times m \} . \]
As is well known, the quotient Grass_{O^m_S} = GL(n)/P is a projective scheme defined over $\mathbb{Z}$ and represents the functor Grass_m, that is,
\[ Hom_{SCH}(S, Grass_{O^m_S}) \cong [\pi : O^a_S \to F] \]
functorially. Of course, if $m = n - 1$ or 1, we have Grass_{O^m_S} \cong \mathbb{P}^{n-1}.

We can generalize this construction slightly: Let $S$ be a scheme. Let $\mathcal{E}/S$ be a locally free sheaf on $S$ of constant rank $n$. Then, for each $S$-scheme $S' \to S$, we define a contravariant functor from $S$-$SCH$ to $SETS$ by
\[
Grass_{n,m}(S') = \left[ \pi : f^*\mathcal{E} \to \mathcal{F} \mid \pi : O_T \text{-linear surjective, } \mathcal{F} \text{ locally } O_{S'} \text{-free of rank } m \right] .
\]
Then covering $S$ by sufficiently small open subschemes $U_i$ so that $\mathcal{E}_{U_i} \cong O^a_{U_i}$, we have Grass_{O^m_{U_i}} represented by Grass_{m/U_i} = Grass_m \times U_i. The gluing data $g_{ij} : O_{U_i \cap U_j} \cong \mathcal{E}_{U_i \cap U_j}$ give rise to a Čech cocycle $g_{ij}$ with values in $GL(n)$. This gluing datum induces a gluing datum of $\{ Grass_{m/U_i} \}$, giving rise to the scheme Grass_{n,m} over $S$ which represents the above functor. One can find a detailed proof (from a slightly different viewpoint) of what we have said here in [EGA] I.9.7.

5.1.2. Flag Varieties. — We can further generalize our construction of the grassmanian to flag varieties. We follow [EGA] I.9.9. We consider the following functor from $S$-$SCH$ to $SETS$:
\[
\text{Flag}_B(S') = \left[ \pi_j : \mathcal{E} \to \mathcal{F}_j \mid \text{Ker}(\pi_{j+1}) \subset \text{Ker}(\pi_j), \text{ and } \mathcal{F}_j \text{ is locally free of rank } n - j \ (1 \leq j \leq n-1) \right] .
\]
Here the subscript $B$ indicates a split Borel subgroup of $GL(n)$, since Flag is represented by a projective scheme $\text{Flag}_B = GL(n)/B$ if $\mathcal{E} = O^a_S$. In general, we can show that
\[
\text{Flag}_B \hookrightarrow \text{Grass}_1 \times_S \text{Grass}_2 \times_S \cdots \times_S \text{Grass}_{n-1}
\]
given by $(\pi_j) \mapsto (\pi_{n-j} \in \text{Grass}_j)$ is a closed immersion ([EGA] I.9.9.3). By Plücker coordinates ([EGA] I.9.8), we can embed Grass_m into the projective bundle of $\bigwedge^n \mathcal{E}$; so, Flag_B is projective.

The Flag variety is basically the quotient of $GL(n)$ by its upper triangular Borel subgroup $B$. We can generalize the construction to the quotient of $GL(n)$ by the unipotent radical of $B$. We consider the following functor:
\[
\text{Flag}_{\phi}(S') = \left[ (\pi_j, \phi_j) \mid (\pi_j) \in \text{Flag}_B(S') \text{ and } \phi_{j+1} : \text{Ker}(\pi_j)/\text{Ker}(\pi_{j+1}) \cong O_{S'} \right] .
\]
Here we understand that $\text{Ker}(\pi_0) = \mathcal{E}$, and $j$ runs over all integers between 0 and $n-1$. If $\mathcal{E} \cong O^a_S$ and $S$ is affine, writing $1 = (\pi_j, \phi_j)$ for the standard flag $\pi_j$:
Projective means that we have a closed immersion $\text{Flag}_i$ is represented by Flag. This way, we can associate a $\kappa$-power of $\text{Flag}_i$. Obviously $\text{Flag}_i$ is a $T$-torsor over $\text{Flag}_B$ for the maximal split torus $T \subset GL(n)$. Here the action of $T$ on $\text{Flag}_U$ is given by $(\pi_j, \phi_j) \mapsto (\pi_j, t_j \phi_j)$ for $(t_1, \ldots, t_n) \in T = \mathbb{G}_m \times \cdots \times \mathbb{G}_m$. See [GME] 1.8.3 about torsors.

Let $\pi : \text{Flag}_U \to \text{Flag}_B$ be the projection: $(\pi_j, \phi_j) \mapsto (\pi_j)$. Then for a character $\kappa$ of $T$, we define a sheaf $\mathcal{E}^\kappa(V) = H^0(\pi^{-1}(V), \mathcal{O}_{\text{Flag}_U}[\kappa])$ for each open subset $V \subset \text{Flag}_B$. Then $\mathcal{E}^\kappa$ is a locally free sheaf on $\text{Flag}_B$. Since $f : \text{Flag}_B \to S$ is proper flat over $S$, we find that $f_* \mathcal{E}^\kappa$ (which we again write $\mathcal{E}^\kappa$) is a locally free sheaf on $S$. In this way, we can associate a $\kappa$-power $\mathcal{E}^\kappa$ of the original locally free sheaf $\mathcal{E}$, which is non-zero if and only if $\kappa$ is dominant weight $\kappa$ of $GL(n)$ with respect to $(B, T)$.

5.1.3. Flat Quotient Modules. --- Let $f : X \to S$ be a flat projective scheme over a (separated) noetherian connected scheme $S$ of relative dimension $n$. Here the word “projective” means that we have a closed immersion $\iota : X/S \hookrightarrow P^n_S$. Thus $X$ has a very ample invertible sheaf $\mathcal{O}_X(1) = \iota^* \mathcal{O}_{P^n}(1)$. The sheaf of graded algebras $\mathcal{A} = \bigoplus_{n=0}^\infty f_*(\mathcal{O}_X(1)^n)$ determines $X$ as $X = \text{Proj}_S(\mathcal{A})$.

For a given coherent sheaf $\mathcal{F}$ on $X$, we write $\mathcal{F}(k)$ for $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X(k)$ and define a sheaf of graded modules $F = \bigoplus_{k \geq 0} F_k$ by $F_k = f_* \mathcal{F}(k)$. Then $F$ is a graded $\mathcal{A}$-module of finite type, and we have $\mathcal{F} = \mathcal{F}$. Removing first finitely many graded pieces of $F$ does not alter $\mathcal{F} = \mathcal{F}$. Thus defining $F(n) = \oplus_{k \geq 0} F(n)_k$ with $F(n)_k = F_{n+k}$, we have $\mathcal{F}(n) = \mathcal{F}(n)$.

We suppose that $\mathcal{F}$ is $\mathcal{O}_S$-flat. For each geometric point $s = \text{Spec}(k(s)) \in S$, there is a polynomial $P_\mathcal{F}(T)$ such that

$$\chi(\mathcal{F}(n)) = \sum_{j=0}^{\dim X(s)} (-1)^j \dim_{k(s)} H^j(X(s), \mathcal{F}(n) \otimes_{\mathcal{O}_S} k(s)) = P_\mathcal{F}(n).$$

For sufficiently large $n$, the ampleness of $\mathcal{O}(1)$ tells us that

$$H^j(X(s), \mathcal{F}(n) \otimes_{\mathcal{O}_S} k(s)) = 0 \quad \text{if } j > 0.$$
We consider the following contravariant functor for each locally noetherian $S$-scheme $\phi : T \rightarrow S$ (inducing $\phi_X : X_T = X \times_S T \rightarrow X$):

$$\text{Quot}_{/X/S}(T) = \{ \pi : \phi_X^* \mathcal{F} \rightarrow \mathcal{M} \mid \mathcal{M} \text{ is a coherent } \mathcal{O}_{X_T}\text{-module flat over } \mathcal{O}_T \}.$$  

The isomorphism between such $\pi$’s are similarly defined as in the case of $\text{Grass}$.

For simplicity, we always assume that schemes $T$ are noetherian. Each $\mathcal{M} \in \text{Quot}(T)$ has its Hilbert polynomial $P_M$, and obviously for $g : T' \rightarrow T$, $g^* \mathcal{M}$ has the same Hilbert polynomial. Thus we can split the functor as

$$\text{Quot}_{/X/S} = \bigsqcup_P \text{Quot}_{/X/S}'(P),$$

where

$$\text{Quot}_{/X/S}'(P) = \{ \pi : \phi_X^* \mathcal{F} \rightarrow \mathcal{M} \mid \mathcal{M} \in \text{Quot}_{/X/S}(T), P_M = P \}.$$  

Here is a theorem of Grothendieck:

**Theorem 5.1.** — Let the notation be as above. Suppose that $X/S$ is projective. Then the functor $\text{Quot}_{/X/S}'$ is representable by a projective scheme $Q^P = \text{Quot}_{/X/S}'$ of finite type over $S$. Thus for any $S$-scheme $T$,

$$\text{Hom}_S(T, Q^P) \cong \{ \pi : f^* \mathcal{F} \rightarrow \mathcal{M} \mid \mathcal{M} \in \text{Quot}_{/X/S}(T), P_M = P \}$$

functorially.

We are going to give a sketch of the proof of this theorem. We recall $X = \text{Proj}_S(A)$ for a sheaf $A$ of graded $\mathcal{O}_S$-algebras generated by $A_1$. We cover $\text{Quot}_{/X/S}'$ by the subfunctor $Q_j$ defined as follows for each positive integer $j$: $Q_j(T)$ consists of isomorphism classes of $\pi : \mathcal{F}_{/X_T} \rightarrow \mathcal{M}_{/X_T}$ satisfying the following three conditions:

(a) $R^n f_{T,*}M(n)/X_T = 0$ for all $i > 0$ and $n \geq j$;
(b) $R^n f_{T,*}K(n)/X_T = 0$ for all $i > 0$ and $n \geq j$, where $K = \text{Ker}(\pi)$;
(c) $A_k f_{T,*}(K(j)) = f_{T,*}(K(j + k))$ for all $k > 0$.

Write $K$ (resp. $M$) for the graded $\phi^*A$-module with $\tilde{K} = K$ (resp. $\tilde{M} = M$). Define $K(j)$ and $M(j)$ as above; so, $K(j) = \oplus_{k \geq 0} K(j)_k$ with $K(j)_k = K(j + k)$. First covering $T$ by affine schemes $\text{Spec}(B_i)$ and writing $B_i$ as a union of noetherian rings, we can reduce proofs to noetherian $T$; so, we may assume that $T$ is noetherian as we remarked already. Then by a theorem of Serre ([EGA] III Section 2), for any coherent sheaf $\mathcal{G}_{/X_T}$, we have the vanishing: $R^n f_{T,*}G(n)/X_T = 0$ for $n \gg 0$. Thus (a) and (b) will be satisfied for a given $\pi$ for a suitable $j$. Since $\mathcal{F}$ is coherent (and $X/S$ is of finite type), it is finitely presented; so, $K$ is finitely generated as $\phi^*A$-modules, because $M$ is finitely presented (cf. [CRT] Theorem 2.6). Thus $K(j)$ is generated by $K_j = K(j)_0$ for some $j$, and the last condition will be fulfilled again if $j \gg 0$. This shows that $\text{Quot}_{/X/S}'(T)$ is covered by $Q_j(T)$ for each $T$.

The Euler characteristic is additive with respect to the exact sequence: $0 \rightarrow K \rightarrow \mathcal{F} \rightarrow \mathcal{M} \rightarrow 0$: $\chi(\mathcal{F}) = \chi(K) + \chi(\mathcal{M})$. Thus the conditions (a) and (b) tell us...
that the Hilbert polynomials \( P_\mathcal{K}(n) \) and \( P_\mathcal{F}(n) \) give exact \( \mathcal{O}_T \)-rank of \( f_{T,*}\mathcal{K} \) and \( f_{T,*}\mathcal{F} \). The vanishing of \( R^1 f_{T,*}\mathcal{X} \to X_T = 0 \) implies that \( f_{T,*}\mathcal{X} = 0 \) (EGA III, [ALG] III.12.10 or [GIT] 0.5); so, the conditions (a) and (b) are stable under base-change. The tensor product is a right exact functor; so, the surjectivity of \( p_k : \mathcal{A}_k \otimes f_{T,*}\mathcal{K}(j) \to f_{T,*}\mathcal{K}(j+k) \) is also kept under base-change; so, \( \mathcal{Q}_j \) is a well defined contravariant functor, and we have

\[
\text{Quot}^P = \bigcup_j \mathcal{Q}_j.
\]

By these three conditions (a-c), \( \pi \in \mathcal{Q}_j(T) \) is determined by \( M_j = f_{T,*}\mathcal{M}(j) \) as a flat quotient of rank \( P(j) \) of \( F_j = f_{T,*}\mathcal{F}(j) \). Thus \( \pi \mapsto (\pi_j : F_j \to M_j) \) induces a functorial injection

\[
\mathcal{Q}_j(T) \hookrightarrow \text{Grass}_{F_j,P(j)}(T).
\]

If \( \pi' : F_j \to M' \) is in the image of \( \mathcal{Q}_j \), then \( \pi' \) satisfies:

(i) \( F_{j+k} / (f^*(\mathcal{A}_k)K') \) for \( K' = \text{Ker}(\pi') \) is locally \( \mathcal{O}_T \)-free of rank \( P(j+k) \) for all \( k \geq 0 \). Here we consider \( f^*(\mathcal{A}_k)K' \) in \( f^*(\mathcal{A}) \cdot F \);

(ii) Define a graded module \( K'' \) by \( f^*(\mathcal{A}) \cdot K' \). We require the associated sheaf \( \mathcal{K}' = K''/X_T \) on \( X_T \) to satisfy (b) and the quotient \( M' = \mathcal{F}/\mathcal{K}' \) to satisfy (a) (these (a) and (b) are open conditions).

For any graded \( f^*(\mathcal{A}) \)-module \( M = \bigoplus_k M_k \), defining \( M^{(t)} = \bigoplus_{k \geq t} M_k \), we have \( \overline{M}^{(t)} \cong \overline{M} \) as already remarked. By this fact, the condition (i) assures that \( f_{T,*}\mathcal{M} \) is locally \( \mathcal{O}_T \)-free, and the image of \( \mathcal{Q}_j(T) \) is characterized by (i) and (ii).

We are going to prove the representability of \( \mathcal{Q}_j \), assuming that \( j = 0 \). The general case follows from the same argument replacing 0 by \( j \) everywhere. Let \( \pi^\text{univ} : F_0/G \to M_0 \) be the universal object defined over \( G = \text{Grass} \). Here we have changed our notation and write \( M_0 \) for the universal quotient of \( F_0/G \) (with rank \( P(0) \)). Thus for any morphism \( \pi' : F_0 \to M'_0 \) on \( T \) with \( M'_0 \) locally-free of rank \( P(0) \), we have a unique \( S \)-morphism \( \phi : T \to \text{Grass} \) such that \( \pi' = f^*(\mathcal{A}) \cdot \phi^\text{univ} \). Let \( K = \text{Ker}(\pi^\text{univ}) \). Write \( g : G = \text{Grass} \to S \) for the structure morphism. We consider the subset:

\[
Z = \left\{ s \in G \left| \dim_{k(s)} (\mathcal{A}_k F_0/G/(\mathcal{A}_k K)) \otimes k(s) = P(k) \text{ for all } k \geq 0 \right. \right\}
\]

Write \( M_k = \mathcal{A}_k F_0/G/(\mathcal{A}_k K) \) and put \( M = \bigoplus_{k \geq 0} M_k \). Then \( M_k = \mathcal{A}_k M_0 \). The \( \mathcal{O}_G \)-module \( M \) is flat on a generic point of \( Z_0 = \text{Supp}(M) = G \). Since flatness is an open condition, we find an open connected subscheme \( V_0 \subset Z_0 \) which is maximal among open subschemes \( V \) over which \( M \) is flat. Repeating this process, replacing \( M \) and \( Z_0 \) by \( M \otimes_{\mathcal{O}_{Z_0}} \mathcal{O}_{Z_1} \) and \( Z_1 = Z_0 - V_0 \), we can split \( Z_0 = \bigsqcup V_i \) into a finite disjoint union of connected subschemes \( V_i \) so that \( M \otimes_{\mathcal{O}_{Z_0}} \mathcal{O}_{V_i} \) is flat over \( V_i \). Then we find a
polynomial $Q_i(n)$ such that
\[ \text{rank}_{\mathcal{O}_{V_i}}(M_n \otimes_{\mathcal{O}_X} \mathcal{O}_{V_i}) = Q_i(n) \text{ if } n \geq n_i. \]

By this fact, the open subscheme
\[ U_N = \left\{ s \in Z_0 \mid \dim_{k(s)}(M_n \otimes_{\mathcal{O}_{Z_0}} k(s)) \leq P(n) \ 0 \leq n \leq N \right\} \]
stabilizes as $N$ grows. Therefore on an open (dense) subscheme $U = U_{\infty}$ of $Z_0$, if $n \geq 0$, we have $\dim_{k(s)}(M_n \otimes_{\mathcal{O}_{Z_0}} k(s)) \leq P(n)$ for all $s \in U$. Then we have an exact sequence:
\[ \mathcal{O}_U^j \xrightarrow{p_k} \mathcal{O}_U^{P(k)} \rightarrow M_k \rightarrow 0, \]
and $Z$ is the closed subscheme of $U$ on which all matrix coefficients of $p_k$ vanishes for all $k \geq 0$. Thus the image of $Q_0$ fall into $Z$. The condition (ii) can be checked to be satisfied on an open subscheme of $Z$. Thus we have

**Theorem 5.2.** — The functor $Q_j$ is represented by a quasi-projective scheme $Q_j$ of finite type over $S$.

Here the word “quasi-projective” means that the scheme has an open immersion into a projective scheme. Since Grass is projective, $Q_j$ is quasi-projective.

The next step is to show that the increasing sequence of quasi-projective schemes $\{Q_j\}_j$ stabilizes after $j \geq N_0$; so, $\text{Quot}^P$ is represented by a quasi-projective scheme. The key point of the argument is to show that for any given set of coherent sheaves $\mathcal{F}$ on $X$, each of whose members appear as a fiber of a coherent sheaf $\mathcal{E}$ of $\check{\text{C}}$ech cohomology with respect to the covering $\mathcal{U}$ of $\mathcal{X}$, then $R^if_*\mathcal{O}(j) = 0$ for all $i > 0$ if $j > n + 1$ by a computation of cohomology groups by $\check{\text{C}}$ech cohomology with respect to the covering $X = \bigcup_{j=0}^\infty D_j$ (see [ALG] III.5).

A version of the argument of Grothendieck for $X = \mathbb{P}^n$ to prove (a) and (b) for sufficiently large $j$ for all $\mathcal{M}$ and $\mathcal{K}$ is as follows: Since $P_{\mathcal{F}} = P_{\mathcal{K}} + P_{\mathcal{M}}$ with $P_{\mathcal{M}} = P$, $P_{\mathcal{K}}$ is determined by $P$. Choosing homogeneous generators $x_1, \ldots, x_r$ of degree $-p$ of $\mathcal{K}$, we have a surjection: $\mathcal{O}(p)^r \rightarrow \mathcal{K}$ taking $(a_1, \ldots, a_r) \mapsto \sum_{i=1}^r a_ix_i$. Here $r$ and $p$ are determined by the first two leading terms of $P_{\mathcal{K}}$ and hence those of $P$. Let $\mathcal{K}_0 = \text{Ker}(\pi)$. Then $rP_{\mathcal{O}(p)} = P_{\mathcal{K}_0} + P_{\mathcal{K}}$. Let $r_0 = r$ and $p = p_0$. The polynomial $P_{\mathcal{K}_0}$ is determined by $P_{\mathcal{K}}$. Thus the first two leading terms of $P_{\mathcal{K}_0}$ are bounded below and above independent of $\mathcal{K}$. 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
Repeating this argument, we find an integer \(N_0 \gg 0\) such that for integers \(p_i > -N_0\) \((i = 0, 1, \ldots, n)\) we have the following exact sequences:

\[
0 \rightarrow \mathcal{K}_i \rightarrow \mathcal{O}(p_i)^{\gamma_i} \rightarrow \mathcal{K}_{i-1} \rightarrow 0
\]

with \(\mathcal{K}_{-1} = \mathcal{K}\). Then by the associated long exact sequence, if \(j \geq N_0 + n + 1\), \(R^{n+i+1}f_\ast\mathcal{K}_i(j) \cong R^n f_\ast\mathcal{K}_{i-1}(j)\). Since cohomological dimension of \(\mathbb{P}^n\) is \(n\) (as easily checked by Čech cohomology), for \(i > 0\),

\[
0 = R^{n+i+1}f_\ast\mathcal{K}_n(j) = R^{n+i}f_\ast\mathcal{K}_{n-1}(j) = \cdots = R^i f_\ast\mathcal{K}(j).
\]

By the same argument, \(R^i f_\ast\mathcal{M}(j) = 0\) for all \(i > 0\) and all \(j > N_0 + n + 1\).

Since \(D_i \cap X\) is affine for any projective scheme \(X_S \subset \mathbb{P}^n_S\), the same argument works for \(X\) in place of \(\mathbb{P}^n\), and \(\text{Quot}^P_{\mathcal{O}_X/X_S}\) is represented by a quasi-projective scheme (see [PAF] pages 261-262). It is customary to identify \(\pi \in \text{Quot}_{\mathcal{O}_X/X_S}(T)\) with a closed immersion of \(\text{Spec}_Q((\text{Im} (\pi)))\) into \(X\); so, \(\text{Quot}^P_{\mathcal{O}_X/X_S}\) represents the following contravariant functor

\[
\text{Hilb}^P_X(T) = \{\text{closed subschemes of } X_T \text{ flat over } T \text{ with Hilbert polynomial } P\}.
\]

This scheme is called the \(\text{Hilbert scheme}\) of \(X\) for the polynomial \(P\).

We now finish the proof of the following theorem.

**Theorem 5.3 (Grothendieck).** — For each projective scheme \(X_S\) for a noetherian connected scheme \(S\) and a numerical polynomial \(P(t) \in \mathbb{Q}[t]\), the functor \(\text{Hilb}^P_X/S\) is represented by a projective scheme \(\text{Hilb}^P_X/S\) over \(S\).

**Proof.** — We only need to prove the projectivity by the valuative criterion. Let \(\pi : \mathcal{O}_X \rightarrow \mathcal{M}/\eta \in \text{Quot}^P_{\mathcal{O}_X/X_S}(\eta)\) for \(\eta = \text{Spec}(K)\) of the field \(K\) of fractions of a discrete valuation ring \(V\). Then we define \(\text{Ker}(\pi)_S\) for \(T = \text{Spec}(V)\) by the largest subsheaf over \(T\) of \(\mathcal{O}_X\) inducing \(\text{Ker}(\pi)\), that is, \(\mathcal{O}_X \cap \text{Ker}(\pi)\), which is a coherent sheaf with quotient \(\mathcal{M}/T\) locally free over \(X_T\) inducing \(\mathcal{M}/\eta\) after tensoring \(K\), because \(V\) is a discrete valuation ring. Thus the point \(\pi \in \text{Quot}^P_{\mathcal{O}_X/X_S}(\eta)\) extends to \(\text{Quot}^P_{\mathcal{O}_X/X_S}(T)\). Since \(\text{Quot}^P = \bigcup_j Q_j\) is quasi-projective, it is separated; so, it is proper. Since \(\text{Quot}^P_{\mathcal{O}_X/X_S}\) is quasi-projective, it has to be projective. \(\square\)

### 5.1.4. Morphisms between Schemes.

In this section, we first consider the contravariant functor \(\text{Sec}_{X/Y,S}, \text{Hom}_S(X, Y) : S\text{-SCH} \rightarrow \text{SETS}\) given by

\[
\text{Sec}_{X/Y,S}(T) = \text{Hom}_T(Y_T, X_T) \quad \text{and} \quad \text{Hom}_S(X, Y)(T) = \text{Hom}_T(X_T, Y_T).
\]

Here for \(\text{Sec}\), \(X\) is supposed to be an \(S\)-scheme over \(Y\). The latter is a special case of the former because

\[
\text{Hom}_T(X_T, Y_T) = \text{Sec}_{Y_X/S}(T) \quad (Y_X = Y \times_S X).
\]
Each section $s : Y \to X$ defines a closed subscheme of $X_S$ isomorphic to $Y$ via the given projection $f : X \to Y$. Write $H = \text{Hilb}_{X/S} = \bigcup_p \text{Hilb}^p_{X/S}$. Then we have the universal closed subscheme $Z$ of $X_H = X \times_S H$ satisfying the commutative diagram:

$$
\begin{array}{ccc}
Z & \overset{\subset}{\longrightarrow} & X_H \\
\downarrow & & \downarrow \\
H & \overset{\cong}{\longrightarrow} & H
\end{array}
$$

such that for any $S$-scheme $T$ and a closed subscheme $W \hookrightarrow X_T$ flat over $T$, we have a unique morphism $\phi_W : T \to H$ over $S$ such that the pull back of the above square by $\phi_W$ is identical to

$$
\begin{array}{ccc}
W & \overset{\subset}{\longrightarrow} & X_T \\
\downarrow & & \downarrow \\
T & \overset{\cong}{\longrightarrow} & T.
\end{array}
$$

We consider $S$-subschemes $U \subset H$ such that $\tilde{f}_U : Z_U \subset X_U \overset{f_U}{\to} Y_U$ for a given $f : X \to Y$ induces an isomorphism $\tilde{f}_U : Z_U \cong Y_U$. From this, it is easy to see that $\text{Sec}_{X/Y/S}$ is represented by the maximal subscheme $U$ of $H_S$ with this property $\tilde{f}_U : Z_U \cong Y_U$. For each closed point $x \in H$, if $\tilde{f}_x$ is an isomorphism, it is an isomorphism on an open neighborhood of $x$; so, $U$ is an open subscheme of $H$. Since $\text{Hilb}^p_{X/S}$ is projective over $S$, and $U \cap \text{Hilb}^p_{X/S}$ is open, each connected component of $U$ is an open-subscheme of the projective scheme $\text{Hilb}^p_{X/S}$ for some $P$; so, each connected component of $U$ is quasi-projective over $S$. Thus we get

**Theorem 5.4.** — Let $X_S$ and $Y_S$ be projective schemes over a connected noetherian scheme $S$. Then the functors $\text{Sec}_{X/Y/S}$ and $\text{Hom}_S(X, Y)$ are representable by schemes $\text{Sec}_{X/S}$ and $\text{Hom}_{X/S}$ over $S$, respectively. Each connected component of $\text{Sec}_{X/S}$ and $\text{Hom}_{X/S}$ is quasi-projective over $S$.

By construction, the scheme representing these functors may not be of finite type over $S$, because $\text{Hilb}^p_{X/S}$ could have infinitely many components. However each connected component of the scheme is of finite type over $S$.

**Corollary 5.5.** — Let the notation and the assumption be as in the theorem. Then the functor: $T_S \mapsto \text{Hom}_T(X_T, X_T)$ is represented by a scheme $E_{X/S}$ over $S$ whose connected components are quasi-projective over $S$.

If a section $s : S \to X$ is given, keeping representability, we can insist an endomorphism $\phi \in E_T(X_T)$ to take $s_T$ to $s_T$. This goes as follows: Consider the functorial map: $E_{X/S}(T) \ni \phi \mapsto \phi(s_T) \in X(T)$ which induces a morphism $\sigma : E_{X/S} \to X$. Then
writing the set of endomorphisms keeping \( s \) as \( E^s_T(X_T) \), the functor: \( T \mapsto E^s_T(X_T) \) is again representable by a scheme
\[
E^s_{X/S} = E_{X/S} \times_{X,s,s} S
\]
over \( S \).

5.1.5. Abelian Schemes. — An abelian scheme \( X/S \) is a smooth geometrically connected group scheme proper over a separated locally noetherian base \( S \).

We can drop “local noetherian” hypothesis, because a smooth geometrically connected and proper group scheme over any base is a base change of such a scheme over a locally noetherian base (cf. [DAV] I.1.2).

We actually suppose that \( S \) is a noetherian scheme for simplicity. Since \( X \) is a group, it has the identity section \( 0 : S \to X \). As in the elliptic curve case, any \( S \)-morphism \( \phi : X \to X' \) of abelian schemes is a homomorphism if \( \phi(0_X) = \phi \circ 0 = 0_{X'} \) (by Rigidity lemma: [ABV] Section 4, [GIT] 6.4 and [GME] 4.1.5). In particular, if \( X \) is an abelian scheme over \( S \), every scheme endomorphism of \( X/S \) keeping the zero section is a homomorphism of group structure. Thus \( E^0_{X/S} \) is a ring scheme associated to the functor: \( T \mapsto E^0_T(X_T) \) with values in the category of rings.

Assume that \( X \) is an abelian scheme over a connected noetherian base \( S \). Take a connected component \( E \subset E^0_{X/S} \). Each connected component of \( E_{X/S} \) is quasi-projective over \( S \). Since \( S \) is noetherian, \( E \) is of finite type over \( S \), because of our construction:
\[
E^0_{X/S} = E_{X/S} \times_{X,0,0} S.
\]
Suppose we have a discrete valuation ring \( A \) with field of fractions \( K \) and a morphism \( \eta : \text{Spec}(K) \to E \) which is over a morphism \( i : \text{Spec}(A) \hookrightarrow S \). In other words, we have the following commutative diagram:

\[
\begin{array}{ccc}
\text{Spec}(K) & \xrightarrow{\eta} & E \\
| & | & | \\
\text{Spec}(A) & \xrightarrow{i} & S.
\end{array}
\]

Then \( \eta \) gives rise to a section of \( E^0_{X/S}(K) \). Since homomorphisms of abelian schemes are kept under specialization (which we call the rigidity of endomorphism; see [GME] Subsections 4.1.5-6 and [DAV] I.2.7), \( \eta \) extends to \( \text{Spec}(A) \) uniquely. By the valuative criterion of properness, we find that \( E \) is projective over \( S \). If \( \phi \) is an endomorphism of the abelian scheme \( X/S \), \( \text{Ker}(\phi) \) is again a group scheme. If \( \dim_S \text{Ker}(\phi) = 0 \), \( \text{Ker}(\phi) \) is a locally-free group scheme of finite rank; in this case, we call \( \phi \) an isogeny. We define the degree \( \text{deg}(\phi) \) of \( \phi \) by the rank of \( \text{Ker}(\phi) \) over \( S \) in this case. If \( \dim_S \text{Ker}(\phi) > 0 \), we simply put \( \text{deg}(\phi) = 0 \). If the connected component \( E \subset E^0_{X/S} \) contains an isogeny, the degree is independent of the point of \( E \). As is well known, for any abelian variety
over a field, the number of isogeny with a given positive degree is finite. Thus \( E \) is projective and quasi-finite; so, \( E \) is finite over \( S \) ([GME] Proposition 1.9.11).

**Corollary 5.6.** — Let \( X/S \) be an abelian scheme over a connected noetherian base \( S \). Then the functor \( T \mapsto \text{End}_T(X_T) \) is represented by a scheme \( \text{End}_{X/S} = E_{X/S}^0 \) over \( S \). Each connected component of \( \text{End}_{X/S} \) is projective over \( S \). If the connected component contains an isogeny, it is finite over \( S \). Here \( \text{End}_T(X_T) \) denotes endomorphisms of \( X \) compatible with group structure on \( X \).

The subscheme \( E \) corresponds to a section \( x_E \) of \( \text{End}_{X/S}(X_E) = \text{End}_E(X_E) \). We assume that \( x_E : X_E \to X_E \) is an isogeny; so, \( \text{Ker}(x_E) \) is a locally-free group scheme over \( E \), and \( E \) is finite over \( S \). We consider the subalgebra \( A_E = \mathbb{Z}[x_E] \subseteq \text{End}(X_E) \). Since \( E \) is connected, \( \text{End}_E(X_E) \) is an algebra free of finite rank over \( \mathbb{Z} \) (see [ABV] Section 19). We suppose that \( B = A_E \otimes_\mathbb{Z} \mathbb{Q} \) is semi-simple and commutative. Thus it is a product of finitely many number fields, and hence the algebra automorphism group \( \text{Aut}(B) \) is finite.

Let us fix a commutative algebra \( A \) free of finite rank over \( \mathbb{Z} \) with semi-simple \( B = A \otimes_\mathbb{Z} \mathbb{Q} \). Suppose that \( A = \mathbb{Z}[x] \) for a single generator \( x \). Let \( \mathbb{E}_x \subseteq \text{End}_{X/S} \) be the collection of all connected components \( E \) such that \( \theta_E : A \cong A_E \text{ with } \theta(x) = x_E \). A priori, the scheme \( \mathbb{E}_{x/S} \) may have infinitely many connected components, although we later see that up to inner automorphisms of \( \text{End}(X_{s/s}) \), the number of components are finite using the fact that \( \text{End}(X_{s/s}) \otimes_\mathbb{Z} \mathbb{Q} \) is a semi-simple algebra of bounded dimension, where \( s \in S \) is a geometric point and \( X_s \) is the abelian variety fiber \( X_s \) over \( s \in S \). Suppose that we have an embedding \( \theta_T : A \hookrightarrow \text{End}_{T/S}(X_T) \) for an \( S \)-scheme \( T/S \). Then by the rigidity of endomorphisms, \( \theta_T(x) \) is a \( T \)-point of \( \mathbb{E}_x \). In other words, \( \mathbb{E}_{x/S} \) represents the following functor

\[
\mathcal{F}_A(T) = \{ \theta_T : A \hookrightarrow \text{End}_T(X_T) \mid \theta_T(1_A) = \text{id}_{X_T} \}
\]

from \( \text{SCH}_{/S} \) into \( \text{SETS} \). On \( \mathcal{F}_A \), the finite group \( \text{Aut}(A) \) of algebra automorphisms acts by \( \theta_T \mapsto \theta_T \circ \sigma \).

We can generalize the above argument to any algebra \( A \) free of finite rank over \( \mathbb{Z} \) with semi-simple \( A \otimes_\mathbb{Z} \mathbb{Q} \). We take a finite set of generators \( \{x_1, \ldots, x_j\} \) and consider \( E = \mathbb{E}_{x_1} \times_S \mathbb{E}_{x_2} \times_S \cdots \times_S \mathbb{E}_{x_j} \). Then we define \( \mathbb{E}_A \) to be the maximal subscheme of \( E \) such that we have an algebra embedding \( \theta : A \hookrightarrow \text{End}_{E_A}(X_{E_A}) \) taking \( x_i \) to \( x_{i,E_A} \) for all \( i \). Then we have

**Corollary 5.7.** — Let \( S \) be a noetherian scheme, and \( X/S \) be an abelian scheme over \( S \). Let \( A \) be an algebra free of finite rank over \( \mathbb{Z} \) with semi-simple \( A \otimes_\mathbb{Z} \mathbb{Q} \). Then the functor

\[
T/S \mapsto [(X_T, \theta : A \hookrightarrow \text{End}_T(X_T)) \mid \theta(1_A) = \text{id}_X]
\]

is representable by a scheme \( \mathbb{E}_A \) over \( S \), and each connected component of \( \mathbb{E}_A \) is finite over \( S \).
A semi-abelian scheme $X/S$ is a smooth separated group scheme with geometrically connected fiber such that each geometric fiber is an extension of an abelian variety by a torus. The toric rank may depend on the fiber. We suppose that $X/S$ is a semi-abelian scheme. It is known that any homomorphism of semi-abelian schemes: $X_U \to Y_U$ defined over an open dense subscheme $U \subset S$ extends uniquely to $X/S \to Y/S$ as long as $S$ is normal (endomorphisms are kept under specialization; a result of M. Raynaud: [DAV] I.2.7). Thus if $X/S$ is an abelian scheme $S$, we have a unique extension of the scheme $\text{End}_{X_U/U}$ over a dense open subscheme $U$ of $S$ to the scheme $\text{End}_{X/S}$ over $S$ which represents the functor in the above corollary. Applying the valuative criterion using this rigidity of endomorphisms, we find that $\text{End}_{X/S}$ has connected components each projective over $S$. Suppose that we have an embedding $\theta_s: A \hookrightarrow \text{End}_{X(s)}(X(s))$ for a geometric point $s \in S$ with abelian variety fiber $X(s)$. Then by the rigidity of endomorphisms, the maximal connected subscheme $Z \subset S$ containing $s$ such that $\theta_s$ extends to the embedding $\theta: A \hookrightarrow \text{End}(X/Z)$ is a closed connected subscheme $Z \subset S$. Thus each connected component of $\text{End}_{X/S}$ is projective over $S$. In the same manner as in the case of an abelian scheme $X/S$, we can prove that each connected component of $\text{End}_{X/S}$ is quasi-finite; so, it is finite over $S$. Thus we get (see [PAF] Corollary 6.10).

**Corollary 5.8.** — Let $X/S$ be a semi-abelian scheme with abelian variety fiber over a dense open subset of $S$. If an abelian variety fiber $X(s)$ has an inclusion $A \hookrightarrow \text{End}_s(X(s))$, then the functor

$$T/S \mapsto [(X_T, \theta: A \hookrightarrow \text{End}_T(X_T))]$$

is represented by a scheme over $S$. Each connected component of this scheme is finite over $S$.

5.2. Mumford Moduli. — We describe the Mumford construction of the moduli over $\mathbb{Z}$ of abelian schemes of dimension $n$ with a given polarization of degree $d^2$.

5.2.1. Dual Abelian Scheme and Polarization. — We consider the following Picard functor:

$$\text{Pic}_{X/S}(T) = \text{Pic}(X_T)/f_T^*\text{Pic}(T) = \text{Ker}(0_T^*)$$

for $f: T \to S$. It is known that $\text{Pic}_{X/S}$ is represented by a (locally noetherian) reduced group scheme (Grothendieck: Bourbaki Sem. Exp. 232, 1961/62; [ABV] Section 13, [GIT] 0.5 and [DAV] I.1). Let $\tilde{X}$ be the identity connected component $\tilde{X}_{X/S}$ of the group scheme representing $\text{Pic}_{X/S}$. Then $\tilde{X}/S$ is an abelian scheme.

Let $\mathcal{L}^{\text{univ}} \in \tilde{X}(\tilde{X}) = \text{Pic}_{\tilde{X} \times_S \tilde{X}}^0 = \text{Ker}(0_{\tilde{X} \times_S \tilde{X}}^*)$ corresponding to the identity. Then the sheaf $\mathcal{L}^{\text{univ}}$ is characterized by the following two properties:

- $0_{\tilde{X} \times_S \tilde{X}}^*\mathcal{L}^{\text{univ}} = \mathcal{O}_{\tilde{X}}$;
Let $T \to S$ be an $S$-scheme. For any invertible sheaf $L$ on $X_T$ algebraically equivalent to $\mathcal{O}_{X_T}$, there exists a unique morphism $\iota_L : T \to \hat{X}$ such that $(\iota \times \text{id}_X)^* \mathcal{L}_{\text{univ}} = L$.

Let $L$ be an invertible sheaf on $X$. For $x \in X$, we define the translation $T_x(y) = x + y$, which is an automorphism of $X$. Then $(T_x^* L) \otimes L^{-1}$ is an element in $\hat{X}$, and we obtain a morphism $\Lambda(L) : X \to \hat{X}$. This $S$-homomorphism is an isogeny (that is, surjective) if and only if $L$ is ample (that is, choosing a base of $H^0(X, L^\otimes n)$ for sufficiently large $n$, one can embed $X$ into $\mathbb{P}^m_S$ locally on $S$; see [ABV] Section 6).

The degree of the polarization is defined to be the square root of the degree of the homomorphism $\Lambda(L)$.

There is another construction of $\Lambda(L)$. Consider $a^*(L) \otimes p_1^*(L) \otimes p_2^*(L)^{-1}$ as an invertible sheaf on $X_X = X \times_S X$, where $a : X \times X \to X$ is the addition on the group scheme $X$. Then this invertible sheaf induces an $X$-valued point of $\text{Pic}_{X/S}(X)$, which factors through $\hat{X}$, because at the identity, this sheaf specializes to the trivial invertible sheaf (so, the image is in the connected component of $\text{Pic}_{X/S}$). We claim this $X$-valued point of $\hat{X}$ is actually $\Lambda(L)$. By specializing this sheaf at $x : S \to X \in X(S)$, we get $T^*_x(L) \otimes L^{-1} \otimes x^*(L)$, which is equivalent in $\text{Pic}(X)/f^* \text{Pic}(S)$ to $T^*_x L \otimes L^{-1}$, as desired.

5.2.2. Moduli Problem. — We fix three positive integers $n$, $d$ and $N$. We consider the following functor over $\text{Spec}(\mathbb{Z}[1/N])$:

$$
\mathcal{A}_{d,N}(S) = [(X, \phi_N : (\mathbb{Z}/N\mathbb{Z})^{2n} \cong X[N], \lambda)/S]
$$

where

1. $X/S$ is an abelian scheme with $\dim_S X = n$,
2. $\phi_N$ is an isomorphism over $S$,
3. $\lambda$ is a polarization, étale locally $\lambda = \Lambda(L)$ on $S$ and $\deg(\lambda) = d^2$.

It is known that if $\lambda$ is locally of the form $\Lambda(L)$, then $2\lambda$ is globally $\Lambda(L^\lambda(\lambda))$ for the invertible sheaf $L^\lambda(\lambda)$ given by $(1_X \times \lambda)^* (\mathcal{L}_{\text{univ}})$ (see [GIT] Proposition 6.10).

Here is a theorem of Mumford:

**Theorem 5.9 (Mumford).** — There exists a quasi-projective scheme $M_N$ over $\mathbb{Z}[1/N]$ such that

1. For any geometric point $s = \text{Spec}(k)$ of $\text{Spec}(\mathbb{Z}[1/N])$, $\mathcal{A}_{d,N}(k) \cong M_N(k)$;
2. If $N \geq 3$, there exists a universal object $(X, \phi_N, \lambda)_{/M_N}$ such that for each triple $(X, \phi_N, \lambda) \in \mathcal{A}_{d,N}(S)$ there exists a unique morphism $\iota : S \to M_N$ such that

$$(X, \phi_N, \lambda) \cong (X, \phi, \lambda) \times_{M_N} S = \iota^*(X, \phi, \lambda),$$

3. The above association: $(X, \phi_N, \lambda)_{/S} \mapsto \iota$ induces a functorial isomorphism: $\mathcal{A}_{d,N} \cong \underline{M}_N$, where $\underline{M}_N(S) = \text{Hom}_{\mathbb{Z}[1/N]}(S, M_N)$.  

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
We are going to give a sketch of the proof of the above theorem. Let $X_{/A}$ be an abelian scheme over a ring $A$. The key idea is that for a given very ample invertible sheaf $L_{/X}$, the embedded image of $X$ under $L$ in $\mathbb{P}^m$ for $m + 1 = \text{rank}_A H^0(X, L)$ is determined just by the choice of basis $b$ of $H^0(X, L)$. In other words, the images of the embeddings associated to different basis are transported each other by an element of $PGL_{m+1}(A)$. Since for an abelian scheme, by the generalized Riemann-Roch theorem (see [ABV] Section 16), we can compute the Hilbert polynomial $P$ of $L$, the moduli functor of $(X, b)$ is a subfunctor of $\text{Hilb}_{\mathbb{P}^m}$. Proving that the image is a quasi-projective subscheme of $\text{Hilb}_{\mathbb{P}^m}$, the moduli of $X$ is constructed as $M_1 = PGL(m + 1) \backslash H$. This an outline of what we are going to do.

5.2.3. Abelian Scheme with Linear Rigidification. — Let $(X, \phi, \lambda)_{/S} \in \mathcal{A}_{d,N}(S)$, and write $f : X \to S$ for the structure morphism. We consider the invertible sheaf $L = f_* L^\Delta(\lambda)^3$ on $S$ of rank $6^d d$. The sheaf $L^\Delta(\lambda)^3$ is very ample, because $L^3$ is very ample if $L$ is ample (see [ABV] Section 17) and ampleness of $L^\Delta(\lambda)$ follows from $\Lambda(L^\Delta(\lambda)) = 2\lambda$ as we already remarked.

Let $\text{Sym}(L)$ be the symmetric algebra: $\bigoplus_{k=0}^{\infty} L^k$ and put $P(L) = \text{Proj}(\text{Sym}(L))$ which is a projective scheme over $S$ locally isomorphic to $\mathbb{P}^m_{/S}$. A linear rigidification is an isomorphism $\iota : P(L) \cong \mathbb{P}^m_{/S}$. Thus $\iota$ is determined by the choice of a base of $L$ up to scalar multiplication.

Since the very ample sheaf $L^\Delta(\lambda)^3$ on an abelian scheme satisfies ([ABV] Section 16):

- $L = f_*(L^\Delta(\lambda)^3)$ is locally free of finite rank;
- $R^i f_*(L^\Delta(\lambda)^3) = 0$ if $i > 0$,

the formation of $f_*(L^\Delta(\lambda)^3)$ as above commutes with base change. Thus the association

$$\mathcal{A}'_{d,N}(S) = [(X, \phi, \lambda, \iota)_{/S}]$$

is a well defined contravariant functor.

The embedding $I : X_{/S} \hookrightarrow P(L) \to P^m_{/S}$ determines the sheaf $L^\Delta(\lambda)^3 = \mathcal{O}(1)$, which in turn determines $\lambda$ because $\Lambda(L) = 6\lambda$ ($\text{Pic}_{X_{/S}}/\hat{X}$ is torsion free). Having $\phi_N$ is equivalent to having $2n$ (linearly independent) sections $\sigma_j = \phi_N(e_j)$ of $X$ over $S$ for the standard base $\{e_1, \ldots, e_{2n}\}$ of $(\mathbb{Z}/N\mathbb{Z})^{2n}$. We write $\sigma_0 = e$ for the identity section of $X$.

We record here what we have seen:

**Proposition 5.10.** — The data $(X, \phi, \lambda, \iota)$ is determined by the embedding

$$(I : X_{/S} \hookrightarrow P^m_{/S}, \sigma_0, \sigma_1, \ldots, \sigma_{2n}).$$

In other words, defining a new functor by

$$\mathcal{H}_{d,N}(S) = [(I : X_{/S} \hookrightarrow P^m_{/S}, \sigma_0, \sigma_1, \ldots, \sigma_{2n})_{/S}],$$

we have an isomorphism of functors: $\mathcal{A}'_{d,N} \cong \mathcal{H}_{d,N}$. 

ASTÉRISQUE 298
5.2.4. Embedding into the Hilbert Scheme. — For simplicity, we just write \( P \) for \( \mathbb{P}^m_{\mathbb{Z}[1/\ell]} \). We write \( S_0 \) for \( \text{Spec} (\mathbb{Z}[1/\ell]) \). We consider the functor \( \text{Hilb}_{\mathbb{P}}^P \) associating to each \( S \) the set of closed subschemes of \( \mathbb{P}_S \) flat over \( S \) with Hilbert polynomial \( P \). As we have already seen, this functor is represented by a projective scheme \( H = \text{Hilb}_{\mathbb{P}}^P \) over \( \mathbb{Z} \). Write \( Z \to H \) for the universal flat family inside \( \mathbb{P}_{\mathbb{P}}^H \) with Hilbert polynomial \( P \). For each subscheme \( V \subset \mathbb{P}_S \) flat over \( S \) having Hilbert polynomial \( P \), we have a unique morphism \( h : S \to H \) such that \( V \) is given by to \( S \times_H Z \subset \mathbb{P}_S \) over \( S \).

By the generalized Riemann-Roch theorem ([ABV] Section 16), the Hilbert polynomial of \( (X,L) \) (or of the image \( I(X) \)) is given by

\[
P(T) = 6^n d T^n.
\]

Thus the image \( I(X) \) induces a unique morphism \( h : S \to H \) such that \( I(X) = S \times_H Z \) in \( \mathbb{P}_S \).

Let \( H^k = \text{Hilb}_{\mathbb{P}}^{P^k} = \mathbb{Z} \times_H Z \times_H \cdots \times_H Z \). Then by the very definition of the fiber product, we get

\[
\text{Hom}_{S_0} (S,H^k) = \{(h,s_1,\ldots,s_k) \mid h \in \text{Hom}_{S_0} (S,H), \ s_j \in \text{Hom}_S (S,Z)\},
\]

where \( h : S \to Z \to H \) for the projection \( p \) of \( Z \) to \( H \) (so, \( h \) is determined by any of \( s_j \)). Thus \( H^k \) classifies all flat closed subschemes of \( \mathbb{P} \) with Hilbert polynomial \( P \) having \( k \) sections over \( S \). The universal scheme over \( H^k \) with \( k \) sections is given by \( Z(k) = Z \times_H H^k \). It has \( k \) sections:

\[
\tau_i : H^k \ni z \mapsto (z_i, z) \in Z(k) \quad (z = (z_1, \ldots, z_k)).
\]

This shows that \( \mathcal{H}_{d,N} \subset H^k \) for \( k = 2n + 1 \). For simplicity, write \( H_0 \) for \( H^k \). Since “smoothness” is an open condition (because it is local; in other words, smoothness at a point \( x \) of a morphism \( f \) follows from formal smoothness of the local ring at \( x \) over the local ring at \( f(x) \)), there is an open subscheme \( H_1 \) of \( H_0 \) over which \( Z \) is smooth. Then \( H_1 \) represents smooth closed subschemes in \( \mathbb{P} \) with Hilbert polynomial \( P \) and \( k \) sections.

Now we use a result of Grothendieck. Abelian varieties have rigidity such that if in a smooth projective family \( X \to S \) for connected locally noetherian \( S \) with a section \( e : S \to X \), if one fiber is an abelian variety with the identity section induced by \( e \), \( X \) itself is an abelian scheme ([GIT] Theorem 6.14). This shows that over \( H_1 \), \( H_1 \) has a closed subscheme \( H_2 \) over which \( Z_2 = Z_{H_2} \) is an abelian scheme with the identity section \( e \) inducing \( \tau_0 \) on \( X \).

Let \( \tau_j \) \( (j = 1, \ldots, 2n) \) be the universal \( 2n \) sections of \( Z(k) \). We have a maximal closed subscheme \( H_3 \subset H_2 \) with \( [N] \circ \tau_i = e \), where \( [N] \) is the multiplication by the integer \( N \).
The relation $\sum_{j=1}^{2n} a_j \tau_j = e$ for a given $a = (a_j) \in (\mathbb{Z}/N\mathbb{Z})^{2n} - \{0\}$ gives a closed subscheme $H_a$ of $H_3$; so, we define $H_4 = H_3 - \bigcup_a H_a$. Thus the abelian scheme $Z_4$ over $H_4$ has $2n$ linearly independent sections of order $N$.

Since $Z_4$ is a subscheme of $\mathbf{P}_{/H_4}$, it has the line bundle $\mathcal{O}_{Z_4}(1)$ which is the restriction of $\mathcal{O}(1)_{/P}$.

Then we define $H_5$ to be the maximal subscheme of $H_4$ such that $p^*L \cong L^3(\lambda)^3$ for a polarization $\lambda : Z_4 \to ^t Z_4$, where $p : Z_5 = Z_4 \times_{H_4} H_5 \to Z_4$ is the inclusion. It is proved in [GIT] proposition 6.11 that the maximal subscheme $H_5$ with the above property exists and is closed in $H_4$.

5.2.5. Conclusion. — By the argument in the previous proposition, the functor $\mathcal{H}_{d,N}$ is represented over $S_0$ by a quasi-projective scheme $H_5$ with the universal abelian scheme $X = Z_5$ over $H_5$. The group $\text{PGL}(m+1)$ acts on $H_5$ by $\iota \mapsto \iota \circ g$ ($g \in \text{PGL}(m+1)$). Then Mumford verifies through his theory of geometric quotient that the quotient quasi-projective scheme exists (see GIT Chapter 3 and Section 7.3):

$$M_{d,N} = \text{PGL}(m+1) \setminus H_5.$$ 

It is easy to check that if $\text{PGL}(m+1)$ has no fixed point, then $H_5$ is a $\text{PGL}(m+1)$-torsor over $M_{d,N}$. This is the case where the structure $(X, \phi_N, \lambda)$ does not have non-trivial automorphisms, which follows if $N \geq 3$ by a result of Serre (see [PAF] pages 281-282 for this point). In this case, $M_{d,N}$ represents the functor $\mathcal{A}_{d,N}$ over $S_0$. Otherwise, $M_{d,N}$ gives a coarse moduli scheme for the functor.

5.2.6. Compactification. — Here we quote a result from Faltings-Chai [DAV] V.2 on the minimal compactification. Let $X = (X_{d,N}, \phi_N, \lambda) \to M_{d,N}$ be the universal abelian scheme of relative dimension $n$ with level $N$-structure $\phi_N$ and the polarization of degree $d^2$. We assume that $N \geq 3$.

Since we have already studied via Tate curves the compactification of $M = M_{d,N}$ when $n = 1$ (the moduli of elliptic curves), we assume here $n > 1$. We then define $\varpi = f_* \Omega_{X/M}$. This is a locally free sheaf over $M$ of rank $n$. We define $\det(\varpi) = \Lambda^n \varpi$.

In [DAV] IV, a smooth toroidal compactification $\overline{M} = \overline{M}_{d,N}$ over $\mathbb{Z}[1/d]$ is made (actually, details are exposed there for $d = 1$ but the argument works for $d > 1$ over $\mathbb{Z}[1/d]$). We shall come back to this topic later with more details. They also proved that $\varpi$ extends to the compactification $\overline{M}$. Then we define a graded algebra

$$\mathcal{G} = \mathcal{G}_{d,N} = \bigoplus_{m=0}^{\infty} H^0(\overline{M}_{d,N}, \det(\varpi)^{\otimes m}) = \bigoplus_{m=0}^{\infty} H^0(M_{d,N}, \det(\varpi)^{\otimes m}),$$

where $\det(\varpi)^{\otimes m}$ is the $m$-th power of the invertible sheaf $\det(\varpi)$. The last identity in the above definition follows from Koecher’s lemma ([DAV] V.1.5) if $n > 1$. It is proven in [DAV] V.2.5 that this graded algebra is finitely generated over $\mathbb{Z}[_N\zeta_N]$ for a fixed primitive $N$-th root $\zeta_N$ of unity, and by the first equality, the graded algebra is normal. Thus we may define $\mathcal{G}_{d,1}$ to be the normalization of $\mathcal{G}_{d,1}$ in the algebra.
\(\mathcal{R}_{d,N}\) defined below. We define the minimal compactification by

\[ M^*_{d,N} = \text{Proj}(\mathcal{G}_{d,N}). \]

It is called “minimal” because any smooth toroidal compactification \(\overline{\mathcal{M}}_{d,N}\) covers canonically \(M^*_{d,N}\).

We can define a sheaf of graded algebras over \(M_{d,N}\) by

\[ \mathcal{R} = \mathcal{R}_{d,N} = \bigoplus_{m=-\infty}^{\infty} f_*(\det(\omega)^{\otimes m}). \]

Then \(\mathcal{M}_{d,N} = \text{Spec}_M(\mathcal{R})\) represents the following functor:

\[ \mathcal{P}_{d,N}(S) = \left\{ \left( X, \phi_N, \lambda, \omega \right)_{/S} \mid (X, \phi_N, \lambda) \in \mathcal{A}_{d,N}(S), H^0(X, \det \Omega_{X/S}) = \mathcal{O}_S\omega \right\}, \]

and we have

\[ M_{d,N} = \mathbb{G}_m \backslash \mathcal{M}_{d,N} \]

as a geometric quotient. Here \(\mathcal{M}_{d,N}\) is the relative spectrum of \(\mathcal{R}\) over \(M\); so, \(\mathcal{M}_{d,N}\) is affine over \(M_{d,N}\). In particular, \(\mathcal{M}_{d,N}\) is a \(\mathbb{G}_m\)-torsor over \(M_{d,N}\) if \(N \geq 3\). Here \(a \in \mathbb{G}_m\) acts on the functor \(\mathcal{P}_{d,N}\) by \((X, \phi_N, \lambda, \omega) \mapsto (X, \phi_N, \lambda, a\omega)\). The relation between the moduli scheme classifying abelian schemes with level structure and the one classifying with an extra information of nowhere vanishing \(n\)-differentials is exactly the same as in the elliptic modular case, which amuses me a bit, and it is also interesting that this is proven only after a hard work of compactifying smoothly the open moduli \(M_{d,N}\).

6. Shimura Varieties

In this lecture, we sketch basic theory of Shimura varieties of PEL type following [Sh3], [D2] and [Ko].

Shimura originally constructed canonical models in the 1950’s to 1960’s as a tower of quasi-projective geometrically connected varieties (over a tower of canonical abelian extensions of the reflex field) with a specific reciprocity law at special algebraic points (in the case of Shimura varieties of PEL-type, they are called CM points carrying an abelian variety of CM type; [Sh3]). His theory includes interesting cases of canonical models of non PEL type (for example, Shimura curves over totally real fields different from \(\mathbb{Q}\)), but in this paper, we restrict ourselves to the case where we have a canonical family of abelian varieties over the canonical model (so, the construction of the models is easier, as was basically done in [Sh2]).

Deligne reformulated Shimura’s tower as a projective limit of (possibly non-connected) models over the reflex field (incorporating theory of motives in its scope). We follow Deligne’s treatment in order to avoid the definition of the canonical fields of definition of the connected components, although by doing this, we may lose some of finer information.
Kottwitz extended the Deligne’s definition of Shimura varieties of PEL type to a projective limit of schemes over a valuation ring of mixed characteristic, when the level is prime to \( p \). Since we are interested in formal completion at \( p \) of the Kottwitz model (and an analogue of the Igusa tower over the Kottwitz model), what we use most is Kottwitz’s formulation.

## 6.1. Shimura Varieties of PEL Type

We construct the moduli of abelian schemes with specific endomorphism algebra.

### 6.1.1. Endomorphisms

Let \( B \) be a finite dimensional simple \( \mathbb{Q} \)-algebra with center \( F \). Let \( S \) be a set of primes of \( F \) over \( p \). We always assume

\[ \text{(unr)} \quad \text{We have an isomorphism } B_p = B \otimes \mathbb{Q}_p \cong \bigoplus_{p \in S} M_n(F_p) \text{ and } F_p/\mathbb{Q}_p \text{ is unramified for all } p \in S. \]

Let \( \ast \) be an involution on \( B \) which satisfies \( \text{Tr}(xx^\ast) > 0 \) for all \( 0 \neq x \in B_{\infty} = B \otimes \mathbb{R} \). We call such an involution a *positive* involution. We fix a maximal order \( O = O_B \) of \( B \) stable under \( \ast \). We assume that the isomorphism in (unr) induces \( O_p = O \otimes \mathbb{Z}_p \cong \bigoplus_{p \in S} M_n(O_{F,p}). \)

We fix a left \( B \)-module \( V \) of finite type and assume that we have a non-degenerate alternating form \( \langle \ , \rangle : V \times V \to \mathbb{Q} \) such that \( \langle bv, w \rangle = \langle v, b^\ast w \rangle \) for all \( b \in B \). Write \( V_p = V \otimes \mathbb{Q}_p \) and \( V_{\infty} = V \otimes \mathbb{R} \). We also assume to have an \( O \)-submodule \( L \subset V \) of finite type such that

1. \( L \otimes \mathbb{Q} = V; \)
2. \( \langle \ , \rangle \) induces \( \text{Hom}_{\mathbb{Z}_p}(L_p, \mathbb{Z}_p) \cong L_p, \) where \( L_p = L \otimes \mathbb{Z}_p. \)

Put \( C = \text{End}_B(V) \), which is a semi-simple \( \mathbb{Q} \)-algebra with involution again denoted by \( \ast \) given by \( (cv, w) = \langle v, c^\ast w \rangle \). Then we define algebraic \( \mathbb{Q} \)-groups \( G \) and \( G_1 \) by

\[ (6.1) \quad G(A) = \left\{ x \in C \otimes A \mid xx^\ast \in A^\times \right\}; \quad G_1(A) = \left\{ x \in G(A) \mid xx^\ast = 1 \right\}. \]

We now take an \( \mathbb{R} \)-algebra homomorphism \( h : \mathbb{C} \to C_{\infty} = C \otimes \mathbb{R} \) with \( h(\overline{z}) = h(z)^\ast \). We call such an algebra homomorphism \( \ast \)-homomorphism. Then \( h(i)^\ast = -h(i) \) for \( i = \sqrt{-1} \) and hence \( x^\ast = h(i)^{-1} x^\ast h(i) \) is an involution of \( C_{\infty} \). We suppose

\( \text{(pos)} \) The symmetric real bilinear form \( (v, w) \mapsto \langle v, h(i)w \rangle \) on \( V_{\infty} \) is positive definite.

The above condition implies that \( \iota \) is a positive involution (e.g. [Ko] Lemma 2.2).

Since \( h : \mathbb{C} \to C_{\infty} \) is an \( \mathbb{R} \)-algebra homomorphism, we can split \( V_C = V \otimes \mathbb{Q} \subset C \) into the direct sum of eigenspaces \( V_C = V_1 \oplus V_2 \) so that \( h(z) \) acts on \( V_1 \) (resp. \( V_2 \)) through multiplication by \( z \) (resp. \( \overline{z} \)). Since \( h(\mathbb{C}) \subset C_{\infty} \), \( h(z) \) commutes with the action of \( B \); so, \( V_1 \) is stable under the action of \( B_C = B \otimes \mathbb{C} \). Thus we get the complex representation \( \rho_1 : B \to \text{End}_C(V_1) \). We define \( E \) for the subfield of \( \mathbb{C} \) fixed by

\[ \left\{ \sigma \in \text{Aut}(\mathbb{C}) \mid \rho_1^\sigma \cong \rho_1 \right\}. \]

The field \( E \) is called the reflex field (of \( B \)). We write \( O_E \) for the integer ring of \( E \). Let \( \mathbb{Z}_{(p)} = \mathbb{Z}_p \cap \mathbb{Q} \) and put \( O_{(p)} = O \otimes \mathbb{Z}_{(p)}. \)
Let $K^{(p)}$ be an open compact subgroup of $G(A^{(p\infty)})$ for

$$K^{(p\infty)} = \{ x \in A \mid x_p = x_{\infty} = 0 \}.$$ 

Let $K_p = \{ g \in G(Z_p) \mid gL_p = L_p \}$, and put $K = K_p \times K^{(p)} \subset G(A^{(\infty)})$. We call an open compact subgroup $K$ of $G(A^{(\infty)})$ of this type an open compact subgroup maximal at $p$.

We study classification problem of the following quadruples: $(X, \lambda, i, \pi^{(p)})$. Here $X$ is a (projective) abelian scheme over a base $S$, $\hat{X} = \text{Pic}^0_{X/S}(X)$ is the dual abelian scheme of $X$, $\lambda : X \to \hat{X}$ is an isogeny with degree prime to $p$ (prime-to-$p$ isogeny geometrically fiber by fiber induced from an ample divisor (polarization), $i : O(p) \hookrightarrow \text{End}^{(p)}_S(X) = \text{End}_S(X) \otimes \mathbb{Z}(p)$, and $\pi^{(p)}$ is the level $K^{(p)}$-structure (see below for the definition of the level structure). The base scheme $S$ is assumed to be a scheme over $\text{Spec}(\mathbb{Z}(p))$. We now explain the meaning of the level $K^{(p)}$-structure $\pi^{(p)}$. We consider the Tate module:

$$T(X) = \lim_{n \to \infty} X[N], \quad T^{(p)}(X) = T(X) \otimes_{\mathbb{Z}} \hat{Z}^{(p)} \quad \text{and} \quad V^{(p)}(X) = T(X) \otimes_{\mathbb{Z}} A^{(p\infty)},$$

where $N$ runs over all positive integers ordered by divisibility, and $\hat{Z}^{(p)} = \prod_{\ell \neq p} \mathbb{Z}_\ell$. This module is equipped with a non-degenerate skew hermitian form induced by the polarization $\lambda$. Fix a base (geometric) point $s \in S$ and write $X_s$ for the fiber of $X$ at $s$. Then the algebraic fundamental group $\pi_1(S, s)$ acts on $V^{(p)}(X_s)$ leaving stable the skew hermitian form up to scalar. Then $\eta^{(p)} : V(A^{(p\infty)}) = V \otimes_{\mathbb{Q}} A^{(p\infty)} \sim V^{(p)}(X_s)$ is an isomorphism of skew hermitian $B$-modules. We write $\pi^{(p)} = \eta^{(p)} \mod K^{(p)}$ and suppose that $\sigma \circ \pi^{(p)} = \pi^{(p)}$ for all $\sigma \in \pi_1(S, s)$ (this is a way of describing that the level structure $\eta^{(p)}$ is defined over $S$). Even if we change the point $s \in S$, everything will be conjugated by an isomorphism; so, the definition does not depend on the choice of $s$ as long as $S$ is connected (see [PAF] 6.4.1). When $S$ is not connected, we choose one geometric point at each connected component.

As examples of $K^{(p)}$ and open compact subgroups $K$ maximal at $p$ of $G(A^{(\infty)})$, we could offer the following subgroups:

$$\hat{\Gamma} = \{ x \in G(A^{(\infty)}) \mid xL = L \}, \quad \hat{\Gamma}^{(p)} = \{ x \in \hat{\Gamma} \mid x_p = 1 \};$$

$$\hat{\Gamma}^{(p)}(N) = \{ x \in \hat{\Gamma}^{(p)} \mid x\ell \equiv \ell \mod NL \text{ for all } \ell \in L \}$$

for an integer $N > 0$ prime to $p$.

Since every maximal compact subgroup of $GL_B(V(A^{(\infty)})) = \text{Aut}_{B_A}(V(A^{(\infty)}))$ is the stabilizer of a lattice $L$ stable under a maximal order, we find a lattice $L$ with $L \otimes_{\mathbb{Z}} \hat{Z}^{(p)}$ stable under $K^{(p)}$, where $\hat{Z}^{(p)} = \prod_{\ell \neq p} \mathbb{Z}_\ell$ for $\ell$ running through all primes different from $p$. Changing $L$ by a sublattice of $p$-power index if necessary, we may assume that $L$ satisfies the conditions (L1-2). We call a quadruple $\hat{X}_{/S} = (X, \lambda, i, \pi^{(p)})_{/S}$ isomorphic to $\hat{X'}_{/S} = (X', \lambda', i', \pi'^{(p)})_{/S}$ if we have an isogeny.
\[ \phi : X \to X' \] defined over \( S \) such that \( p \nmid \deg(\phi) \), \( \hat{\phi} \circ \lambda' \circ \phi = c\lambda \) with \( c \in \mathbb{Z}_{(p)}^\times \), \( \phi \circ i \circ \phi^{-1} = i' \) and \( \eta'(p) = \phi \circ \overline{\eta}(p) \). In this case, we write \( X \approx X' \). We write \( X \approx X' \) if the isogeny is an isomorphism of abelian schemes, that is, \( \deg(\phi) = 1 \).

Let \( S_0 = \text{Spec}(O_E \otimes \mathbb{Z}(p)) \). We take the fiber category \( \mathcal{C} = \mathcal{C}_B \) of the quadruples \((X, \lambda, i, \overline{\eta}(p))_S \) over the category \( S_0 - \text{SCH} \) of \( S_0 \)-schemes and define

\[
(6.2) \quad \text{Hom}_{\mathcal{C}/S}((X, \lambda, i, \overline{\eta}(p))/S, (X', \lambda', i', \overline{\eta}'(p))/S) = \left\{ \phi \in \text{Hom}_S(X, X') \otimes \mathbb{Z}(p) \left| \begin{array}{l}
\hat{\phi} \circ \lambda' \circ \phi = c\lambda \\
\phi \circ i = i' \circ \phi \\
\eta'(p) = \phi \circ \overline{\eta}(p)
\end{array} \right. \right\}.
\]

We consider the functor \( \mathcal{P}_K^{(p)} : S_0 - \text{SCH} \to \text{SETS} \) given by

\[
\mathcal{P}_K^{(p)}(S) = \mathcal{P}_K^{(p)}(S) = \left\{ X_S = (X, \lambda, i, \overline{\eta}(p))/S \left| X \text{ satisfies (det) below} \right. \right\}/\sim.
\]

This functor is representable by the Shimura variety \( \text{Sh}_K^{(p)} \) defined over \( S_0 \) as we will see later. Here the determinant condition is given as follows: We fix a \( \mathbb{Z}(p) \)-base \( \{\alpha_j\}_{1 \leq j \leq \ell} \) of \( O(p) \) and consider a homogeneous polynomial

\[
f(X_1, \ldots, X_\ell) = \det(\alpha_1 X_1 + \cdots + \alpha_\ell X_\ell|_{V_1}).
\]

Then \( f(X) \in O_E \otimes \mathbb{Z}(p)[X_1, \ldots, X_\ell] \) and coefficients of \( f(X) \) generates \( E \) over \( \mathbb{Q} \). Here \( O_E \) is the integer ring of \( E \). For a given quadruple \( X = (X, \lambda, i, \overline{\eta}(p))/S \), we have the Lie algebra \( \text{Lie}(X) \) of \( X \) over \( O_S \), which is a \( O(p) \otimes \mathbb{O}_S \)-module via \( i \). Then we can think of \( g(X_1, \ldots, X_\ell) = \det(\alpha_1 X_1 + \cdots + \alpha_\ell X_\ell|_{\text{Lie}(X)}) \in O_S[X_1, \ldots, X_\ell] \). We impose

\[
(\text{det}) \quad g(f(X_1, \ldots, X_\ell)) = g(X_1, \ldots, X_\ell),
\]

where \( j : O_E \otimes \mathbb{Z}(p) \to O_S \) is the structure homomorphism. Over a field of characteristic 0, one can characterize representations of a semi-simple algebra by its trace, but over a general base, we need, by the Brauer-Nesbitt theorem, the entire characteristic polynomial to determine a given representation; so, the determinant has to be fixed as above.

Allowing any closed subgroup \( K \subset G(\mathbb{A}(\infty)) \) (not necessarily maximal at \( p \)), replacing isogenies of degree prime to \( p \) by (any) isogenies and imposing one more condition (pol) below, we may consider the functor \( \mathcal{P}_K : E - \text{SCH} \to \text{SETS} \) given by

\[
\mathcal{P}(S) = \mathcal{P}_K(S) = \left\{ X_S = (X, \lambda, i, \overline{\eta})/S \left| X \text{ satisfies (det) and (pol)} \right. \right\}/\sim,
\]

where \( \eta : V(\mathbb{A}(p)) = V \otimes \mathbb{A}(\infty) \cong V(X) = T(X) \otimes \mathbb{A}(\infty), \overline{\eta} = (\eta \bmod K) \), \( K \) is any closed subgroup of \( G(\mathbb{A}(\infty)) \), and \( (X, \lambda, i, \overline{\eta})/S \sim (X', \lambda', i', \overline{\eta}')/S \) if the two quadruples are equivalent to each other under an isogeny (not necessarily of degree prime to \( p \)). Here is the condition (pol):
There exists an $B$-linear isomorphism $f : V \cong H_1(X, \mathbb{Q})$ such that $f^{-1} \circ h_X \circ f$ is a conjugate of $h$ under $G(\mathbb{R})$, $(f \otimes 1_{\mathbb{A}_F}) \in \mathcal{P}$ and $E_X(f(x), f(y)) = \alpha(x, y)$ up to $\alpha \in \mathbb{Q}^\times$,

where $E_X$ is the Riemann form on $H_1(X, \mathbb{Q})$ (see [ABV] Section 1) and $h_X : \mathbb{C} \hookrightarrow C_\infty$ is the $\mathbb{R}$-algebra homomorphism induced by the complex structure on $V_\mathbb{R} \cong H_1(X, \mathbb{R})$ induced by the complex structure of $X(\mathbb{C})$. This functor is representable by the Shimura variety $Sh_K$ defined over $E$. The scheme $Sh_{K/E}$ is the model Deligne studied.

We have a canonical inclusion $i : Sh_{K/E} \hookrightarrow Sh_K^{(p)} \times_{S_\mathbb{R}} E$ if $K$ is maximal at $p$. The isomorphism class of $G$ over each local field is determined by the level structure $\eta$, but $G$ is not uniquely determined globally without the condition (pol). In other words, if $G$ does not satisfy the Hasse principle, even if $K$ is maximal at $p$, the inclusion $i$ may not be an isomorphism. As verified by Kottwitz (see [Ko] Section 8 and [PAF] Theorem 7.5), if $G$ is either an inner form of the symplectic group (type $C$ groups) or an inner form of a quasi-split unitary group $U(n, n)$ (type $A$) or $F$ is an imaginary quadratic field, we have $Sh_{K/E} \cong Sh_K^{(p)} \times_{S_\mathbb{R}} E$. Otherwise, the situation is more subtle (see [Ko] Section 8).

6.1.2. Construction of the Moduli. — Here is a brief outline of how to show the representability of the functor $P_K^{(p)}$ for $K$ maximal at $p$. If $K^{(p)}$ is sufficiently small so that $\text{Aut}_S(X) = \{1_X\}$, the prime-to-$p$ isogeny giving the isomorphism $X \approx X'$ in the definition of $P^{(p)}(\cdot)$ can be taken to be an isomorphism by changing $X'$ in the isomorphism class under “$\approx$” (and insisting $\eta^{(p)}(L^{(p)}) = T^{(p)}(X)$; see the argument below and [D1] 4.10). Therefore we have $P_K^{(p)}(S) \cong P'_K(S)$, where

$$P'_K(S) = \left\{ X/S \mid X \text{ with } (\det), X \approx \exists X' \in P^{(p)}(S) \text{ and } \eta^{(p)}(L^{(p)}) = T^{(p)}(X) \right\},$$

where $\left\{ \cdots \right\} / \cong$. Under this setting, we change the morphism set of $\mathcal{C}/S$ from $\text{Hom}_{\mathcal{C}/S}$ to $\text{Isom}_{\mathcal{C}/S}$:

$$\text{Isom}_{\mathcal{C}/S}((X, \lambda, i, \overline{\eta}^{(p)})/S, (X', \lambda', i', \overline{\eta}'^{(p)})/S)$$

$$= \left\{ \phi \in \text{Isom}_S(X, X') \mid \begin{array}{c} \hat{\phi} \circ \lambda' \circ \phi = c \lambda \text{ with } 0 < c \in \mathbb{Z}_\mathbb{R}^\times, \\ \phi \circ i = i' \circ \phi \text{ and } \overline{\eta}'^{(p)} = \phi \circ \overline{\eta}^{(p)}. \end{array} \right\}.$$
We now supplement the above outline with details. We are going to show that we can replace “$\cong$” by “$\equiv$” in the definition of the functor $P_K^{(p)}$, imposing an additional condition. Let $\mathbb{A}_{/S} = (A, \lambda, i, \eta^{(p)})_{/S}$. Then $\eta^{(p)}$ induces $V(\mathbb{A}^{(p\infty)}) = V \otimes \mathbb{A}^{(p\infty)} \cong V^{(p\infty)}(A)$, and therefore, we have $\eta^{(p)}(L^{(p)}) \subset V^{(p\infty)}(A)$ for $L^{(p)} = L \otimes_{\mathbb{Z}} \mathbb{Z}^{[p]}$.

If $\eta^{(p)}(L^{(p)}) \subset T^{(p)}(A)$, the cokernel is an étale group subscheme $C \subset \mathbb{A}_{/S}$ locally-free over $S$ of rank prime to $p$. Make the quotient abelian scheme $A' = A/C$ over $S$ (see [ABV] Section 12 and [GME] Proposition 1.8.4), and then we have a prime-to-$p$ isogeny $\phi : A \to A'$ with $\text{Ker}(\phi) = C$. We then consider $A' = (A', \lambda', i', \eta^{(p)'})_{/S}$ given by $\lambda = \hat{\phi} \circ \lambda' \circ \phi$, $i' = \phi \circ i(\alpha) \circ \phi^{-1}$ and $\eta^{(p)'} = \phi \circ \eta^{(p)}$. Then $A'$ satisfies $\eta^{(p)'}(L^{(p)}) = T^{(p)}(A)$.

If $\eta^{(p)}(L^{(p)}) \supset T^{(p)}(A)$, we can find a prime-to-$p$ isogeny $A' \to A$ such that $\phi \circ \eta^{(p)} = \eta^{(p)'}$, $\lambda' = \hat{\phi} \circ \lambda \circ \phi$, $i' = \phi^{-1} \circ i(\alpha) \circ \phi$ and $\eta^{(p)'}(L^{(p)}) = T^{(p)}(A')$. This fact follows from the canonical identification: $T(A) = \pi_1(A, 0)$ for the origin $0$ of $A$ (see [ABV] Section 18).

If neither $\eta^{(p)}(L^{(p)}) \subset T^{(p)}(A)$ nor $\eta^{(p)}(L^{(p)}) \supset T^{(p)}(A)$, we can find two prime-to-$p$ isogenies: $A \to A''$ and $A' \to A''$ for $A'' = (A'', \lambda'', i'', \eta^{(p)''})_{/S}$ so that $\eta^{(p)}(L^{(p)}) = \eta^{(p)}(L^{(p)} \cap T^{(p)}(A''))$ and $\eta^{(p)'}(L^{(p)}) = T^{(p)}(A')$. Thus always we can find in the prime-to-$p$ isogeny class of a given $\mathbb{A}_{/S}$, a quadruple $A'_S$ with $\eta^{(p)}(L^{(p)}) = T^{(p)}(A')$.

If $\phi : A_{/S} \to A_{/S}'$ is a prime-to-$p$ isogeny with $\eta^{(p)}(L^{(p)}) = T^{(p)}(A)$ and $\eta^{(p)'}(L^{(p)}) = T^{(p)}(A')$, then $\deg(\phi) = 1$ and $A_{/S} \cong A_{/S}'$ by $\phi$.

Thus insisting $\eta^{(p)}(L^{(p)}) = T^{(p)}(A)$, we can replace $\approx$ by $\equiv$ in order to define the functor $P_K^{(p)}$ (see [D1] 4.12). In other words,

$$P_K^{(p)}(S) \cong \{ \mathbb{A}_{/S} \mid (\det) \text{ and } \eta^{(p)}(L^{(p)}) = T^{(p)}(A) \}/ \cong .$$

The functor defined in this way can be proven to be representable by an $S_0$-scheme $M(G_0, X_0)/K$ by works of Deligne, Mumford and Shimura (cf. [Ko] and [PAF] 7.1.2).

Since $\mathbb{G}^{(p\infty)}(N)$ for $N$ prime to $p$ gives a fundamental system of neighborhoods of the identity in $G(\mathbb{A}^{(p\infty)})$, we may assume that $K = \hat{G}(N) = \hat{G}^{(p\infty)}(N) \times \hat{G}_p$. We only need to show that $C$ is relatively representable over the Mumford moduli $M_Q$ given by Theorem 5.9. Let $P^B_K$ be the functor with respect to $K$ and $B$. Write $2d = \dim_Q(V)$. Then $d = \dim_S X$ for $X_{/S} \in P^B_K(S)$, which is therefore independent of the choice of $X_{/S}$ by (det). For simplicity, we assume that the polarization pairing $\langle \cdot, \cdot \rangle$ in (L2) gives the self-duality of $L$. Then we can identify the similitude group of $\langle \cdot, \cdot \rangle$ acting on $L$ with $GSp_{2d}(\mathbb{Z})$. In this way, we let $GSp_{2d}(A)$ act on $V \otimes \mathbb{A}$. Write $K_0 = \text{Ker}(G(\mathbb{A}^{(p\infty)}))$ for the maximal compact subgroup of $GSp_{2d}(\mathbb{A}^{(p\infty)})$ preserving $L$ and principal level $N$ structure. Then $K_0 \cap G(\mathbb{A}^{(p\infty)}) = K$ and $K = \hat{G}(N)$ with respect to $B = Q$.

As described in Theorem 5.9, $P^Q_{K_0}$ is representable by a quasi-projective scheme $M = M_Q = M_N$ defined over $\mathbb{Z}(p)$. Let $\mathbb{X}_{/M}$ be the universal quadruple over $M$ and $A \in \mathbb{X}$ be the universal abelian scheme. We consider the functor from $M$-SCH into...
SETS:

\[ T/S \mapsto \left[ (\mathcal{A}_T, i : O_B \rightarrow \text{End}_T(\mathcal{A}_T)) \mid i(1_B) = \text{id}_A \right] \]

This functor is representable by a scheme \( M_B/M \) basically by Corollary 5.7 (see Corollary 6.11 in [PAF] for the version of Corollary 5.7 which is necessary to prove this fact). Since the level structure \( \mathfrak{P}^p \) on \( \mathcal{A} \) gives rise to a level structure \( \mathfrak{P}(p) \) of \( (\mathcal{A}_T, i) \), we have a triple \( \mathfrak{X} = (\mathcal{A}_T, i, \mathfrak{P}(p))_T \). Thus \( \mathcal{P}^p_B \) is a subfunctor of the above functor.

Again by the rigidity of endomorphisms under specialization, \( \mathcal{P}^B_K \) is represented by a closed subscheme \( \text{Sh}_{p}(\mathcal{P})_K \) of \( M_B \) whose connected components are (each) finite over \( M_Q \) (see [PAF] 7.1.2 for more details). We are going to show that \( \text{Sh}_{p}(\mathcal{P})_K \) is of finite type over \( M_Q \) (so it is projective and finite over \( M_Q \)). Take a geometric point \( x \in M_Q \); suppose that we have \( i : O_B \leftrightarrow \text{End}_A \) as above, which gives rise to a geometric point \( y \in \text{Sh}_{p}(\mathcal{P})_K \). For a given \( T \), if \( T \) is connected, \( \text{End}((\mathcal{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q}) \) is a semi-simple algebra of finite dimension with positive involution (see [ABV] IV). Thus the number of embedding \( B \hookrightarrow \text{End}(\mathcal{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q} \) is finite up to inner automorphism. Moreover the number is bounded by a constant only depending on the dimension of \( A \), that is \( d \), because \( \text{dim}((\text{End}(\mathcal{A}_T) \otimes_{\mathbb{Z}} \mathbb{Q})) \) is bounded by \( 4d^2 \) (e.g. [GME] Theorem 4.1.19). If one changes \( i \) by an inner automorphism induced by \( \alpha \in \text{End}(\mathcal{A}_T) \) and if we suppose that \( (\mathcal{A}_T, \alpha \alpha^{-1}, \lambda, \overline{\eta}(p)) \) is still an element of \( \mathcal{P}_K^B(T) \), it is easy to show, by the condition that \( \eta(p)(T^{(p)}(X)) = L^{(p)} \) combined with (L1-2) that \( \alpha \) has to be an automorphism of \( A \).

Since automorphisms of an abelian variety preserving a given polarization are only finitely many by the positivity of polarization, there are only finitely many possibilities of having \( i : O_B \leftrightarrow \text{End}(\mathcal{A}_T) \) which gives rise to an element of \( \mathcal{P}_K^B(T) \). Thus \( \text{Sh}_{p}(\mathcal{P})_K \rightarrow M_Q \) is quasi finite. Then the projectivity of each connected component of \( \text{Sh}_{p}(\mathcal{P})_K \) over \( M_Q \) implies the finiteness of the map: \( \text{Sh}_{p}(\mathcal{P})_K \rightarrow M_Q \). Actually, one can show that the morphism: \( \text{Sh}_{p}(\mathcal{P})_K \rightarrow M_Q \) is a closed immersion (over \( Q \)) if \( N \) is sufficiently large (cf. [DI] 1.15 and [PAF] 8.4.2).

Again by the rigidity of endomorphism of abelian schemes (and semi-abelian schemes) over a normal base under specialization ([DAV] I.2.7), for \( N \) sufficiently large, \( \mathcal{P}_K^{B} \) is represented by the schematic closure of \( \text{Sh}_{p}(\mathcal{P})_K \) in \( M_{Q/S_0} \), and hence \( C_B \) for general \( B \) is a representable by a scheme \( \text{Sh}_{p}(\mathcal{P})_K \) projective over \( M_{Q/S_0} \) if \( K^{(p)} \) is sufficiently small.

Although we assumed that \( L \) is self dual, replacing \( GSp_{2d} \) by its suitable conjugate in \( GL(2d) \), we can easily generalize the above argument to a given polarization of degree prime to \( p \).

In exactly the same way, we may conclude \( \mathcal{P}_K \cong \mathcal{P}^{B}_K \) over \( E \) (not over \( S_0 \)) even if \( K \) is not maximal at \( p \); so, we get the representability of \( \mathcal{P}_K \) by the Shimura variety \( \text{Sh}_{K/E} \) and the inclusion \( \text{Sh}_{K/E} \hookrightarrow \text{Sh}_{K}^{(p)} \times_{S_0} E \) if \( K \) is maximal at \( p \). Hereafter, if confusion is unlikely, we remove the superscript “\((p)\)” from the notation \( \text{Sh}_{K}^{(p)} \), and if we consider the Shimura variety \( \text{Sh}_{K} \) over \( S_0 \)-scheme, we implicitly assume...
Sh_{K/E} = Sh_{K}^{(p)} \times_{S_0} E$, that $K$ is maximal at $p$ and that the model is the integral Kottwitz model $Sh_{K}^{(p)}$. As we already remarked, $Sh_{K/E} = Sh_{K}^{(p)} \times_{S_0} E$ holds if $G$ is a type C group or $F$ is an imaginary quadratic field ([PAF] Theorem 7.5).

In the non-compact case, in [DAV], depending on the data at the cusps governing how toroidal compactification is done, a semi-abelian scheme $G_{/\overline{M}_Q}$ (universal under the data) is constructed. Then a similar argument using Corollary 5.8 (applied to $G_{/\overline{M}_Q}$ in place of $A_{/M_0}$) gives a projective scheme over $\overline{M}_Q$ for a toroidal compactification $\overline{M}_Q$ of the Mumford moduli (by Chai and Faltings). Since the endomorphism algebra of an abelian variety $X_k$ for an algebraically closed field $k$ (after tensoring $\mathbb{Q}$) is semi-simple, there is only finitely many possibility of embedding $B$ into $\text{End}_{k}(X) \otimes \mathbb{Z} \mathbb{Q}$ up to conjugation. Thus the morphism $\overline{M}_B \to \overline{M}_Q$ has finite geometric fiber everywhere, that is, the morphism is quasi-finite. Since the scheme $\overline{M}_B$ is proper over $\overline{M}_Q$ (see Corollary 5.8), it has to be finite. Thus writing $\omega = f_! \Omega_{A/M_B}$ for $f : A \to M_B$ and defining $\det(\omega)$ by its maximal exterior product, we can define a graded algebra:

$$G^K = G^K_B = \bigoplus_{n=0}^{\infty} H^n(\overline{M}_B, \det(\omega)^n).$$

Moreover, as seen in the last subsection of Section 5, $M^*_Q = \text{Proj}(G^*_Q)$ and hence we have the minimal compactification of $\text{Sh}_K$ defined by $\text{Sh}_K^* = \text{Proj}(G^*_B)$, which is finite over the minimal compactification $M^*_Q$ of the Mumford moduli.

If one shrinks enough the group $K$ outside $p$, any endomorphism of the semi-abelian scheme sitting over the cusp of $\overline{M}_K$ extends uniquely to infinitesimal neighborhood of the image of the cusp of $\overline{M}_B$ in $\overline{M}_Q$; so, $\overline{M}_B$ is étale around the cusp over the image of $\overline{M}_B$ in $\overline{M}_Q$. The smoothness of $\overline{M}_B$ at cusps for a well chosen cuspidal datum was shown by Fujiwara for $C$ of type $A$ and $C$ ([F]). If one choose the cuspidal data for $GSp(2d)$ and $G$ so that they are compatible (in other words, so that the pull back of the semi-abelian scheme over $\overline{M}_Q$ is the semi-abelian scheme over $\overline{M}_B$ associated to the cuspidal data for $G$), this guarantees that the $q$-expansion parameter is well defined over $S_0$ and projectivity for $\overline{M}_B$ of level prime to $p$, because it is finite over $\overline{M}_Q$.

Even if $K$ is not very small, we always have a coarse moduli scheme $\text{Sh}_K$ representing the functor $\mathcal{P}_K^{(p)}$ or $\mathcal{P}_K$ over $S_0$ or $E$ accordingly. The above arguments all work well. We write $\overline{\text{Sh}}_K$ for a toroidal compactification of $\text{Sh}_K$ and $\overline{\text{Sh}}_K^*$ for the minimal compactification. Since the natural morphisms:

$$\overline{\text{Sh}}_K^* \to M^*_Q \text{ and } \overline{\text{Sh}}_K \to \overline{M}_Q$$

are quasi-finite and projective, they are finite. Let $V$ be the image of $\overline{\text{Sh}}_K^*$ in $M^*_Q$. Then $V = \text{Proj}(G^*_K)$ for a graded algebra $G^*$ which is the quotient of $G^*_Q$. Then, assuming the existence of a smooth toroidal compactification of $\text{Sh}_K$, we have

$$(\text{Proj}) \quad \overline{\text{Sh}}_K^* = \text{Proj}(G^*_K).$$
Here $G^K_B$ is the integral closure of $G^K_Q$ for the Mumford moduli in the algebraic closure of the total quotient ring of $G$ if $K$ is sufficiently small. This follows from the fact that $\text{Sh}_K^*$ is smooth outside cusps, and at the cusps, if $K$ is sufficiently small, it is finite over $\mathcal{M}_Q$ (and normal over $V$). The graded algebra $G^K_B$ is the graded algebra of automorphic forms on $G$ if $\dim \text{Sh}_K > 1$.

We have formulated the moduli problem for the similitude group $G$. But we can impose polarization $\lambda$ without ambiguity modulo $\mathbb{Z}_{(p)} = \mathbb{Z}_p \cap \mathbb{Q}$. Then we automatically obtain the moduli problem for $G_1$; so, we do not describe the moduli problem and Shimura varieties for $G_1$, although our theorems are valid also for $G_1$ with some obvious modification.

6.2. Shimura Variety of Unitary Similitude Groups. — We could think of the Shimura variety of the unitary similitude group given by

\[(6.4) \quad \text{GU}(A) = \{ x \in C \otimes Q \; A \mid xx^* \in (A \otimes Q F_0)^x \}, \]

where $F_0$ is the subfield of $F$ fixed by the involution $a^*$. Thus we have $\text{GU} \supset G \supset G_1$.

To define the moduli problem of abelian schemes associated to $\text{GU}/Q$, we need to modify slightly the morphisms of the fiber category $C_B$: We define the fiber category $\mathcal{A} = \mathcal{A}_B$ over $\text{SCH}_{/S_0}$ to be the category of quadruples $\mathcal{X}/S = (X, \lambda, i, \pi^p)$ for $\pi^p = \eta^p$ mod $K$, where $K \subset \text{GU}(\mathbb{A}^{\infty})$ is a closed subgroup maximal at $p$. Write $O_0$ for the integer ring of $F_0$. Then we define

\(6.5\) \quad \text{Hom}_{A_B} (\mathcal{X}/S, \mathcal{X}'/S) = \left\{ \phi \in \text{Hom}(X, X') \otimes \mathbb{Z}_{(p)} \mid \hat{\phi} \circ \lambda' \circ \phi = \lambda \circ i(a) \text{ with } a \in (O_0 \otimes \mathbb{Z}_{(p)})_+ \right\}, \]

where $(O_0 \otimes \mathbb{Z}_{(p)})_+$ indicates the group of totally positive units in $(O_0 \otimes \mathbb{Z}_{(p)})^x$. We then consider the functor

\[p^{(p)}_{K} : \mathcal{A}_B \mapsto (\mathcal{X}/S \mid X \text{ satisfies } (\det))^/ \approx, \]

where $\approx$ indicates isomorphism classes in $\mathcal{A}_B$. The above functor can be proved to be representable if $K$ is sufficiently small by the same argument as in the case of $G$ (see [PAF] 7.1.3), and its generic fiber gives the Shimura variety over $E$ (defined adding a requirement analogous to $(\text{pol})$; see [PAF] Theorem 7.5). The compactification of the moduli space $M_{K/S_0}$ can be also done as described above. The only point we need to make explicit is that if the class $\lambda$ of polarizations modulo multiplication by totally positive element in $(O_0 \otimes \mathbb{Z}_{(p)})^x$ is defined over $S$, we can always find a representative $\lambda$ defined over $S$. Indeed, picking one symmetric polarization $\lambda$, the pull back by $1 \times \lambda$ of the universal line bundle over $X \times S^{1}X/X$ (the Poincaré bundle) is always ample and is equal to $2\lambda$ (see [GIT] Proposition 6.10); so, in the class $\lambda$, we can always find a polarization globally defined over $S$. 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
6.2.1. Classification of $G$. — Let $F_0$ be the subfield of $F$ fixed by the involution "*$". We define for $F_0$-algebras $A$,

$$G_0(A) = \{ x \in C \otimes_{F_0} A \mid xx^* = 1 \}.$$ 

Then we have $G_1 = \text{Res}_{F_0/Q} G_0$. The involution "*" either induces a non-trivial involution on $F$ (a positive involution of the second kind) or the identity map on $F$ (a positive involution of the first kind). If "*" is of second kind, $F$ is a totally imaginary quadratic extension over a totally real field $F_0$ (a CM field over $F_0$), "*" coincides on $F$ the unique non-trivial automorphism over $F_0$ (complex conjugation; see [Sh1] and [ABV] Section 21). Then $G_0$ is an inner form of a quasi split unitary group over $F_0$. We call this case Case A and call the group $G$ type A.

When "*" induces the identity map on $F$, then $F = F_0$ is totally real, and the group $G_0$ is an inner form of either the symplectic group (Case C and the group of type C) or an orthogonal group of even variable (Case D and the group of type D).

We have

$$C_\infty \cong \begin{cases} 
M_n(\mathbb{C})^{I_0} \text{ and } x^* = I_{s,t}^t I_{s,t} & \text{in Case A}, \\
M_{2n}(\mathbb{R})^{I_0} \text{ and } x^* = J_{n,t}^t J_{n,t} & \text{in Case C}, \\
M_n(\mathbb{H})^{I_0} \text{ and } x^* = -i\overline{x} & \text{in Case D},
\end{cases}$$

where $I_0$ is the set of all field embeddings of $F_0$ into $\mathbb{R}$, $\mathbb{H} = \mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$ is the algebra of Hamilton quaternions, $J_n = \left( \begin{smallmatrix} 0 & -1_n \\ 1_n & 0 \end{smallmatrix} \right)$, $I_{s,t} = \left( \begin{smallmatrix} 1_s & 0 \\ 0 & -1_t \end{smallmatrix} \right)$ for the $t \times t$ identity matrix $1_t$, and $x \mapsto \overline{x}$ is either complex conjugation or quaternion conjugation.

Suppose that $p > 2$ if we are in Case D. When $K^{(p)}$ is sufficiently small, $\text{Sh}_K$ is smooth over $S_0$. This follows from the fact that the deformation ring of a quadruple $(X, \lambda, t, \mathfrak{p}(p))$ is always formally smooth (cf. [GIT] Proposition 6.15, and [K]). If $C = \text{End}_B(V)$ is a division algebra, the big division algebra $B$ sitting inside $\text{End}_S(X) \otimes \mathbb{Q}$ for $S = \text{Spec}(K)$ with $K = \text{Frac}(V)$ for a valuation ring $V$ forces reduction of $X$ modulo the maximal ideal $\mathfrak{m}_V$ to be an abelian variety; so, by the valuative criterion of properness, $\text{Sh}_K/S_0$ is proper. Since $\text{Sh}_K$ is projective over the Mumford moduli $M_Q$ which is quasi-projective over $S_0$, $\text{Sh}_K/S_0$ has to be projective ([Ko] Section 5).

We now briefly describe the complex points of $\text{Sh}_K$. We can define the symmetric domain $\mathcal{X}$ as the collection of $h : C \mapsto C_\infty$ satisfying the positivity, etc., we described above. Since the stabilizer $C_h$ of a fixed $h$ in $G(\mathbb{R})$ is the product of the center and a maximal compact subgroup, the connected component of $\mathcal{X}$ is isomorphic to the symmetric domain $D = G(\mathbb{R})/C_h$. An explicit form of $D$ as a classical bounded matrix domain is given in [Sh1] (see also [ACM] Chapter VI for the domains in Case A and C), along with an explicit method of constructing all possible analytic families of abelian varieties over the domain. We have computed $D$ for unitary groups (that is, groups of type $A$) already in Section 4. The complex analytic space $\text{Sh}_K(\mathbb{C})$ is given by $G(\mathbb{Q}) \backslash G(A)/KC_h$, and its connected component is given by $\Gamma \backslash D$ for the
congruence subgroup $\Gamma = (gKg^{-1}G(\mathbb{R})_+) \cap G(\mathbb{Q})$ with a suitable $g \in G(\mathbb{A}^{(\infty)})$, where $G_+(\mathbb{R})$ is the identity component of the Lie group $G(\mathbb{R})$.

7. Formal Theory of Automorphic Forms

In this lecture, we describe the theory of false automorphic forms. The theory we describe is a generalization of the work of Deligne-Katz in the elliptic modular case (see [K1] Appendix III). The main purpose of this lecture is threefold:

1. Approximate $p$-adic automorphic forms by finite sums of classical forms;
2. Define the $p$-ordinary projector;
3. Find a set of (axiomatic) conditions which guarantees the VCT.

7.1. True and False Automorphic Forms. — In our application, we remove super-singular locus from the moduli $M/W$ of abelian schemes of PEL-type and write $S/W$ for $M[1_E]$ for a lift $E$ of the Hasse invariant. In this setting, sections in $H^0(S, \omega^\kappa)$ are called “false” automorphic forms. On the other hand, sections in $H^0(M, \omega^\kappa)$ are called “true” or “classical” automorphic forms.

7.1.1. An analogue of the Igusa tower. — Let $W$ be a mixed characteristic complete discrete valuation ring with residue characteristic $p$. Let $\varpi$ be a uniformizing parameter. Write $W_m = W/p^mW$. Let $S$ be a flat $W$-scheme. We put $S_m = S \times_W W_m$. Then $S_m$ is a sequence of flat $W_m$-schemes, given with isomorphisms:

$$S_{m+1} \otimes_{W_{m+1}} W_m \cong S_m.$$ 

Let $P$ be a rank $g$ $p$-adic étale sheaf on the $S_m$'s; thus, $P/S_{m+1}$ induces $P/S_m$, $P = \varprojlim_n P/p^n P$, and $P_n = P/p^n P$ is a twist of the constant sheaf $(\mathbb{Z}/p^n \mathbb{Z})^g$. We write $S_\infty$ for the formal completion of $S$ along $S_1$; so, $S_\infty = \varprojlim_m S_m$.

We can slightly generalize our setting and could suppose that there exists a finite extension $F/\mathbb{Q}$ with integer ring $O = O_F$ and a homomorphism: $O \hookrightarrow \text{End}_{S_m}(P)$ such that $P_n \cong (O/p^n O)^g$ for all $n$ locally under étale topology. Since we can transfer any of our results to this slightly general situation, just replacing $\mathbb{Z}_p$ by $O_p = O \otimes \mathbb{Z}_p$, we state our result only for $P$ with $P_n \cong (\mathbb{Z}/p^n \mathbb{Z})^g$. This simplification also allows us to save some symbols.

Let $\omega_m$ be the vector bundle $P \otimes_{}\mathbb{Z}_p O_{S_m}$. We define

$$\pi_{m,n} : T_{m,n} = \text{Isom}_{S_m}(P_n, (\mathbb{Z}/p^n \mathbb{Z})^g) \rightarrow S_m$$

to be a finite étale $S_m$-scheme which represents the following functor on $\text{SCH}/S_m$:

$$(\pi : X \rightarrow S_m) \mapsto \{\text{isomorphisms } \psi_n : P_n/X \cong (\mathbb{Z}/p^n \mathbb{Z})^g_X \}.$$ 

The representability follows from the theory of Hilbert schemes as we have seen. By definition, $T_{m,n}/S_m$ is étale. Since each geometric fiber of $T_{m,n}$ over $S_m$ is isomorphic
to $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ everywhere, it is faithfully flat and finite. Therefore $T_{m,n}$ is affine over $S_m$. We define $V_{m,n} = H^0(T_{m,n}, \mathcal{O}_{T_{m,n}})$.

The group $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on $T_{m,n}$ freely by $\psi \mapsto g\psi$ for $g \in GL_g(\mathbb{Z}/p^n\mathbb{Z})$, and we have $T_{m,n}/T_{n,n'} \cong T_{m,n'}$ for all $n' \leq n$, where $\Gamma_{n,n'} = \{ x \in GL_g(\mathbb{Z}/p^n\mathbb{Z}) \mid x \equiv 1 \mod p^{n'} \}$.

Then we have a tower:

$$V_{m,0} \subset V_{m,1} \subset \cdots \subset V_{m,n}$$

with $V_{m,0} = H^0(S_m, \mathcal{O}_{S_m})$. We put $V_{m,\infty} = \bigcup_n V_{m,n}$ and $T_{m,\infty} = \lim_{n \to \infty} T_{m,n}$.

7.1.2. Rational representations and vector bundles. — For a given ring $A$ or a sheaf of rings $A$ over a scheme, we look at the projective scheme $\mathcal{F}_A = \text{Flag}_{B/A}$ of all maximal flags in $A^g$ (cf. [PFA] II.6.13). We write $B \subset GL(g)$ for the upper triangular Borel subgroup. Let $U$ be its unipotent radical, and put $T = B/U$ for the torus. Then $\mathcal{F} \cong GL(g)/B$. We define $\mathcal{H}_A = \text{Flag}_{U/A} = GL(g)/U$. Write 1 for the origin of $\mathcal{H}$ represented by the coset $U$. Then

$$R_A = H^0(\mathcal{H}, \mathcal{O}_\mathcal{H}) = \bigoplus_{\kappa \in X(T)_+} R_A[\kappa]$$

for the space $R_A[\kappa]$ of weight $\kappa$. Here

$$R_A[\kappa] = \{ f : GL(g)/U \to A^1 \in \Gamma(GL(g)/U, \mathcal{O}_{GL(g)/U}) \mid f(ht) = \kappa(t)f(h) \}$$

for $t \in T$ for the diagonal torus $T \cong B/U \cong ^tB/^tU$. The pull-back action of $GL(g)$ on $R_A[\kappa]$: $f(x) \mapsto \rho(h)f(x) = f(h^{-1}x)$ gives a representation $\rho = \rho_\kappa$ such that $R_A[\kappa]^U \cong A$ on which $T$ acts by $-w_0\kappa$, where $w_0$ is the longest element of the Weyl group of $T$. The dual $R_A[\kappa] = \text{Hom}_A(R_A[\kappa], A)$ is the universal representation of highest weight $\kappa$ (cf. [RAG] II.2.13). Thus the coinvariant space $R_A[\kappa]_U$ (on which $T$ acts by $-\kappa$) is $A$-free of rank 1, and there is a unique $U$-invariant linear form $\ell_{can} : R_A[\kappa] \to A$ (up to $A$-unit multiple), which generates $(R_A[\kappa]^*)^U$. We can normalize $\ell_{can}$ so that it is the evaluation of $\phi \in R_A[\kappa]$ at the origin $1 \in GL(g)/U$. Then we have a tautological embedding $R_A[\kappa] \hookrightarrow \Gamma(GL(g)/U, \mathcal{O}_{GL(g)/U})$ given by

$$\phi \mapsto \{ h \mapsto \ell_{can}(\rho(h^{-1})\phi) \}.$$ 

If $h^{-1} \in M_{g \times g}$ for the $g \times g$ matrix algebra $M_{g \times g}$ as a multiplicative semi-group scheme, the action of $\rho(h)$ is well defined on $R_A[\kappa]$ for any $A$.

In [RAG], a slightly different module is considered:

$$H^0(-\kappa) = \text{ind}_{T^U}^{GL(g)} A(-\kappa) = \{ f : GL(g)/U \to A^1 \mid f(xt) = t^{-\kappa}f(x) \forall t \in T \}.$$ 

The action of $GL(g)$ is given by $hf(x) = f(h^{-1}x)$. In this context, $-\kappa$ is a positive weight with respect to $^tB$, and the $H^0(-\kappa)^U$ contains the highest weight vector. Using conjugation by $w_0$, we can remove the use of the lower triangular Borel subgroup $^tB$, but we need to modify the results of [RAG] accordingly, when we quote them (this will be done without further warning).
Let $f \in R_{A^{\infty}}[\kappa]$. By definition, $f$ induces a function on $GL_g(\mathbb{Z}_p)$ by $f(h) = \rho(\kappa)h(1)$. Therefore we see that $h \mapsto \ell_{\text{can}} \circ f(h)$ is an element in $R_{A^{\infty}}[\kappa]$ by tautology. This shows the following fact:

(c) We have a canonical map $R_{A}[\kappa] \rightarrow \mathcal{C}(GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p), A)[\kappa]$, which is injective if $A$ is flat over $\mathbb{Z}_p$. Here $\mathcal{C}(GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p), A)$ is the space of $(p$-adic) continuous functions with values in $A$ on $GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p)$, and $\text{”}[\kappa]\text{”}$ indicates the $\kappa$-eigenspace under the right action of $T(\mathbb{Z}_p)$ on $GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p)$. The cokernel of the map (c) is large, because it is the continuous induction from $B(\mathbb{Z}_p)$ to $GL_g(\mathbb{Z}_p)$ for a $p$-adic ring $A$ if $P_n$ is constant. When $A$ is a finite ring, the space of continuous functions $\mathcal{C}(GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p), A)$ is equal to the space of locally constant functions $\mathcal{C}(GL_g(\mathbb{Z}_p)/U(\mathbb{Z}_p), A)$, and we use $\mathcal{C}$ instead of $\mathcal{C}$ when $A$ is finite.

7.1.3. Weight of automorphic forms and representations. — We define a coherent sheaf $\omega_{m,n}$ on $T_{m,n}$ by $(\pi_{m,n}^*P_m) \otimes \mathcal{O}_{T_{m,n}}$. On $T_{m,n}$ with $n \geq m$, we have the universal isomorphism

$I_{\text{can}} : \pi_{m,n}^*P_m \cong (\mathbb{Z}/p^n\mathbb{Z})^{\phi}$

so we have an action of $\text{Gal}(T_{m,\infty}/S_m)$ on $\pi_{m,n}^*P_m$, and

$\omega_{\text{can}} = I_{\text{can}} \otimes \text{id} : \omega_{m,n} \cong \mathcal{O}_{T_{m,n}}$ is an isomorphism. Then we can identify $H/T_{m,n}$ with

$p_Y : Y = Y_{m,n} = GL(\omega_{m,n})/U_{\text{can}} \rightarrow T_{m,n}$ on $V_{m,n}$ ($n \geq m$), where $U_{\text{can}}$ is the pull back of $U$ under $\omega_{\text{can}}$. Thus $\omega_{\text{can}}$ induces an isomorphism:

$\omega_{\text{can}}^\kappa : p_Y^*(\mathcal{O}_Y[\kappa]) \cong R_{V_{m,n}}[\kappa]$. We write $\omega_{m,n}^\kappa$ for the sheaf $p_Y^*(\mathcal{O}_Y[\kappa])$ on $T_{m,n}$. By definition, $GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on $Y$ on the left. The Galois group $\text{Gal}(T_{m,\infty}/S_m) = GL_g(\mathbb{Z}/p^n\mathbb{Z})$ acts on $\omega_{m,n}^\kappa$ via the rational structure given from $\omega_{m,n}$, and we then descend the sheaf to $\omega_{m,n}^\kappa$ on $S_m$. In other words, for an $\mathcal{O}_{T_{m,n}}$-algebra $A$, $f \in H^0(\text{Spec}(A), \omega_{m,n}^\kappa)$ is a functorial rule assigning $f(X, \psi) \in R_{A}[\kappa]$ to $X_{/A}$ and $\psi : P_{n/X} \cong (\mathbb{Z}/p^n\mathbb{Z})^{\phi}$ on $X$. We let $h \in GL_g(\mathbb{Z}/p^n\mathbb{Z}) = \text{Gal}(T_{m,n}/S_m)$ act on $f$ by $f \mapsto \{(X, \psi) \mapsto \rho(h)^{-1}f(X, h\psi)\}$. Then for any $\mathcal{O}_{T_{m,n}}$-algebra $A$,

$A \rightarrow H^0(\text{Gal}(T_{m,n}/S_m), H^0(\text{Spec}(A) \times S_m T_{m,n} ; \omega_{m,n}^\kappa))$

defines a coherent sheaf on $S_m$ (by the Hochschild-Serre spectral sequence), which we write $\omega_{m,n}^\kappa$. We have

$H^0(S_m, \omega_{m,n}^\kappa) = \{ f \in H^0(T_{m,n}, R_{V_{m,n}}[\kappa]) | f(X, h\psi) = \rho(h)f(X, \psi) \text{ for } \forall h \in GL_g(\mathbb{Z}/p^n\mathbb{Z}) \}$.
There is another description of $\omega^\kappa_{m/S_m}$. Since $P_m \cong (\mathbb{Z}/p^m\mathbb{Z})^g$ on $T_{m,m}$, the action of $\text{Gal}(T_{m,n}/S_m)$ on $P_m$ extends to an action of the Galois group on $R_{\mathbb{Z}/p^m\mathbb{Z}}[\kappa]$, which determines an étale torsion sheaf $\mathcal{P}_m^\kappa$ over $S_m$. Then we have
\[ \omega^\kappa_m = \mathcal{O}_{S_m} \otimes_\mathbb{Z} \mathcal{P}_m^\kappa. \]
In this construction, we have $\det^k(\omega_m) = (\wedge^g \omega)^\otimes k$ and $\text{Sym}^k(\omega_m) = \omega_m^{k\omega_1}$ for the first standard dominant weight $\omega_1$.

By our definition, for each $f \in H^0(S_m, \omega_m^\kappa)$,
\[ \varphi(f) = \omega^\kappa_{\text{can}}(f) \in H^0(T_{m,m}, R_{T_{m,m}}[\kappa]), \]
which can be regarded as a functorial rule assigning each test object $(X/S_m, \psi : P_m/X \cong (\mathbb{Z}/p^m\mathbb{Z})^g)$ a value $\varphi(f)(X, \psi) \in H^0(X, R_{O_X}[\kappa])$ such that $\varphi(f)(X, h\psi) = \rho(h)\varphi(f)(X, \psi)$ for all $h \in GL_g(\mathbb{Z}/p^n\mathbb{Z})$ and $\varphi(f)(Y, \varphi^* \psi) = \varphi^*(\varphi(f)(X, \psi))$ for any morphism $f : Y \to X$ of $S_m$-schemes. Similarly, $\varphi \in V_{m,n}$ is a functorial rule assigning $(X, \psi)$ a value $\varphi(X, \psi) : P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g \ni h \to \varphi(Y, h\psi)$ for all morphisms $f : Y \to X$ of $S_m$-schemes. Thus we have a natural map of $H^0(T_{m,m}, R_{T_{m,m}}[\kappa])$ into $V_{m,n}$ associated to each linear form $\ell \in R_{V_{m,n}}[\kappa]^*$. The map associates $f \in H^0(T_{m,m}, R_{T_{m,m}}[\kappa])$ with a rule: $(X, \psi) \mapsto \ell(\varphi(f)(X, \psi))$, which is a matrix coefficient of $\varphi(f)(X, \psi)$.

We let $\ell \in GL_g(\mathbb{Z}_p)$ act on test objects by $(X, \psi) \mapsto (X, h\psi)$. In this way, we identify $GL_g(\mathbb{Z}_p)$ with $\text{Gal}(T_{m,\infty}/S_m)$. For the Borel subgroup $B \subset GL(g)$, we put $T_{m,n}^B$ for the quotient $T_{m,n}/B(\mathbb{Z}/p^n\mathbb{Z})$. Thus $V_{m,n}^B = H^0(T_{m,n}^B, \mathcal{O}_{T_{m,n}^B})$ is made of a functorial rule $(X, \psi) \mapsto \varphi(X, \psi) \in H^0(X, \mathcal{O}_X)$ such that $\varphi(X, b\psi) = \varphi(X, \psi)$ for all $\psi$ and $b \in B(\mathbb{Z}_p)$. We define similarly $V_{m,n}^U$ and $T_{m,n}^U$ for the unipotent subgroup $U \subset B$.

Let $e_1, \ldots, e_g$ be the standard base $e_j = \{0, \ldots, 0, 1, 0, \ldots, 0\}$ of $(\mathbb{Z}/p^n\mathbb{Z})^g$, and we consider the standard filtration $1_n : (\mathbb{Z}/p^n\mathbb{Z})^g = L_g \supset L_{g-1} \supset \cdots \supset L_0 = \{0\}$ given by $L_j = \sum_{i=1}^j (\mathbb{Z}/p^n\mathbb{Z}) e_i$. Then $\psi_n(1_n)$ gives a (full) filtration $\text{fil} = \text{fil}_n$ of $P_n$ and all full filtrations $P_n = P_n^{(g)} \supset P_n^{(g-1)} \supset \cdots \supset P^{(0)} = \{0\}$ of $P_n$ are given in this way. Since the stabilizer of $1_n$ is $B(\mathbb{Z}/p^n\mathbb{Z})$, we may regard $\varphi \in V_{m,n}^B$ as a functorial rule assigning a value $\varphi(X, \text{fil}_n) \in H^0(X, \mathcal{O}_X)$ to a test object $(X, \text{fil}_n)$. To describe $V_{m,n}^U$ in this way, we need to bring in an isomorphism of graded modules: $\phi_n : \text{gr}(\text{fil}_n) \cong \oplus_{j=1}^g (\mathbb{Z}/p^n\mathbb{Z})$ inducing $\phi_n^{(j)} : P^{(j)}/P^{(j-1)} \cong (\mathbb{Z}/p^n\mathbb{Z})$. In other words, $T_{m,n}^U$ classifies triples $(X, \text{fil}_n, \phi_n)$. Since we pulled back the filtration $1_n$ by $\psi_n$, $h \in GL_g(\mathbb{Z}_p)$ acts on $(X, \text{fil}_{\psi_n, \phi_n})$ by $\text{fil}_n \mapsto \psi^{-1}h^{-1}\psi\text{fil}_n = (h\psi_n)^*1_n$ and $\phi_n \mapsto \psi^{-1}h^{-1}\psi\phi_n$.

We can think of the image of $R_{V_{m,m}}[\kappa]$ inside $V_{m,n}^U[\kappa]$, which is the homomorphic image of $H^0(S_m, \omega_m^\kappa)$ under $f \mapsto \ell_{\text{can}} \circ \varphi(f)$. Thus we have a natural map
\[ \beta : H^0(S_m, \omega_m^\kappa) \to V_{m,\infty}[\kappa], \]
\[ \beta(f) = \omega_{\text{can}}^\kappa(f) \in H^0(T_{m,m}, R_{T_{m,m}}[\kappa]). \]
where \( V_{m,m}^U [\kappa] \) is the \( \kappa \)-eigenspace of the right action of \( T \). The above map is injective if \( m = \infty \). Then we define

\[
R' = \bigoplus_{\kappa \in \mathcal{X}(T)_+} \mathcal{H}^0(S_m, \omega^\kappa).
\]

Here \( \gg \) implies sufficiently regular. See [PAF] 5.1.3 for a definition of regularity.

We assume to have a locally free sheaf \( \omega_{/S} \) of finite rank such that \( \omega \otimes_W W_m = \omega_m \) for all \( m \). From \( \omega \), we can create \( \omega^\kappa_P S \) as \( \pi_* \mathcal{O}_{\text{Flag}_U}(\omega) \) for \( \pi : \text{Flag}_U \to S \). The global sections \( \mathcal{H}^0(S, \omega) \) inject into \( \mathcal{H}^0(S_{\infty}, \omega_{\infty}) = \lim_{\leftarrow m} \mathcal{H}^0(S_m, \omega_m) \). We define

\[
\mathcal{R}' = \bigoplus_{\kappa} \mathcal{H}^0(S, \omega^\kappa) \hookrightarrow \mathcal{R}_{\infty} = \lim_{\leftarrow m} \mathcal{R}'_m.
\]

We call an element of \( \mathcal{H}^0(S, \omega^\kappa_S) \) a false automorphic form of weight \( \kappa \). A true automorphic form is a global section in \( \mathcal{H}^0(M, \omega^\kappa_M) \) for a compactification \( M \supset S \) of \( S \) we will specify later. In other words, false automorphic forms are meromorphic sections over \( M \) with a specified location of their poles.

7.1.4. Density theorems. — We suppose now that for all \( \kappa \gg 0 \), the short exact sequence:

\[
0 \longrightarrow \omega^\kappa \longrightarrow p^m \omega^\kappa \longrightarrow \omega^\kappa_m \longrightarrow 0
\]

gives rise to an exact sequence:

\[
(\text{Hyp1}) \\
0 \longrightarrow \mathcal{H}^0(S, \omega^\kappa) \longrightarrow p^m \mathcal{H}^0(S, \omega^\kappa) \longrightarrow \mathcal{H}^0(S_m, \omega^\kappa_m) \longrightarrow 0;
\]

\[
(\text{Hyp2}) \\
V^U_{m,\infty} = V^U / p^m V^U.
\]

This condition is obviously satisfied when \( S_m \) is affine. From this, we have

\[
R'/p^m R' \cong R'_m \quad \text{and} \quad \mathcal{H}^0(S, \omega^\kappa)/p^m \mathcal{H}^0(S, \omega^\kappa) \cong \mathcal{H}^0(S_m, \omega^\kappa_m).
\]

We now define a homomorphism

\[
\beta(m) : R'_m \longrightarrow V^U_{m,m}
\]

in the following way. Over \( T_{m,m} \), we have a canonical isomorphism \( \omega_{\text{can}} = I_{\text{can}} \otimes \text{id} : \omega_{m,m} \cong \mathcal{O}_{T_{m,m}}^\beta \). Then

\[
\beta(m)(\sum_{\kappa \gg 0} f_\kappa) = \{(X/T_{m,m}, \psi) \mapsto \sum_\kappa \ell_{\text{can}}(\omega^\kappa_{\text{can}}(f_\kappa(X, \psi))\}
\]

for \( f_\kappa \in \mathcal{H}^0(S_m, \omega^\kappa_m) \). Here, the image of \( \beta(m) \) actually falls in \( V^U_{m,n} \) because \( \ell_{\text{can}} \circ \rho_\kappa (u) = \ell_{\text{can}} \) for all \( u \in U(\mathbb{Z}_p) \), and \( \omega^\kappa_{\text{can}}(f_\kappa) \in R'_{m,m}[\kappa] \). By construction, \( \beta(n) \) mod \( p^m = \beta(m) \) for all \( n > m \). Thus taking the projective limit, we have

\[
\beta(\infty) : R'_{\infty} \longrightarrow V^U = \lim_{\leftarrow m} V^U_{m,\infty}.
\]

Since \( S_m \) is flat over \( W_m = W/p^m W \) and \( T^U_{m,n} = T_{m,n}/U(\mathbb{Z}/p^n \mathbb{Z}) \) is étale over \( S_m \), \( V^U_{m,\infty} \) is flat over \( W_m \). Therefore, \( V^U \) is a \( W \)-flat \( GL_g(\mathbb{Z}_p) \)-module. This is a subtle point. If \( \omega \) extends to the compactification \( M_m \), assuming \( M_m \) to be \( W_m \)-flat,
$H^0(M_m, \omega^\kappa)$ is also $W_m$-flat. It is easy to create an example in the Hilbert modular case such that $H^0(M_1, \omega^\kappa)/H^0(M_1, \omega_1) \neq 0$ for the interior $M$ of the Satake compactification of $S$ if $\kappa$ is not parallel but $\kappa \mod [(O/pO)_S]$ is parallel. By the Koecher principle, if $\omega^\kappa$ extends to the Satake compactification $M^\ast$ as a line bundle, we have $H^0(M^\ast, \omega^\kappa) = H^0(M^\ast, \omega_1)$; so, we cannot expect the good base-change property.

Since $B$ normalizes $U$, we can think of the action of $T = B/U$ on $V^U$ and the $\kappa$-eigenspace $V^U[\kappa]$ of $V^U$. By definition, $\beta = \beta(\infty)$ induces

$$\beta = \beta_\kappa : H^0(S, \omega^\kappa) \xrightarrow{\lim_{m}} H^0(S_m, \omega^\kappa_m) \longrightarrow V^U[\kappa].$$

**Proposition 7.1.** — Suppose (Hyp1,2) for $S$. The above map $\beta_\kappa$ is an injection.

**Proof.** — Since $T_{m,n}$ is faithfully flat and étale over $S_m$, we may make a base-change: $T^U_{m,n}/S_m$ to $T^U_{m,n}/S_m \times S_m T^U_{m,m}$, and hence we may suppose that $P$ is constant. Then $V^U_{m,\infty}$ is made up of locally constant functions on $GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p)$ with values in $V_{m,0}$. By taking the limit, $V^U = \text{the space of locally constant functions on } GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p)$ of continuous function on $GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p)$ with values in $V_{\infty,0}$.

Since $H^0(S, \omega^\kappa)$ is inside the limit of global sections of $\lim_m R_{1,m,0}[\kappa]$, which injects into $\mathcal{C}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), V_{\infty,0})[\kappa]$. This shows the assertion. 

We now put, for $\beta = \beta(\infty)$

$$D' = \beta(R')_p \cap V^U = \beta\left( \bigoplus \kappa \gg 0 H^0(S, \omega^\kappa) \right)_p \cap V^U.$$

**Theorem 7.2.** — Suppose (Hyp1-2) for $S$. The inclusion $\beta = \beta(\infty) : D'/p^m D' \to V^U$ induces an isomorphism

$$D'/p^m D' \cong V^U/p^m V^U$$

for all $m$. In other words, $D'$ is $p$-adically dense in $V$.

**Proof.** — The injectivity of $D'/p^m D' \to V^U/p^m V^U$ follows from the definition (see [K1] Appendix III) (or as easily seen after faithfully flat extension to $T^U_{m,m}$).

We thus need to prove that $D'/pD' \to V^U/pV^U = V^U_{1,\infty}$ is surjective. Since $T_{m,n}/S_m$ is étale finite, replacing $S_m$ by $T^U_{m,\infty}$, we may assume that $P$ is constant (see [K1] Appendix III pages 364-5), because we can recover the global sections of $\omega^\kappa$ over $S_m$ as Galois invariants of that over $T^U_{m,m}$. Then

$$\mathcal{O}_{V^U_{1,m}} = \mathcal{O}_{S_1} \otimes_W W[GL_2(\mathbb{Z}/p^n \mathbb{Z})/U(\mathbb{Z}/p^n \mathbb{Z})] = \mathcal{O}_{V^U_{1,0}} [GL_2(\mathbb{Z}/p^n \mathbb{Z})/U(\mathbb{Z}/p^n \mathbb{Z})].$$

This shows $V^U_{1,m} = \mathcal{L}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), V_{1,0})$, where $\mathcal{L}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), V_{1,0})$ is the space of locally constant functions on the $p$-adic analytic space $GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p)$ with values in $H^0(S_1, \mathcal{O}_{S_1}) = V_{1,0}$. Writing $V^U_{1,0}$ as a union of $W$-free modules $X$ of finite rank, we have $\mathcal{L}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), V^U_{1,0}) = \bigcup_X \mathcal{L}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), X)$. Thus we need to prove that

$$\mathcal{L}(GL_2(\mathbb{Z}_p)/U(\mathbb{Z}_p), X/pX) = \mathcal{D}_X/p \mathcal{D}_X,$$
where $\mathcal{D}_X$ is the space of polynomial functions of homogeneous degree $\geq 0$ (with coefficients in $K = \text{W}[rac{1}{p}]$ on the flag manifold Flag$_U$) which has values in $X$ over $GL_g(\mathbb{Z}_p)$. This last fact follows from Mahler’s theorem of the density of the linear span of the binomial polynomials in the space of continuous functions on $\mathbb{Z}_p$ with values in $\mathbb{Z}_p$ (see [PAF] Theorem 8.3 for more details of the use of Mahler’s theorem).

We now assume that there exists a proper flat scheme $M/W$ such that $S \subset M$, and $M - S$ is a proper closed subscheme of codimension $\geq 1$. We further assume that $ω_{/S}$ extends to $M$. Then automatically $ω^n$ extends to $M_m$ by the theory of flag varieties.

The sheaf $ω_{/M}$ is uniquely determined by $\{ω_{m/M_m}\}_{m=1,2,...}$ by the formal existence theorem of Grothendieck [EGA] III.5.1.4. By the properness of $M$, $H^1(M, ω^n)$ is a $W$-module of finite type. Thus taking the projective limit with respect to $m$ of the exact sequences:

$$0 \to H^0(M, ω^n) \otimes_W W_m \to H^0(M_m, ω^n_m) \to H^1(M_m, ω^n)[p^n] \to 0,$$

we get $\lim_{m} H^0(M_m, ω^n_m) = H^0(M, ω^n)$. Let $R_m = \bigoplus_{\kappa \geq 0} H^0(M_m, ω^n_m)$ and $R = \bigoplus_{\kappa \geq 0} H^0(M, ω^n)$. Then we know that $R$ is $p$-adically dense in $R_\infty = \lim_m R_m$. By definition, $R \subset R'$. Note that $det(ω)^{-1}$ is trivial on $S_1$. Let $a \in H^0(S_1, det(ω)^{-1})$ be the section corresponding to $1 \in det(ω)^{-1} \cong O_{S_1}$. We assume that $a$ extends to $M_1$ so that it vanishes outside $S_1$. Suppose that we have a section $E \in H^0(M, det(ω)^{t(p-1)})$ such that $E \mod ω = a^t$. By further raising power, that is, replacing $E$ by $E^{p^m}$, we may assume that $E \mod p = a^t$. Then by definition,

$$H^0(S_m, ω^n_m) = \lim_n \frac{H^0(M_m, ω^n_m) \otimes det^{nt(p-1)}(ω^n_m)}{E^n}.$$
Put

\[ D = \beta \left( \bigoplus_{\kappa > 0} H^0(M, \omega^\kappa) \right)[1/p] \cap V^U. \]

Then \( D \) is \( p \)-adically dense in \( V^U \).

**7.1.5. \( p \)-Ordinary automorphic forms.** — We now suppose to have a projector \( e \) (so \( e^2 = e \)) acting (continuously) on \( V^U \), which projects down \( V^U[\kappa] \) onto a \( W \)-free module of finite rank (for all \( \kappa \gg 0 \)). We put \( V_U = \varprojlim_{n} V^U/p^nV^U = \varprojlim_{n} V^U_{n, \infty} \).

We have \( V_U[\kappa] = \varprojlim_{m} V^U_{m, m}[\kappa] \). Since \( V^U \) is \( W \)-flat, \( V_U \) is \( p \)-divisible, and its direct summand \( eV_U \) is \( p \)-divisible.

In practice, the projector \( e \) will be constructed so that it brings \( V_U[\kappa] \) down onto \( eH^0(M/W, \omega^\kappa \otimes \mathbb{T}_p) \) for \( \kappa \gg 0 \), where \( \mathbb{T}_p = \mathbb{Q}_p/\mathbb{Z}_p \). This implies

\[(7.4) \quad eV_U[\kappa] = eH^0(S/W, \omega^\kappa \otimes \mathbb{T}_p) = eH^0(M/W, \omega^\kappa \otimes \mathbb{T}_p) \]

if \( \kappa > 0 \). By (Hyp1), \( H^0(S/W, \omega^\kappa \otimes \mathbb{Z}_p \mathbb{T}_p) \) is \( p \)-divisible. By assuming (7.4), \( eV_U[\kappa] \) is \( p \)-divisible. Since \( H^0(M/W, \omega^\kappa) \) is a \( W \)-module of finite type, \( eV_U[\kappa] \) is a \( p \)-divisible module of finite corank. In any case, we just assume that \( eV[\kappa] \) is \( p \)-divisible and of finite corank for \( \kappa \gg 0 \).

Let \( V_U^* \) be the Pontryagin dual module of \( V_U \). Since \( V_U \) is a discrete \( T(\mathbb{Z}_p) \)-module, \( V_U^* \) is a compact \( W[[T(\mathbb{Z}_p)]] \)-module. Let \( T(\mathbb{Z}_p) = \Gamma_T \) be the \( p \)-profinite part of \( T(\mathbb{Z}_p) \). Then \( T(\mathbb{Z}_p) = \Gamma_T \times \Delta \) for a finite group \( \Delta \) of order prime to \( p \). We fix a character \( \chi : \Delta \to \mathbb{F}^\times \) for \( \mathbb{F} = W/\varpi W \). Then we write \( \chi : \Delta \to W^\times \) for the Teichmüller lift of \( \chi \). We write \( X_{\chi} \subset X(T) \) for the set of algebraic characters \( \kappa : T \to \mathbb{G}_m \) such that \( \kappa \equiv \chi \mod \varpi \) and \( \kappa \) is sufficiently regular so that the above equation (7.4) holds. Then \( X_{\chi} \) is Zariski-dense in \( \text{Spec}(W[[\Gamma_T]])(W) \). We write \( V_{\text{ord}}^* \) for \( eV_U^* \). Let us decompose

\[ V_{\text{ord}}^* = \bigoplus_{\tilde{\chi} \in \Delta} V_{\text{ord}}^*(\tilde{\chi}) \]

into the direct sum of the \( \tilde{\chi} \)-eigenspaces under the action of \( \Delta \). The \( \tilde{\chi} \)-eigenspace \( V_{\text{ord}}^*(\tilde{\chi}) \) is a compact module over \( W[[\Gamma_T]] \). Then by (7.4), \( V_{\text{ord}}^*(\tilde{\chi}) \otimes_{W[[\Gamma_T]]} W \) is \( W \)-free of finite rank \( s(\tilde{\chi}) \) for \( \kappa \in X_{\tilde{\chi}} \). Thus, by topological Nakayama’s lemma, \( V_{\text{ord}}^*(\tilde{\chi}) \) is a \( W[[\Gamma_T]] \)-module of finite type with minimum number \( s(\tilde{\chi}) \) of generators. Since \( X_{\tilde{\chi}} \) is Zariski-dense in \( \text{Spec}(W[[\Gamma_T]]) \), we see that \( V_{\text{ord}}^*(\tilde{\chi}) \) is \( W[[\Gamma_T]] \)-free of rank \( s(\tilde{\chi}) \).

Thus we have, assuming (7.4) for the middle equality,

\[(7.5) \quad \text{rank}_{W[[\Gamma_T]]} V_{\text{ord}}^*(\tilde{\chi}) = \text{rank}_{W} V_{\text{ord}}^*(\tilde{\chi}) \otimes_{W[[\Gamma_T]]} W = \text{rank}_{W} (eH^0(M, \omega^\kappa) \otimes \mathbb{T}_p)^* = \text{rank}_{W} eH^0(M, \omega^\kappa) \]

for all \( \kappa \in X_{\tilde{\chi}} \). Therefore we get
Theorem 7.4. — Suppose (Hyp1-2), the existence of the idempotent \( e : V^U \to V^U \) as above and the assumptions of Corollary 7.3. Then \( V_{\text{ord}}^* \) is a well controlled \( W[[T(Z_p)]] \)-projective module of finite type. If we assume (7.4), this means that

\[
V_{\text{ord}}^* \otimes W[[T(Z_p)]] \cdot \kappa \cong \text{Hom}_W(eH^0(M, M^\kappa), W)
\]

canonical if \( \kappa \) is sufficiently regular. For each \( \tilde{\chi} \)-component, \( V_{\text{ord}}^*[\tilde{\chi}] \) is free of finite rank over \( W[[\Gamma_T]] \) for the maximal \( p \)-profinite subgroup \( \Gamma_T \) of \( T(Z_p) \).

7.1.6. Construction of the projector \( e_{GL} \). — We are going to construct an approximation \( e_{GL} \) of the projector \( e \). In the paper [H02] Section 2.6, we wrote: “\( e_{GL} \) is constructed using solely local data of the Galois group Gal\( (T_m,\infty/S_m) = GL_n(Z_p) \), while the projector \( e \) will be constructed as \( e = e_{GL}e_G \) for a global projector \( e_G \) depending on the group \( G \).” This statement is misleading. We actually need a global input. To explain this point, let us introduce the expanding semi-group of \( GL_n(Q_p) \).

Let diag\( [X_1, \ldots, X_j] \) denote the diagonal matrix whose diagonal blocks are given by \( X_1 \) to \( X_j \) from the top. We first look at the semi-group given by

\[
D = D_{GL(g)} = \{ \text{diag}[p^{e_1}, \ldots, p^{e_g}] \mid e_1 \leq e_2 \leq \cdots \leq e_g \}.
\]

Then \( \Delta_0^g = I_{\gamma, n} DU_{\gamma, n} \Delta_0^{U_{\gamma, n}} = U(Z_p)DU(Z_p) \) and \( \Delta_\infty^{B_{\gamma, n}} = B(Z_p)DB(Z_p) \) are semi-groups, and we call them expanding semi-groups. If confusion is unlikely, we simply write \( \Delta \) for one of these semi-groups.

The global input we need comes from the fact that \( T_{m,n} \) in our application classes not just trivializations of \( P_n \) but abelian varieties \( X \) with \( X[p^n]^{\text{et}} \cong P_n \). In other words, each \( g \in GL_n(Q_p) \) acts on \( S_m \) by an appropriate isogeny of abelian varieties classified, and it acts not only the étale quotient of the \( p \)-divisible group of the abelian variety but also on the connected component of the \( p \)-divisible group. The action changes the isomorphism class of the abelian varieties, and hence it acts on \( S_m \) through endomorphisms (not necessarily through automorphisms).

Since at this point, we do not assume that \( S_m \) classifies abelian varieties, we instead assume to have such an action of the expanding semi-group (as defined below) on \( S_m \) which is at worst “radiciel” mod \( p \); so, it does not affect the étale trivialization \( P_n \). This action \( \delta : S_m \to S_m \) sends an \( S_m \)-scheme \( X \xrightarrow{f} S_m \) to \( \delta \cdot X = X \times_{S_m, \delta} S_m \).

We consider the following triples:

\[
\mathcal{X} = (X/S_m, \text{til}_n, \phi_n : \oplus_{j=1}^g Z/p^nZ \cong \text{gr(\text{til}_n)}),
\]

where \( \text{til}_n : P_n/X = P_n^{(g)} \supset P_n^{(g-1)} \supset \cdots \supset P_n^{(0)} = \{0\} \) with \( \phi_j : Z/p^nZ \cong P_n^{(j)}/P_n^{(j-1)} \) for \( j = 1, \ldots, g \). If \( P \) is constant, the space classifying the above test objects over \( S_m \) is given by \( T^U_{m,n} = T_{m,n}/U(Z/p^nZ) \). Similarly, the classifying space of couples \((X, \text{til}_n)\) over \( S_m \) is given by \( T^B_{m,n} = T_{m,n}/B(Z/p^nZ) \). On test objects over \( T_{m,n} \),
we have a natural action of $h \in GL_n(\mathbb{Z}_p)$ given by $(X, \psi) \mapsto (X, h\psi)$. Writing $\text{fil}_n = \psi^{-1}1_n$, we then see that $\psi^{-1}h^{-1}1_n = \psi^{-1}1_n = h \cdot \text{fil}_n$. Thus the Galois action on filtrations is given by $h \cdot \text{fil} = (\psi)^{-1}h^{-1}\psi P_\alpha^{(j)}$ and $h \cdot \phi = (\psi)^{-1}h^{-1}\psi \phi$, where $\psi : P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g$ such that $\psi^*1_n \phi_\alpha = (\text{fil}_n, \phi_n)$ for the standard identification $\text{id} : \text{gr}(1_n) \cong (\mathbb{Z}/p^n\mathbb{Z})^g$. Thus these test objects are always invariant under $U(\mathbb{Z}_p)$.

The new test objects $(X, \text{fil}_n, \phi_n)$ are useful in defining an isogeny action of $\delta \in \Delta$ and in constructing the idempotent $e_{GL}$, although we may stick to the test objects $(X, \psi_n : P_n \cong (\mathbb{Z}/p^n\mathbb{Z})^g)$ if we want. We assume that

\begin{enumerate}
\item[(d1)] $\delta$ induces an isomorphism $\delta^*\text{fil}_{n/\delta X} \cong \text{fil}_{n/X}$ compatible with the action of the semi-group on $\text{fil}_n$ (this holds if $\delta$ mod $p$ is radiciel), where the action of $\delta \in \Delta$ on $\text{fil}_n$ is the multiplication by $\delta$ up to scalars (as we specify later);
\item[(d2)] $h \cdot X = X$ if $h \in GL_n(\mathbb{Z}_p)$.
\end{enumerate}

Here is how to create the idempotent $e_{GL}$ using $p$-Hecke operators (modulo $p^m$). We study Hecke operators $t_j (j = 1, \ldots, g)$ acting on $V^U$ and $V^\alpha_{m, n}$ for $? = B$ and $U$.

We can thus think of the Hecke ring $R(I_{t,n}, \Delta)$ ($n = 1, 2, \ldots, \infty$) made of $\mathbb{Z}$-linear combinations of double cosets $I_{t,n}\delta I_{t,n}$ for $\delta \in \Delta^B_\alpha$. These two algebras are commutative and all isomorphic to the polynomial ring $\mathbb{Z}[t_1, \ldots, t_g]$ for $t_j = I_{t,n}\alpha_j I_{t,n}$ with $\alpha_j = \text{diag}(1_{g-j}, p1)$. A key to getting this isomorphism is that once we choose a decomposition: $U(\mathbb{Z}_p)\alpha_j U(\mathbb{Z}_p) = \bigcup_j U(\mathbb{Z}_p)\delta_j$, then $I_{t,n}\alpha_j I_{t,n}$ for any $n$ and $? = 0$ is decomposed in the same way: $I_{t,n}\alpha_j I_{t,n} = \bigcup_j I_{t,n}\delta_j$ (see [PAF] (5.3)). We have for $\alpha = \prod_{j=1}^{\gamma-1} \alpha_j$

\begin{equation}
I_{t,n+1}\backslash I_{t,n+1}\alpha I_{t,n} \cong I_{t,n}\backslash I_{t,n}\alpha I_{t,n} \cong I_{t,n+1}\backslash I_{t,n+1}\alpha I_{t,n+1}.
\end{equation}

For $\delta \in \Delta^B_\alpha$, the action $\rho(\delta^{-1})$ is well defined on $R_A[\kappa]$ for any $p$-adic ring $A$, because $\rho(\delta^{-1})\phi(y) = \phi(\delta y)$ for $y \in GL(g)/U$. Decompose $I_{t,n}\delta I_{t,n} = \bigcup_j I_{t,n}\delta_j$ and regarding $f \in \text{H}^0(T^B_{m,n, \kappa})$ as a function of test objects $\mathcal{X}/T_{m, \infty}$, we define

\begin{equation}
f[I_{t,n}\delta I_{t,n}](\mathcal{X}) = \sum_j \rho_{\kappa}(\delta_j^{-1})f(\delta_j \mathcal{X}),
\end{equation}

where $\delta \mathcal{X} = (\delta \cdot X, \delta(\delta^*\text{fil}_n))$. The sum above is actually “heuristic”, because if the action of $\delta$ on $S$ is wildly ramified (that is, purely inseparable in characteristic $p$), we need to replace the sum by the trace as already described in 3.1.3. We will clarify this point in 8.2.1 more carefully; so, for the moment, we content ourselves with this heuristic action.

Although we have not yet specified the action of the element $\delta \in \Delta^B_\alpha$ on $\delta^*\text{fil}_n$, if it exists, then the operator is well defined independent of the choice of $\delta_j$ because for $u \in I_{t,n}$,

$$\rho((u\delta_j)^{-1})f(u\delta_j \mathcal{X}) = \rho((\delta_j)^{-1})\rho(u)^{-1}f(u\delta_j \mathcal{X}) = \rho(\delta_j^{-1})f(\delta_j \mathcal{X}).$$

Further, by (7.7),

\begin{equation}
\text{for } f \in \text{H}^0(T^B_{m,n, \kappa}), \ f|t(p)^{n-1} \in \text{H}^0(T^B_{m,1, \kappa}),
\end{equation}

ASTÉRISQUE 298
where $t(p) = \prod_{j=1}^{s-1} t_j$. When $P_n$ comes from a universal abelian scheme, we have a natural isogeny action on test objects, and in this way, we can define Hecke operator on $H^0(S_m, \omega^\infty)$.

Since $\text{fil}_n$ is an element of the flag variety of $(\mathbb{Z}/p^n\mathbb{Z})^g$, to study the action of $\Delta$ on filtrations, we study general flag varieties. For each commutative ring $A$, we consider the free module $L = L(A) = A^g$ and the flag space

$$y(A^g) = \{(L_i) \mid L = L_g, L_i \supset L_{i-1}, L_i/L_{i-1} \cong A \text{ for } i = 1, \ldots, g\},$$

$$\mathcal{Y}(A^g) = \{(L_i, \phi_i) \mid (L_i) \in y(A), \phi_i : A \cong L_i/L_{i-1} \text{ for } i = 1, \ldots, g\}.$$ 

We can extend the above definition to $P_n$ over $S_m$: We define for each scheme $T/S_m$

$$y(P_n)/T = \left\{(P_n^{(i)}) \mid P_n = P_n^{(g)}, P_n^{(i)} \supset P_n^{(i-1)}, \frac{P_n^{(i)}}{P_n^{(i-1)}} \cong \mathbb{Z}/p^n\mathbb{Z} \text{ (} i = 1, \ldots, g \text{)}\right\},$$

$$\mathcal{Y}(P_n)/T = \left\{(P_n^{(i)}, \phi_i) \mid (P_n^{(i)}) \in y(P_n)/T, \phi_i : \mathbb{Z}/p^n\mathbb{Z} \cong \frac{P_n^{(i)}}{P_n^{(i-1)}} \text{ (} i = 1, \ldots, g \text{)}\right\}.$$ 

After a finite étale extension to $T/S_m$, the spaces $y(P_n)$ and $\mathcal{Y}(P_n)$ get isomorphic to $y((\mathbb{Z}/p^n\mathbb{Z})^g) \times S_m T$ and $\mathcal{Y}((\mathbb{Z}/p^n\mathbb{Z})^g) \times S_m T$. Writing the standard base of $L$ as $e_1, \ldots, e_g$, we define $1 = (\sum_{j=1}^{g} Ae_i, \phi_i = \text{id}) \in \mathcal{Y}(A^g)$, which we call the origin. We may let $GL_q(A)$ act on $\mathcal{Y}(A^g)$ and $y(A^g)$ by $x((L_i), \phi_i) = (xL_i, x \circ \phi_i)$. Then $GL_q(A)/U(A) \cong \mathcal{Y}(A^g)$ by $xU(A) \mapsto x1$. Now we assume that $A$ to be a $p$-adic ring, that is, $A = \lim_{\leftarrow n} A/p^n A$. We then define

$$\mathcal{Y}_n(A^g) = \{(L_i, \phi_i)|\{L_i/p^n L_i\} = 1 \in y(A/p^n A)\}.$$ 

Similarly, we define $\mathcal{Y}_{n/T}(P_n')$ for $n \leq n' \leq \infty$. We note that $\mathcal{Y}_n(\mathbb{Z}_p^g) \cong I_{U,n}/U(\mathbb{Z}_p)$ via $1 \mapsto x$ and similarly $y_n(\mathbb{Z}_p^g) \cong I_{B,n}/U(\mathbb{Z}_p)$. So we have the conjugate action of $\Delta$ on these spaces introduced in Section 4.

We now write down explicitly the conjugate action of the semi-group $\Delta$ on $\mathcal{Y}_n(A^g)$. Since $y(\mathbb{Z}_p^g) = y(\mathbb{Q}_p^g)$ (because $y = \text{Flag}_B$ is projective), the group $GL_q(\mathbb{Q}_p)$ acts naturally on $y(\mathbb{Q}_p^g)$. This action is described as follows: Take $x \in GL_q(\mathbb{Q}_p)$. Then $x(L_i) = (xL_i \otimes Q_p \cap L(\mathbb{Z}_p)) \in y(\mathbb{Z}_p^g)$. We write $x(L_i) = (x \cdot L_i)$, that is, $x \cdot L_i = xL_i \otimes Q_p \cap L(\mathbb{Z}_p)$. We now define an action of the semi-group $\Delta_n^B = I_{B,n}D_{GL(g)}I_{B,n}$ on $\mathcal{Y}_n(\mathbb{Z}_p^g)$. For each $u^\Delta u' \in \Delta_n^B$ with $u, u' \in I_{B,n}$ and $\delta \in D_{GL(g)}$. We write $\delta = \text{diag}[p^{e_1(\delta)}, \ldots, p^{e_g(\delta)}]$. Then for $(L_i, \phi_i) \in \mathcal{Y}_n(\mathbb{Z}_p^g)$, $\delta \cdot (L_i)$ is a surjective isomorphism as shown in [H95] page 438. Since $I_{B,n}$ acts naturally on flag varieties, the above action of $D_{GL(g)}$ extends an action of the semi-group $\Delta_n^B$.

For a given $\mathcal{X} = (X, \text{fil}, \phi) = (X, \psi^{-1}1)$, $\psi$ brings “fill” to 1, and hence the action of $\Delta_n^B$ defined on the neighborhood of 1 (after conjugation by $\psi$) is enough to get an association: $\mathcal{X} \mapsto \{\delta \mathcal{X}\}$. By this, after a change of the base scheme $S_m$ (for example to $T = T_m, \infty$) to trivialize $\mathcal{Y}(P)$, we have an action of $\Delta_n^B$ on $\mathcal{Y}(P)$. However this is sufficient to define the Hecke operators $[I_{T,n} \delta I_{T,n}]$ acting on $H^0(T_{m,n}/S_m, \omega^\infty)$ by the following reason: After extending scalar, define $f[I_{T,n} \delta I_{T,n}]$ by (7.8). The formation
of $f[I_n; \delta I_{1,n}]$ commutes with the base-change, in other words, it commutes with the Galois action of the base: $\text{Gal}(T/S_m)$; so, $f[I_n; \delta I_{1,n}]$ is actually defined over the original base scheme $T^\circ_{m,n}/S_m$. This justifies the contraction property (7.9).

Let $t_j = U(Z_p)\alpha_j U(Z_p)$ in $R(U(Z_p), \Delta_\infty)$ with $\alpha_j = \text{diag}[1, p_{1-j}]$, and define $t(p) = \prod_{j=1}^g t_j$. As shown in [H95] Lemma 3.1, for $\alpha = \prod_{j=1}^g \alpha_j$ contracts $y_n(Z/p^{n+1}Z)$ to the origin $1_{n+1}$. Identifying $y_1(Z/p^nZ)$ with $I_{1,n}/I_{1,n}$, if the filtration $\text{fil}_n$ corresponds to $x \in I_{B,n}$, then the filtration is given by $\sum_{j=1}^g (Z/p^nZ)x_j$ for the $j$-th column vector $x_j$ of $x$. Choose a representative set $U(Z_p)\alpha U(Z_p) = \bigcup_{u \in R} U(Z_p)\alpha_u$. Then we have $I_{B,n+1}\alpha I_{U,n} = \bigcup_{u \in R} I_{B,n+1}\alpha_u$, and $\alpha_u x^{-1} = x_u^{-1}\alpha_u$ for some $u' \in R$. This $x^{-1} \mapsto x_u^{-1}$ coincides with the action of $\alpha_u$ on the flag variety $Y_1((Z/p^nZ)^g)$ if one identifies elements in $I_{B,1}$ with a flag. Here we need to use $x^{-1}$ instead of $x$, because the action of $h \in GL_g(Z_p)$ on filtrations is given by $fil \mapsto h \cdot fil = (\psi^{-1} h^{-1} \psi)fil$ as already explained. The element $x_u$ gives rise to a couple $(\alpha_u X, \text{fil}_{n+1,u} = \alpha_u^*(\text{fil}_{n+1}))$, which is uniquely determined independently of the choice of $\alpha_u$. We then define for $f \in V^B_{m,n+1}$ \( f(t(p))(X, \text{fil}_n) = \sum_{u \in R} f(\alpha_u X, \text{fil}_{n+1,u}) \). Similarly, if we start from $f \in V^U_{m,n}$, by the same process, we get $x_u \in I_{U,n}/U(Z_p)$ corresponding to $(X, \text{fil}_{n,n}, \phi_u)$, because we still have $I_{U,n}\alpha I_{U,n} = \bigcup_{u \in R} I_{U,n}\alpha_u$. We then define $f(t(p))(X, \text{fil}_n, \phi_u) = \sum_{u \in R} f(\alpha_u X, \text{fil}_{n,n}, \phi_u)$ and define the idempotent $e_{GL}$ by $e_{GL} = \lim_{n \to \infty} t(p)^n$ whenever it is well defined.

As we have seen in Section 4, $\kappa(\alpha)t(p) = t(p)$ on $H^0(S_{\infty}, \omega_\infty^\kappa)$, because on $\omega_\kappa$, we used the action of $\Delta$ coming from schematic induction.

7.1.7. Axiomatic control result. — In this subsection, we describe a simple prerequisite to have the control theorem relating false automorphic forms (sections over $S$) to true automorphic forms (sections over $M$). Later we will verify the requirement for automorphic vector bundles on Shimura varieties of PEL type.

Since in this general situation, $S_m$ is not supposed to classify anything; so, we cannot define Hecke operators acting on $H^0(S_m, \omega_1^\kappa)$ in this generality. Anyway, we suppose to have a Hecke operator $\tau(p)$ acting on $H^0(S_m, \omega_1^\kappa)$ and $H^0(M_m, \omega_1^\kappa)$ such that $\tau(p) \equiv t(p)$ on $H^0(S_1, \omega_1^\kappa)$ if $\kappa > 0$. We define $e_0 = \lim_{n \to \infty} \tau(p)^n$.

Now suppose that there exist further two projectors $e_0^G$ acting on $V^U_{m,\infty}$ and $e_0^G$ on $H^0(M_m, \omega_1^\kappa)$ for $\kappa > 0$ (depending on the reductive group $G$) such that $e_0^G e_0^G = e_0^G$, $e_0^G e_0^G = e_0^G$, and $e_0^G \equiv e_0^G$ mod $p$, that is, they are equal each other on $e_0^G H^0(S_1, \omega_1^\kappa)$. In addition to the above conditions, writing $K$ for the field of fractions of $W$, we suppose the following two conditions:

(C) \( e_G(Ef) = E(e_G f) \) for $f \in H^0(S_1, \omega_1^\kappa)$,

(F) \( \dim_K e_0^G H^0(M_n, \omega_1^\kappa \otimes \det^k(\omega)_K) \) is bounded independent of $k$.

Let $f_1, f_2, \ldots$ be a sequence of linearly independent elements in $e_G H^0(S_1, \omega_1^\kappa)$ over $W_1$. Since $H^0(S/W, \omega_1^\kappa_W) \otimes_W W_1 = H^0(S_1, \omega_1^\kappa)$ (Hyp1), we can lift $f_i$ to $f_i \in H^0(S/W, \omega_1^\kappa)_W$ so that $f_i = (f_i \mod p)$. Then for any given integer $N > 0$, we
can find a sufficiently large integer \( m \) such that \( E^m f_i \in H^0(M/W, \omega^m) \). Since multiplication by \( E \) is an isomorphism on \( S_1 \) (by definition of \( S_1 = M_1[1/p] \)), by (C) and \( e_G \equiv e_G^0 \mod p \), \( \{ (e_G^0(E^m f_i) \mod p) \}_{i=1,...,N} \) are linearly independent over \( W_1 \); so, \( \{ e_G^0(E^m f_i) \}_{i=1,...,N} \) are linearly independent over \( W \). This implies

\[
\dim_K e_G^0 H^0(M/K, \omega^m \otimes \det^t(p-1)m(\omega)/K) \geq N.
\]

If \( \text{rank}_{W_1} e_G H^0(S_1, \omega^m) = \infty \), we can take \( N \) to be arbitrarily large, which contradicts the boundedness (F) of the dimension. Thus \( \text{rank}_{W_1} e_G H^0(S_1, \omega^m) \) has to be finite, and \( \text{rank}_{W_1} e_G H^0(S_1, \omega^m \otimes \det^k(p-1)(\omega)) \) is independent of \( k \). Thus the existence of the desired projector follows from (F), (C) and (7.4).

The condition (F) can be proven in our application via group cohomology using the (generalized) Eichler-Shimura isomorphism combined with the \( p \)-adic density of \( D_{\text{cusp}} \) in \( D'_{\text{cusp}} \) (see [H95] for such boundedness for forms of \( GL(n) \), [TiU] for inner forms of \( GSp(2n) \) and [Mo] for more general groups).

The condition (C) can be proven either by \( q \)-expansion or the fact that Hasse invariant does not change after dividing an abelian variety by an étale subgroup.

8. Vertical Control for Projective Shimura Varieties

8.1. Deformation Theory of Serre and Tate. — Let \( W \) be a complete discrete valuation ring of mixed characteristic with residue field \( \mathbb{F} \) of characteristic \( p \). We suppose that \( \mathbb{F} \) is an algebraic closure over \( \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \). In this section, we describe deformation theory of abelian schemes over local \( W_m \)-algebras for \( W_m = W/p^m W \). We follow principally Katz’s exposition [K].

8.1.1. A Theorem of Drinfeld. — Let \( R \) be a local \( W_m \)-algebra. Let \( G : R-LR \rightarrow AB \) be a covariant functor into the category \( AB \) of abelian groups. When \( m = \infty \), the category \( R-LR \) (resp. \( W_\infty-LR \)) is made of \( p \)-adically complete local \( R \)-algebras \( A = \varprojlim_n A/p^n A \) and morphisms are supposed to be \( p \)-adically continuous. For simplicity, we always assume that rings we consider are noetherian. Thus if we regard \( G \) as a functor from the category of affine \( R \)-schemes (or formal schemes), it is contravariant. Suppose that, for any faithfully flat extension of finite type \( A \rightarrow B \) of \( R \)-algebras,

- (1) The group \( G(A) \) injects into \( G(B) \), that is, \( G(A) \hookrightarrow G(B) \);

- (2) Let \( B' = B \otimes_A B \) and \( B'' = B \otimes_A B \otimes_A B \). Write \( \iota_i : B \hookrightarrow B' \) (\( i = 1, 2 \)) two inclusions (that is, \( \iota_1(r) = r \otimes 1 \)) and \( \iota_{ij} : B' \hookrightarrow B'' \) be three inclusions (i.e. \( \iota_{12}(r \otimes s) = r \otimes s \otimes 1 \)). If \( x \in G(B) \) satisfies \( y = G(\iota_1)(x) = G(\iota_2)(x) \) and \( G(\iota_{12})(y) = G(\iota_{23})(y) = G(\iota_{13})(y) \), then \( x \) is in the image of \( G(A) \).

Such a \( G \) is called an abelian sheaf on \( R-LR \) with the \( fppf \)-topology (or simply abelian \( fppf \)-sheaf). If \( X/R \) is an abelian scheme or a torus (a multiplicative group, like \( \mathbb{G}_m \)), then \( G(A) = X(A) = \text{Hom}_S(\text{Spec}(A), X) \) (\( S = \text{Spec}(R) \) or \( \text{Spf}(R) \)) is an \( fppf \)-sheaf.
We call $G$ \emph{$p$-divisible} if for any $x \in G(A)$, there exists a finite faithfully flat extension $B$ of $A$ and a point $y \in G(B)$ such that $x = py$. If $G$ comes from an abelian scheme $X$, it is $p$-divisible (e.g. [GME] Corollary 4.1.18). This also shows that $X[p^\infty] = \bigcup_n X[p^n]$ for $X[p^n] = \ker(p^n : X \to X)$ is $p$-divisible.

Let $R$ be a local $W_m$-algebra and $I$ be an ideal of $R$ such that $I^{p+1} = 0$ and $NI = 0$ for a power $N$ of $p$. We define a new functor $G_I$ and $\hat{G}$ by

$$G_I(A) = \ker(G(A) \to G(A/I)) \quad \text{and} \quad \hat{G}(A) = \ker(G(A) \to G(A/\mathfrak{m}_A)),$$

where $\mathfrak{m}_A$ is the maximal ideal of $A$. When $\hat{G}(A) = \text{Hom}_{R/LR}(\mathcal{R}, A)(= G(A))$ for $\mathcal{R} = R[[T_1, \ldots, T_n]]$ (that is $G_0 = \text{Spf}(\mathcal{R})/R$) and the identity element $0$ corresponding to the ideal $(T_1, \ldots, T_n)$, we call $G$ a formal group. If $G$ is formal, $G_I(A) = \{(t_1, \ldots, t_n) \in I\}$ by $\text{Hom}_{R/LR}(\mathcal{R}, A) \ni \phi \mapsto (\phi(T_1), \ldots, \phi(T_n))$.

Suppose that $G_I$ is formal. Then multiplication by $[N]$ induces a continuous algebra homomorphism $[N] : R \to R$. Then on the tangent space at the origin: $t_0 = (T_1, \ldots, T_n)(T_1, \ldots, T_n)^2$, the addition induced by the group law of $G$ coincides with the addition of the tangent vectors (cf. [ABV] Section 11). Thus $[N](T_i) \equiv NT_i \mod (T_1, \ldots, T_n)^2$, and $[N](G_I(A)) = G_{I+1}(A)$ because $NI = 0$. Similarly, we have inductively, $[N](G_{I^0}(A)) = G_{I^{0+1}}(A)$, and $[N^r]G_I = G_0 = \{0\}$. We get

$$G_I \subset \ker([N^r] : G \to G) \quad \text{if} \quad G \text{ is formal.}$$

\textbf{Theorem 8.1 (Drinfeld).} — Let $G$ and $H$ be abelian $\text{fpf}$-sheaf over $R/LR$ and $I$ be as above. Let $G_0$ and $H_0$ be the restriction of $G$ and $H$ to $R/I-LR$. Suppose

(i) $G$ is $p$-divisible;

(ii) $\hat{H}$ is formal;

(iii) $H(A) \to H(A/J)$ is surjective for any nilpotent ideal ($H$ is formally smooth).

Then

(1) $\text{Hom}_{R/GP}(G, H)$ and $\text{Hom}_{R/I/GP}(G_0, H_0)$ is $p$-torsion-free, where “$\text{Hom}_{X/GP}$” stands for the homomorphisms of abelian $\text{fpf}$-sheaves over $X-LR$;

(2) The natural map, so-called

\[ \text{"reduction mod } I": \text{Hom}_{R/GP}(G, H) \to \text{Hom}_{R/I/GP}(G_0, H_0) \]

is injective;

(3) For any $f_0 \in \text{Hom}_{R/I/GP}(G_0, H_0)$, there exists a unique $\Phi \in \text{Hom}_{R/GP}(G, H)$ such that $\Phi \mod I = N^r f_0$. We write $\tilde{N}^r f$ for $\Phi$ even if $f$ exists only in $\text{Hom}_{R/GP}(G, H) \otimes \mathbb{Z}/\mathbb{Q}$;

(4) In order that $f \in \text{Hom}_{R/GP}(G, H)$, it is necessary and sufficient that $\tilde{N}^r f$ kills $G[N^r]$.

\textbf{Proof.} — The first assertion follows from $p$-divisibility, because if $pf(x) = 0$ for all $x$, taking $y$ with $py = x$, we find $f(x) = pf(y) = 0$ and hence $f = 0$. 

\textbf{ASTÉRISQUE 298}
We have an exact sequence: $0 \to H_1 \to H \to H_0 \to 0$; so, we have another exact sequence:

$$0 \to \text{Hom}(G, H_1) \to \text{Hom}(G, H) \xrightarrow{\text{mod} I} \text{Hom}(G, H_0) = \text{Hom}(G_0, H_0),$$

which tells us the injectivity since $H_1$ is killed by $N^\nu$ and $\text{Hom}(G, H)$ is $p$-torsion-free.

To show (3), take $f_0 \in \text{Hom}(G_0, H_0)$. By surjectivity of $H(A) \to H_0(A/I)$, we can lift $f_0(x \mod I)$ to $y \in H(A)$. The class $y \mod \text{Ker}(H \to H_0)$ is uniquely determined. Since $\text{Ker}(H \to H_0)$ is killed by $N^\nu$, for any $x \in G(A)$, therefore $N^\nu y$ is uniquely determined; so, $x \mapsto N^\nu y$ induces functorial map: $\tilde{N}^\nu f : G(A) \to H(A)$. This shows (3).

The assertion (4) is then obvious from $p$-divisibility of $G$. The uniqueness of $f$ follows from the $p$-torsion-freeness of $\text{Hom}(G, H)$.

8.1.2. A Theorem of Serre-Tate. — Let $\mathcal{A}_R$ be the category of abelian schemes defined over $R$. We consider the category $\text{Def}(R, R/I)$ of triples $(X_0, D, \varepsilon)$, where $X_0$ is an abelian scheme over $R/I$. $D = \bigcup_n D^{(n)}$ with finite flat group scheme $D^{(n)}$ over $R$ with inclusion $D^{(n)} \hookrightarrow D^{(n+1)}$, which is $p$-divisible, and $\varepsilon : D_0 \cong X_0[p^\infty]$. We have a natural functor $\mathcal{A}_R \to \text{Def}(R, R/I)$ given by $X \mapsto (X_0 = X \mod I, X[p^\infty], \text{id})$.

**Theorem 8.2 (Serre-Tate).** — The above functor: $\mathcal{A}_R \to \text{Def}(R, R/I)$ is a canonical equivalence of categories.

**Proof.** — By the Drinfeld theorem applied to $X[p^\infty]$ and $X$ (both abelian fppf-sheaf), the functor is fully faithful (see [K] for details).

For a given triple, $(X_0, D, \varepsilon)_R$, we need to create $X/R$ which gives rise to $(X_0, D, \varepsilon)_R$. It is known that we can lift $X_0$ to an abelian scheme $Y$ over $R$. This follows from the deformation theory of Grothendieck ([GIT] Section 6.3). When $R/I$ is a finite field, by a theorem of Tate, $X_0$ has complex multiplication. By the theory of abelian varieties with complex multiplication, $X_0$ can be lifted to a unique abelian scheme $Y$ over $R$ with complex multiplication (because the isomorphism classes of such abelian varieties of CM type corresponds bijectively to the lattice in a CM field). Thus we have an isomorphism $\alpha_0^{(p)} : Y_0[p^\infty] \to X_0[p^\infty]$. Then we have a unique lifting (by the Drinfeld theorem) $f = \tilde{N}_N^\nu \alpha_0^{(p)} : Y[p^\infty] \to D$ of $N^\nu \alpha_0^{(p)}$. Since the special fiber is an isogeny having inverse $(\alpha_0^{(p)})^{-1}$, $f$ is an isogeny, whose (quasi) inverse is the lift of $N^\nu(\alpha_0^{(p)})^{-1}$. Thus $\text{Ker}(f)$ is a finite flat group subscheme of $Y$. The geometric quotient of $Y$ by a finite flat group subscheme exists (see [ABV] Section 12) and is an abelian scheme over $R$. Then dividing $Y$ by $\text{Ker}(f)$, we get the desired $X/R \in \mathcal{A}_R$.

8.1.3. Deformation of an Ordinary Abelian Variety. — Let $A$ be a ring of characteristic $p$ and $(X, \omega)$ be a pair of an abelian variety over $S = \text{Spec}(A)$ of relative dimension $g$ and a base $\omega$ of $H^0(X, \Omega_{X/A})$ over $A$. We have the absolute Frobenius endomorphism $F_{\text{abs}} : X/A \to X/A$. Let $T_{X/S}$ be the relative tangent bundle; so,
$H^0(X,T_{X/S})$ is spanned by the dual base $\eta = \eta(\omega)$. For each derivation $D$ of $\mathcal{O}_{X,0}$, by the Leibnitz formula, we have

$$D^p(xy) = \sum_{j=0}^p \binom{p}{j} D^{p-j} x D^j y = x D^p y + y D^p x.$$ 

Thus $D^p$ is again a derivation. The association: $D \mapsto D^p$ induces an $F_{\text{abs}}$-linear endomorphism $F^*$ of $T_{X/S}$. Then we define $H(X,\omega) \in A$ by $F^* \wedge^g \eta = H(X,\omega) \wedge^g \eta$. Since $\eta(\lambda\omega) = \lambda^{-1} \eta(\omega)$ for $\lambda \in GL_g(A)$, we see

$$H(X,\lambda\omega) \wedge^g \eta(\lambda\omega) = F^* \wedge^g \eta(\lambda\omega) = F^* (\text{det}(\lambda)^{-1} \wedge^g \eta(\omega))$$

$$= \text{det}(\lambda)^{-p} F^* \wedge^g \eta(\omega) = \text{det}(\lambda)^{-p} H(X,\omega) \wedge^g \eta(\omega)$$

$$= \text{det}(\lambda)^{-p} H(X,\omega) \text{det}(\lambda) \wedge^g \eta(\lambda\omega) = \text{det}(\lambda)^{-1-p} H(X,\omega) \wedge^g \eta(\lambda\omega).$$

Thus we get

$$H(X,\lambda\omega) = \text{det}(\lambda)^{1-p} H(X,\omega).$$

We call $X$ ordinary if $X[p] \cong (\mathbb{Z}/p\mathbb{Z})^g \times \mu_p^g$ étale locally. In the same manner as in the elliptic curve case, we know

$$H(X,\omega) = 0 \iff X \text{ is not ordinary.}$$

Let $F$ be an algebraic closure of $\mathbb{F}_p$. Let $R$ be a pro-artinian local ring with residue field $F$. Write $CL_{/R}$ be the category of complete local $R$-algebras with residue field $F$. We fix an ordinary abelian variety $X_0/R$. Write $\hat{X}_{/R}$ for the dual abelian scheme of an abelian scheme $X_{/R}$. We write $TX[p\infty]^\text{et}$ for the Tate module of the maximal étale quotient of $X[p\infty]$. We consider the following deformation functor: $\hat{P} : CL_{/R} \to \text{SETS}$ given by

$$\hat{P}_{X_0}(A) = [(X_{/A},\iota_X) \mid X \text{ is an abelian scheme over } A \text{ and } \iota_X : X \otimes_A F \cong X_0].$$

Here $f : (X_{/A},\iota_X) \cong (X_{/A}',\iota_{X'})$ if $f : X \to X'$ is an isomorphism of abelian schemes with the following commutative diagram:

$$\begin{array}{ccc}
X & \xrightarrow{f_0} & X' \\
\downarrow{\iota_X} & & \downarrow{\iota_{X'}} \\
X_0 & = & X_0.
\end{array}$$

**Theorem 8.3 (Serre-Tate).** — We have

1. A canonical isomorphism $\hat{P}(A) \cong \text{Hom}_{\mathbb{Z}_p}(TX_0[p\infty]^\text{et} \times T\hat{X}_0[p\infty]^\text{et}, \hat{G}_m(A))$ taking $(X_{/A},\iota_X)$ to $q_{X_0/A}(\iota_X)$.

2. The functor $\hat{P}$ is represented by the formal scheme

$$\text{Hom}_{\mathbb{Z}_p}(TX_0[p\infty]^\text{et} \times T\hat{X}_0[p\infty]^\text{et}, \hat{G}_m) \cong \hat{G}_m^g.$$
(3) \( q_{X/A}(x, y) = q_{X/A}(y, x) \) under the canonical identification: \( \hat{X} = X \).

(4) Let \( f_0 : X_0/Y \to Y_0/Y \) be a homomorphism of two ordinary abelian varieties with the dual map: \( f_0 : \hat{Y}_0 \to \hat{X}_0 \). Then \( f_0 \) is induced by a homomorphism \( f : X/A \to Y/A \) for \( X \in \hat{P}_{X_0}(A) \) and \( Y \in \hat{P}_{Y_0}(A) \) if and only if \( q_{X/A}(x, f_0(y)) = q_{Y/A}(f_0(x), y) \).

Proof. — We are going to give a sketch of the construction of \( q_{X/A} \).

We prepare some facts. Let \( f : X \to Y \) be an isogeny; so, \( \text{Ker}(f) \) is a finite flat group scheme over \( S \). Pick \( x \in \text{Ker}(f) \), and let \( \mathcal{L} \in \text{Ker}(\hat{f}) \subset \hat{Y} \) be the line bundle on \( Y \) with \( \mathcal{L} = \mathcal{O}_S \) (\( S = \text{Spec}(A) \)) for an artinian \( R \)-algebra \( A \). Thus \( f^* \mathcal{L} = \mathcal{O}_X \).

Cover \( Y \) by affine subsets \( U_i = \text{Spec}(A_i) \) so that \( \mathcal{L}|_{U_i} = \phi_i^{-1} \mathcal{O}_{U_i} \). Since \( \mathcal{O}_Y \mathcal{L} = \mathcal{O}_S \), we may assume that \( \phi_i/\phi_j \circ \mathcal{O}_Y = 1 \). Since \( f : X \to Y \) is finite, it is affine. Write \( V_i = f^{-1}(U_i) = \text{Spec}(B_i) \). Then \( f^* \mathcal{L}|_{V_i} = \phi_i^{-1} \mathcal{O}_{V_i} \) with \( \phi_i = \phi_i \circ f \), and we have, regarding \( x : S \to \text{Ker}(f) \),

\[
\frac{\phi_i \circ x}{\phi_j \circ x} = \frac{\phi_i \circ f \circ x}{\phi_j \circ f \circ x} = \frac{\phi_i \circ \mathcal{O}_Y}{\phi_j \circ \mathcal{O}_Y} = 1.
\]

Thus \( \phi_i \circ x \) glue into a morphism \( \{x, \mathcal{L}\} : S \to \mathcal{G}_m \), and in this way, we get a pairing

\[
e_f : \text{Ker}(f) \times \text{Ker}(\hat{f}) \to \mathcal{G}_m.
\]

Since \( X \) is a \( \text{Ker}(f) \)-torsor over \( Y \), we have \( X \times_Y X \cong \text{Ker}(f) \times_Y Y \). Thus for any homomorphism \( \zeta : \text{Ker}(f) \to \mathcal{G}_m \), we can find a morphism \( \phi : \text{Ker}(f) \times_Y Y \to \mathbb{P}^1 \) such that \( \phi(y + t) = \zeta(t)\phi(y) \) for \( t \in \text{Ker}(f) \). This function \( \phi \) gives rise to a divisor \( D \) on \( Y_X = Y \times_X X \). By definition \( f^* \mathcal{L}(D) = \mathcal{O}_{X_X} \), and \( e_f(x, \mathcal{L}(D)) = \zeta(x) \). Thus, over \( X \), \( e_{f/X} : \text{Ker}(f)/X \times \text{Ker}(\hat{f})/X \to \mathcal{G}_m \) is a perfect pairing. Since \( X \to S \) is faithfully flat, we find that the original \( e_f \) is perfect.

We apply the above argument to \( f = [p^n] : X \to X \), write the pairing as \( e_n \) and verify the following points (e.g. [GME] 4.1.5):

- (P1) \( e_n(\alpha(x, y) = e_n(x, \alpha(y)) \) for \( \alpha \in \text{End}(X/A) \);
- (P2) Write \( X_0[p^n] = \mu_{p^n}^m \subset X_0[p^n] \). Then \( e_n \) induces an isomorphism of group schemes: \( X_0[p^n] \cong \text{Hom}(\hat{X}_0[p^n]^{\text{ét}}, \mathcal{G}_m) \);
- (P3) Taking limit of the above isomorphisms with respect to \( n \), we find

\[
X^\circ \cong \text{Hom}(T \hat{X}[p^n]^{\text{ét}}, \mathcal{G}_m) \cong \text{Hom}(T \hat{X}_0[p^n]^{\text{ét}}, \mathcal{G}_m)
\]

as formal groups. We denote the induced pairing by

\[
E_X : X^\circ \times T \hat{X}_0[p^n]^{\text{ét}} \to \mathcal{G}_m.
\]

In particular \( X^\circ = \mathcal{G}_m^q \).

The structure of the \( p \)-divisible group \( X[p^\infty] \) is uniquely determined by the extension class of:

\[
0 \to \text{Hom}(T \hat{X}_0[p^n]^{\text{ét}}, \mathcal{G}_m)[p^n] \to X[p^n] \xrightarrow{\pi} X_0[p^n]^{\text{ét}} \to 0
\]
for \( n = 1, 2, \ldots, \infty \). Take \( x = \lim_{n \to \infty} x_n \in TX_0[p^n]^\text{et} \) for \( x_n \in X[p^n]^\text{et} \). Lift \( x_n \) to \( v_n \in X[p^n] \) so that \( \pi(v_n) = x_n \). Then \( q_n(x) = \hat{p}^nv_n \in \text{Hom}(TX_0[p^n]^\text{et}, \hat{\Gamma}_m)[p^n] \). Take the limit of \( q_n \) to get \( q(x) \in \text{Hom}(TX_0[p^n]^\text{et}, \hat{\Gamma}_m(A)) \). This \( q(x) \) completely determines the extension class of (8.2) so the deformation \( X_A \) because it is determined by \((X_0, X[p^n])\) by the Serre-Tate theorem in the previous subsection. Then we define \( q_{X/A}(x, y) = q(x)(y) \).

It is known that for any given \( q(x, y) \) as above an extension (8.2) exists by the theory of Barsotti-Tate groups studied by Messing (see [CBT] Appendix). This shows the assertions (1) and (2). All other assertions follows from (P1-3) easily. \( \square \)

8.1.4. Symplectic Case. — We now fix a polarization \( \lambda_0 : X_0 \to \hat{X}_0 \) of degree prime to \( p \). We consider the functor

\[
\widehat{P}_{X_0, \lambda_0}(A) = \left\{ (X_A, \iota_X, \lambda) \mid (X, \iota_X) \in \widehat{P}_{X_0}(A) \text{ and } \lambda \text{ induces } \lambda_0 \right\}.
\]

Here we call \( f : (X, \lambda_X, \iota_X) \to (Y, \lambda_Y, \iota_Y) \) an isomorphism if \( f : (X, \iota_X) \cong (Y, \iota_Y) \) and \( \hat{f} \circ \lambda_Y \circ f = \lambda_X \). Note that by Drinfeld theorem, \( \text{End}(X/A) \) is torsion-free, and hence, \( \text{End}(X/A) \to \text{End}^\text{et}(X_A) = \text{End}(X_A)[p] \). We write \( \alpha^* = \lambda_0^{-1} \circ \hat{\alpha} \circ \lambda_0 \) for \( \alpha \in \text{End}(X_0) \otimes \mathbb{Z}/p \mathbb{Q} \). Since \( \text{End}(X/A) \subset \text{End}(X_0) \) again by Drinfeld’s theorem, the involution keeps \( \text{End}^\text{et}(X/A) \) stable (because on \( \text{End}^\text{et}(X_A) \), it is given by \( \alpha^* = \lambda^{-1} \circ \hat{\alpha} \circ \lambda \)). The involution \( \alpha \mapsto \alpha^* \) is known to be positive (see [ABV] Section 21). The polarization \( \lambda_0 \) induces an isomorphism \( \lambda_0 : X[p^n]^\text{et} \cong \hat{X}[p^n]^\text{et} \). We identify \( TX_0[p^n]^\text{et} \) and \( T\hat{X}_0[p^n]^\text{et} \) by \( \lambda_0 \). Then the involution \( \alpha \mapsto \hat{\alpha} \) is replaced by the positive involution “*”. Then it is clear from the previous theorem that

\[
\widehat{P}_{X_0, \lambda_0}(A) \cong \text{Hom}_{\mathbb{Z}/p}(\text{Sym}^2 TX_0[p^n]^\text{et}, \hat{\Gamma}_m(A)) \cong \hat{\Gamma}_m^{g(g+1)/2}(A).
\]

8.2. Proof of the VCT in the Co-compact Case. — We first describe the deformation space in the unitary case, and then we prove the VCT for such groups.

8.2.1. Unitary Case. — We fix a division algebra \( B \) with positive involution “*”. The center of \( B \) is either a CM field \( F \) (inducing complex conjugation on \( F \)) or a totally real field on which * is trivial. We fix a \( B \)-module \( V \) with *-hermitian alternating form \( \langle \, , \, \rangle \) satisfying conditions (L1-2) in Section 6. Out of these data, we define the group

\[
G_1(A) = \{ x \in C \otimes \mathbb{Q} A \mid xx^* = 1 \},
\]

where \( C = \text{End}_B(V) \) and \( \langle xv, w \rangle = \langle v, x^*w \rangle \). For simplicity, we suppose that \( F = \mathbb{Q}[\sqrt{-D}] \) for a positive integer \( D \) (we suppose that \(-D \) is the discriminant of \( F/\mathbb{Q} \)). In particular, we have \( \text{Sh}_{K/F} \cong \text{Sh}_{K/F}^{(p)} \times s_0E \) for \( K \) maximal at \( p \) ([PAF] Theorem 7.5). The group \( G_1 \) is an inner form of a unitary group of signature \((m, n)\). Let \( \varepsilon = \text{diag}[1, 0, \ldots, 0] \in O_{B,p} \). By the condition (det), the representation of \( F \) on \( \varepsilon(\text{Lie}(X)) \) for \((X, \lambda, i, \eta^{(p)}) \in \mathcal{P}(A) \) \((A \in \text{W-CL})\) is \( m \text{id} + nc \) for \( \text{id} : O_F \to W \) and non-trivial automorphism \( c \) of \( F \). We fix an \( O_B \)-lattice \( L \) of \( V \) such that \( \langle \, , \, \rangle \) induces
a self duality of $L_p = L \otimes \mathbb{Z} \mathbb{Z}_p$. We suppose that $p = \mathbb{p}\mathbb{p}$ in $F$; so, $O_{B,p} = O_{B} \otimes \mathbb{Z} \mathbb{Z}_p \cong M_r(O_{F,p}) \oplus M_r(O_{F,p})$. Supposing that $X_{/\mathbb{F}}$ is ordinary, we have $\text{Lie}(X_{/\mathbb{F}}) \cong \text{Lie}(X[p])$, where $X[p]$ is the connected component of $X$. Thus we may assume that $T_p X[p^\infty] \otimes \mathbb{Z}_p \cong M_{r \times n}(O_{F,p})$ and $T_p X[p^\infty] \otimes \mathbb{Z}_p \cong M_{r \times n}(O_{F,p})$.

For an artinian local $W$-algebra $A$ with residue field $\mathbb{F} = W/\mathfrak{m}W = \mathbb{F}_p$ and $(X, \lambda, i, \eta(p)) \in \mathcal{P}(A)$, we consider $D_X = \varepsilon(X[p^\infty])$. Since $X[p^\infty] \cong D_X$ as Barsotti-Tate $p$-divisible groups, the abelian scheme $X$ as a deformation of $X_0 = X \otimes_A \mathbb{F}$ is completely determined by $D_X$.

Suppose that $X_0$ is ordinary. We write the $O_{F,p}$-component of $T_p D_X^\text{et} = T_p D_{X_0}^\text{et}$ as $T_p D_{X_0}^\text{et}$. Then the symmetric pairing

$$q_{X/A}(\cdot, \cdot) : T_p X_0[p^\infty] \otimes T_p X_0[p^\infty] \longrightarrow \hat{\mathbb{G}}_m(A)$$

induces a homomorphism:

$$q_{X/A} : T_p D_{X_0}^{\text{et}} \otimes_{\mathbb{Z}_p} T_p D_{X_0}^{\text{et}} \longrightarrow \hat{\mathbb{G}}_m(A),$$

because the pairing is $c$-hermitian (that is, the involution $\ast$ induces complex conjugation $c$). Since the level $N$-structure outside $p$ lifts uniquely to deformations, we can ignore the level structure while we study deformations of $(X_0, \lambda_0, i_0, \eta_0(p))_{/\mathbb{F}}$. So we consider the functor

$$\hat{\mathcal{P}}_{X_0, \lambda_0, i_0}(A) = \{(X/A, i_X, \lambda, i) \mid (X, i_X, \lambda) \in \hat{\mathcal{P}}_{X_0, \lambda_0}(A) \text{ and } i \text{ induces } i_0\}.$$  

Then the above argument combined with the theorem of Serre-Tate (Theorem 8.3) shows

$$\hat{\mathcal{P}}_{X_0, \lambda_0}(A) \cong \text{Hom}_{Z_p}(T_p D_{X_0}^{\text{et}} \otimes_{O_{F,p}} T_p D_{X_0}^{\text{et}}, \hat{\mathbb{G}}_m(A)) = \hat{S}(A),$$

because the symmetric ($c$-hermitian) form on

$$(T_p D_{X_0}^{\text{et}} \times T_p D_{X_0}^{\text{et}}) \otimes (T_p D_{X_0}^{\text{et}} \times T_p D_{X_0}^{\text{et}})$$

is determined by its restriction on $(T_p D_{X_0}^{\text{et}} \times \{0\}) \times (\{0\} \times T_p D_{X_0}^{\text{et}})$.

8.2.2. Hecke Operators on Deformation Space. — Let $O_C = \{x \in C \mid x L \subset L\}$. We write $G_1(\hat{\mathbb{Z}})$ for $G \cap G_1(\mathbb{A}(\infty))$, where $O_C = O_C \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$. We fix an open compact subgroup $K \subset G_1(\hat{\mathbb{Z}})$ such that $K = K_p \times K(p)$ with $K_p \cong GL_{m+n}(\mathbb{Z}_p)$ via the projection to $p$-factor.

We consider $p$-ordinary test objects $\underline{X} = (X, \lambda, i, \eta(p))_{/A}$ over a local artinian $W$-algebra $A$. Since the pairing $q_{X/A} \in \hat{S}(A)$ is actually determined by its restriction to $q_{X/A} : T_p D_{X_0}^{\text{et}} \times T_p D_{X_0}^{\text{et}}$, we only look into this restriction. We study the $O_{F,p}$-linear endomorphism algebra $\text{End}_{BT}(\varepsilon X[p^\infty]_{/A})$ of the Barsotti-Tate group $\varepsilon X[p^\infty]_{/A}$. Write each endomorphism $\alpha$ as $(a_{\alpha}, b_{\alpha})$ with

$$a_{\alpha} \in \text{End}_{BT}(\varepsilon X[p^\infty]^{c}), b_{\alpha} \in \text{Hom}_{BT}(\varepsilon X[p^\infty]^{c}, \varepsilon X[p^\infty]^{c})$$

and

$$c_{\alpha} \in \text{End}_{O_{F,p}}(\varepsilon T_p^{\text{et}}(X)).$$
If $A$ is an algebraically closed field of characteristic $p$, the étale-connected exact sequence $\varepsilon X_0[p^\infty]^c \hookrightarrow \varepsilon X_0[p^\infty] \rightarrow \varepsilon X_0[p^\infty]^c$ is (non-canonically) split. In any case, $\alpha$ acts on $T_{X,\beta} = T_p D_X^\infty \oplus T_p D_X^\infty$ diagonally via $a_\alpha$ and $d_\alpha$. We regard $T = \mathbb{G}_m^n$ as a maximal split torus of $GL_m(O_{F,p}) \times GL_n(O_{F,p})$, which is the automorphism group of $\varepsilon X[p^\infty]^{c\et} \times \varepsilon X[(p^c)^{c\et}]$.

Let $X_{/\text{Sh}_K}$ be the universal abelian scheme. We write $P_t = \varepsilon X_{/S}[p^\infty]^{c\et}[p^t]$ and $P'_t = \varepsilon X_{/S}[p^\infty]^{c\et}[p^t]$ and apply the theory developed in Section 7 to each piece $P_t$ and $P'_t$, so we obtain the theory of false automorphic forms for $GL_n(O_{F,p}) \times GL_m(O_{F,p})$ $(O_{F,p} \cong O_{F,\overline{F}} \cong \mathbb{Z})$. Since $p$ is unramified in $F$, $O_F \otimes_{\mathbb{Z}_p} W \cong W^I$ for the set of embedding $I = \{\sigma = \text{id}, c\}$ of $F$ into $\overline{F}$. Then we consider filtrations $\text{fil}_\sigma$ and $\text{fil}_c$ of $\varepsilon X(p^t)^{c\et}$ and $\varepsilon X(p^s)^{c\et}$, and consider the following test objects: $\{X, \lambda, i, \tau(p^c), \text{fil}_\sigma, \text{fil}_c\} / A$. Let $M = \text{Sh}_K$ and $S = M[\frac{1}{p}]$, where $E$ is a lift of the Hasse invariant. We write $T_{t,s}/S_t$ for the étale covering over $S_t = S \otimes_W W_t$ ($W_t = W/pW$) classifying the above test objects. Similarly, $T_{t,t,s}$ classifies

$$(X, \lambda, i, \tau(p^c), (O_{F,p}^x)^n) \times (O_{F,\overline{F}}^x)^m \cong \varepsilon X[p^{t+c}] \times \varepsilon X[p^{s+c}].$$

The covering $T_{t,t,s}/S_t$ is an étale Galois covering with Galois group isomorphic to $GL_m(O_{F,\overline{F}}) \times GL_n(O_{F,p})$.

We had an action of the expanding semi-group on filtrations of $P_t$ ($0 < t \in \mathbb{Z}$) for an étale sheaf $P_t \cong (\mathbb{Z}/p^{t+c})^n$. If we have a $p$-isogeny $\beta : P_\infty = \bigcup_t P_t \rightarrow P_\infty$ preserving a filtration of $P_\infty$, we may assume that the matrix form of $\beta$ is given by $\beta_j = \left(\begin{array}{cc}1 & -j \\ 0 & p^t\end{array}\right)$ with respect to a base compatible with the filtration. Then the action of $\beta_j$ is to give a new filtration on $P_\infty$.

Since we cannot separate $P_t$ and $P'_t$ (which is sitting in the single universal abelian scheme $X$), we define $\alpha_{m+j} = \text{diag}[\beta_j, \text{fil}_c]$ and consider an isogeny of type $\alpha_{m+j}$. We can thus interpret the operator action in terms of the quadruple: $(X, \lambda, i, \tau(p), \text{fil}_\sigma, \text{fil}_c)$ as follows: Take an isogeny $\alpha : X \rightarrow X_\alpha$ of type $\alpha_{m+j}$ as above (inducing $\beta$ on $P_\infty$ and multiplication by $p$ on $P'_\infty$). Then we get a new filtration $\beta(\text{fil}_\sigma)$. The $p$-isogeny is insensitive to the level $N$-structure, and $X_\alpha$ has an induced polarization, still written as $\lambda$. Then we have

$$f|t_j(p) \times t_m(\overline{p})|(X, \lambda, i, \tau(p), \text{fil}_\sigma, \text{fil}_c) = \sum \tilde{\rho}_n(\beta^{-1})f(X, \lambda, i, \tau(p), \beta(\text{fil}_\sigma), \text{fil}_c).$$

Thus we have a $GL(n) \times GL(m)$-Hecke operator $t_j(p) \times t_m(\overline{p})$ acting on the coherent cohomology $H^0(S_{t}, \omega^c)$. This is actually an over-simplified version. The exact sequence:

$$0 \rightarrow \text{Hom}(P'_t, \mathbb{G}_m) \rightarrow X[p^t] \rightarrow P_t \rightarrow 0$$

may not split over $S_t$; so, the isogeny $\alpha$ can be defined only over a finite flat extension $S_{t'}^{\text{finite}}$ of $S_t$ (which is radicial over an étale extension of $S_t$). In other
words, if we replace the term: 

$$f(X_n, \lambda, i, \overline{\eta}^{(p)}, \beta(\text{fil}_c), \text{fil}_c)$$

in (8.3) by the trace

$$\text{Tr}_{\mathcal{O}_{S_{n+m+1}}/\mathcal{O}_{S_n}}(f(X_{\beta}, \lambda, i, \overline{\eta}^{(p)}, \beta(\text{fil}_c), \text{fil}_c))$$

we can relate $t_j$ to a global Hecke operator $U_j(p)$ which is divisible by the degree of $S^n_{X, \text{fil}}$ over the maximal étale cover of $S_n$ under $S^n_{X, \text{fil}}$. The operator $t_j$ is not well defined on coherent cohomology, although it is well defined on $\ell$-adic étale cohomology (because $\ell$-adic étale cohomology is insensitive to radiciel base-change).

For the moment, we pretend that the over-simplified version (8.3) is valid (and we later justify our argument). Thus for a while, our argument is just heuristic.

Since $G_1(\mathbb{Q}_p) \cong GL_{m+n}(\mathbb{Q}_p)$, we can embed $GL(n) \times GL(m)$ into $G_1(\mathbb{Q}_p)$ by $(x, y) \mapsto \text{diag}[x, y]$. This implies the $p$-isogeny whose kernel sits in $X[p^\infty]^{\text{ét}}$ (resp. $X[p^\infty]^{\text{ét}}$) corresponds to $x$ (resp. $y$). Write $P = P_{m,n}$ for the upper triangular parabolic subgroup of $G_1$ whose Levi subgroup is given by the image of $GL_n \times GL_m$. Let $U_{m,n}$ be the unipotent radical of $P_{m,n}$.

Write $\alpha_j = \alpha_j(p) \in G(\mathbb{Q}_p)$ whose projection to $C_p = C \otimes_F \mathbb{F}_p$ is given by $\text{diag}[1_{m+n-j}, p \cdot 1_j]$ and $\nu(\alpha_j) = p$. We then have Hecke operators

$$U_j(p) = U(\alpha_j(p)) = U_{g} \alpha_j U_{g},$$

where $U_{g}$ is the upper unipotent subgroup of $G_1(\mathbb{Z}_p)$. Since we identify $G_1(\mathbb{Q}_p)$ with $GL_{m+n}(\mathbb{Q}_p)$ by projecting down $C \otimes \mathbb{Q}_p$ to the first component $C_p$, as a double coset, we see (symbolically)

$$U_j(p) = \left\{ \begin{array}{ll}
\bigcup_{u \in U_{m,n}/\alpha_j^{-1}U_{m,n} \alpha_j} U_{g}(t_{j-n}(p) \times t_m(\mathcal{F}))u & \text{if } j > n, \\
\bigcup_{u \in U_{m,n}/\alpha_j^{-1}U_{m,n} \alpha_j} U_{g}(t_0(p) \times t_j(\mathcal{F}))u & \text{if } j \leq n,
\end{array} \right.$$

where we mean, for example, by $(t_{j-n}(p) \times t_m(\mathcal{F}))$, the double coset:

$$\tilde{U}(\text{diag}[1_{2n-j}, p \cdot 1_{j-n}] \times p \cdot 1_m) \tilde{U}$$

in $GL_n(F_p) \times GL_m(F_p)$ for the upper triangular unipotent subgroup $\tilde{U}$. This shows that the Hecke operator $U_j(p)$ induces

$$[U_{m,n} : \alpha_j^{-1}U_{m,n} \alpha_j](t_{j-n}(p) \times t_m(\mathcal{F})) \text{ or } [U_{m,n} : \alpha_j^{-1}U_{m,n} \alpha_j](t_0(p) \times t_m(\mathcal{F}))$$

according as $j > n$ or not. By computation, we get the following heuristic multiplicity formula:

$$[U_{m,n} : \alpha_j(p)^{-1}U_{m,n} \alpha_j(p)] = \mu_{m,n}(\alpha_j) = \left\{ \begin{array}{ll}
|p|^{j-n(m+n-j)} & \text{if } j > n, \\
|p|^{m-j} & \text{if } j \leq n.
\end{array} \right.$$

This formula suggests us that $U_j(p)$ is divisible by $\mu_{m,n}(\alpha_j)$, which we will justify later.

Since the universal deformation space of $(X, \lambda, i, \overline{\eta}^{(p)}, \text{fil}_p, \text{fil}_p)$/$\mathcal{I}_p$ is isomorphic to

$$\hat{S} = \text{Hom}(T_pD_X^\text{ét} \otimes_{\mathcal{O}_{F_{0,p}}} T_{\mathcal{F}^{-1}}D_X^\text{ét}, \mathbb{G}_m),$$

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
as already seen, we can think of the effect of the isogeny \( \beta : \hat{X}_{/\hat{S}} \to \hat{X}'_{/\hat{S}} \) of type \( \alpha_j \) on the universal deformation space \( \hat{X}_{/\hat{S}} \), which sends

\[
\text{Hom}(T_p D_{X}^{\text{ét}} \times T_p D_{X}^{\text{ét}}, \hat{G}_m) \ni q(x, y) \mapsto q(\alpha(x), \alpha(y))/p.
\]

We need to divide by \( p \) as above by the following reason: Since \( q \in \hat{S} \) measures the depth of non-splitting of the exact sequence \( \text{Hom}(P_t, \hat{G}_m) \hookrightarrow X[p] \to P_t \), and the sequence for \( t = 1 \) is split if \( q \) is a \( p \)-power. Thus the isogeny \( \alpha \) exists over \( \hat{S}^{1/p} \).

Here we have written the group structure on \( \hat{G}_m \) additively; so, “division by \( p \)” would become “taking \( p \)-th root” if we had formulated the group structure multiplicatively.

The isogeny is defined over a smaller covering \( \hat{S}^{(q \circ \alpha)^{1/p}} = \text{Spf}(\hat{O}_S[(q \circ \alpha)^{1/p}]) \) by definition; so, \( \hat{S}^{\alpha} / \hat{S} \) is given by \( \hat{S}^{(q \circ \alpha)^{1/p}} \). At this point, we are taking \( p \)-th roots, and hence pure inseparability (we pretended not to have) comes in. Then the action of the isogeny \( \alpha \) of type \( \alpha_j \) on \( \hat{S} \) only depends on its effect on \( T_p D_{X}^{\text{ét}} \) and \( T_p D_{X}^{\text{ét}} \) not on the individual choice \( \alpha \). This means that the covering \( \hat{S}^{\alpha} \) over \( S \) carrying the isogeny \( \alpha \) only depends on the image of \( \alpha \) in the Levi-quotient of \( P \). Indeed, taking a base \((x_i)_i \) of \( T_p D_{X}^{\text{ét}} \) and \((y_k)_k \) of \( T_p D_{X}^{\text{ét}} \) so that the matrix of the isogeny is exactly \( \alpha_j \), the effect on \( t = (T_{k,l}) = (e(x_k, y_i)) \) is given by

\[
(1_m T_0) \mapsto \alpha_j(p) \left( \begin{array}{cc} 1_m & T_0 \\ 0 & 1_n \end{array} \right) \alpha_j(p)^{-1},
\]

and \( S^\alpha_t \) has degree of pure inseparability given by the value in (8.4). Hereafter we write \( S^{\alpha_j}_t \) for \( S^\alpha_t \).

Here is the justification of our argument. Write the multiplicative variable on \( \hat{S} \) as an \( m \times n \) variable matrix \( t = (t_{k,l}) \). The conjugation: \( \left( \begin{array}{cc} 1_m & T_0 \\ 0 & 1_n \end{array} \right) \mapsto \alpha_j \left( \begin{array}{cc} 1_m & T_0 \\ 0 & 1_n \end{array} \right) \alpha_j^{-1} \) induces \( T_{k,l} \mapsto p^{-1} T_{k,l} \) for some indices \((k, l)\). We split the set of indices \((k, l)\) into a disjoint union \( I \cup J \) of two subset so that the conjugation by \( \alpha_j \) induces \( T_{k,l} \mapsto p^{-1} T_{k,l} \) if and only if \((k, l) \in I \). The covering \( \hat{S}^{\alpha_j} \) is given by \( \text{Spf}(\hat{O}_S[t_{k,l}^{-1}])_{(k,l) \in I} \). Thus a formal function on \( \hat{S} \) has expansion \( \sum \xi \, a_{\xi} \xi \in W[t_{k,l}]^{-1} \) for \( \xi \in \mathbb{Z}^{I \cup J} \). Writing \( \xi(I) \) for the \( I \)-part of the index \( \xi \), a formal function \( f \) on \( \hat{S}^{\alpha_j} \) has expansion \( f = \sum \xi \, a_{\xi} \xi^{\xi(I)/p+\xi(J)} \), and we have

\[
\text{Tr}_{\hat{S}^{\alpha_j}/\hat{S}}(f) = \mu_{m,n}(\alpha_j) \sum_{\xi: \xi(I) \equiv 0 \mod p} a_{\xi} \xi^{\xi(I)/p+\xi(J)},
\]

because \( \mu_{m,n}(\alpha_j) \) is the degree of the (purely) wildly ramified covering \( \hat{S}^{\alpha_j} \to \hat{S} \) and \( \text{Tr}(t_{k,l}^{1/p}) = \mu_{m,n}(\alpha_j)t_{k,l}^{1/p} \) or 0 according as \( p | i \) or not. Thus by replacing the term:

\[
f(X, \lambda, i, \eta^{p}, \alpha(\varphi), \alpha(\beta), \alpha(\delta)),
\]

in (8.3) by the trace

\[
\text{Tr}_{\hat{S}^{\alpha_j}/\hat{S}}(f(X, \lambda, i, \eta^{p}, \alpha(\varphi), \alpha(\beta), \alpha(\delta))),
\]

we get the \( p \)-divisibility of the operator \( U_j(p) \) as the (heuristic) multiplicity formula (8.4) suggests. This justifies the heuristic argument we gave (the heuristic argument is actually valid for \( \ell \)-adic étale cohomology with \( \ell \neq p \) as already explained).
Let \( \hat{S}_\ell = \hat{S} \otimes_W W_\ell \). On the universal deformation \( \hat{X}_S \), the sheaf \( \varepsilon(\text{Lie}(\hat{X}))_{/\hat{S}_\ell} \) is given by \( \varepsilon(\text{Lie}(\hat{X}[p^\ell])) \). By duality, \( \omega_{\hat{S}_\ell} = \mathcal{O}_{\hat{S}_\ell} \otimes_{\mathbb{Z}_p} \hat{X}[p^\ell]^{\text{et}}, \) which again only depends on \( X_0[p^\ell]^{\text{et}} \); so, the Hecke operator \( U_j(p) \) is still divisible by \( \mu_{m,n}(\alpha_j) = [U_p : \alpha_j^{-1} U_p \alpha_j] \) on \( \omega_{\hat{S}_\ell}^m \) for all \( \kappa > 0 \). Thus the action of the correspondence of characteristic 0 on \( H^0(S_\infty, \omega^\kappa) \) is exactly a multiple by the number in (8.4) of the operator induced by the mod \( p \) correspondence, which is an integral operator. From this our claim is clear.

In any case, we can divide the action of \( U_j(p) \) by the number in (8.4) keeping the integrality of the operator on \( \omega^\kappa \).

**Lemma 8.4.** — Let the notation be as above. We have a well defined integral operator \( [U_p : \alpha_j^{-1} U_p \alpha_j]^{-1} U_j(p) \) on \( H^0(T_{m,n}, \omega^\kappa \otimes \Omega_{S/W}) \).

We then define

\[
e_G = \lim_{n \to \infty} (U(p))^{n!},
\]

where

\[
U(p) = \prod_{j=1}^{n+m} (U_p : \alpha_j^{-1} U_p \alpha_j)^{-1} U_j(p).
\]

As for \( T_j(p) \), if \( \kappa \geq \mu_{m,n} \) (that is, \( \kappa - \mu_{m,n} \) is in the Weyl chamber), \( T_j(p) \equiv U_j(p) \) mod \( p \) on \( H^0(M_\ell, \omega^\kappa \otimes \omega^\mu_{M/W}) \) for \( M_\ell = \text{Sh}_{K/W} \otimes_W W_\ell \). The operator \( T_j(p) \) is well defined on \( \omega^\kappa \) over \( M \) as a linear operator, using moduli theoretic interpretation.

Let \( \tilde{U} \) be the upper unipotent subgroup of \( \text{Gal}(T_{t,\infty}/S_\ell) \). Following Lecture 7, we can define the space of \( p \)-adic automorphic forms \( V^{\tilde{U}} \) on \( S_\infty \) (which is the formal completion of \( S \) along \( S_1 \)). Thus

\[
V = \lim_{\ell} \left( \lim_{t,s} V_{t,\ell,s}^{\tilde{U}} \right)
\]

for \( V_{t,\ell,s} = H^0(T_{t,\ell,s}, \mathcal{O}_{T_{t,\ell,s}}) \). We also define

\[
V = \lim_{\ell} \left( \lim_{t,s} V_{t,\ell,s}^{\tilde{U}} \right)
\]

The boundedness condition (F) in Section 7 is verified in [H95] in Case A because \( G_1(\mathbb{Q}_p) \cong GL_{m+n}(\mathbb{Q}_p) \). The hypotheses (Hyp1-2) are clear because \( S = \text{Sh}_{K}[1/\ell] \) is an affine scheme in the cocompact case. The value \( f[U_j(p)(X, \lambda, i, \eta^{(p)}), \text{fil}, \omega] \) is the sum (more precisely, the trace) of \( f(X/C, \lambda', i', \eta^{(p)}), \text{fil}', \omega' \) for étale subgroups \( C \) of \( X[p] \). Since the Hasse invariant is insensitive to étale isogeny (by its definition), the commutativity condition (C) in Section 7 holds. Then we have

**Theorem 8.5.** — Let \( W \) be a \( p \)-adic completion of the integer ring of the Galois closure of \( F/\mathbb{Q} \). Suppose that \( M/W = \text{Sh}_{K/W} \) is proper over \( W \). Let \( \overline{B} \) be the upper triangular Borel subgroup of \( GL_m(O_F, \overline{p}) \times GL_n(O_F, \overline{p}) \) and \( \tilde{U} \) is the unipotent radical of \( \overline{B} \). Let
\( T = \tilde{B}/\tilde{U} \), and regard it as a diagonal torus of \( G_1(\mathbb{Q}_p) \). We say \( \kappa \in X(T) \) positive if \( \kappa \) is positive with respect to the opposite Borel subgroup of \( \tilde{B} \). We write \( X_+(T) \) for the set of positive weights \( \kappa \).

(1) There exists a canonical inclusion for \( A = \mathbb{Z}_p \) and \( \mathbb{Q}_p/\mathbb{Z}_p \)

\[
\beta : \bigoplus_{\kappa \in X_+(T)} H^0(M/W, \omega^\kappa \otimes \mathbb{Z}_p A) \hookrightarrow V \otimes \mathbb{Z}_p A.
\]

(2) \( \text{Im}(\beta)[\mathbb{L}] \cap V \) is dense in \( V \);
(3) Write \( U(p) = \prod_{1 \leq j \leq m+n} u_j(p) \) for the standard Hecke operators at \( p \) of level \( p^\infty \), that is,

\[
u_j(p) = \frac{U_j(p)}{[U_p : \alpha_j^{-1}U_p \alpha_j]}
\]

for the unipotent radical \( U_p \) of the upper triangular maximal parabolic subgroup of \( GL_{m+n}(O_F, p) \) with Levi-subgroup isomorphic to \( GL(m) \times GL(n) \), and define the ordinary projector \( e = \lim_{n \to \infty} U(p)^n \) on \( V \). Then

\( eH^0(S, \omega^\kappa \otimes \mathbb{Z}_p \mathbb{Q}_p/\mathbb{Z}_p) \) is a \( p \)-divisible module with finite corank.

(4) If \( \kappa > \mu_{m,n} \) is sufficiently regular, \( eH^0(S, \omega^\kappa \otimes \mathbb{Z}_p \mathbb{Q}_p/\mathbb{Z}_p) \) (resp. \( eH^0(S_\infty, \omega^\kappa) \)) is isomorphic to \( H^0_{\text{ord}}(M, \omega^\kappa \otimes \mathbb{Z}_p \mathbb{Q}_p/\mathbb{Z}_p) \) (resp. \( H^0_{\text{ord}}(M, \omega^\kappa) \));
(5) Let \( \mathcal{V}_{\text{ord}}^* \) be the Pontryagin dual \( W[[T(\mathbb{Z}_p)]] \)-module of \( eV \) (which is isomorphic to the \( W \)-dual of \( eV \)). Then \( \mathcal{V}_{\text{ord}}^* \) is \( W[[\Gamma_T]] \)-free of finite rank, where \( \Gamma_T \) is the maximal \( p \)-profinite subgroup of \( T(\mathbb{Z}_p) \);
(6) If \( \kappa \in X_+(T) \),

\[
\mathcal{V}_{\text{ord}}^* \otimes W[[T(\mathbb{Z}_p)]] \cong \text{Hom}_W(eH^0(S_\infty, \omega^\kappa), W).
\]

Although we restricted ourselves to cocompact unitary cases here, a similar result can be obtained in more general settings of cusp forms on a non-compact Shimura varieties of unitary groups and symplectic groups (see [H02]). In [H02], we have given the heuristic argument for the divisibility of \( U(p) \), but it can be justified using the trace (in place of the sum of the values) from (wildly ramified) finite flat covering (carrying specified \( p \)-isogeny of the universal abelian scheme) over the Shimura variety as we did; so, the final result in [H02] is intact.

9. Hilbert Modular Forms

We shall give concrete examples in the non-co-compact case. These are Hilbert modular varieties. We give a sketch of the proof of the vertical control theorems. More details can be found in Chapter 4 of [PAF].
9.1. Hilbert Modular Varieties. — We first recall the toroidal compactification of the Hilbert-Blumenthal moduli space. Main references are [C], [K2] and [Ra] (and [HT], [DT]).

Let $A = \lim \frac{A}{p^n} A$ be a $p$-adic ring. Let $F$ be a totally real field with integer ring $O_F$ and $N$ be an integer $\geq 3$ prime to $p$. So our groups are given by $G = \text{Res}_{O_F/\mathbb{Z}} GL(2)$ and $G_1 = \text{Res}_{O_F/\mathbb{Z}} SL(2)$. We write $T$ for the diagonal torus of $G_1$ defined over $\mathbb{Z}$; thus, we have $T(A) = (O_F \otimes_{\mathbb{Z}} A)^\times$. We consider a triple

$$(X, \lambda, \phi : (\mathfrak{d}^{-1} \otimes_{\mathbb{Z}} \mu_N) \oplus (O_F \otimes_{\mathbb{Z}} \mathbb{Z}/N\mathbb{Z}) \cong X[N])$$

over a scheme $S$ made of an abelian variety with real multiplication by $O_F$ (an AVRM). This means that $X/S$ is an abelian scheme with $O_F \hookrightarrow \text{End}(X/S)$ such that $\text{Lie}(X)$ is free of rank 1 over $O_S \otimes_{\mathbb{Z}} O_F$. Here $\mathfrak{d}$ is the absolute different of $F$, and $\lambda$ is a $c_+$-polarization for a fractional $F$-ideal $c$. This means that $\lambda : \tilde{X} \cong X \otimes_{O_F} c$. The word $c_+$-polarization means that the set of symmetric morphisms induced (fiber by fiber) by ample invertible sheaves: $P_+(X) \subset \text{Hom}(X, \tilde{X})$ is isomorphic to $c_+$ by $\lambda$. This notion only depends on the strict ideal class of $c$. Thus hereafter we assume that $c \subset O_F$.

Tensoring $X$ over $O_F$ with the following exact sequence:

$$0 \rightarrow c \rightarrow O_F \rightarrow O_F/c \rightarrow 0,$$

we get another exact sequence:

$$0 \rightarrow \text{Tor}_1(O_F/c, X) \rightarrow X \otimes c \rightarrow X \rightarrow 0.$$

Thus the above condition on polarization can be stated as

$$\tilde{X}/\tilde{X}[c] \cong X$$

for $\tilde{X}[c] = \{x \in (X \otimes c)| cx = 0\}$. We also note that

$$X = X'/X'[a] \iff X' = X \otimes a \iff X = X' \otimes a^{-1},$$

which will be useful.

To describe the toroidal compactification, let

$$C = \{\xi \in F_\infty \mid \xi^\sigma > 0 \text{ for all } \sigma : F \hookrightarrow \mathbb{R}\}$$

be the cone of totally positive numbers in $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$. Choose a cone decomposition $C = \bigsqcup \sigma$ such that

1. $\sigma$ is a non-degenerate open rational polyhedral cone without containing any entire line. Here the word “rational” implies that the cone is generated by a finitely many elements in $F_+ = F \cap C$ over $\mathbb{R}_+$;

2. the set of cones $\{\sigma\}$ is permuted under multiplication of $T(\mathbb{Z})(N)^2$, where

$$T(\mathbb{Z})(N) = \{\varepsilon \in O_F^\times \mid \varepsilon \equiv 1 \text{ mod } N\}.$$ 

There are only finitely many cones modulo $T(\mathbb{Z})(N)$, and $\varepsilon(\overline{\sigma}) \cap \overline{\sigma} \neq \{0\}$ implies that $\varepsilon = 1$ (see [LFE] Theorem 2.7.1 for an exposition of such decomposition);
(3) $\sigma$ is smooth (that is, generated by a part of a base of $O_F$);
(4) $\{\sigma\}$ is sufficiently fine so that the toroidal compactification is projective (see [C] and [DAV] IV.2.4 for an exact condition for projectivity).

Let $\sigma^\vee$ be the dual cone:

$$\sigma^\vee = \{ x \in F_\infty \mid \text{Tr}_{F/Q}(x\sigma) \geq 0 \}.$$  

Then $C = \bigcap_{\sigma} \sigma^\vee$.

Here is an oversimplified description of how to make the toroidal compactification over $W/\mathbb{Z}_p$, where $W$ is the discrete valuation ring we took as the base ring. Each Hilbert modular form $f$ (defined over a ring $W$) has a $q$-expansion $f(q) = \sum_{\xi \in a \cap C} a(\xi)q^j$ for an ideal $a$. Thus we convince ourselves that the formal stalk of the minimal compactification at the cusp $\infty$ is the ring $R_\infty(a) = W[[q^j]_{a \cap C} = W[[a \cap C]]$, which is the completion of the monoid ring of the semi-group $a \cap C$. Thus $R_\infty(a) = \bigcap_{\sigma} R_{\sigma}(a)$ where $R_{\sigma}(a) = W[[q^j]_{a \cap \sigma^\vee} = W[[a \cap \sigma^\vee]]$. For each $\sigma$ as above, by enlarging $a$ if necessary, we may assume that $a \cap \sigma$ is generated over $\mathbb{Z}$ by $t_1, \ldots, t_r$ ($0 < r \leq [F : \mathbb{Q}] = g$). Then we have a base $\xi_1, \ldots, \xi_g$ of $\sigma^\vee$ so that $\text{Tr}(t_i \xi_j) = \delta_{ij}$ for $1 \leq i < j \leq r$ and $\text{Tr}(t_i \xi_j) = 0$ ($j > 0$). Then each $\xi \in a \cap \sigma^\vee$ can be uniquely written as $\xi = \sum_i m_i \xi_i$ with $m_i \in \mathbb{Z}$ and $m_j \geq 0$ if $j < r$. Thus writing $T_j = q^{\xi_j}$, we find

$$R_{\sigma}(a) = W[[T_1, \ldots, T_g]][\frac{1}{T_{r+1}}, \ldots, \frac{1}{T_g}].$$  

Thus $\text{Spf}(R_{\sigma}(a)) = \widehat{A}^r \times \widehat{\mathbb{Q}}_{p^{-r}}$ which is a compactification of $\text{Spf}(R_{\tau}(a))$ for each face $\tau$ of $\sigma$. Thus we can glue $\{\text{Spf}(R_{\sigma}(a))\}_\sigma$ on the ring in the common intersection of the $\sigma^\vee$’s, and getting a formal scheme $\mathcal{X}$ on which $T(\mathbb{Z})(N)$ acts by translation. Then make a quotient $\mathcal{X}/T(\mathbb{Z})(N)$. The algebraization of the quotient is the toroidal compactification at the infinity cusp.

We consider the moduli space $\mathcal{M}_{c,N}/W$ of test objects $(X, \lambda, \phi)_A$ for $W$-algebras $A$, where $W$ is a discrete valuation ring containing all conjugates of $O_F$. We assume that $W$ is unramified over $\mathbb{Z}_p$ and that $Nc$ is prime to $p$. From the above data, we get a unique toroidal compactification $M = M_{c,N}$ of $\mathcal{M}_{c,N}$, which carries a (universal) semi-AVRM $\mathcal{G} = \mathcal{G}_{c,N}$ with a level structure $\mathfrak{d} \otimes \mu_N \hookrightarrow \mathcal{G}[N]$. The semi-AVRM coincides with the universal abelian scheme $A = A_{c,N}$ over $\mathcal{M}_{c,N}$. Let $M_{\infty}$ be the formal completion of $M$ along $M_1 = M \otimes W W_1$. Write $S_{\infty} \subset M_{\infty}$ for the ordinary locus, that is, $S_{\infty}$ is the maximal formal subscheme of $M_{\infty}$ on which the connected component $\mathcal{G}[p]_\circ$ of $\mathcal{G}[p]$ is isomorphic to $\mu^d_p$ locally under étale topology, and thus $S_{\infty}$ is the formal completion of $S = M[\frac{1}{p}]$ along $S_1 = S \otimes W W_1$, where $E$ is a lift of Hasse invariant. Then we put $S_m = S \times W W_m$. Let

$$T_{m,n}/W_m = \text{Isom}_{O_F}(\mathfrak{d}^{-1} \otimes \mu_{p^n}, \mathcal{G}[p^n]_\circ) \cong \text{Isom}_{O_F}(\widehat{\mathcal{G}[p^n]_\circ}, O_F/p^nO_F).$$  

Then $T_{m,n}/S_m$ is an étale covering with Galois group $T(\mathbb{Z}/p^n\mathbb{Z}) = (O_F/p^n)^{\times}$ for $T = \text{Res}_{O_F/\mathbb{Z}} \mathcal{G}_m$. By a result of K. Ribet [Ri] (see also [PAF] Theorem 4.21 and [DT] Section 12), $T_{m,n}$ is irreducible.

ASTÉRISQUE 298
The sheaf $\omega_\infty / S_\infty = O_{S_\infty} \otimes_{\mathbb{Z}_p} \lim_{\leftarrow n} \mathcal{G}[p^n]^\wedge$ is isomorphic to the dual of $f_* \text{Lie}(\mathcal{G}/M)$ for $f : \mathcal{G} \to M$. In other words, $\omega_{\mathcal{G}/M} = \text{Hom}(f_* \text{Lie}(\mathcal{G}/M), O_M)$ is the algebraization of the formal sheaf $\omega_\infty$ on $S_\infty$ (which is unique). Identifying $X(T) = \text{Hom}_{\text{alg-gr}}(T, \mathcal{G}_m/W)$ with $\mathbb{Z}[I]$ for the set $I$ of embeddings of $O_F$ into $W$, we write $\omega^k$ for the sheaf associated to $k \in X(T)_+$. We then define

$$M^* = \text{Proj} \left( \bigoplus_{j \geq 0} H^0(M, \omega^{j!}) \right),$$

where $t = \sum_{\sigma \in T} \sigma$. Then $S^* \subset M^*$ is defined by

$$S^* = \text{Spec} \left( \bigoplus_{j \geq 0} H^0(M, \omega^{j!})/(E - 1) \right)$$

for the lift of the Hasse invariant $E$. Write $\pi : M \to M^*$.

The only thing we need to verify is (Hyp1):

$$H^0(S, \omega_k) \otimes_W W_m = H^0(S_m, \omega_k \otimes_W W_m)$$

for the sheaf $\omega_k \subset \omega^k$ of cusp forms of weight $k$. Since $H^0(S, \omega_k) = H^0(S^*, \pi_*(\omega_k))$ and $S^*$ is affine, we need to verify

$$\pi_*(\omega_k/S) \otimes_W W_m = \pi_*(\omega_k/W \otimes W_m).$$

We shall do this stalk by stalk. Outside the cusps, the two sheaves are the same; so, nothing to prove.

Now we have for each cusp $x$ associated to the ideals $ab^{-1} = c$ and a $p$-adic $W$-algebra $A = \lim_{\leftarrow m} A/p^m A$:

$$\pi_*(\omega^k/A)_x \cong H^0(T(\mathbb{Z})(N)^2, A[[\frac{1}{N}(ab)_+]])$$

$$= \left\{ \sum_{\xi \in \frac{1}{N}(ab)_+} a(\xi)q^\xi \in A[[\frac{1}{N}(ab)_+]] \mid a(\xi) \equiv a(\xi) ~ \forall \xi \in T(\mathbb{Z})(N) \right\},$$

where $(ab)_+ = \{ \xi \in ab \mid \xi \geq 0 \} \cup \{0\} = C \cap ab$, and $T(\mathbb{Z})(N) = \{ u \in O_F^\times \mid u \equiv 1 \text{ mod } N \}$ acts on $A[[\frac{1}{N}(ab)_+]]$ by $\varepsilon \sum_{\xi \in (ab)_+} a(\xi)q^\xi = \sum_{\xi \in (ab)_+} \varepsilon^{-k} a(\varepsilon^2 \xi)q^\xi$. When $N \geq 3$, for each $\varepsilon^2 \in T(\mathbb{Z})(N)^2$, there is a unique $\varepsilon \in T(\mathbb{Z})(N)$; so, there is no ambiguity of $(\pm \varepsilon)^{-k}$ in the above formula. We define $\omega^k$ by requiring its stalk at every cusp is given by those $q$-expansions vanishing at the cusp. The group cohomology $H^0(T(\mathbb{Z})(N)_x, X)$ commutes with $\otimes_W W_m$ if $X$ is $A[T(\mathbb{Z})(N)^2]$-free. Then from the above fact, we get

$$\pi_*(\omega^k/W)_x \otimes_W W_m \cong \pi_*(\omega^k/W_m)_x.$$
We put
\[ V_{m,n}^{\text{cusp}} = H^0(T_{m,n}, \mathcal{O}_{T_{m,n}}(-D)), \quad V_{\text{cusp}} = \lim_{m} V_{m,\infty}^{\text{cusp}} \]
\[ V_{\text{cusp}} = \lim_{m} V_{m,\infty}^{\text{cusp}}, \quad R_{\ell}^{\text{cusp}} = \bigoplus_{k > \ell} H^0(M, \omega_k) \]
\[ D_{\ell}^{\text{cusp}} = \beta(R_{\ell}^{\text{cusp}})_{[\frac{1}{\ell}]} \cap V_{\text{cusp}}. \]

Here \( k > \ell \) means that \( k_\sigma > \ell_\sigma \) for all \( \sigma \in I \), and \( D = \pi^{-1}(\sum_{x \text{cusp}} x) \) is the cuspidal divisor on the toroidal compactification.

In this \( GL(2) \)-case, it is known that we have two Hecke operators \( U(p) \) acting on cusp forms of level divisible by \( p \) and \( T(p) \) acting on cusp forms of level prime to \( p \), normalized as in Lemma 8.4 to keep integrality of \( \omega_k \). The operator \( U(p) \) has its effect on \( q \)-expansion \( a(\xi, f|U(p)) = a(\xi p, f) \) and decreases the level to the minimum as long as it is \( p^n \) for \( n > 0 \), and if \( k > 2t \), then \( T(p) \equiv U(p) \mod p \). Let \( e \) (resp. \( e^\circ \)) be the idempotent attached to \( U(p) \) (resp. \( T(p) \)). We attach a subscript or superscript “ord” to the object after applying the idempotent \( e \) or \( e^\circ \) (depending on the setting). From this, we conclude

**Theorem 9.1.** — Let \( F \) be a totally real field of degree \( d \) and \( N \) be an integer \( N \geq 3 \). Suppose that \( p \) is prime to \( NN_F/\mathbb{Q}(d) \). Then we have the following facts:

1. \( D_{\ell}^{\text{cusp}} \) is dense in \( V_{\text{cusp}} \);
2. The Pontryagin dual \( V_{\text{cusp}}^{\text{ord},*} \) (which is isomorphic to \( \text{Hom}_W(V_{\text{cusp}}^{\text{ord}}, W) \)) of \( V_{\text{cusp}}^{\text{ord}} \) is a projective \( W[[T(z_p)]] \)-module of finite type;
3. \( V_{\text{cusp}}^{\text{ord},*} \otimes_W[T(z_p)], k W \cong \text{Hom}_W(H^0_{\text{ord}}(S_{\infty, \omega_k}), W) \) if \( k \geq 3t \);
4. If \( k \geq 3t \) (\( t = \sum_{\sigma} \sigma \)), \( e \) induces an isomorphism
   \[ H^0_{\text{ord}}(S_{\infty, \omega_k}) \cong H^0_{\text{ord}}(M, \omega_k), \]
   where \( H^0_{\text{ord}}(S_{\infty, ?}) = eH^0(S_{\infty, ?}) \) and \( H^0_{\text{ord}}(M, ?) = e^\circ H^0(M, ?) \).

We shall give a very brief sketch of the proof (see [PAF] Theorem 4.8 for more details).

**Proof.** — The assertions (1) and (2) follows from the general argument, using the theory of false modular forms (Section 7). Then the assertions (3) and (4) follow for sufficiently large \( k \). It is known that \( \dim H^0_{\text{ord}}(M, \omega_k) \) depends only on \( k|T(F_p) \) if \( k \geq 3t \) (see [H88] Theorems 2.1 and 8.1 and [PAF] Theorem 4.37). From this, the assertion (3) and (4) for small \( k \) follows.

9.1.1. **Moduli problem of \( \Gamma_1(N) \)-type.** — Let \( \bar{\Gamma}_1(N) \) be an open compact subgroup in \( GL_2(\mathcal{O}_F) (\mathcal{O}_F = O_F \otimes_{\mathbb{Z}} \mathbb{Z}) \) consisting of elements congruent to upper triangular matrices of the form \( \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \) modulo \( N \). Let \( \hat{\Gamma}(N) \) be the subgroup of \( \bar{\Gamma}_1(N) \) consisting of matrices congruent to 1 modulo \( N \).
In place of the full level $N$-structure, we could have started with the moduli problem classifying test objects $(X, \lambda, \phi : \mathcal{O}^{-1} \otimes_{\mathbb{Z}} \mu_N \hookrightarrow X[N])_{/A}$. As long as the moduli problem is representable (that is, $N$ is sufficiently deep), we get the same assertions as in Theorem 9.1 replacing $M_{\epsilon,N}$ by the moduli $M_{\epsilon,\pi_1(N)}$ for weight $k$ with $k_\sigma \equiv k_\tau \mod 2$ for all $\sigma, \tau : F \hookrightarrow \overline{\mathbb{Q}}$. This parity condition is necessary to define $\varepsilon \mapsto \varepsilon^{k/2}$ for totally positive units $\varepsilon \in T(\mathbb{Z})$ (since $\varepsilon^{k} = 1$ for such units). In this subsection, we only consider the moduli problem of $\tilde{\Gamma}_1(N)$-type, and accordingly, we define $V$, $\nu_{\text{cusp}}$, and $\nu_{\text{cusp}}^*$ for $\tilde{\Gamma}_1(N)$. For simplicity, we hereafter assume that $k$ is even (so we write $2k$ instead of $k$), since the general case is exposed already in [H96] Chapter 2. Since these spaces actually depend on the choice of the ideal $\mathfrak{c}$, we add subscript $\mathfrak{c}$ if we want to make explicit the dependence on $\mathfrak{c}$.

We consider a test object $(X, \lambda, \phi)_{/A}$ of level $\tilde{\Gamma}_1(N)$. For any ideal $\mathfrak{a}$ prime to $Np$, we make a quotient $X' = X/X[\mathfrak{a}] = X \otimes_{\mathbb{Z}} \mathfrak{a}$; thus, $X = X' \otimes \mathfrak{a}$. Then $X'[N] \cong X[N]$ canonically; so, $\phi$ induces a level $\tilde{\Gamma}_1(N)$-structure $\phi'$. Let $P(X) = \{ \lambda \in \text{Hom}(X, \tilde{X}) \mid \tilde{\lambda} = \lambda \}$ and $P_+(X) \subset P(X)$ is the subset made of polarizations. Then we have an isomorphism: $\lambda : P \cong \mathfrak{c}$ of $O_F$-modules taking $P_+$ onto the subset $\mathfrak{c}_+$ of totally positive elements of $\mathfrak{c}$. Dualizing the exact sequence:

$$0 \rightarrow X[\mathfrak{a}] \rightarrow X \rightarrow X' \rightarrow 0,$$

we get another exact sequence:

$$0 \rightarrow \tilde{X}'[\mathfrak{a}] \rightarrow \tilde{X}' \rightarrow \tilde{X} \rightarrow 0,$$

because $\tilde{X}'[\mathfrak{a}]$ is the Cartier dual of $X[\mathfrak{a}]$. This shows $\tilde{X}' \cong \tilde{X} \otimes \mathfrak{a} \cong X \otimes \mathfrak{a} \cong X' \otimes \mathfrak{a}^2$, $\lambda$ induces $\lambda' : P(X)' \cong (\mathfrak{a}^2)'_+$. Thus $(X, \lambda, \phi) \mapsto (X', \lambda', \phi')$ induces $[\mathfrak{a}] : V_{\mathfrak{a}^2} \cong V_{\mathfrak{c}}$.

We identify $V_{\mathfrak{c}}$ and $V_{\mathfrak{a}^2}$ by $[\mathfrak{a}]$. Thus $V_{\mathfrak{c}}$ only depends on the strict ideal class of $\mathfrak{c}$ (and also modulo square ideal classes).

We then define

$$(9.3) \quad \mathcal{V}_{\text{cusp}}(\tilde{\Gamma}_1(N)) = \bigoplus_{\mathfrak{c}} \mathcal{V}_{\text{cusp},\mathfrak{c}}^\text{ord} \quad \text{and} \quad \mathcal{V}_{\text{cusp}}^\ast(\tilde{\Gamma}_1(N)) = \bigoplus_{\mathfrak{c}} \mathcal{V}_{\text{cusp},\mathfrak{c}}^\ast,$$

where $\mathfrak{c}$ runs over strict equivalence classes of ideals modulo square classes; thus, it runs over the group $\text{Cl}_F^+/(\text{Cl}_F^+)^2$, where $\text{Cl}_F^+$ is the strict ideal class group.

Note that

$$PGL_2(F_\mathfrak{c}) = \bigcup_{a \in \text{Cl}_F^+/(\text{Cl}_F^+)^2} PGL_2(F) (a \ 0 \ 0 \ 1) \Gamma_1(N) PGL_2(F_\infty),$$

where $a$ runs over a complete representative set for $\text{Cl}_F^+/(\text{Cl}_F^+)^2$ in $F_\mathfrak{c}^+$; $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$; $PGL_2^+(F_\infty)$ is the identity connected component of $PGL_2(F_\infty)$, and $\Gamma_1(N)$ is the image of $\tilde{\Gamma}_1(N)$ in $PGL_2(F_\mathfrak{c}^{(\infty)})$. Thus we may regard $\mathcal{V}_{\text{cusp},\mathfrak{c}}^\ast(\tilde{\Gamma}_1(N))$ as the $W$-dual of the space of $p$-adic cusp forms of level $\tilde{\Gamma}_1(N)$ on $PGL_2(F_\mathfrak{c})$. For a given modular form $f = (f_\mathfrak{c})$ the above spaces, say in $\mathcal{V}_{\text{cusp},\mathfrak{c}}$, it has $q$-expansion at the cusp.
∞ = (a = OF, b = c−1):

\[ f_\xi = \sum_{\xi \in \mathcal{E}_\xi^{-1}} a(\xi; f_\xi)q^\xi. \]

If \( f \in \mathcal{V}_{\text{cusp}, \chi}[2k] \), as we have already seen for level \( N \)-modular forms, \( a(\varepsilon^2 \xi; f) = \varepsilon^{2k} a(\xi; f) \) for \( \varepsilon \in T(\mathbb{Z}) \subset SL_2(OF) \). Since we only have level \( \hat{\Gamma}_1(N) \)-structure, \( f \) satisfies invariance under the matrix \( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \) for totally positive units \( \varepsilon \) in addition to the invariance under \( \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix} \) for arbitrary units \( \varepsilon \). Thus we actually have

\[ a(\varepsilon^2 \xi; f) = \varepsilon^{2k} a(\xi; f) \quad \text{for all totally positive units } \varepsilon. \]

Choose a finite idele \( c \) so that \( cOF = c \) and \( cp = 1 \). For finite integral ideles \( y \) and \( f \in \mathcal{V}_{\text{cusp}, \chi} \), we may define a continuous function \( y \mapsto a_p(y; f) \in K/W \) for the field of fractions \( K \) of \( W \) as follows: Choose ideal representatives (prime to \( Np \)) \( c \) and \( a \) so that \( yOF = \xi ca^2 \) for \( \xi \in (ca^2)^{-1} \), and write \( y = u \xi ca^2 \) for ideles \( u, c, a \) with \( ap = cp = 1, cOF = c, aOF = a \) and \( ap = cp = 1 \). Then we define

\[ a_p(y; f) = a_p^k(\xi; f_\xi | [a]). \]

We can verify that \( a_p(uw; f) = a_p(y; f) \) for \( u \in \hat{O}_F^\times \) with \( u_p = 1 \) and if \( f \in \mathcal{V}_{\text{cusp}, \chi}[2k] \), then \( a_p(uw; f) = a_p^k(u; f) \) for \( u \in \hat{O}_F^\times \). Thus \( a_p \) is well defined independently of the choice of \( c \), and for an integral ideal \( n \) prime to \( p \), choosing a finite idele \( n \) so that \( nOF = n \) and \( n_p = 1 \), \( a_p(ny; f) \) is well defined independent of the choice of \( n \). We write \( a_p(y; f) = a_p(ny; f) \).

We extend the function \( a_p \) outside integral ideles by defining it to be 0 and extend it to general \( f \in \mathcal{V}_{\text{cusp}, \chi} \) using the fact that \( \mathcal{V}_{\text{cusp}, \chi} = \sum_{2k} \mathcal{V}_{\text{cusp}, \chi}[2k] \). By the \( q \)-expansion principle due to Ribet (which we will prove in a more general setting in the last lecture: Section 10), the \( p \)-adic modular form is determined by the function \( a_p \) on integral ideles. An important fact (see [H96] 2.4) is the following formula for integral ideals \( n \) prime to \( p \) and the Hecke operator \( T(n) \):

\[ a_p(y; f|T(n)) = \sum_{l \supseteq n+yOF} N(l)^{-1} a_p(yn/T^2; f). \]

For \( w \in OF_p \cap F_p^\infty \), we write \( T(w) \) for the normalized Hecke operator corresponding to the double coset \( U(Z_p) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} U(Z_p) \). Then we have

\[ a_p(y; f|T(w)) = ap(yw; f). \]

**Lemma 9.2.** — Let \( h \) be the subalgebra of \( \text{End}(\mathcal{V}_{\text{cusp}}^{\text{ord}}(\hat{\Gamma}_1(N))) \) generated topologically by \( T(w) \) for \( w \in OF_p \cap F_p^\infty \) and \( T(n) \) for integral ideals \( n \) prime to \( p \). Then \( h \cong \mathcal{V}_{\text{cusp}}^{\text{ord}, s}(\hat{\Gamma}_1(N)) \) as \( W[[T(Z_p)]] \)-modules.

**Proof.** — We shall give a sketch of a proof. We consider the following pairing: \( (\ , \ ) : \mathcal{V}_{\text{cusp}}^{\text{ord}}(\hat{\Gamma}_1(N)) \times h \to K/W \) given by \( (f, h) = a_p(1, f|h) \). Then \( (T(w)T(n), f) = a_p(wn; f) \) by (9.4) and (9.5), and hence, by the \( q \)-expansion principle, if \( (f, h) = 0 \), then \( f = 0 \). By the perfectness of the Pontryagin duality, we thus have a surjective
\(h\)-linear morphism: \(h \rightarrow \mathcal{V}_{\text{ord}}^*(\hat{\Gamma}_1(N))\) of Hecke modules. Since \(h\) acts faithfully on \(\mathcal{V}_{\text{cusp}}^\ast(\hat{\Gamma}_1(N))\), we conclude the injectivity.

Since a similar duality holds between the weight \(2k\) Hecke algebra \(h_{2k}(\hat{\Gamma}_1(N); W)\) acting on \(\bigoplus \mathcal{H}^0_{\text{ord}}(M_{\hat{\Gamma}_1(N)}; \mathcal{Z}_p)\) and \(\mathcal{V}_{\text{cusp}}^\ast(\hat{\Gamma}_1(N))[2k]\), Theorem 9.1 implies the control result for the Hecke algebra:

\[
h \otimes W[[T(\mathbb{Z}_p)]]_{2k} W \cong h_{2k}(\hat{\Gamma}_1(N); W) \quad \text{for all} \quad 2k \geq 3t.
\]

We can extend this result to \(GL(2)\) (from \(PGL(2)\)). Let \(Z = \text{Cl}^p(p^\infty)\) be the ray class group modulo \(p^\infty\), that is, \(\lim_{r \rightarrow \infty} \text{Cl}^p(p^r)\). We decompose \(Z = \Gamma_Z \times \Delta_Z\) so that \(\Gamma_Z\) is \(p\)-profinite and \(\Delta_Z\) has order prime to \(p\).

Since the universal nearly \(p\)-ordinary Hecke algebra for \(p > 2\) on \(\text{Res}_{F/\mathbb{Q}} GL(2)\) is the Pontryagin dual of \(\mathcal{C}(Z, \mathcal{V}_{\text{cusp}}^\ast)\), the Hecke algebra is isomorphic to \(\mathcal{V}_{\text{cusp}}^\ast(\hat{\Gamma}_1(N)) \otimes W[[Z]]\) as \(W[[T(\mathbb{Z}_p)]]\)-modules (see [MFG] Theorem 5.6.1 for a proof when \(p > 2\) and \(N = 1\) and [PAF] 4.2.12 for more general results). Thus we have the following facts when \(N\) is sufficiently deep so that the \(\hat{\Gamma}_1(N)\)-moduli problem is representable:

**Corollary 9.3.** — Let \(p \nmid 2NN_{F/\mathbb{Q}}(\mathfrak{d})\) be a prime. Suppose either \(p \geq 5\) or that \(N\) is sufficiently deep so that the \(\hat{\Gamma}_1(N)\)-moduli problem is representable. Then we have

1. The universal \(p\)-nearly ordinary Hecke algebra of auxiliary level \(\hat{\Gamma}_1(N)\) is \(W[[\Gamma_Z \times \Gamma_T]]\)-free of finite rank;

2. The specialization of the universal Hecke algebra at each arithmetic point \(P \in \text{Spec}(W[[Z \times T(\mathbb{Z}_p)]])(\mathcal{O}_p)\) inducing weight \(k > 2t\) in \(X_+\) produces the nearly ordinary Hecke algebra of level \(\hat{\Gamma}_1(N)\) and weight \(P\) without any error terms. When \(k \geq 2t\), the specialization produces the Hecke algebra of weight \(k\) with level \(\hat{\Gamma}_1(N) \cap \hat{\Gamma}_0(p)\).

See [PAF] 4.2.12 for the proof when \(p \geq 5\).

**9.2. Elliptic \(\Lambda\)-adic Forms Again.** — We describe how to view \(\Lambda\)-adic forms as \(p\)-adic modular forms defined over \(\Lambda\). Once this is done, we can evaluate \(\Lambda\)-adic forms at elliptic curves, which gives us a convenient method of constructing and analyzing \(p\)-adic \(L\)-functions. Then, we shall give a short account of the \(\Lambda\)-adic Eisenstein series and examples of \(\Lambda\)-adic \(L\)-functions.

All arguments presented here can be generalized to Hilbert modular case, Siegel-Hilbert modular case and quasi-split unitary cases, which will be treated in a forthcoming work.
9.2.1. Generality of $\Lambda$-adic forms. — For simplicity, we assume that $p > 2$ and only consider the $\Lambda$-adic forms of level $p^{\infty}$. Let $\Lambda = \mathbb{Z}_p[[T]]$. In the third lecture, we introduced the space $G(\chi; \Lambda)$ of $p$-ordinary $\Lambda$-adic forms, which is a free $\Lambda$-module of finite rank with

$$G(\chi; \Lambda) \otimes_{\Lambda, k} \mathbb{Z}_p \cong G_k^{\text{ord}}(\Gamma_0(p), \chi \omega^{-k}; \mathbb{Z}_p)$$

for all $k \geq 2$. Here $k : \Lambda \to \mathbb{Z}_p$ is the evaluation at $u^k - 1$ of the power series. If we identify $\Lambda$ with the Iwasawa algebra $\mathbb{Z}_p[[1 + p\mathbb{Z}_p]]$ (via $1 + T \mapsto u \in 1 + p\mathbb{Z}_p$), $k$ is induced by the character $1 + p\mathbb{Z}_p \ni z \mapsto z^k \in \mathbb{Z}_p^\times$.

We write $G(\Lambda)$ for the $\Lambda$-module made of formal $q$-expansions

$$\Phi = \sum_{n \geq 0} a(n; \Phi)(T)q^n \in \Lambda[[q]]$$

such that $\Phi(u^k - 1) \in V[k]$ for infinitely many $k$. Thus we have $\bigoplus_k G(\chi; \Lambda) \subset G(\Lambda)$, where $\chi$ runs over (actually even) powers of Teichmüller characters.

We now consider the space of $p$-adic modular forms $V_{/\Lambda}$ over $\Lambda$ of level $p^{\infty}$. In other words, we shall make base-change $T_{m,n/\mathbb{Z}_p}$ to $T_{m,n/\Lambda} = T_{m,n/\mathbb{Z}_p} \otimes_{\mathbb{Z}_p} \Lambda$ and consider $p$-adic modular forms over $\Lambda$. The functions in $V_{/\Lambda} = V \otimes_{\mathbb{Z}_p} \Lambda$ classify couples: $(E, \phi : \mu_{p^{\infty}} \hookrightarrow E[p^{\infty}])_{/R}$ defined over $p$-adic $\Lambda$-algebras $R$, and $f \in V_{/\Lambda}$ is a functorial rule assigning the value $f(E, \phi) \in R$ for each couple $(E, \phi)_{/R}$ as above.

This space has two $\Lambda$-module structures: One coming from the base ring $\Lambda$ and another coming from the action of $\text{Gal}(T_{m,\infty}/S_m) = \mathbb{Z}_p^\times$ by diamond operators $(z)$. Let $\nu : 1 + p\mathbb{Z}_p \to \Lambda^\times$ be the universal character given by $\nu(z) = [z] \in 1 + p\mathbb{Z}_p$. Then we can define

$$(9.7) \quad \mathcal{G}(\Lambda) = \{ f \in V_{/\Lambda} \mid f(z) = \nu(z)f \ \forall z \in 1 + p\mathbb{Z}_p \}.$$

Each $\Phi \in \mathcal{G}(\Lambda)$ has a $q$-expansion at $\infty$: $\Phi(T, q) = \sum_{n \geq 0} a(n; \Phi)(T)q^n$. By definition, we have a natural map:

$$V_{/\Lambda} \otimes_{\Lambda, s} \mathbb{Z}_p \to V_{/\mathbb{Z}_p}$$

for each $s : \Lambda \to \mathbb{Z}_p$ taking $\Phi(T)$ to $\Phi(u^s - 1)$ for $s \in \mathbb{Z}_p$. Here the tensor product is taken using $\Lambda$-module structure induced by the diamond operators. The map is injective by the $q$-expansion principle. Since on $\mathcal{G}(\Lambda)$, the two $\Lambda$-module structures coincide, this map brings $\Phi \in \mathcal{G}(\Lambda)$ to a $p$-adic modular form of weight $s$. Therefore, $\Phi$ is a $\Lambda$-adic form.

Conversely, starting from a $\Lambda$-adic form $\Phi$, we regard $\Phi$ as a bounded measure on $1 + p\mathbb{Z}_p$ having values in $V_{/\mathbb{Z}_p}$. Here we use the fact that $\mathbb{Z}_p[[1 + p\mathbb{Z}_p]]$ is canonically isomorphic to the measure space on $1 + p\mathbb{Z}_p$ by $a(T) \mapsto \int x^a da = a(u^s - 1)$. Thus $\Phi$ is a bounded $\mathbb{Z}_p$-linear map of $C(1 + p\mathbb{Z}_p, \mathbb{Z}_p)$ into $V_{/\mathbb{Z}_p}$. Then for each test object $(E, \phi)_{/R}$ for a $p$-adic $\Lambda$-algebra $R$, regarding $R$ as a $p$-adic $\mathbb{Z}_p$-algebra, we can evaluate $\int \phi d\Phi \in V_{/\mathbb{Z}_p}$ at $(E, \phi_p)_{/R}$, getting a bounded $\mathbb{Z}_p$-linear form from the space $C(1 + p\mathbb{Z}_p, \mathbb{Z}_p)$ into $R$, which we write $\Phi(E, \phi)(T) \in R \otimes_{\mathbb{Z}_p} \Lambda = R[[T]]$. Since $R$ is already a $\Lambda$-algebra, the $\Lambda$-module structure $\Lambda \otimes R \to R$ given by $\lambda \otimes r = \lambda r$ induces a surjective algebra
homomorphism $m : R \otimes_{\mathbb{Z}} \Lambda \to R$. We then define $\Phi(E, \phi)$ by $m(\Phi(E, \phi)(T))$. Then the assignment: $(E, \phi) \mapsto \Phi(E, \phi)$ satisfies the axiom of the $p$-adic modular forms defined over $\Lambda$. It is easy to check that this $p$-adic modular from is in $G(\Lambda)$ having the same $q$-expansion at $\infty$ as $\Phi$. Thus we have found:

**Theorem 9.4.** — The subspace $G(\Lambda) \subset V/\Lambda$ is isomorphic to the space $G(\Lambda)$ of all $\Lambda$-adic forms via $q$-expansion at the cusp $\infty$. In particular, we have

$$\bigoplus_{\chi} G(\chi; \Lambda) \cong e(G(\Lambda))$$

for the $p$-ordinary projector $e : V/\Lambda \to V^\text{ord}$.

Let $(E, \omega)_W$ be an elliptic curve with complex multiplication by an imaginary quadratic field $F = \mathbb{Q}[\sqrt{-D}]$. We suppose that $\omega$ is defined over $\mathbb{Q}$ fixing an embedding $i : \mathbb{Q} \hookrightarrow \mathbb{Q}_p$. Suppose that $p$ splits in $F$ and write $p = p_1 p_2$. Under this assumption, $E$ has ordinary good reduction modulo $p$; so, $E$ is $p$-ordinary. We may assume that $W = W(\mathbb{F}_p)$ and $E[p_\infty]$ is the etale part of $E[p^\infty]$ over $W$. Thus we have $\phi : \mu_{p^\infty} \cong E[p^\infty]$. In this way, we can evaluate a given $\Lambda$-adic form $\Phi$ at $(E, \phi)$.

**Corollary 9.5.** — If $\Phi(E, \phi) = 0$ for infinitely many distinct $E$ with complex multiplication, then $\Phi = 0$. There exists a finitely many elliptic curves $(E_i, \phi_i)_W$ such that any given linear form $G(\chi; \Lambda) \to \Lambda$ is a $\Lambda$-linear combination of evaluation at $(E_i, \phi_i)$.

If $\Phi \in G(\chi; \Lambda)$ and further if $f = \Phi(u^k - 1) \in G_k(\Gamma_0(p), \chi \omega^{-k}; W)$, then $f(E, \omega) \in W$. The morphism $\mu_{p^\infty} \hookrightarrow E$ induces a canonical differential $\omega_{\text{can}} = \phi_* dt/t$. Then $\omega = \Omega_p \omega_{\text{can}}$, and we have a result of Katz [K2] Chapter II:

$$f(E, \omega) = \frac{f(E, \omega_{\text{can}})}{\Omega_p^k} = \frac{\Phi(E, \phi)(u^k - 1)}{\Omega_p^k} \in W \subset \overline{W}.$$  

We may assume that $E(\mathbb{C}) = \mathbb{C}/O_F$. Let $w$ be the variable of $\mathbb{C}$. Then $dw$ induces a canonical differential $\omega_\infty$ on $E(\mathbb{C})$. Then $\omega = \Omega_\infty \omega_\infty$, and we get a result of Shimura [Sh4]:

$$f(E, \omega) = \frac{f(E, \omega_\infty)}{\Omega_\infty^k} = \frac{f(E, \omega_{\text{can}})}{\Omega_p^k} \in W \subset \overline{W}.$$  

The lattice $O_F = H_1(E, \mathbb{Z}) \subset \mathbb{C}$ is generated over $O_F$ by a single element $\gamma = 1$ and

$$\Omega_\infty = \int_\gamma \omega,$$

because $\int_\gamma dw = 1$. 

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
9.2.2. Some $p$-adic $L$-functions. — For simplicity, we assume that $p > 2$ and only consider the $Λ$-adic Eisenstein series of level $p^∞$. Let us fix an even power $χ = ω^a$ of the Teichmüller character. For simplicity, we choose $a \neq 0 \mod p - 1$; so, $χ$ is non-trivial. Then we consider the Kubota-Leopoldt $p$-adic $L$-function $−\frac{1}{2}L_p(1−s, χ) = a_0(u^s−1)$ $(u = 1 + p)$ with $a_0 \in Z_p[[T]]$ (cf. [LFE] 3.4-5). Then we have

$$L_p(1−k, χ) = (1 − χω^k(p)p^{k−1})L(1−k, χω^k)$$

for positive integers $k$. Then we define an element $E χ ∈ G(χ; Λ)$ by

$$a(n, E χ) = \sum_{0<d|n, p∤d} χ(d)d^{−1}(1 + T)^{log(d)/log(u)}$$

and $a(0, E χ) = a_0(T)$.

We want to relate $E χ(u^k − 1)$ to the following classical Eisenstein series:

$$E_k(E, ω) = \frac{1}{2} \sum_{(m,n)≠(0,0)} \frac{1}{(mw_1 + nw_2)^k},$$

where $(E, ω)/C$ corresponds to the lattice $L = Zω_1 + Zω_2$ by Weierstrass theory (that is, $E(C) = C/L$ and $ω = dw$ for the variable $w ∈ C$). As is well known (e.g. [LFE] 5.1), for even $k > 2$

$$E_k = c(k)\left(−\frac{1}{2}\zeta(1−k) + \sum_{n>0} \left(\sum_{0<d|n} d^{k−1}\right)q^n\right)$$

for $c(k) = Γ(k)/(2π\sqrt{−1})^k$. Thus shows that if $ω^k = χ$, then

$$E χ(u^k − 1) = c(k)^{-1}(E_k − p^{k−1}E_k(pz)).$$

If we take the elliptic curve $(E, ω)$ defined by $y^2 = 1 − x^4$ with $ω = dx/y$, then it has complex multiplication by $Q[\sqrt{−1}]$ and for $k > 2$ with $ω^k = χ$, we have

$$\frac{1}{2}E_k(E, ω) = \frac{L(k, λ_k)}{Ω^k_∞}.$$
form defined by $\Psi = \ell(\Psi)\Phi + x$ for $x \in \mathbb{X}$. We consider two copies of $\Lambda$, say $\mathbb{Z}_p[[S]]$ and $\mathbb{Z}_p[[T]]$. Take two Hecke eigenforms $\Phi \in G(\chi; \mathbb{Z}_p[[T]])$ and $\Psi \in G(\psi; \mathbb{Z}_p[[S]])$. Extend linearly $\ell$ to $G(\chi; \mathbb{Z}_p[[T]]) \otimes_{\mathbb{Z}_p, \mathbb{Z}_p} \mathbb{Z}_p[[S]] \to L \otimes_{\mathbb{Z}_p, \mathbb{Z}_p} \mathbb{Z}_p[[S]]$. Then we define

$$L_p(S, T) = \ell(e(\Psi(S))\mathcal{E}_{\chi\psi^{-1}}((1 + T)(1 + S)\overline{S}^{-1} - 1)),$$

where $e : V/\mathbb{Z}_p[[T]] \to V^{\text{ord}}/\mathbb{Z}_p[[T]]$ is the $p$-ordinary projector. Then we see

$$\Psi(S)\mathcal{E}_{\chi\psi^{-1}}((1 + T)(1 + S)\overline{S}^{-1})|_{S = u^m - 1, T = u^k - 1} = \Psi(u^m - 1)\mathcal{E}_{\chi\psi^{-1}}(u^{k-m} - 1).$$

Thus $\Psi(u^m - 1)\mathcal{E}_{\chi\psi^{-1}}(u^m(1 + T) - 1) \in G(\chi; \mathbb{Z}_p[[T]])$ and hence $L_p(u^m - 1, u^k - 1)$ is the coefficient of $\Psi(u^m - 1)E_{k-m}$ in $\Phi(u^{k-1})$ for a suitable Eisenstein series $E_{k-m}$ of weight $k - m$. As is shown by Shimura, this coefficients can be computed by the Rankin product value

$$D(k - 1, \Phi(u^{k-1}), \Psi(u^m - 1)) \quad (k > m)$$

for the Petersson inner product $\langle \cdot, \cdot \rangle$ up to an explicit constant; so, $L_p$ gives $p$-adic interpolation of the Rankin product. For an explicit evaluation formula for $L_p(S, T)$, see [LFE] Chapter 7 and 10 and [H96] Chapter 6.

What I would like to emphasize is that the we have used almost everywhere are:

(1) Vertical Control Theorem;
(2) The $q$-expansion principle (irreducibility of the Igusa tower).

10. Igusa Towers

We sketch a proof of irreducibility of the generalized Igusa tower by using the determination of the automorphism group of the arithmetic automorphic function field by Shimura and his students. The method is classical and goes back to works of Deuring [Du] and Igusa [I]. By this result, the $q$-expansion principle holds for $p$-adic modular forms on symplectic groups, and for unitary groups, one need to modify it in an appropriate way. We can construct, as Panchishkin did for Siegel modular forms, the $p$-adic Eisenstein measure for quasi-split unitary groups. The difference of our result from Panchishkin’s treatment is that our measure has values in the space of $p$-adic automorphic forms (not just in the formal $q$-expansion ring in Panchishkin’s work), since we dispose the $q$-expansion principle. A detailed proof of the result presented here and a further generalization are in [PAF] Section 8.4.

10.1. Automorphism Groups of Shimura Varieties. — Let the notation be as in Section 6. For a number field $X$, we write $I_X$ for the set of all field embeddings of $X$ into the algebraic closure $\overline{Q}$ of $Q$ in $C$. Let $W$ be the ring of Witt vectors $W(F)$ for an algebraic closure $F$ of $F_p$, and we identify $W$ with a subring of the $p$-adic completion of an algebraic closure $\overline{Q}_p$ of $Q_p$. We fix an embedding $i_p : \overline{Q} \hookrightarrow \overline{Q}_p$, and write $W$ for the pull back image $W$ under $i_p$. We write $m_W$ (resp. $m$) for the maximal
ideal of \( \mathcal{W} \) (resp. \( W \)). Recall the setting in Section 6 that \( F_0 \) is a totally real finite extension of \( \mathbb{Q} \), that \( F = F_0 \) in Case C and D and that in Case A, \( F \) is a totally imaginary quadratic extension of a totally real field \( F_0 \). Let the algebraic group \( G \) be as in Section 6 and also assume that we are in Case A or C. Presumably Case D can be treated similarly, but the results of Shimura we need are often formulated only for groups of type A and C. Towards the end of this lecture, we assume for simplicity that \( F_0 = \mathbb{Q} \) and \( G \) in Case A is quasi-split over \( \mathbb{Q} \).

10.1.1. Automorphism Groups of Automorphic Function Fields. — For the moment, we do not assume that \( F_0 = \mathbb{Q} \). The group \( G \) is indefinite at \( \infty \), that is, \( G(\mathbb{R}) \) is not compact modulo its center \( \mathbb{Z}(\mathbb{R}) \). We use the formulation of \( \text{Sh}_K \) described in Section 6 which represents the functor \( \mathcal{P} \) classifying quadruples \((X, i, \lambda, \eta)\) for \( E \)-schemes \( S \), where \( E \) is the reflex field. Thus \( E \) is the minimal field of definition of the complex representation of \( B \) on \( V_1 \) in Section 6. Take a finite Galois extension \( F'/\mathbb{Q} \) containing \( F \). When we are in Case A, writing formally the signature of \( G \) as \( s = \sum \sigma m(\sigma) \sum \tau \in \text{Gal}(Q'/Q) \, \tau \sigma \) for embeddings \( \sigma : F \hookrightarrow F' \) and for \( \mathfrak{R}' = Gal(F'/F) \), \( E \) is the fixed field of \( \mathfrak{R}' \). Then we can define \( m'(\sigma) \) for \( \sigma \in I_E \) by

\[
\sum_{\sigma \in I_E} m(\sigma) \sum_{\tau \in \mathfrak{R}'} (\tau \sigma)^{-1} = \sum_{\sigma \in I_E} m'(\sigma) \sum_{\tau \in \mathfrak{R}'} \tau \sigma.
\]

Then \( \theta = \sum_{\sigma \in I_E} m'(\sigma)\sigma \) can be regarded as a character of \( \text{Res}_{E/Q} \mathbb{G}_{m/E} \) with values in \( \text{Res}_{E/Q} \mathbb{G}_{m/F} \) (see \([\text{Sh}3]\) Section 1). Then \( \theta(E_K^\infty) = \theta(\text{Res}_{E/Q} \mathbb{G}_{m/E}(\hat{\mathbb{A}})) \) is a closed subgroup of \( F_K^\times = \text{Res}_{F/Q} \mathbb{G}_{m/F}(\hat{\mathbb{A}}) \).

Kottwitz formulated the Shimura variety over \( O_E \otimes \mathbb{Z}(p) \), but we only need Deligne’s models over \( E \) to define the automorphic function field \( \mathcal{R} \). We then take a tower \( \{V_K\} \) (allowing \( K \) not necessarily maximal at \( p \)) of the geometrically irreducible component of \( \text{Sh}_K \) so that \( V_K(\mathbb{C}) = \Gamma_K \backslash \mathfrak{3} \) for \( \Gamma_K = KG(\mathbb{R})_+ \cap G(\mathbb{Q}) \) and \( V_K \) is covered by \( V'_K \) if \( K' \subset K \), where \( \mathfrak{3} \) is the symmetric hermitian domain of \( G(\mathbb{R})_+ \). Then the union of \( \mathcal{R} \) of the function field \( \mathbb{Q}(V_K) \) of \( V_K \) is independent of the choice of the tower (up to isomorphisms), since \( V_K \) is the canonical model in Shimura’s sense (\([\text{ACM}] \) and \([\text{AAF}] \) Chapters I and II). Since the group \( G(\mathbb{A}(\infty)) \) acts on the functor \( \mathcal{P} \) by isogenies, we let \( G(\mathbb{A}) \) act on \( \mathcal{P} \) through the projection \( G(\mathbb{A}(\infty)) \). Let \( \mathcal{G}_+ \subset G(\mathbb{A}) \) be the stabilizer of the tower \( \{V_K\} \). Since \( V_K(\mathbb{C}) = \Gamma_K \backslash \mathfrak{3} \), the closure of \( \bigcup_K \Gamma_K G(\mathbb{R})_+ \) is contained in \( \mathcal{G}_+ \).

We now suppose that \( G \) is an inner form of \( GSp(2n)/\mathbb{Q} \) in Case C and in Case A

\[
G(\mathbb{Q}) = \{ \alpha \in GL_{2n}(F) \mid \langle J_n \alpha \rangle = \nu(\alpha) J_n \text{ for } \nu(\alpha) \in \mathbb{Q} \}
\]

with \( J_n \) as in (6.6) for an imaginary quadratic field \( F \). Therefore \( E = F_0 = \mathbb{Q} \), \( \theta = id : \mathbb{Q} \hookrightarrow F \), \( B \) is either \( \mathbb{Q} \) or a quaternion algebra over \( \mathbb{Q} \) in Case C, and in Case A, \( B = F \) and \( G \) in Case A is quasi-split over \( \mathbb{Q} \) (any quasi-split unitary group acting
on a hermitian space of dimension $2n$). In this case, we have an explicit description of $G_+$ by a work of Shimura ([ACM] 26.8, [AAF] 8.10, [Mik], [Mit] and [MiS]):

(Sh1) $G_+ = \psi^{-1}((\mathbb{F}^\times)^{1-c}\mathbb{R}_c^\times) in G(\mathbb{A})$ for $\psi = \det /\nu^n : G(\mathbb{A}) \to \mathbb{F}_c^\times$ in Case A; $G_+ = G(\mathbb{A})_+ = \{x \in G(\mathbb{A}) \mid \nu(x_\infty) > 0\}$ in Case C; so, we have $G_+ \supset SG(\mathbb{A})$ $(SG(\mathbb{A}) = \{x \in G(\mathbb{A}) \mid p(\det(x) = \nu(x) = 1)\}$ and $\nu(G_+) = \mathbb{A}_c^\times = (\mathbb{A}(\infty))^{\times} \times \mathbb{R}_c^\times$;

(Sh2) (Shimura’s reciprocity map) Let $Z \subset G$ be the center. Then we have a canonical exact sequence:

$$1 \longrightarrow Z(\mathbb{Q})G(\mathbb{R})_+ \longrightarrow G_+ \longrightarrow \frac{\tau}{\tau} Aut(\mathbb{R}) \longrightarrow 1,$$

and $\tau$ is continuous and open under the Krull topology on $\mathbb{R}$ (see [IAT] 6.3 for the topology and [PAF] Theorem 7.7 for a description of $\tau$).

(Sh3) The maximal abelian extension $\mathbb{Q}_{ab}$ of $\mathbb{Q}$ is the field of scalars of $\mathbb{R}$, that is, $\mathbb{R} \supset \mathbb{Q}_{ab}$ and $\mathbb{R}$ and $\mathbb{C}$ are linearly disjoint over $\mathbb{Q}_{ab}$. In particular, $\tau(x)$ acts on $\mathbb{Q}_{ab}$ through the image of $\nu(x)$ under the projection: $\mathbb{A}_c^\times \to \mathbb{A}_c^\times /\mathbb{R}_c^\times \mathbb{Q}^\times \cong \text{Gal}(\mathbb{Q}_{ab}/\mathbb{Q})$ (Artin reciprocity map).

(Sh4) The subfield $E_K$ of $\mathbb{Q}_{ab}$ fixed by $\nu(K)$ is the field of definition of $V_K$, that is, $E_K$ is isomorphic to the algebraic closure of $\mathbb{Q}$ in $\mathbb{Q}(V_K)$.

(Sh5) The extension $\mathbb{R}/Q(V_K)$ is a Galois extension with

$$\text{Gal}(\mathbb{R}/Q(V_K)) = \tau(KG(\mathbb{R})_+).$$

The first three terms of the exact sequence of (Sh2) are proven in [Sh3] and [Mik] along with finiteness of the cokernel of $\tau$. The surjectivity of $\tau$ can be shown, using the result in [Mit] (see [MiS]). When $F_0 \neq \mathbb{Q}$, we need to replace $\mathbb{Q}^\times G(\mathbb{R})_+$ by the adelic closure $F^\times G(\mathbb{R})_+$ in (Sh2) and $\text{Coker}(\tau)$ is non-trivial (basically given by $\text{Aut}(F_0)$), and the notation $\mathbb{R}_c^\times$ is often used in place of $G_+$ in the literature we quoted.

We suppose the following condition:

(ord) $p$ split in $F$ (in Case A).
(spt) $G$ in Case C is split over $\mathbb{Q}_p$.

Thus, identifying $G(\mathbb{Q}_p)$ with the symplectic or unitary similitude group of $J_n$, we have the parabolic subgroup $P_n \subset G$ given by $\{(a_+^\times) \mid * is of size n \times n\}$.

We fix a place $\mathfrak{p}$ of $\mathbb{Q}_{ab}$ over $p$. For an open compact subgroup $K = K_p \times K(p)$ with $K_p = GL_g(\mathbb{Z}_p) \times \mathbb{Z}_p^\times$ in Case A and $GSp_g(\mathbb{Z}_p)$ in Case C, we know that $Sh^{(p)}_K$ (and hence $V_K$) has good reduction at $\mathfrak{p}$ and hence $\mathfrak{V}_K = (V_K \bmod \mathfrak{p})$ is irreducible (as described in Section 6). Recall that such an open compact subgroup is called maximal at $p$. Let $v$ be a valuation of $\mathbb{R}$ over $\mathfrak{p}$ such that the residue field of $v$ restricted to $\mathbb{Q}(V_K)$ is the function field of $V_K \bmod \mathfrak{p}$ for any open compact subgroup maximal at $p$. In other words, the field $\bigcup_{K_{\text{max}} at p} \mathfrak{p}_p(V_K)$ for $K$ maximal at $p$ is the residue field of $v$ restricted to $\mathfrak{R}(p) = \bigcup_{K_{\text{max}} at p} \mathbb{Q}(V_K) \subset \mathfrak{R}$. The valuation $v|_{\mathfrak{R}(p)}$ is unique and is discrete, because $\mathfrak{R}(p)$ is the function field of a smooth model.
Thus in Case A or C with $F$ Theorem 10.1, $\tau$ (under $G$ in Case A) is algebraic, $\overset{\sim}{\longrightarrow}$ \text{lim}_{K: \text{max at } p \ V_{K/W}} \text{over the discrete valuation ring } W$. Since $R/\overset{\sim}{\longrightarrow} R(p)$ is algebraic, $v|_{R(p)}$ extends to a valuation $v$ on $R$ (which is not discrete).

Let

$$\mathcal{D} = \{ \sigma \in \text{Aut}(R) \mid v \circ \sigma = v \}.$$ 

Thus $\mathcal{D}$ is the decomposition (or monodromy) group of $v$ inside $\text{Aut}(R)$. Since $R/\overset{\sim}{\longrightarrow} R(p)$ is algebraic, $\mathcal{D}$ is unique up to conjugations in $\text{Aut}(R)$.

We now state our main theorem:

**Theorem 10.1.** — Let the notation and assumptions be as above. Suppose that we are in Case A or C with $F_0 = E = Q$. In addition to (ord) and (spt), we suppose that $G$ in Case A is quasi split isomorphic to $U(n,n)$. Then the group $\mathcal{D}$ is the image (under $\tau$) of a conjugate in $G_+$ of

$$\mathcal{P} = \left\{ x \in \left( \prod_{P_n \in \text{p}} G \left( \mathbb{A}^{(p)} \right) \right) \cap G_+ \mid \nu(x) \in \mathbb{Q}_p^+ \times \mathbb{R}_+^\times \right\},$$

where $\mathbb{R}_+^\times$ is the identity connected component of $\mathbb{R}^\times$.

We will prove the theorem in Case A in the following section. See [PAF] 6.4.3 and Section 8.4 for the proof valid for more general Shimura varieties in Cases A and C.

Suppose that $K$ is maximal at $p$ and $K^{(p)}$ is sufficiently small. Let $S = Sh_K[\mathbb{A}^{(p)}]$ for a lift $E$ of the Hasse invariant $H$. Let $S^0$ be a geometrically connected component of $S_{/W}$. Since $S$ is smooth over $W$, by the existence of the projective compactification of $Sh_K^{(p)}$, as described at the end of 6.1.2, $S^0 = S^0 \otimes_W \mathbb{F}$ is geometrically connected. Let $T_{1,\infty}/S_1$ be the Igusa tower as in Sections 7 and 8. Since we only care $T_{1,\infty}$, we simply write $T_\infty$ for $T_{1,\infty}$. Let $L_n$ be the Levi subgroup of $G_1 \cap P_n$. Thus $L_n(\mathbb{Z}_p)$ is isomorphic to $GL_n(\mathbb{Z}_p)$ in Case C and to $GL_n(\mathbb{O}_p) \times GL_n(\mathbb{O}_p)$ in Case A, writing $O = O_F$ for the integer ring of $F$. By construction, $L_n(\mathbb{Z}_p)$ acts transitively on the set of geometrically connected components of $T_\infty$ over $S_1$. Thus $\text{Gal}(T_\infty/S_1)$ for a geometrically connected component $T_\infty$ of $T_\infty$ is a subgroup of $L_n(\mathbb{Z}_p)$. In Case A, by (ord), we have $\Sigma = \{ \mathfrak{p}, \mathfrak{F} \}$. We define a subgroup $\mathfrak{G}$ of $L_n(\mathbb{Z}_p)$ by

$$\mathfrak{G} = \left\{ \left\{ (g_p, g_\mathfrak{F}) \in GL_n(\mathbb{O}_p) \times GL_n(\mathbb{O}_\mathfrak{F}) \mid \det(g_p) = \det(g_\mathfrak{F}) \right\} \text{ in Case A,} \right\} \left\{ GL_n(\mathbb{Z}_p) \right\} \text{ in Case C.}$$

Let $\omega_\sigma$ be the $\sigma$-eigenspace of the action of $O$ on $\omega$, where $\sigma : O \hookrightarrow W$ is an embedding. For the moment we suppose that we are in the unitary case. Extending scalar to $\mathbb{C}$ (from $W$), the automorphic factor $j_\sigma(g,z)$ defining $\omega_\sigma$ satisfies

$$\det(j_\sigma(g,z)) = \det(g)^{-1} \det(j_\sigma(g,z)).$$

In Subsection 4.2, $j_\sigma(g,z)$ (resp. $j_\sigma(g,z)$) is written as $h(g,z)$ (resp. $j(g,z)$). These sheaves are actually defined over $W$, and the difference (which is $\det(g)$) factors through the map $\tau_{|_{\mathfrak{o}_{ab}}}$ (because basically $\det = \nu^0$ on $G_+$). Thus the two sheaves $\det(\omega_\sigma)$ and $\det(\omega_\sigma)$ are equivalent over $W = W(\mathbb{F}_p)$. 

\text{ASTÉRIQUE 298}
Corollary 10.2. — Let the assumption be as in the theorem. The Galois group \( \text{Gal}(T_\infty^\infty/S_1) \) is equal to the above group \( \mathfrak{G} \). In the symplectic case, \( T_{1,m} \times S S_1^\infty \) is geometrically irreducible. In the unitary quasi-split case, each geometrically irreducible components of the Igusa tower \( T_\infty^\infty \) has Galois group over \( S_1^\infty \) isomorphic to \( \mathfrak{G} \) as

Corollary 10.3. — Let the assumption be as in the theorem. We assume that \( G = GSp(2n)/\mathbb{Q} \). Then a \( p \)-adic automorphic form (in \( V^U \)) on \( G \) is determined by its \( q \)-expansion at the infinity (or any other cusps unramified over \( \text{Sh}_K \)). If \( f \) and \( g \) in \( V^U/W \) have congruences \( a(\xi; f) \equiv a(\xi; g) \mod p^k \), then \( f = g \) in \( V^U/W = V^U/p^kV^U \).

10.1.2. \( q \)-Expansion Principle for Quasi-split Unitary Groups. — Hereafter we assume that \( F \) is an imaginary quadratic field with \( (p) = \mathfrak{P} \mathfrak{P} \) and that \( G \) is given by the quasi-split group \( GU(n,n) \). As stated in Corollary 10.2, the original Igusa tower is not irreducible; so, to get the \( q \)-expansion principle, we need to take a smaller tower.

Let us explain how to define a smaller (irreducible) tower. Let \( \mathbb{X} \) be the universal abelian scheme over \( S_1^\infty \). Then we write \( P_m \) (resp. \( \overline{P}_m \)) for the étale quotient \( \mathbb{X}[p^m]^{\text{ét}} \) of \( \mathbb{X}[p^m] \) (resp. \( \mathbb{X}[\overline{p}^m]^{\text{ét}} \)). The original tower \( T_m/S_1^\infty \) represents the functor \( \text{Isom}(O/p^m)^n \times (O/\overline{p}^m)^n, P_m \times \overline{P}_m \) taking an \( S_1^\infty \)-scheme \( T \) to the set of \( O \)-linear isomorphisms \( \psi : (O/p^m)^n \times (O/\overline{p}^m)^n \cong P_m \times \overline{P}_m \). By the shape of \( \mathfrak{G} \) in (10.1), we find that

\[
Q_m/S_1^\infty = \bigwedge (P_m \oplus \overline{P}_m) \cong \left( \wedge^n P_m \right) \otimes \left( \wedge^n \overline{P}_m \right)
\]

is constant over \( S_1^\infty \) because \( \mathfrak{G} \) acts trivially on \( Q_m/F \). Thus fixing an isomorphism

\[
\iota_m : (\mathbb{Z}/p^m\mathbb{Z}) \times S_1^\infty \cong Q_m \quad (m = 1, 2, \ldots)
\]

over \( F \) so that \( \iota_{m+1} \) induces \( \iota_m \), the irreducible component \( T^\infty_m(\iota_m)/S_1^\infty \) (corresponding to \( \iota_m \)) represents a subfunctor

\[
T \mapsto \{ \psi \in \text{Isom}_T((O/p^m)^n \times (O/\overline{p}^m)^n, P_m \times \overline{P}_m) \mid \wedge^{2n}\psi = \iota \}
\]

of \( \text{Isom}(O/p^m)^n \times (O/\overline{p}^m)^n, P_m \times \overline{P}_m \). Considering the tower \( T^\infty_m(\iota_m) \) over \( W_\alpha \), we can think of the ring of global sections \( V^\infty_{\alpha,m}(\iota_m) = H^0(T^\infty_{\alpha,m}(\iota_m)/W_\alpha, \mathfrak{O}_{T^\infty_{\alpha,m}}) \), and
define
\[ V^U(\iota_\infty) = \lim_{\alpha} V_{\infty,\alpha}(\iota_\infty)^U. \]
This space of \( p \)-adic modular form is a subspace of \( V^U \) we considered before.

The formal scheme \( T_{\infty,m} = \lim_{\alpha} T_{\alpha,m} \) is étale over the formal completion \( S^\dagger_\infty \) and extends to a unique toroidal compactification \( \mathcal{T}_{m,\infty} \) étale over the toroidal compactification \( \mathcal{T}_\infty \) of \( S^\dagger_\infty \). In other words, taking the semi-abelian scheme \( G/\mathcal{T}_\infty \) extending the universal abelian scheme \( \mathcal{X}/\mathcal{S}_\infty \), we have \( \mathcal{T}_{\infty,m} = \text{Isom}(O/p^mO, G[p^m]^\circ) \), where \( G[p^m]^\circ \) is the Cartier dual of the connected component \( G[p^m]^\circ \) (which naturally extends \( \mathcal{X}[p^m]^\text{ét} = P_m \oplus \mathcal{T}_m \) by the duality). Since \( \mathcal{S}_\infty \) contains the infinity cusp, we have a well chosen infinity cusp of \( \mathcal{T}_{\infty,m} \).

**Corollary 10.4.** — Let the assumption be as in the theorem. Suppose that \( G \) is given by \( GU(n,n) \) for an imaginary quadratic field \( F \) where \( p \) splits. Then a \( p \)-adic automorphic form (in \( V^U(\iota_\infty) \)) on \( G \) is determined by its \( q \)-expansion at the infinity (or any other cusps unramified over \( \text{Sh}_K \)). If \( f \) and \( g \) in \( V^U(\iota_\infty)_W \) have congruences \( a(\xi; f) \equiv a(\xi; g) \mod p^k \), then \( f = g \) in \( V^U_{k,\infty}(\iota_\infty) = V^U(\iota_\infty)/p^kV^U(\iota_\infty) \).

**10.2. Quasi-split Unitary Igusa Towers.** — We shall give a sketch of a proof of the theorem in the quasi-split unitary case of even dimension at the end of this lecture. The proof in the split symplectic case is basically the same and actually easier (see [PAF] 6.4.3).

**10.2.1. Preliminaries.** — First we describe necessary ingredients of the proof. Recall that \( J_n = (\begin{smallmatrix} 1_n & 0 \\ 0 & -1_n \end{smallmatrix}) \). Then \( G = GU(n,n) \) can be identified with the following group functor:

\[ G(A) = \left\{ \alpha \in GL_{2n}(A \otimes \mathbb{Q} F) \mid \alpha J_n \alpha^c = \nu(\alpha) J_n, \ \nu(\alpha) \in A^\times \right\}. \]

Here \( c \) is the non-trivial automorphism of \( F/\mathbb{Q} \) extended to \( A \otimes \mathbb{Q} F \) for each \( \mathbb{Q} \)-algebra \( A \). We consider the \( F \)-vector space \( V \) of dimension \( 2n \) and the alternating pairing \( \langle x, y \rangle = xJ_n'y^c \). Then \( \langle bx, y \rangle = \langle x, b'y \rangle \); so, the positive involution \( * \) on \( B = F \) is given by \( c \). Then \( C = \text{End}_F V = M_{2n}(F)^{\text{opp}} \); in other words \( M_{2n}(F) \) acts on \( V \) by the right multiplication. Let \( L \subset V \) be a \( O \)-lattice with \( L_p \cong \text{Hom}_{\mathbb{Z}_p}(L_p, \mathbb{Z}_p) \) under \( \langle , \rangle \). We take \( h : \mathbb{C} \hookrightarrow C_\infty = C \otimes \mathbb{Q} \mathbb{R} \) to be \( h(i) = -J_n \otimes 1 \). In this case, the representation of \( O \) on \( V_1 \) is just a multiple of the regular representation of \( O \); so, its \( (p \text{-adic}) \) isomorphism class is just 1 under (ord). We consider the following moduli problem for an integer \( N > 0 \) prime to \( p \): To each \( W \)-scheme \( S \), we associate the set of isomorphism classes: \( [(X, i, \lambda, \eta^{(p)})_S \mid (\text{det})] \) such that

- \( i : O \hookrightarrow \text{End}(X/S) \) taking 1 to \( i_{\text{det}} \);
- \( \eta^{(p)} \) is made of a pair of \( \eta^{(p)}_1 : T^{(p)}(X_s) \cong \hat{L}^{(p)} \) modulo \( \hat{G}^{(p)} \) as \( \hat{G}^{(p)} \)-modules for any geometric point \( s \in S \) and \( \eta_N : L/NeL \cong X[N] \), where for \( \hat{L} = L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}} \), \( \hat{G} = \hat{G}_L = \{ x \in G(A(\infty)) \mid x\hat{L} = \hat{L} \} \).

**ASTÉRISQUE 298**
– $\lambda : X \to \hat{X}$ is a polarization which induces $\langle \ , \rangle$ on $\hat{L}^{(p)}$ under $\eta_1$ and of degree prime to $p$.

The open compact subgroup of $G(\mathbb{A}^{(\infty)})$ corresponding to this moduli problem is:

$$\hat{\Gamma}(N) = \hat{\Gamma}_L(N) = \{ x \in \hat{\Gamma}_L \mid (x - 1)\hat{L} = N\hat{L} \}.$$ 

Suppose that $N$ is sufficiently large so that the moduli problem has a solution, that is, we have a fine moduli scheme $M$.

We want to know the exact objects the generic fiber $M_\eta$ classifies. For a given quadruple $(X, i, \lambda, \eta^{(p)})$, if it has a generic fiber $X_\eta$, $T(X_\eta)$ can be embedded (as skew hermitian $O$-modules) into $V \otimes_{\mathbb{Q}} \hat{\mathbb{A}}^{(\infty)}$ so that the embedding coincides with $\eta$ at $\ell$ for each prime $\ell \mid p$. Thus we know the isomorphism class of the localizations of $H_1(X_\eta, \mathbb{Z})$ as skew hermitian $O$-lattice in $V$ outside $p$. Let $L'$ be the image of $H_1(X_\eta, \mathbb{Z})$ in $V$.

For any given $O$-lattice $\Lambda \subset V$, we define

- $\mu(\Lambda)$ to be the ideal of $\mathbb{Z}$ generated by $\langle x, x \rangle$ for all $x \in \Lambda$;
- $\mu_0(\Lambda)$ to be the $O_F$-ideal generated by $\langle x, y \rangle$ for all $x, y \in \Lambda$.

If $\Lambda$ is maximal among lattices having the same $\mu$, we call it maximal. By the self duality at $p$ of $L'$, $\mu_0(L'_p) = O_{F,p}$. It is easy to see that $\mu(\Lambda)O \subset \mu_0(\Lambda) \subset \mathfrak{d}^{-1}\mu(\Lambda)$ for the relative different $\mathfrak{d}$ of $F/\mathbb{Q}$. If $L''_p \supset L'_p$ with $\mu(L''_p) = \mu(L'_p)$, then by (ord), we find

$$\mu(L'_p)O_{F,p} \subset \mu_0(L'_p) \subset \mu_0(L''_p) \subset \mu(L''_p)O_{F,p}.$$ 

Thus $L'_p = L''_p$ and hence $L'_p$ is maximal with $\mu(L'_p) = \mathbb{Z}_p$. By the same argument, $L_p$ is maximal with $\mu(L_p) = \mathbb{Z}_p$. Then by a lemma of Shimura proven in the 1960’s ([EPE] 5.9 or [Ko] Corollary 7.3), there exists $x_p \in G_1(\mathbb{Q}_p)$ so that $L'_p = x_pL_p$. By self duality of $L_p$ and $L''_p$, we see that $x^2_p \in \hat{\Gamma}_p$ and hence $x_p \in \hat{\Gamma}_p$. Thus we find that $L'_p = L''_p$ as skew hermitian $O_{F,p}$-modules.

This shows that there are only finitely many isomorphism classes of hermitian $O$-lattices in the genus class of $L$ (approximation theorem). Thus the generic (geometrically) irreducible component of $M$ classifies $(X, i, \lambda, \eta_1^{(p)})$ satisfying the following conditions:

- $i : O \hookrightarrow \text{End}(X_S)$ taking 1 to $\text{id}_X$;
- $\eta_1^{(p)}$ is made of a pair of $\eta_1 : H_1(X_s, \mathbb{Z}) \cong L$ up to isomorphisms as skew hermitian $O$-modules for any geometric point $s \in S$ and $\eta_N : L/NL \cong X[N]$.

$\lambda : X \to \hat{X}$ is a polarization which induces $\langle \ , \rangle$ on $L$ under $\eta_1$.

This type of moduli problem has been studied over the reflex field $E$ by Shimura (see, for example, [ACM] Section 26, [AAF] Chapters I and II and [Sh2]). In the formulation of [AAF] Section 4, the above conditions are summarized into a PEL type: $\Omega = (V, \Psi, L, J_n, t_1, \ldots, t_{2n})$, where $t_j$ are generators of $L/NL$ over $O$ and $\Psi$ is the isomorphism class over $\mathbb{Q}$ of the representation of $F$ on $V_1 \subset V \otimes_{\mathbb{Q}} \mathbb{C}$ on which $h(\sqrt{-1})$ acts by the multiplication by $\sqrt{-1}$. A quadruple $(X, i, \lambda, \eta_1^{(p)})$ over $\mathbb{C}$ is called

---

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
of type $\Omega$ if we have a real analytic isomorphism $V_\infty \xrightarrow{\xi} X(\mathbb{C})$ with $\text{Ker}(\xi) = L$ such that

- $\xi$ induces an identification of $V_1 \cong \text{Lie}(X)$ as complex vector space on which $F$ acts by $\Psi$,
- $\xi$ induces the polarization $\langle \cdot, \cdot \rangle$ on $V$ (up to positive rational multiple). This means that $\langle x, y \rangle = \text{Tr}_{F/\mathbb{Q}}(xJ_n y^t)$,
- $\xi(tx) = i(a)\xi(t)$ for $t \in V$ and $a \in O$,
- $\overline{\Psi}: (O/NO)^{2n} \to X[N]$ given by $\overline{\Psi}(a_1, \ldots, a_{2n}) = \sum_j a_j \xi(t_j).

The condition on $\Psi$ is equivalent to $(\det)^n$ over $E = \mathbb{Q}$. We can think this moduli problem over $E$ for an arbitrary $N$ and get a tower of moduli space $M_N$. We now take $L$ to be $O^{2n}$. Then each geometrically irreducible component of $M_N$ is defined over $\mathbb{Q}[\zeta_N]$. The component $V_N/\mathbb{Q}[\mu_N] = V_{\tilde{T}^\ell(N)/\mathbb{Q}[\mu_N]}$ classifies quadruples $(X, i, \lambda, \eta)$ over $\mathbb{Q}[\mu_N]$ under an extra condition that $e_N(t_i, \tilde{t}_j) = \zeta_N^{N(t_i, \tilde{t}_j)}$ for the duality pairing $e_N: X[N] \times \tilde{X}[N] \to \mu_N$ and the dual base $t_j$ of $\tilde{t}_j$ under $\langle \cdot, \cdot \rangle$ localized at $N$. We then consider the union of the tower of fields $\mathbb{Q}[\mu_N](V_N) = \mathbb{Q}(V_N)$, and write the field as $\mathfrak{R}$. Naturally the group $x \in G(\mathbb{A}^{(\infty)})$ acts on $M = \varprojlim N M_N$ by changing $L$ to $xL$ and $t_j$ to $xt_j$, and if $x \in G_1(\mathbb{A}^{(\infty)})$, $x$ keeps $\langle \cdot, \cdot \rangle$. Let

$$H(A) = \{ x \in G(\mathbb{A}) \mid \det(x) = \nu(x)^n, \nu(x) \in \mathbb{A}^\times \}.$$ 

Then we have the following explicit description ([AAF] 8.8):

$G_+ = H(\mathbb{A}^{(\infty)})G(\mathbb{Q})_+G(\mathbb{R})_+ = (\tilde{T}^\ell(N) \cap H(\mathbb{A}^{(\infty)}))\iota(\tilde{\mathbb{Z}}^\times)G(\mathbb{Q})_+G(\mathbb{R})_+$,

where $G(\mathbb{R})_+$ is the identity connected component of $G(\mathbb{R})$, $G(\mathbb{Q})_+ = G(\mathbb{Q}) \cap G(\mathbb{R})_+$, and $\iota(s) = \text{diag}[1_n, s^{-1}1_n] \in H(\mathbb{A})$ for $s \in \mathbb{A}^\times$. To see this, we write $\psi = \det/\nu^n : G(\mathbb{A}) \to F_A^\times$. Then $H(\mathbb{A}) = \text{Ker}(\psi)$. By (Sh1), for a given $x \in G_+$, we have $\psi(x) = \zeta u$ for $u \in \mathbb{R}_+^\times$ and $\zeta = \xi^{1-\varepsilon}$ for $\xi \in M^\times$ ($\iff \zeta^\varepsilon = 1$). Taking $\alpha \in G(\mathbb{Q})_+$ with $\psi(\alpha) = \zeta$, we find that $\alpha^{-1-h} \in H(\mathbb{A})$, which shows the first equality of the above expression.

For the second equality, we refer the reader to [AAF] 8.8. Since $F$ is imaginary quadratic, it has only finitely many units; so $(F^\times)^{1-\varepsilon}\mathbb{R}_+^\times$ is a closed subgroup of $F_A^\times$, and hence $G_+$ is a closed subgroup of $G(\mathbb{A})$. This is the reason why we do not need to take closure of $(F^\times)^{1-\varepsilon}\mathbb{R}_+^\times$ in $F_A^\times$ in the definition of $G_+$ in (Sh1).

For $p \nmid N$, we have

$$\text{Gal}(\mathbb{Q}[\mu_{Np^n}, V_{Np^n}]/\mathbb{Q}[\mu_{Np^n}, V_N]) \cong \Gamma(N/p^n)/\Gamma(N) \cong \text{SG}(\mathbb{Z}/p^n\mathbb{Z})$$

for $N$ sufficiently large by (Sh3,5) (and the strong approximation theorem). Here $\text{SG}(A) = SL_2n(A \otimes \mathbb{Z}) \cap G_A(1)$ and

$$\Gamma(N) = \{ \gamma \in \text{SG}(\mathbb{Q}) \mid \gamma L = L \text{ and } (\gamma - 1)L \subset NL \forall L \forall \ell \mid \ell \nmid N \}.$$ 

The moduli variety $M_{Np^n}$ classifies quadruple $(X, \overline{\Psi}, \overline{\Phi}) : (O/p^nO)^{2n} \to X[p^n]$ for a level $\tilde{T}^\ell(N)$-structure $\overline{\Psi}$. Thus taking the universal abelian variety $X$ over $M_N$, we
have, for each $\text{Sh}_{\Gamma_L(N)}$-scheme $T$

$$M_{N_{p^n}}(T) = \{ \phi : (O/p^nO)_{/T}^{2n} \cong X[p^n]/T \}.$$ 

The action of $g \in G(\mathbb{Z}_p)$ on $M_{N_{p^n}}$ is induced by the action on the level structure $\phi \mapsto \phi \circ g$. Thus taking $2n$-th exterior power $\wedge^{2n} X[p^n]$, we find that $g \in G(\mathbb{Z}_p)$ acts by $\det(g) \in O_F^\times$. Then by the description of the stabilizer $G_+$ of a geometrically irreducible component $Sh$, we find that the action is trivial on $\wedge^{2n} X[p^n]$ if $g \in G_+ \cap SG(\mathbb{A}^{(\infty)})$. Thus the group scheme

$$\wedge^{2n}_O X[p^n] = (\wedge^{2n} X[p^n]) \oplus (\wedge^{2n} X[\mathcal{F}])$$

is constant on each geometrically irreducible component $V_N$ of $M_N$. In other words, $\wedge^{2n} X[p^n]_{/V_N}$ is a base-change of $\mu_{p^n} \times (\mathbb{Z}/p^n\mathbb{Z})^n$ from $\overline{\mathbb{Q}}$ to $V_{N/\overline{\mathbb{Q}}}$.

Now we look into the Kottwitz model $\text{Sh}^{(p)}_{\Gamma(N)/\mathbb{W}}$, for $N$ prime to $p$. In the rest of the paper, we always suppose that $N$ is prime to $p$. Since each geometrically irreducible component of $\text{Sh}^{(p)}_{\Gamma(N)/\mathbb{W}} = M_{N/\mathbb{W}}$ is defined over $\mathbb{W}[\mu_N]$ in the sense of Weil, it remains irreducible after taking spacial fiber modulo $\mathfrak{m}W$ (Zariski’s connectedness theorem combined with the existence of a smooth projective compactification). Thus we can talk about geometrically irreducible component $V_{N/W}^{(p)}$ of $\text{Sh}^{(p)}_{\Gamma(N)/\mathbb{W}}$, whose generic fiber is $V_{N/\mathbb{W}}$ and whose special fiber is the special fiber of the schematic closure of $V_N$ in $\text{Sh}^{(p)}_{\Gamma(N)/\mathbb{W}}$.

Since the universal abelian scheme $\mathcal{X}_{/V_N}$ extends to the universal abelian scheme of the Kottwitz model $V_{N/\mathbb{W}}^{(p)}$,

$$Q_m = \wedge^{2n}_O X[p^n]_{\text{et}} = (\wedge^{2n} X[p^n]_{\text{et}}) \otimes (\wedge^{2n} X[\mathcal{F}]_{\text{et}})$$

is constant over $S^\circ_{1/\mathbb{F}}$. Since the Igusa tower $T_m$ over $S^\circ_{1/\mathbb{F}}$ is given by

$\text{Isom}((O/p^nO)_{/V_{\mathbb{W}}}, X[p^n]_{\text{et}})$,

$T_m$ cannot be irreducible, and each irreducible component of $T_m/\mathbb{F}$ is contained in $T_m(t_m)$ for an isomorphism $t_m : \mathbb{Z}/p^n\mathbb{Z} \cong \wedge^{2n} X[p^n]_{\text{et}} = Q_m$. Thus, for a geometrically irreducible component $T_\infty$ of $T_{\infty}$, the Galois group $\text{Gal}(T_\infty[T^\circ_{1/\mathbb{F}}])$ is a subgroup of $S$ in (10.1). We reached the same conclusion before stating Corollary 10.2 by looking into vector bundles $\mathcal{V}_m$. In any case, we need to show that

$$\text{Gal}(T_\infty[T^\circ_{1/\mathbb{F}}]) = \mathcal{G}$$

to prove Corollary 10.2.

Since $p$ splits in $\mathbb{F}$, we have $SG(\mathbb{Z}/p^n\mathbb{Z}) \cong SL_{2n}(\mathbb{Z}/p^n\mathbb{Z})$. Since we have a smooth model of $M_N$ over $\mathbb{W}$, we take the valuation $v$ of $\mathfrak{a}_N = \mathbb{Q}[\mu_N](V_N)$ corresponding to the generic point of $V_N$ mod $\mathfrak{P} = V_N \otimes_\mathbb{W} \mathbb{F}$ containing the infinity cusp. Since $M_N^* = \text{Proj}([G_{\Gamma(N)}]$) under the notation in Section 6, we can write the Satake compactification.
of $V_{N/W}$ as $\text{Proj}(R)$ for $R = \bigoplus_{j \geq 0} R_j$ with $R_j = H^0(V_N; \det(\omega)^j/W)$. By $q$-expansion at $\infty$, we can embed $R$ into $W[[q^\xi]]_{\xi \in M_n(F)_+}$, where

$$M_n(F)_+ = \{ [a^c = x \in M_n(F) \mid x \text{ is totally non-negative}] \}.$$ 

and the symbol $A[[q^\xi]]_{\xi \in M_n(F)_+}$ indicates the completion by the augmentation ideal of the monoid algebra of the additive semi-group $M_n(F)_+$ with $q^\xi$ indicating the element represented by $\xi$. Each $f \in R_j$ has $q$-expansion $\sum_\xi a(\xi; f) q^\xi \in W[[q^\xi]]_{\xi \in M_n(F)_+}$. Replacing $q^\xi$ by $\exp(2\pi i \text{Tr}(\xi z))$, $z \in \mathbb{F}$, we get the Fourier expansion at $\infty$ of $f$ (regarding $W \hookrightarrow \mathbb{C}$).

We take a valuation $v$ of $\mathfrak{A}$ which is induced by a valuation $\nu$ on $R$ given by

$$v(\sum_\xi a(\xi; f) q^\xi) = \inf_\xi \text{ord}_p(a(\xi; f)), $$

where $\text{ord}_p$ is the discrete valuation of $W$ with $\text{ord}_p(p) = 1$. Here we used the existence of the smooth toroidal compactification of $V_N$ ($p \nmid N$) worked out by Fujiwara ([F]) and the $q$-expansion principle for $f \in R_j$ on $\Gamma(N)$ with $p \nmid N$ to assure that the residue field of $\nu$ restricted to $\Gamma(N)$ for $p \nmid N$ is the function field of $V_N$ mod $\mathfrak{A}$. Since the Satake compactification of $M_{NP^m}/\mathfrak{A}$ is again given by $\text{Proj}(R)/\langle \mu_{NP^m} \rangle$ for $R = \bigoplus_{j \geq 0} R_j$ with $R_j = H^0(M_{NP^m}/\mathfrak{A}, \det(\omega)^j/\mathfrak{A})$, we can extend the valuation $v$ to $R$ by the same formula in terms of the unique extension of $\text{ord}_p$ to $W[\mu_{NP^m}]$. This extension induces a valuation on $Q(V_{NP^m}) = \mathfrak{R}_{NP^m}$ and on $\mathfrak{A} = \bigcup_m \mathfrak{R}_{NP^m}$. We are going to show that the decomposition group $D_\nu$ of $\nu$ in $\text{Aut}(\mathfrak{A})$ contains $P_n(\mathfrak{A}) \cap \mathcal{G}_+$ and $G_1(\mathfrak{A}(\mu_{NP^m}))$.

10.2.2. Proof of the irreducibility theorem. — Let $L \subseteq V$ be an $O$-lattice satisfying (L1-2) of Section 6 and recall

$$\hat{\Gamma}_L = \{ x \in G(\mathfrak{A}(\infty)) \mid x\hat{L} = \hat{L} \},$$

$$\widehat{\Gamma}_L(N) = \{ x \in \hat{\Gamma}_L \mid (x - 1)\hat{L} = N\hat{L} \},$$

where $\hat{L} = L \otimes Z \hat{\mathbb{Z}}$. Let $X_N$ be the universal abelian scheme over $V_N \subset \text{Shf}(N)$ for $N$ sufficiently large. We have the following specification of the action of $\mathcal{G}_+$ (see [AAF] Theorem 8.10):

(1) $x \in \mathcal{G}_+$ acts on the maximal abelian extension $Q_{ab}$ of $Q$ by the image of $\nu(x)$ under the reciprocity map of class field theory.

(2) If $\gamma \in G(\mathbb{Q})_+ = G(\mathbb{Q}) \cap G(\mathfrak{A}_+)$, $\gamma$ regarded as an element of $G(\mathbb{Q})_+ \subset \mathcal{G}_+$ satisfies $f^{\gamma}(\gamma) = f \circ \gamma$.

(3) If $x = \left( \begin{array}{cc} 1 & 0 \\ 0 & s_n \end{array} \right)$ with $s \in \hat{\mathbb{Z}}^\times$, then $a(\xi; f^{\gamma}(x)) = \sigma(a(\xi; f))$ for all $\xi \in M_n(F)_+$, where $\sigma$ is the automorphism of $Q_{ab}$ corresponding to $s^{-1}$ by class field theory. Strictly speaking, writing $f = h/g$ for $g, h \in R_j$, we have $f^{\gamma}(x) = h^\sigma/g^\sigma$ with $a(\xi; x^\sigma) = \sigma(a(\xi; x))$.

(4) The natural action of $\hat{\Gamma}_L(N)\langle p \rangle$ on $V_{NP^m} = \text{Isom}(L/tL/V_N, X_N[t]/V_N)$ induces the action of $\hat{\Gamma}_L(N)\langle p \rangle \cap \mathcal{G}_+$ on $E(V_{\hat{\Gamma}_L(N)t})$. 


By (1), $\mathcal{D} = D_v$ is contained in the image (under $\tau$) of
\[ \{ x \in \mathcal{G}_+ \mid \nu(x) \in \mathbb{Q}_p^\times \mathbb{Q}^\times \mathbb{R}^\times \}. \]

Let $U_n$ be the unipotent radical of $P_n$. By (2), we have $\tau(P_n(\mathbb{Q})) \subset \mathcal{D}$, since $(\frac{a}{d}) \in P_n(\mathbb{Q})$ acts on $q$-expansion just by $q^\xi \mapsto \exp(2\pi i \text{Tr}(\xi d^{-1}))q^{-1} \xi^a$. Then by density of $\mathbb{Q}$ in $\mathbb{A}$, we conclude that
\[ \tau(U_n(\mathbb{A})) \subset \mathcal{D}. \]

By the strong approximation theorem, cusps of $\Gamma_K = K \cap SG(\mathbb{Q})$ are in bijection to $K \setminus SG(\mathbb{A}^{(\infty)})/P_n(\mathbb{Q})$. Choosing $K$ to be maximal, by Iwasawa decomposition: $SG(\mathbb{A}^{(\infty)}) \subset KP_n(\mathbb{A}^{(\infty)})$, the above set of cusps is embedded into $K \cap P_n(\mathbb{A}^{(\infty)})\setminus P_n(\mathbb{A}^{(\infty)})/P_n(\mathbb{Q})$. We have $G(\mathbb{A}^{(\infty)}) = \bigcup_{b \in B} G(\mathbb{Q})bK$ for a finite set $B \subset P_n(\mathbb{A}^{(p)})$. From this, $K \setminus P_n(\mathbb{A}^{(\infty)}) \cap SG(\mathbb{A}^{(p)})/P_n(\mathbb{Q})$ is a finite set indexed by ideal classes. Thus the action of $b \in P_n(\mathbb{A}^{(p)})$ brings the $q$-expansion of $f \in \mathcal{R}_j$ to its $q$-expansion at other cusps. If $K_p$ is maximal, $V_K$ is smooth over the valuation ring of $\mathfrak{p}$, and hence the action preserves $v$ restricted to $\beta(H^0(V_K, \omega_\mathfrak{p})) (p \nmid N)$, where $\beta$ is the embedding into the space $V_U$ of $p$-adic modular functions (with respect to $\hat{\Gamma}(N)$) we studied in Section 7. Note that the integral closure $\bar{R}$ of the graded algebra $R(V_K = \text{Proj}(R))$ in $\mathcal{R}_U(\bar{L})$ for $U(L) = U_n(\mathbb{Z}_p) \times \hat{\Gamma}(N)/(p)$ is contained in $V_U$ by definition.

Let $\omega_\mathfrak{p} \subset \omega_\mathfrak{p}^{\infty}$ be the sheaf of cuspidal forms. Since $D_{\text{cusp}} = \left( \bigoplus_K H^0(V_K, \omega_\mathfrak{p}/\mathbb{Q}) \right) \cap V_{\text{cusp}}^U$ is $p$-adically dense in $V_{\text{cusp}}^U$ (the density theorem in Section 7), we conclude that the action of $P_n(\mathbb{A}^{(p)}) \cap P$ preserves $v$ restricted to $\mathfrak{p}(p) = \bigcup_{\mathfrak{p} | N} \mathbb{Q}_\mathfrak{p}(V_N)$ and also $\mathfrak{r}_{U_n}(\mathfrak{p}) = \bigcup_{\mathcal{L}} \mathfrak{r}_U(L) \subset \mathfrak{r}(p)$ for the unipotent radical $U_n$ of $P_n$, because $\mathfrak{r}_U(L)$ is generated by ratios $f/g$ of cuspidal forms $f$ and $g$ in $\bar{R} \supset D_{\text{cusp}}$. Here $L$ runs over all lattices satisfying (L1-2) in Lecture 6. Thus $\mathcal{D} \cdot \tau(U_n(\mathbb{Z}_p)) / \tau(U_n(\mathbb{Z}_p)) \subset \text{Aut}(\mathfrak{r}_{U_n}(\mathfrak{p}))$ contains the image of $(P_n(\mathbb{A}^{(p)}) \cap P) \times U_n(\mathbb{Z}_p)$ in $\text{Aut}(\mathfrak{r}_{U_n}(\mathfrak{p}))$.

Then by (U), we conclude $\mathcal{D}$ contains the image under $\tau$ of
\[ P \cap (P_n(\mathbb{A}^{(p)}) \times U_n(\mathbb{Z}_p)). \]

By the same argument applied to $K^{(p)}$, we find that $\tau(P \cap K^{(p)}) \subset \mathcal{D}$. Note that
\[ P \cap \left( \bigcup_L P_n(\mathbb{A}^{(p)}) \cap P_n(\mathbb{Q}) \hat{\Gamma}_L(N)/(p) \right) \]

is dense in $P$ and hence $\mathcal{D} \supset \tau(P)$ (see the proof of Theorem 6.27 in [PAF] for a different argument giving this inclusion).

Since $g = \left( \begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix} \right) \in P \cap G(\mathbb{Z}_p)$ acts on $T_m$ through its diagonal entries $(a, d)$, we find that $\text{Gal}(T_m^\infty/S_1^\infty)$ has to contain $\mathfrak{S}$ because the matrices $(a, d)$ fills $\mathfrak{S}$. This proves Corollary 10.2.

The reverse inclusion: $\tau(P) \supset \mathcal{D}$ follows from the existence of the exact sequence:
\[ 0 \longrightarrow \mathcal{X}[p^\infty]_0/S_1^\infty \longrightarrow \mathcal{X}[p^\infty]/S_1^\infty \longrightarrow \mathcal{X}[p^\infty]_1/S_1^\infty \longrightarrow 0. \]
See [PAF] Theorem 6.28 and 8.4.3 for more details of how to prove the reverse inclusion from the above exact sequence. This finishes the proof of Theorem 10.1 in Case A.

**References**

**Books**


**Articles**


H. HIDA, Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, U.S.A.
E-mail : hida@math.ucla.edu


NEWTON POLYGONS AND $p$-DIVISIBLE GROUPS:
A CONJECTURE BY GROTHENDIECK

by

Frans Oort

Abstract. — In my talk in 2000 I discussed a conjecture in 1970 by Grothendieck concerning deformations of $p$-divisible groups; a proof of this conjecture give access to finding properties of Newton polygon strata in the moduli spaces of polarized abelian varieties in positive characteristic.

Résumé (Polygones de Newton et groupes $p$-divisibles: une conjecture de Grothendieck)
En 1970 Grothendieck a formulé une conjecture concernant les déformations de groupes $p$-divisibles (groupes de Barsotti-Tate). Nous décrivons une démonstration de cette conjecture. Cela donne une information sur des strates définies par le polygone de Newton dans les espaces de modules des variétés abéliennes en caractéristique positive.

Introduction

0.1. We consider $p$-divisible groups (also called Barsotti-Tate groups) in characteristic $p$, abelian varieties, their deformations, and we draw some conclusions.

For a $p$-divisible group (in characteristic $p$) we can define its Newton polygon. This is invariant under isogeny. For an abelian variety the Newton polygon of its $p$-divisible group is "symmetric". We are interested in the strata defined by Newton polygons in local deformation spaces, or in the moduli space of polarized abelian varieties.

2000 Mathematics Subject Classification. — 14K15, 14L05, 14G15, 14G35.
Key words and phrases. — Abelian varieties, Newton polygons, $p$-divisible groups, Barsotti-Tate groups, moduli spaces, Grothendieck conjecture.

Talk on 27-IV-2000 in the “automorphic semester”, Centre Émile Borel at Institut Henri Poincaré.

© Astérisque 298, SMF 2005
0.2. Grothendieck showed that Newton polygons “go up” under specialization, see [4], page 149, see [11], Th. 2.3.1 on page 143; we obtain Newton polygon strata as closed subsets in the deformation space of a $p$-divisible group or in the moduli space of polarized abelian varieties in positive characteristic.

In 1970 Grothendieck conjectured the converse. In [4], the appendix, we find a letter of Grothendieck to Barsotti, and on page 150 we read: “...The wishful conjecture I have in mind now is the following: the necessary conditions [...] that $G'$ be a specialization of $G$ are also sufficient. In other words, starting with a BT group $G_0 = G'$, taking its formal modular deformation [...] we want to know if every sequence of rational numbers satisfying [...] these numbers occur as the sequence of slopes of a fiber of $G$ as some point of $S$.”

0.3. In this talk we study this conjecture by Grothendieck for $p$-divisible groups, for abelian varieties, for quasi-polarized $p$-divisible groups and for polarized abelian varieties. Then we draw conclusions for NP-strata. These results can be found in [8, 21, 23].

0.4. We give a proof of this conjecture by Grothendieck. This is done by combining various methods (below we explain the string of ideas leading to this proof); in various stages of the process we need quite different ideas and methods. Hence, in spirit, the proof of a straight statement is not uniform. We have not been able to unify these in one straightforward method. We wonder what Grothendieck would have substituted for our proof.

1. Notations

1.1. We fix some notations. All base fields will be of characteristic $p > 0$. The $p$-divisible group of an abelian variety $X$ will be denoted by $X[p^\infty]$. We will use covariant Dieudonné modules.

We follow [15] by writing $G_{m,n}$ for the following $p$-divisible group (defined over $\mathbb{F}_p$, and considered over every field of characteristic $p$): this is a $p$-divisible group of dimension $m$, with Serre-dual of dimension $n$; here $m, n \in \mathbb{Z}_{\geq 0}$ are coprime integers; we have $G_{1,0} = G_m[p^\infty]$, and we write $G_{0,1}$ for its Serre dual; for coprime $m, n \in \mathbb{Z}_{> 0}$ the formal $p$-divisible group $G_{m,n}$ is given in the covariant Dieudonné module theory by

$$\mathbb{D}(G_{m,n}) = W[[F,V]]/W[[F,V]] \cdot (F^m - V^n)$$

(in this ring $W[[F,V]]$ we have the relations $FV = p = VF$ and for all $a \in W = W_\infty(K)$, where $K$ is a perfect field, we have $Fa = a^\sigma F$ and $Va^\sigma = aV$; in case $K = \mathbb{F}_p$ this results in a commutative ring). We use $H_{m,n}$ as in [9], 5.3; this is a $p$-divisible group isogenous with $G_{m,n}$; it can be characterized by saying that moreover its endomorphism ring over $\mathbb{F}_p$ is the maximal order in its endomorphism algebra.
We need some combinatorial notation concerning Newton polygons:

Throughout the paper we fix a prime number $p$. We apply notions as defined and used in [22], and in [9]. For a $p$-divisible group $G$, or an abelian variety $X$, over a field of positive characteristic we use its Newton polygon, abbreviated by NP, denoted by $N(G)$, respectively $N(X)$. For dimension $d$ and height $h = d + c$ of $G$ (respectively dimension $g = d = c$ of $X$) this is a lower convex polygon in $\mathbb{R} \times \mathbb{R}$ starting at $(0,0)$ ending at $(h,c)$ with integral break points, such that every slope is non-negative and at most equal to one. We write $\beta \prec \gamma$ if every point of $\gamma$ is on or below $\beta$ (the locus defined by $\gamma$ contains the one defined by $\beta$). For further details we refer to [22]. For example, the Newton polygon $N(G_{m,n})$ consists of $m + n$ slopes equal to $n/(m + n)$.

1.2. We use the following notation: we fix integers $h \geq d \geq 0$, and we write $c := h - d$. We consider Newton polygons ending at $(h,c)$. For a point $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ we write $(x, y) \prec \beta$ for the property "the point $(x, y)$ is on or above the Newton polygon $\beta$". For a Newton polygon $\beta$ we write:

$$T(\beta) = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} | y < c, y < x, (x, y) \prec \beta\},$$

and we define

$$\dim(\beta) := \#(T(\beta)).$$

Note that for the "ordinary" Newton polygon $\rho := d \cdot (1,0) + c \cdot (0,1)$ the set of points $T = T(\rho)$ is a parallelogram; this explains our notation. Note that $\#(T(\rho)) = d \cdot c$.

1.3. We fix an integer $g$. For every symmetric Newton polygon $\xi$ of height $2g$ we define:

$$\Delta(\xi) = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} | y < x \leq g, (x, y) \prec \xi\},$$

and we write

$$\text{sdim}(\xi) := \#(\Delta(\xi)).$$

For the ordinary symmetric Newton polygon $\rho = g \cdot ((1,0) + (0,1))$ indeed $\Delta = \Delta(\rho)$ is a triangle; this explains our notation. But you can rightfully complain that the "triangle" $\Delta(\xi)$ in general is not a triangle. Note that $\#(\Delta(\rho)) = g(g + 1)/2$. 
1.4. A theorem by Grothendieck and Katz, see [12], 2.3.2, says that for any family $\mathcal{G} \to S$ of $p$-divisible groups (in characteristic $p$) and for any Newton polygon $\gamma$ there is a unique closed set $W \subset S$ containing all points $s$ at which the fiber has a Newton polygon equal to or lying above $\gamma$:

$$s \in W \iff N(\mathcal{G}_s) \prec \gamma.$$ 

This set will be denoted by

$$W_\gamma(\mathcal{G} \to S) \subset S.$$

In case of symmetric Newton polygons we write

$$W_\gamma(A_g \otimes \mathbb{F}_p) =: W_\gamma$$

for the Newton polygon stratum given in the moduli space of polarized abelian varieties in characteristic $p$. We will study this mainly inside $A := A_{g,1} \otimes \mathbb{F}_p$, the moduli space of principally polarized abelian varieties in characteristic $p$.

1.5. We study formal abelian schemes, and formal $p$-divisible groups over formal schemes, and we study abelian schemes and $p$-divisible groups. Without further comments we use the following ideas.

*Formal $p$-divisible groups.* — As finite group schemes are “algebraizable”, the same holds for certain limits; if $\mathcal{G} \to \text{Spf}(A)$ is a formal $p$-divisible group, it comes from a $p$-divisible group over $\text{Spec}(A)$, see [6], 2.4.4. We use the passage from formal $p$-divisible groups over $\text{Spf}(A)$ to $p$-divisible groups over $\text{Spec}(A)$ without further comments (here $A$ is a complete local ring).

*Serre-Tate theory.* — Suppose given an abelian variety $X_0$ over a field $K$, and its $p$-divisible group $G_0 := X_0[p^\infty]$. A theorem by Serre and Tate gives an equivalence between formal deformations of (polarized) abelian schemes and the corresponding (quasi-polarized) $p$-divisible groups, see [12], Th. 1.2.1: any formal deformation of $G_0$ induces uniquely a formal deformation of $X_0$. 

\[\text{ASTÉRISQUE 298}\]
Formal abelian schemes. — In general a formal abelian scheme $\mathcal{A} \to \text{Spf}(A)$ is not algebraizable. However polarized abelian schemes can algebraized: by the Chow-Grothendieck algebraization method for polarized formal schemes (“formal GAGA”), see [5], III 1.5.4, it follows that from a polarized formal abelian scheme $(X, \mu) \to \text{Spf}(A)$ we obtain an actual polarized abelian scheme over $\text{Spec}(A)$.

1.6. Suppose $X \to S$ is a scheme over an integral scheme $S$. Let $\eta \in S$ be its generic point, and let $0 \in S$ be a closed point. In this situation we will say that “$X_\eta$ specializes to $X_0$”, and we say that “$X_0$ deforms to $X_\eta$” sometimes without specifying the base scheme $S$ and the family $X \to S$.

1.7. Displays. — Given a Dieudonné module $M$ of a $p$-divisible group, and a $W$-base for the $W$-free module, the map $F : M \to M$ is given by a matrix, called a display. Mumford showed that deformations of certain $p$-divisible groups can be given by writing out a display over a more general base ring. What we need is contained in [30], [29]; also see [17], [18]. Below we construct deformations of local-local $p$-divisible groups. We shall write out the display, and use several times (without further mention) that this defines a deformation, see [30], Chapter 3, in particular his Corollary 3.16. Deformations of polarized formal $p$-divisible groups can be described with the help of displays, see [18], Section 1.

1.8. For an abelian variety $X$ over a field $K$ we write $f = f(X)$ for its $p$-rank, i.e. the integer such that $\text{Hom}(\mu_p, X \otimes k) \cong (\mathbb{Z}/p)^f$, where $k = \overline{K}$.

For a group scheme $N$ over a field $K \supset \mathbb{F}_p$ we write $a(N) = \dim_L \text{Hom}(\alpha_p, N \otimes L)$, where $L \supset K$ is a perfect field containing $K$.

Note that there exist examples in which

$$\dim_K(\text{Hom}(\alpha_p, G)) < \dim_L(\text{Hom}(\alpha_p, G_L)).$$

However, if $a(G) = 1$, then $\dim_K(\text{Hom}(\alpha_p, G)) = 1 = \dim_L(\text{Hom}(\alpha_p, G_L))$.

Note that $\text{Hom}(\alpha_p, G) \neq 0$ iff the local-local part of $G$ is non-trivial, i.e. iff $G$ is not ordinary. Hence if we write $a(G) \leq 1$ we intend to say: either $G$ is ordinary, or $a(G) = 1$.

We use the notation $k$ for an algebraically closed field (of characteristic $p$).

2. Results: deformations of $p$-divisible groups

2.1. Theorem (conjectured by Grothendieck, Montreal 1970). — Let $K$ be a field of characteristic $p$, and let $G_0$ be a $p$-divisible group over $K$. We write $N(G_0) := \beta$ for its Newton polygon. Suppose given a Newton polygon $\gamma$ “below” $\beta$, i.e. $\beta \prec \gamma$. Then there exists a deformation $G_\eta$ of $G_0$ such that $N(G_\eta) = \gamma$. 

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
We write Def($G_0$) for the universal deformation space (in equal characteristic $p$) of a $p$-divisible group $G_0$.

2.2. Theorem (properties of Newton polygon strata). — Suppose given a $p$-divisible group $G_0$ over a field $K$. Let $\gamma$ be a Newton polygon with $\gamma \succ N(G_0) =: \beta$. Consider the closed formal subset $W_\gamma(\text{Def}(G_0)) =: V_\gamma \subset \text{Def}(G_0)$. The dimension of every component of $V_\gamma$ equals $\dim(\gamma) = \#(\triangle(\gamma))$ and generically on every component of $V_\gamma$ the Newton polygon is $\gamma$ and the $a$-number generically is at most one.

(In fact on $V_\gamma$ the $a$-number generically is equal to one iff $\gamma \neq \rho := d \cdot (1, 0) + c \cdot (0, 1)$.)

3. Results: deformations of polarized $p$-divisible groups and of abelian varieties

3.1. Theorem (the principally polarized analog of the conjecture by Grothendieck)

Let $K$ be a field of characteristic $p$, and let $(G_0, \lambda_0)$ be a principally quasi-polarized $p$-divisible group over $K$. We write $N(G_0) = \beta$ for its Newton polygon. Suppose given a symmetric Newton polygon $\gamma$ “below” $\beta$, i.e. $\beta \prec \gamma$. Then there exists a deformation $(G_\eta, \lambda)$ of $(G_0, \lambda_0)$ such that $N(G_\eta) = \gamma$.

3.2. Corollary. — Let $K$ be a field of characteristic $p$, and let $(X_0, \lambda_0)$ be a principally polarized abelian variety over $K$. We write $N(X_0) = \beta$ for its Newton polygon. Suppose given a symmetric Newton polygon $\gamma$ “below” $\beta$, i.e. $\beta \prec \gamma$. Then there exists a deformation $(X_\eta, \lambda)$ of $(X_0, \lambda_0)$ such that $N(X_\eta) = \gamma$.

Indeed, using the Serre-Tate theorem we deduce this corollary from the previous theorem.

3.3. Theorem. — Suppose given a principally quasi-polarized $p$-divisible group $(G_0, \lambda_0)$ over a field $K$. Let $\gamma$ be a symmetric Newton polygon with $\gamma \succ N(G_0) =: \beta$. Consider the closed formal subset $W_\gamma(\text{Def}(G_0, \lambda_0)) =: V_\gamma \subset \text{Def}(G_0, \lambda_0)$. The dimension of every component of $V_\gamma$ equals $\text{sdim}(\gamma) = \#(\triangle(\gamma))$ and generically on every component of $V_\gamma$ the Newton polygon is $\gamma$ and the $a$-number generically is at most one.

3.4. Theorem (see [20]). — For every $p$, and $g$ and every symmetric Newton polygon $\beta$ we have:

(a) For every irreducible component $W$ of $W_\beta := W_\beta(A) \subset A_{g,1} \otimes \mathbb{F}_p$, we have $N(-, W \subset A) = \beta$ and $a(-, W \subset A) \leq 1$, i.e. generically on $W$ the Newton polygon of $X \to W$ equals $\beta$, and generically the $a$-number is at most one.

(b) The dimension of every irreducible component $W$ of $W_\beta$ equals $\text{sdim}(\beta) = \#(\triangle(\beta))$. 

ASTÉRISQUE 298
3.5. **Corollary (a conjecture by Manin, see [15], p. 76).** — Suppose given a prime number $p$ and a symmetric Newton polygon $\xi$. Then there exists an abelian variety $X$ defined over $\mathbb{F}_p$ with $N(X) = \xi$.

This was proved in the Honda-Serre-Tate theory (via reduction modulo $p$ of a well-chosen CM abelian variety in characteristic zero, see [28], page 98). Here we give a proof which uses only geometry in characteristic $p$.

3.6. **Remark.** — Theorem 3.4 applied to the locus $\mathcal{S}_{g,1}$ of principally polarized supersingular abelian varieties says that this locus is pure of dimension $\triangle(\sigma)$; this number turns out to be equal to $\triangle(\sigma) = [g^2/4]$ (were $[\ ]$ indicates the integral part); moreover on every component of $\mathcal{S}_{g,1}$ the $a$-number generically equals one. These results where conjectured by T. Oda and the present author, see [19], pp. 615/616. These results where proved by T. Katsura and the present author for $g=3$, see [10], and by K.-Z. Li and the present author for all $g$; for results and references see [14]. The method described here, and given in the three publication mentioned in the introduction, provides a new, different proof for these results.

3.7. **Remark.** — In 3.1 and in 3.2 the condition that the (quasi-)polarization is principal is essential. In fact, in [10], 6.10.b, we find an example of a component $V$ of dimension 3 of the moduli space of polarized abelian 3-folds, such that every point of $V$ corresponds with a supersingular abelian variety; we see that for such a point $v = [(X_0, \lambda_0)]$ with $N(X_0) = 3 \cdot (1, 1)$, i.e. $X_0$ is supersingular, there is no deformation as polarized abelian variety to $(X, \lambda)$ with $N(X) = (2, 1) + (1, 2)$. In [10], 6.10.c we find an example where generically the $a$-number is not one on an irreducible component of a supersingular Newton polygon stratum. More examples can be found in [14], Section 10 and Section 12. In fact it seems that much more is true, that “many” counter-examples to the analog of Grothendieck’s conjecture can be given in the non-principally polarized case, see 7.8.

4. **Methods: deformations to $a \leq 1$**

4.1. **Theorem (the “Purity theorem”).** — If in a family of $p$-divisible groups (say, over an irreducible scheme) the Newton polygon jumps, then it already jumps in codimension one.

See [9], Th. 4.1. This very non-trivial result will be one of the main tools.

4.2. **Catalogues.** Let us fix a prime number $p$, and coprime $m, n \in \mathbb{Z}_{>0}$. We try to “classify” all $p$-divisible groups isogenous with $G_{m,n}$.

In general there is no good theory of moduli spaces for $p$-divisible groups (and there are various ways to remedy this). We use the (new) notion of a “catalogue”. In our case this is a family $\mathcal{G} \to S$, i.e. a $p$-divisible group over some base scheme $S$, such
that every $G \sim G_{m,n}$ defined over an algebraically closed field appears as at least one geometric fiber in $G \to S$. You can rightfully complain that this is a rather vague notion, that a catalogue is not unique (e.g. the pull back by a surjective morphism again is a catalogue), etc. However this notion has some advantages:

**4.3. Theorem.** — Suppose given $p, m, n$ as above. There exists a catalogue $G \to T$ over $\mathbb{F}_p$ for the collection of $p$-divisible groups isogenous with $G_{m,n}$ such that $T$ is geometrically irreducible.

See [9], Theorem 5.11.

**4.4. Theorem.** — Suppose $G_0$ is a $p$-divisible group; there exists a deformation to $G_\eta$ such that $N(G_0) = N(G_\eta)$ and $a(G_\eta) \leq 1$.

**4.5.** We sketch a proof of 4.3, using 4.1, see [9]. We write $H = H_{m,n}$. We write $r := (m-1)(n-1)/2$. We see that for every $G \sim G_{m,n}$ there exists an isogeny $\varphi : H \to G$ of degree exactly $\text{deg}(\varphi) = p^r$, see [9], 5.8. We construct $G \to T$ as the representing object of isogenies $\varphi : H \times S \to G/S$ of this degree (it is easy to see that such a functor is representable).

Using this definition we see that the formal completion at $[(G_0,\varphi)] = s \in T$ embeds in $\text{Def}(G_0)$, i.e. $T^{\text{\text{\text{n}}}} \to \text{Def}(G_0)$. Furthermore we compute the longest chain of Newton polygons between $N(G_{m,n})$ and the ordinary one: this equals $mn - r$ (an easy combinatorial fact). From these two properties, using 4.1, we deduce: every component of $T$ has dimension at least $r$.

We make a stratification of $T$ (using combinatorial data, such a thing like “semi-modules”). We show (using explicit equations) that every stratum is geometrically irreducible, and that there is one stratum, characterized by $a(G) = 1$, of dimension $r$, and that all other strata have dimension less than $r$. These considerations do not contain deep arguments, but the proofs are rather lengthy and complicated.

From these two aspects the proof follows: any component of $T$ on which generically we would have $a > 1$ would have dimension strictly less than $r$, which contradicts “Purity”. Hence the locus where $a = 1$ is dense in $T$, and we see that $T$ is geometrically irreducible.

**4.6.** We sketch a proof of 4.4, see [24]. By 4.3 we conclude this deformation property 4.4 for iso-simple groups. Then we study groups filtered by iso-simple subfactors, and deformation theory of such objects. By the previous result we can achieve a deformation where all iso-simple subfactors are deformed within the isogeny class to $a \leq 1$. Then we write down an explicit deformation (“making extensions between iso-simple subfactors non-trivial”) in order to achieve $a(G_\eta) \leq 1$, see [24], Section 2 for details.
4.7. Remark. — This method of catalogues for $p$-divisible groups works fine for simple groups. However the use of “catalogues” for non-isoclinic groups does not seem to give what we want; it is even not clear that nice catalogues exist in general. Note that we took isogenies of the form $\varphi: H \times S \to G/S$; however over a global base scheme monodromy groups need not be trivial, and this obstructs the existence of one catalogue which works in all cases (to be considered in further publications).

5. Methods: Cayley-Hamilton

This section is taken entirely from [22]. In general it is difficult to read off from a description of a $p$-divisible group (e.g. by its Dieudonné module) its Newton polygon. However in the particular case that its $a$-number is at most one this can be done. This we describe in this section. The marvel is a new idea which produces for a given element in a given Dieudonné module a polynomial (in constants and in $F$) which annihilates this element (but, in general, is does not annihilate other elements of the Dieudonné module). This idea for constructing this polynomial comes from the elementary theorem in linear algebra: every endomorphism of a vector spaces is annihilated by its characteristic polynomial. As we work in our case with an operator which does not commute with constants things are not that elementary. The method we propose works for $a(G_0) = 1$, but it breaks down in an essential way in other cases.

5.1. Theorem (of Cayley-Hamilton type). — Let $G_0$ be a $p$-divisible group over an algebraically closed field $k \supset \mathbb{F}_p$ with $a(G_0) \leq 1$. In $\mathcal{D} = \text{Def}(G_0)$ there exists a coordinate system $\{t_j \mid j \in \mathbb{T}(\rho)\}$ and an isomorphism $\mathcal{D} \cong \text{Spf}(k[[t_j \mid j \in \mathbb{T}(\rho)]])$ such that for any $\gamma \succ \mathcal{N}(G_0)$ we have

$W_\gamma(\mathcal{D}) = \text{Spf}(R_\gamma)$, with $R_\gamma := k[[t_j \mid j \in \mathbb{T}(\gamma)]] = k[[t_j \mid j \in \mathbb{T}(\rho)]]/(t_j \mid j \notin \mathbb{T}(\gamma))$.

5.2. Corollary. — Let $G_0$ be a $p$-divisible group over a field $K$ with $a(G_0) \leq 1$. In $\text{Def}(G_0)$ every Newton polygon $\gamma \succ \mathcal{N}(G_0)$ is realized.

5.3. These methods allow us to give a proof for the Grothendieck conjecture. In fact, starting with $G_0$ we use 4.4 in order to obtain a deformation to a $p$-divisible group with the same Newton polygon and with $a \leq 1$. For that group the method 5.1 of Cayley-Hamilton type can be applied, which shows that it can be deformed to a $p$-divisible group with a given lower Newton polygon. Combination of these two specializations shows that the Grothendieck conjecture 2.1 is proven.

6. Polarized abelian varieties and quasi-polarized $p$-divisible groups

6.1. Methods described in the previous two sections for $p$-divisible groups also work (in almost the same way) for principally quasi-polarized $p$-divisible groups and for principally polarized abelian varieties. In fact we have the following tools:
6.2. Theorem. — Let \((G_0, \lambda_0)\) be a principally quasi-polarized \(p\)-divisible group over a field \(K\), and let \(\xi\) be a symmetric Newton polygon, \(\xi \succ \mathcal{N}(G_0)\). There is a deformation \((G_\eta, \lambda_\eta)\) of \((G_0, \lambda_0)\) such that \(\xi = \mathcal{N}(G_\eta)\) and \(a(G_\eta) \leq 1\).

6.3. Theorem (of Cayley-Hamilton type). — Let \((G_0, \lambda)\) be a principally quasi-polarized \(p\)-divisible group over an algebraically closed field \(k \supset \mathbb{F}_p\) with \(a(G_0) \leq 1\). In \(\mathcal{D} := \text{Def}(G_0, \lambda)\) there exists a coordinate system \(\{t_j \mid j \in \triangle(\rho)\}\) and an isomorphism

\[\mathcal{D} \cong \text{Spf}(k[[t_j \mid j \in \triangle(\rho)]])\]

such that for any symmetric \(\xi \succ \mathcal{N}(X_0)\) we have

\[W_\xi(\mathcal{D}) = \text{Spf}(R_\xi), \quad \text{with } R_\xi := k[[t_j \mid j \in \triangle(\xi)]] = k[[t_j \mid j \in \triangle(\rho)]]/(t_j \mid j \not\in \triangle(\xi)).\]

6.4. By the Serre-Tate theorem these results imply the analogous statements for principally polarized abelian varieties. This provides proofs for 3.1, 3.3, 3.2, and 3.4.

6.5. Remark. — The conjecture by Manin that every symmetric Newton polygon appears as the Newton polygon of an abelian variety, see [15], page 76, see 3.5 follows from the polarized version of the Cayley-Hamilton method, see 6.3 (and we do not need the much deeper results of Section 4). In Section 5 of [22] this is described. Here are the essentials of that proof. We observe that the Manin conjecture holds for the supersingular Newton polygon \(\sigma = g \cdot (1, 1)\) (here is the first algebraization fact). Then we observe that there exists a principally polarized supersingular abelian variety of given dimension \(g\); this follows from [14], (4.9), but that difficult result is not necessary to prove this rather easy result, see [22], Section 4. Then, by 6.3 we show that its quasi-polarized formal group can be deformed to achieve a given symmetric Newton polygon. Then, by the Serre-Tate theorem, and by the Chow-Grothendieck algebraization (here is the second algebraization fact) we conclude the same for deformations of principally polarized abelian schemes starting from the supersingular \(\sigma\), arriving at a a given symmetric \(\xi\), thus proving the Manin conjecture.

6.6. Remark. — In many cases points of \(W_\beta(a > 1)\) are singular points of \(W_\beta\), and (if enough level structure is considered) the open set \(W_\beta(a \leq 1) \subset W_\beta\) consists of non-singular points. Hence what we are doing in 4.4 is: move from a point in \(W_\beta\) to the regular interior. Then 5.1 tells us that the NP-strata around such a point are nested as coordinate hyperspaces, and we see that in the neighborhood of such a point every lower Newton polygon does appear (and we derive dimension statements). This explains our strategy: deformation theory in a singular point of the problem is difficult in general (and we could only progress via “Purity-catalogues”, and not via deformation theory directly); then we arrive at regular points of the strata (in our case this is ensured by \(a = 1\)), and a fairly general argument (of Cayley-Hamilton type) finishes the proof. An analogous remark holds for the proof of the Grothendieck conjecture 2.1 and of 2.2 via 4.4 and 5.1.
7. Some questions and conjectures

7.1. For every Newton polygon $\beta$ (and every $g$ and every $p$) we obtain $W_\beta \subset \mathcal{A} = \mathcal{A}_{g,1} \otimes \mathbb{F}_p$. For $\beta = \sigma$, the supersingular Newton polygon, this locus has “many” components (for $p >> 0$; in fact this number is a class number, asymptotically going to $\infty$ with $p \to \infty$).

Conjecture. — Given $p$, $g$, and $\beta \neq \sigma$ we conjecture that the locus $W_\beta$ is geometrically irreducible.

7.2. We consider complete subvarieties of moduli spaces. It is known that for any field $K$, and any complete subvariety $W \subset \mathcal{A} = \mathcal{A}_{g,1} \otimes \mathbb{F}_p$, the dimension of $W$ is at most $(g(g+1)/2) - g$, see [3], Coroll. 2.7 on page 70. We wonder is this maximum ever achieved? If yes, in which cases?

Conjecture. — Let $g \geq 3$. Suppose $W$ is a complete subvariety $W \subset \mathcal{A} = \mathcal{A}_{g,1} \otimes \mathbb{F}_p$ of dimension equal to $(g(g+1)/2) - g$ (the maximal possible dimension for complete subvarieties). We expect that under these conditions $W$ is equal to the locus $V_0$ of principally polarized abelian varieties with $p$-rank equal to zero. (This locus is complete and has the right dimension.)

If this is true, we have a proof for:

7.3. Conjecture. — Let $g \geq 3$. Let $W \subset \mathcal{A}_g \otimes \mathbb{C}$ be a complete subvariety. We expect that under these conditions:

$$\dim(W) < (g(g+1)/2) - g.$$ 

Hecke orbits are dense in $\mathcal{A}_g \otimes \mathbb{C}$. Chai proved the same for Hecke orbits of ordinary polarized abelian varieties in positive characteristic, see [1]. In his case only $\ell$-power isogenies need to be considered for one prime $\ell \neq p$.

7.4. Conjecture. — Fix a polarized abelian variety $[(X, \lambda)] = x \in \mathcal{A}_g \otimes \mathbb{Z} \mathbb{F}_p$. Consider the Hecke orbit of $x$. We conjecture that this Hecke orbit is everywhere dense in the Newton polygon stratum $W_\mathcal{N}(X)$.

This will be studied in [2].

7.5. Conjecture (Foliations, see [25]). — We expect that the following facts to be true. For every Newton polygon $\beta$ there should exist integers $i(\beta)$, $c(\beta), \in \mathbb{Z}_{\geq 0}$, and for every point $[(X, \lambda)] = x \in \mathcal{A} = \mathcal{A}_{g,1} \otimes \mathbb{F}_p$ with $\mathcal{N}(X) = \beta$ there should exist a closed subset $x \in I(x) = I_\beta(x) \subset W_\beta \subset \mathcal{A}$, contained in $W_\beta^0$, and a closed subset $x \in C(x) = C_\beta(x) \subset W_\beta^0 \subset \mathcal{A}$ in the open Newton polygon stratum $W_\beta^0$ such that:

- $\dim(I(x)) = i(\beta)$ and $\dim(C(x)) = c(\beta)$. 

---

SOCIÉTÉ MATHEMATIQUE DE FRANCE 2005
For every geometric point $[(Z, \zeta)] = z \in \mathcal{C}(x)(k)$ there is an isomorphism $(\mathbb{Z}[p^\infty], \zeta) \cong (X[p^\infty], \lambda)$. All irreducible components of the locally closed set $\mathcal{C}(x)$ contain $x$, and it is the maximal closed set with this and the property just mentioned.

For every geometric point $[(Y, \mu)] = y \in \mathcal{I}(x)$ there is a Hecke-correspondence using only iterates of $\alpha_p$-isogenies relating $[(X, \lambda)]$ and $[(Y, \mu)]$. All irreducible components of the closed set $\mathcal{I}(x)$ contain $x$, and it is the maximal closed set with this and the property just mentioned.

The dimensions are complementary: $i(\beta) + c(\beta) = \text{sdim}(\beta)$, and locally at $x$ their intersection is zero dimensional.

For $\beta \not\sim \gamma$ we have $i(\beta) \geq i(\gamma)$ and $c(\beta) < c(\gamma)$.

If moreover $a(X) \leq 1$, the (locally) closed sets $\mathcal{I}(x)$ and $\mathcal{C}(x)$ are regular at $x \in \mathcal{A}$, intersect transversally at $x$, and together their tangent spaces span the tangent space of $x \in W_\beta$.

Examples:

for the supersingular locus we have $i(\sigma) = \text{sdim}(\sigma) = [g^2/4]$ and $c(\sigma) = 0$;

for the ordinary locus we have $i(\rho) = 0$, and $c(\rho) = \text{sdim}(\sigma) = (g(g + 1))/2$;

for the case the $p$-rank equals one, i.e. $\beta = g \cdot (1, 0) + (1, 1) + g \cdot (0, 1)$ we have $i(\beta) = 0$, and $c(\beta) = \text{sdim}(\sigma) = ((g(g + 1))/2) - 1$.

We have: $p$-rank $f(\beta) < g - 1$ iff $i(\beta) > 0$.

We have: $\beta \neq \sigma$ iff $c(\beta) > 0$.

There is an easy combinatorial argument by which the numbers $i(\beta)$ and $c(\beta)$ can be read off from the Newton polygon diagram of $\beta$.

The sets $\mathcal{I}(x)$ will be called “isogeny leaves”, and the $\mathcal{C}(x)$ will be called “central leaves”.

7.6. Conjecture. — Let $\ell$ be a prime number different from $p$, and $[(X, \lambda)] = x \in \mathcal{A} = \mathcal{A}_{g, 1} \otimes \mathbb{F}_p$. The closure of the Hecke-$\ell$-orbit of $[(X, \lambda)] = x$ inside $W_{X,\lambda}(\mathbb{F}_p)$ equals $\mathcal{C}(x)$.

If this conjecture is true, then it follows that the conjecture 7.4 is true.

7.7. In general $G[p]$ does not determine a $p$-divisible group $G$. But in some cases it does. Let $\beta$ be a symmetric Newton polygon. For a pair of relatively prime integers $(m, n)$ we have defined in [9], Section 5 a $p$-divisible group $H_{m, n}$; it is characterized by: $H_{m, n} \sim G_{m, n}$, and for an algebraically closed field $k \supset \mathbb{F}_p$, the ring $\text{End}(H_{m, n} \otimes k)$ is a maximal order in $\text{End}^0(G_{m, n} \otimes k)$. We define $H_\beta$ to be the direct sum of all $H_{m, n}$ ranging over all slopes of $\beta$. We expect:

Conjecture. — Suppose $G$ is a $p$-divisible group over an algebraically closed field $k$, such that $G[p] \cong H_\beta[p]$; then (?) we should conclude $G \cong H_\beta$. 

ASTÉRISQUE 298
Note that in the special cases $\beta = \rho$ (the ordinary case), and $\beta = \sigma$ (supersingular) this conjecture is true; the conjectural statement above seems the natural generalization of this. Special cases have been proved.

7.8. Conjecture (Newton polygon strata, the non-principally polarized case)

Let $\xi$ be a symmetric Newton polygon and consider all possible polarized abelian varieties, where the polarization need not be principal. This gives a stratum $W_\xi(\mathcal{A}_g \otimes \mathbb{F}_p)$. Let $f = f(\xi)$ be the $p$-rank of $\xi$, i.e. this Newton polygon has exactly $f$ slopes equal to zero. We expect: under these conditions, there is an irreducible component

$$W \subset W_\xi(\mathcal{A}_g \otimes \mathbb{F}_p) \quad \text{with} \quad \dim(W) = ((g(g+1)/2) - (g - f)),$$

i.e. we expect that there is a component of every Newton polygon stratum which is a whole component of its $p$-rank stratum.

If this is the case, we see that there are “many” pairs of polarized abelian variety $(X, \lambda)$ and a Newton polygon $\gamma > N(X)$ such that there exist no deformation of $(X, \lambda)$ to a polarized abelian variety with Newton polygon equal to $\gamma$, namely consider $\beta < \gamma$ with $\beta \neq \gamma$ and $f(\beta) = f(\gamma)$.

7.9. Postscript, November 2004. — After my talk in 2000, several of the conjectures above were proved. Here is a survey of relevant information which I know now.

Conjecture 7.1 has been proved by C-L. Chai and F. Oort. Details will appear in [2].

Conjecture 7.2 seems still unproven. However, Conjecture 7.3 has been proved by S. Keel and L. Sadun; see [13].

All statements in 7.5 have been established and published, see [25].

Conjecture 7.4 and conjecture 7.6 have been proved by C-L. Chai and F. Oort; details will appear in [2].

Conjecture 7.7 has been established; see [27]; also see [26].

References


F. OORT, Mathematisch Instituut, Budapestlaan 6, NL-3508 TA Utrecht, The Netherlands
E-mail: oort@math.uu.nl
A GUIDE TO THE REDUCTION MODULO $p$ OF SHIMURA VARIETIES

by

Michael Rapoport

Abstract. — This is a survey of recent work on the reduction of Shimura varieties with parahoric level structures.

Résumé (Un guide à la réduction modulo $p$ des variétés de Shimura). — Cet article est un survol de résultats sur la réduction des variétés de Shimura à structure de niveau parahorique.

This report is based on my lecture at the Langlands conference in Princeton in 1996 and the series of lectures I gave at the semestre Hecke in Paris in 2000. In putting the notes for these lectures in order, it was my original intention to give a survey of the activities in the study of the reduction of Shimura varieties. However, I realized very soon that this task was far beyond my capabilities. There are impressive results on the reduction of “classical” Shimura varieties, like the Siegel spaces or the Hilbert-Blumenthal spaces, there are deep results on the reduction of specific Shimura varieties and their application to automorphic representations and modular forms, and to even enumerate all these achievements of the last few years in one report would be very difficult. Instead, I decided to concentrate on the reduction modulo $p$ of Shimura varieties for a parahoric level structure and more specifically on those aspects which have a group-theoretic interpretation. Even in this narrowed down focus it was not my aim to survey all results in this area but rather to serve as a guide to those problems with which I am familiar, by putting some of the existing literature in its context and by pointing out unsolved questions. These questions or conjectures are of two different kinds. The first kind are open even for those Shimura varieties which are moduli spaces of abelian varieties. Surely these conjectures are the most urgent
and the most concrete and the most tractable. The second kind are known for these special Shimura varieties. Here the purpose of the conjectures resp. questions is to extend these results to more general cases, e.g. to Shimura varieties of Hodge type.

As a general rule, I wish to stress that I would not be surprised if some of the conjectures stated here turn out to be false, especially in cases of very bad ramification. But I believe that even in these cases I should not be too far off the mark, and that a suitable modification of these conjectures gives the correct answer. My motivation in running the risk of stating precise conjectures is that I wanted to point out directions of investigation which seem promising to me.

The guiding principle of the whole theory presented here is to give a group-theoretical interpretation of phenomena found in special cases in a formulation which makes sense for a general Shimura variety. This is illustrated in the first section which treats some aspects of the elliptic modular case from the point of view taken in this paper. The rest of the article consists of two parts, the local theory and the global theory. Their approximate contents may be inferred from the table of contents below.

I should point out that the development in these notes is very uneven and that sometimes I have gone into the nitty gritty detail, whereas at other times I only give a reference for further developments. My motivation for this is that I wanted to give a real taste of the whole subject — in the hope that it is attractive enough for a student, one motivated enough to read on and skip parts which he finds unappealing.

In conclusion, I would like to stress, as in the introduction of [R2], the influence of the ideas of V. Drinfeld, R. Kottwitz, R. Langlands and T. Zink on my way of thinking about the circle of problems discussed here. In more recent times I also learned enormously from G. Faltings, A. Genestier, U. Görtz, T. Haines, J. de Jong, E. Landvogt, G. Laumon, B.C. Ngô, G. Pappas, H. Reimann, H. Stamm, and T. Wedhorn, but the influence of R. Kottwitz continued to be all-important. I am happy to express my gratitude to all of them. I also thank T. Ito, R. Kottwitz and especially T. Haines for their remarks on a preliminary version of this paper. I am grateful to the referee for his careful reading of the paper and his helpful remarks.

During the semestre Hecke I benefited from the financial support of the Ministère de la Recherche and the Humboldt Foundation (Prix Gay-Lussac/Humboldt).

Contents

1. Motivation : The elliptic modular curve .................273
Part I. Local theory
2. Parahoric subgroups ............................................277
3. $\mu$-admissible and $\mu$-permissible set ......................280
4. Affine Deligne-Lusztig varieties .................................288
5. The sets $X(\mu, b)_K$ ...........................................291
6. Relations to local models ........................................297
1. Motivation: The elliptic modular curve

In this section we illustrate the problem of describing the reduction modulo $p$ of a Shimura variety in the simplest case. Let $G = GL_2$ and let $(G, \{h\})$ be the usual Shimura datum. Let $K \subset G(\mathbb{A}_f)$ be an open compact subgroup of the form $K = K_p \cdot K_p$ where $K_p$ is a sufficiently small open compact subgroup of $G(\mathbb{A}_p)$. Let $G = G \otimes_{\mathbb{Q}} \mathbb{Q}_p$. We consider the cases where $K_p$ is one of the following two parahoric subgroups of $G(\mathbb{Q}_p)$,

(i) $K_p = K_p^{(i)} = GL_2(\mathbb{Z}_p)$ (hyperspecial maximal parahoric)

(ii) $K_p = K_p^{(ii)} = \{g \in GL_2(\mathbb{Z}_p); g \equiv (\ast \ast) \mod p\}$ (Iwahori)

The corresponding Shimura variety $\text{Sh}(G, h)_K$ is defined over $\mathbb{Q}$. It admits a model $\text{Sh}(G, h)_K$ over $\text{Spec} \mathbb{Z}_p$ by posing the following moduli problem over $(\text{Sch}/\mathbb{Z}_p)$:

(i) an elliptic curve $E$ with a level-$K_p$-structure.

(ii) an isogeny of degree $p$ of elliptic curves $E_1 \to E_2$, with a level-$K_p$-structure.

The description of the point set $\text{Sh}(G, h)_K(\overline{\mathbb{F}}_p)$ takes in both cases (i) and (ii) the following form,

\begin{equation}
\text{Sh}(G, h)_K(\overline{\mathbb{F}}_p) = \coprod_{\varphi} I_\varphi(\mathbb{Q}) \backslash X(\varphi)_{K_p} \times X_p/K_p.
\end{equation}

Here the sum ranges over the isogeny classes of elliptic curves and $I_\varphi(\mathbb{Q}) = \text{End}_{\mathbb{Q}}(E)^\times$ is the group of self-isogenies of any element of this isogeny class. Furthermore, $X_p/K_p$ may be identified with $G(\mathbb{A}_p^\times)/K_p$, with the action of $I_\varphi(\mathbb{Q})$ defined by the $\ell$-adic representation afforded by the rational Tate module. The set $X(\varphi)_{K_p}$ is the most interesting ingredient.

Let $\mathcal{O} = W(\mathbb{F}_p)$ be the ring of Witt vectors over $\mathbb{F}_p$ and $L = \text{Fract} \mathcal{O}$ be its fraction field. We denote by $\sigma$ the Frobenius automorphism of $L$. Let $N$ denote the rational Dieudonné module of $E$. Then $N$ is a 2-dimensional $L$-vector space, equipped with a $\sigma$-linear bijective endomorphism $F$ (the crystalline Frobenius). Then in case (i)
(hyperspecial case), the set $X(\varphi)_{K_p^{(\iota)}}$ has the following description

$$X(\varphi)_{K_p^{(\iota)}} = \{\Lambda; \ p\Lambda \subsetneq FA \subsetneq \Lambda\}$$

(1.2) $$= \{\Lambda; \ inv(\Lambda, FA) = \mu\}.$$  

Here $\Lambda$ denotes a $\mathcal{O}$-lattice in $N$. The set of $\mathcal{O}$-lattices in $N$ may be identified with $G(L)/G(\mathcal{O})$. We have used the elementary divisor theorem to establish an identification

$$\text{inv}: G(L)\backslash[G(L)/G(\mathcal{O}) \times G(L)/G(\mathcal{O})] = G(\mathcal{O})\backslash G(L)/G(\mathcal{O}) \simeq \mathbb{Z}^2/S_2.$$  

Furthermore $\mu = (1,0) \in \mathbb{Z}^2/S_2$ is the conjugacy class of one-parameter subgroups associated to $\{h\}$.

In case (ii) (Iwahori case), the set $X(\varphi)_{K_p^{(\iota)}}$ has the following description,

$$X(\varphi)_{K_p^{(\iota)}} = \{p\Lambda_2 \subsetneq \Lambda_1 \subsetneq \Lambda_2; \ p\Lambda_1 \subsetneq FA_1 \subsetneq \Lambda_1, p\Lambda_2 \subsetneq FA_2 \subsetneq \Lambda_2\}.$$  

Here again $\Lambda_1, \Lambda_2$ denote $\mathcal{O}$-lattices in $N$.

In either case $X(\varphi)_{K_p}$ is equipped with an operator $\Phi$ which under the bijection (1.1) corresponds to the action of the Frobenius automorphism on the left hand side.

Let us describe the set $X(\varphi)_{K_p^{(\iota)}}$ in the manner of the second line of (1.2). The analogue in this case of the relative position of two chains of inclusions of $\mathcal{O}$-lattices in $N$, $p\Lambda_2 \subsetneq \Lambda_1 \subsetneq \Lambda_2$ and $p\Lambda'_2 \subsetneq \Lambda'_1 \subsetneq \Lambda'_2$ is given by the identification analogous to (1.3),

$$\text{inv}: G(L)\backslash[G(L)/G_0(\mathcal{O}) \times G(L)/G_0(\mathcal{O})] = G_0(\mathcal{O})\backslash G(L)/G_0(\mathcal{O}) \xrightarrow{\sim} \mathbb{Z}^2 \times S_2.$$  

Here $G_0(\mathcal{O})$ denotes the standard Iwahori subgroup of $G(\mathcal{O})$ and on the right appears the extended affine Weyl group $\tilde{W}$ of $GL_2$. It is now a pleasant exercise in the Bruhat-Tits building of $PGL_2$ to see that

$$\{p\Lambda_2 \subsetneq \Lambda_1 \subsetneq \Lambda_2, \ p\Lambda'_2 \subsetneq \Lambda'_1 \subsetneq \Lambda'_2; \ p\Lambda_1 \subsetneq \Lambda'_1 \subsetneq \Lambda_1, \ p\Lambda_2 \subsetneq \Lambda'_2 \subsetneq \Lambda_2\}$$

$$= \{(g, g') \in (G(L)/G_0(\mathcal{O}))^2; \ inv(g, g') \in \text{Adm}(\mu)\}.$$  

Here $\text{Adm}(\mu)$ is the following subset of $\tilde{W}$,

$$\text{Adm}(\mu) = \{t_{(1,0)}, t_{(0,1)}, t_{(1,0)} \cdot s\}.$$  

Here $t_{(1,0)}$ and $t_{(0,1)}$ denote the translation elements in $\tilde{W} = \mathbb{Z}^2 \times S_2$ corresponding to $(1,0)$ resp. $(0,1)$ in $\mathbb{Z}^2$, and $s$ denotes the non-trivial element in $S_2$.

For $w \in \tilde{W}$ and any $\sigma$-linear automorphism $F$ of $N$, let us introduce the affine Deligne-Lusztig variety,

$$X_w(F) = \{g \in G(L)/G_0(\mathcal{O}); \ inv(g, Fg) = w\}.$$  

Then we may rewrite (1.4) in the following form, where $F_\varphi$ denotes the crystalline Frobenius associated to $\varphi$,

$$X(\varphi)_{K_p^{(\iota)}} = \bigcup_{w \in \text{Adm}(\mu)} X_w(F_\varphi).$$
This is analogous to the second line in (1.2) which may be viewed as a \textit{generalized} affine Deligne-Lusztig variety corresponding to the hyperspecial parahoric $K_p^{(i)}$. It should be pointed out that in this special case the union (1.9) is spurious: only one of the summands is non-empty, for a fixed isogeny class $\varphi$. For more general Shimura varieties this is no longer true.

The model $\text{Sh}(G, h)_K$ is smooth over $\text{Spec} Z_p$ in the hyperspecial case, but it has bad reduction in the Iwahori case. In the latter case there is the famous picture of the special fiber where two hyperspecial models meet at the supersingular points.

Such a global picture is not known in more general cases. The nature of the singularities in the special fiber in the Iwahori case can be understood in terms of the associated \textit{local model}.

We consider the lattice chain $p\Lambda_2 \subset \subset \Lambda_1 \subset \subset \Lambda_2$ in $\mathbb{Q}_p$, where $\Lambda_2 = \mathbb{Z}_p^2$ and $\Lambda_1 = p\mathbb{Z}_p \oplus \mathbb{Z}_p$. Let $\mathcal{M}^{\text{loc}}(G, \mu)_{K_p^{(ii)}}$ be the join of $\mathcal{P}(\Lambda_1)$ and $\mathcal{P}(\Lambda_2)$ over $\mathbb{Z}_p$ ($= \text{scheme-theoretic closure of the common generic fiber}$ $\mathcal{P}^{\text{loc}}_{\mathbb{Q}_p}$ in $\mathcal{P}(\Lambda_1) \times_{\text{Spec} \mathbb{Z}_p} \mathcal{P}(\Lambda_2)$). Then we obtain a diagram of schemes over $\text{Spec} \mathbb{Z}_p$,

\begin{equation}
\begin{array}{ccc}
\widetilde{\text{Sh}}(G, h)_K & \xrightarrow{\pi} & \mathcal{M}^{\text{loc}}(G, \mu)_{K_p^{(ii)}} \\
\downarrow & & \downarrow \\
\text{Sh}(G, h)_K & \xrightarrow{\lambda} & \mathcal{M}^{\text{loc}}(G, \mu)_{K_p^{(ii)}}.
\end{array}
\end{equation}

Here $\pi$ is the principal homogeneous space under the group scheme $\mathcal{G}$ over $\text{Spec} \mathbb{Z}_p$ attached to $K_p^{(ii)}$ (with $\mathcal{G}(\mathbb{Z}_p) = K_p^{(ii)}$), which adds to the isogeny of degree $p$, $E_1 \to E_2$ over a base scheme $S$, and its level-$K^p$-structure, a trivialization of the DeRham homology modules,

\begin{equation}
\begin{array}{ccc}
H_{\text{DR}}(E_1) & \xrightarrow{\lambda} & H_{\text{DR}}(E_2) \\
\downarrow & & \downarrow \\
\Lambda_1 \otimes \mathbb{Z}_p \mathcal{O}_S & \xrightarrow{\lambda} & \Lambda_2 \otimes \mathbb{Z}_p \mathcal{O}_S \\
\end{array}
\end{equation}

The morphism $\tilde{\lambda}$ is given by the Hodge filtration of the DeRham homology.

The diagram (1.10) can be interpreted as a relatively representable morphism of algebraic stacks

\begin{equation}
\lambda : \text{Sh}(G, h)_K \longrightarrow [\mathcal{M}^{\text{loc}}(G, \mu)_{K_p^{(ii)}}/\mathcal{G}].
\end{equation}

This morphism $\lambda$ is smooth of relative dimension $\text{dim} \mathcal{G} = 4$. The analogue of $\lambda$ in the hyperspecial case is a smooth morphism of relative dimension 4,

\begin{equation}
\lambda : \text{Sh}(G, h)_K \longrightarrow [\mathcal{P}^{1}_{\mathbb{Z}_p}/\text{GL}_2 \mathbb{Z}_p].
\end{equation}

At this point we have met in this special case all the main actors which will appear in the sequel: the admissible subset of the extended affine Weyl group, affine Deligne-Lusztig varieties, the sets $X(\varphi)_{K_p}$ (later denoted by $X(\mu, b)_K$), local models.
etc. These definitions can be given purely in terms of the \( p \)-adic group \( \text{GL}_2 \) and its parahoric subgroup \( K_p \). This will be the subject matter of the local part (sections 2–6). On the other hand, the enumeration of isogeny classes and the description (1.1) of the points in the reduction are global problems. These are addressed in the global part (sections 7–10).

We conclude this section with the definition of a Shimura variety of PEL-type. The guiding principle of the theory is to investigate the moduli problems related to them and then to express these findings in terms of the Shimura data associated to them.

Let \( B \) denote a finite-dimensional semi-simple \( \mathbb{Q} \)-algebra, let \( * \) be a positive involution on \( B \), let \( V \neq (0) \) be a finitely generated left \( B \)-module and let \( \langle \cdot, \cdot \rangle \) be a non-degenerate alternating bilinear form \( \langle \cdot, \cdot \rangle : V \times V \to \mathbb{Q} \) of the underlying \( \mathbb{Q} \)-vector space such that \( \langle bv, w \rangle = \langle v, b^*w \rangle \) for all \( v, w \in V, b \in B \). We denote by \( G \) the group of \( B \)-linear symplectic similitudes of \( V \). This is an algebraic group \( G \) over \( \mathbb{Q} \). We assume that \( G \) is a connected, hence reductive, algebraic group (this essentially excludes the orthogonal case). We let \( h : \mathbb{C} \times \to G(\mathbb{R}) \) be an algebraic homomorphism which defines on \( V_{\mathbb{R}} \) a Hodge structure of type \((-1, 0) + (0, -1)\) and which satisfies the usual Riemann conditions with respect to \( \langle \cdot, \cdot \rangle \), compare [W], 1.3.1. These data define by Deligne a Shimura variety \( \text{Sh}(G, h) \) over the Shimura field \( E \).

We now fix a prime number \( p \). Let \( G = G \otimes_{\mathbb{Q}} \mathbb{Q}_p \). We consider an order \( O_B \) of \( B \) such that \( O_B \otimes \mathbb{Z}_p \) is a maximal order of \( B \otimes \mathbb{Q}_p \). We assume that \( O_B \otimes \mathbb{Z}_p \) is invariant under the involution \( * \). We also fix a multichain \( \mathcal{L} \) of \( O_B \otimes \mathbb{Z}_p \)-lattices in \( V \otimes \mathbb{Q}_p \) which is self-dual for \( \langle \cdot, \cdot \rangle \), [RZ2]. Let \( K = K_p \) be the stabilizer of \( \mathcal{L} \) in \( G(\mathbb{Q}_p) \). Most of the time \( K \) is a parahoric subgroup.

Finally we fix an open compact subgroup \( K_p \subset G(\mathbb{A}_f^p) \), which will be assumed sufficiently small. Let \( \mathbf{K} = K^p \cdot K_p \).

We fix embeddings \( \overline{\mathbb{Q}} \to \mathbb{C} \) and \( \overline{\mathbb{Q}} \to \mathbb{Q}_p \). We denote by \( p \) the corresponding place of \( E \) over \( p \) and by \( E = E_p \) the completion and by \( \kappa \) the residue field of \( \mathcal{O}_E \). We then define a moduli problem \( \text{Sh}(G, h) \mathbf{K} \) over \( \text{Spec} \mathcal{O}_E \), i.e., a set-valued functor, as follows. It associates to a scheme \( S \) over \( \mathcal{O}_E \) the following data up to isomorphism ([RZ2], 6.9).

1. An \( \mathcal{L} \)-set of abelian varieties up to prime-to-\( p \) isogeny, \( A = \{ A_\Lambda; \Lambda \in \mathcal{L} \} \).
2. A \( \mathbb{Q} \)-homogeneous principal polarization \( \overline{\lambda} \) of the \( \mathcal{L} \)-set \( A \).
3. A \( K^p \)-level structure

\[
\overline{\eta} : H_1(A, \mathbb{A}_f^p) \simeq V \otimes \mathbb{A}_f^p \mod K^p,
\]
which respects the bilinear forms on both sides up to a constant in \( (\mathbb{A}_f^p)^\times \).

We require an identity of characteristic polynomials for each \( \Lambda \in \mathcal{L} \),

\[
\text{char}(b; \text{Lie } A_\Lambda) = \text{char}(b; V_h^{0,-1}), \quad b \in O_B.
\]
This moduli problem is representable by a quasi-projective scheme whose generic fiber is the initial Shimura variety $\text{Sh}(G, h)_{\text{K}}$ (or at least a finite union of isomorphic copies of $\text{Sh}(G, h)_{\text{K}}$).

However, contrary to the optimistic conjecture in [RZ2], this does not always provide us with a good integral model of the Shimura variety, e.g. flatness may fail. However, if the center of $B$ is a product of field extensions which are unramified at $p$ and excluding the orthogonal case as above, then $\text{Sh}(G, h)_{\text{K}}$ is a good integral model of the Shimura variety [G1], [G2]. For most of the remaining cases there is a closed subscheme of the above moduli space which is a good model [PR1], [PR2]. However, these closed subschemes cannot be defined in terms of the moduli problem of abelian varieties. Still, they can be analyzed and can serve as an experimental basis for the predictions which are the subject of this report.

PART I
LOCAL THEORY

2. Parahoric subgroups

Let $G$ be a connected reductive group over a complete discretely valued field $L$ with algebraically closed residue field. Kottwitz [K4] defines a functorial surjective homomorphism
\[
\tilde{\kappa}_G : G(L) \longrightarrow X^*(\hat{Z}(G)^I).
\]
Here $I = \text{Gal}(\overline{L}/L)$ denotes the absolute Galois group of $L$ and $\hat{Z}(G)$ denotes the center of the Langlands dual group. For instance, if $G = GL_n$, then the target group is $\mathbb{Z}$ and $\tilde{\kappa}_G(g) = \text{ord det } g$; if $G = GSp_{2n}$, then again the target group is $\mathbb{Z}$ and $\tilde{\kappa}_G(g) = \text{ord } c(g)$, where $c(g) \in L^\times$ is the multiplier of the symplectic similitude $g$.

Let $B = B(B_{\text{ad}}, L)$ denote the Bruhat-Tits building of the adjoint group over $L$. Then $G(L)$ acts on $B$.

Definition 2.1. — The parahoric subgroup associated to a facet $a$ of $B$ is the following subgroup of $G(L)$,

\[
K_a = \text{Fix}(a) \cap \text{Ker } \tilde{\kappa}_G.
\]

If $a$ is a maximal facet, i.e. an alcove, then the parahoric subgroup is called an Iwahori subgroup.

Remarks 2.2

(i) We have

\[
K_{ga} = gK_ag^{-1}, \quad g \in G(L).
\]

In particular, since all alcoves are conjugate to each other, all Iwahori subgroups are conjugate.
(ii) This notion of a parahoric subgroup coincides with the one by Bruhat and Tits \cite{BT2}, 5.2.6., cf. \cite{HR}. Let $\mathcal{O}_L$ be the ring of integers in $L$. There exists a smooth group scheme $\mathcal{G}_a$ over $\text{Spec} \mathcal{O}_L$, with generic fiber equal to $G$ and with connected special fiber such that

$$K_a = \mathcal{G}_a(\mathcal{O}_L).$$

(iii) Let $G = T$ be a torus. Then there is precisely one parahoric subgroup $K$ of $T(L)$. Then

$$K = T^0(\mathcal{O}_L).$$

Here $T^0$ denotes the identity component of the Néron model of $T$. For the group of connected components of the special fiber of the Néron model one has\(1\)

$$\pi_0(\mathcal{T}) = X_*(T)_I.$$

2.3. Let $S$ be a maximal split torus in $G$ and $T$ its centralizer. Since by Steinberg’s theorem $G$ is quasi-split, $T$ is a maximal torus. Let $N = N(T)$ be the normalizer of $T$. Let

\begin{equation}
\tilde{\kappa}_T : T(L) \longrightarrow X^*({\hat{T}}^I) = X_*(T)_I
\end{equation}

be the Kottwitz homomorphism associated to $T$ and let $T(L)_1$ be its kernel. The factor group

\begin{equation}
\tilde{W} = N(L)/T(L)_1
\end{equation}

will be called the Iwahori Weyl group associated to $S$. It is an extension of the relative Weyl group

\begin{equation}
W_0 = N(L)/T(L).
\end{equation}

Namely, we have an exact sequence induced by the inclusion of $T(L)_1$ in $T(L)$,

\begin{equation}
0 \longrightarrow X_*(T)_I \longrightarrow \tilde{W} \longrightarrow W_0 \longrightarrow 1.
\end{equation}

The reason for the name given to $\tilde{W}$ comes from the following fact \cite{HR}.

**Proposition 2.4.** — Let $K_0$ be the Iwahori subgroup associated to an alcove contained in the apartment associated to the maximal split torus $S$. Then

$$G(L) = K_0 \cdot N(L) \cdot K_0,$$

and the map $K_0 n K_0 \mapsto n \in \tilde{W}$ induces a bijection

$$K_0 \backslash G(L)/K_0 \simeq \tilde{W}.$$

More generally, let $K$ and $K'$ be parahoric subgroups associated to facets contained in the apartment corresponding to $S$. Let

$$\tilde{W}^K = (N(L) \cap K) / T(L)_1, \text{ resp. } \tilde{W}^{K'} = (N(L) \cap K') / T(L)_1.$$

\(1\)See Notes at the end, n° 1

ASTÉRISQUE 298
Then

\[ K \backslash G(L)/K' \cong \tilde{W}^K \backslash \tilde{W}/\tilde{W}^K'. \]

For the structure of \( \tilde{W} \) we have the following fact [HR].

**Proposition 2.5.** — Let \( x \) be a special vertex in the apartment corresponding to \( S \), and let \( K \) be the corresponding parahoric subgroup. The subgroup \( W^K \) of \( \tilde{W} \) projects isomorphically to the factor group \( W_0 \) and the exact sequence (2.5) presents \( \tilde{W} \) as a semidirect product,

\[ \tilde{W} = W_0 \rtimes X_*(T)_I. \]

Sometimes for \( \nu \in X_*(T)_I \) we write \( t_\nu \) for the corresponding element of \( \tilde{W} \).

Let \( S_{sc} \) resp. \( T_{sc} \) resp. \( N_{sc} \) be the inverse images of \( S \cap G_{der} \) resp. \( T \cap G_{der} \) resp. \( N \cap G_{der} \) in the simply connected covering \( G_{sc} \) of the derived group \( G_{der} \). Then \( S_{sc} \) is a maximal split torus of \( G_{sc} \), and \( T_{sc} \) resp. \( N_{sc} \) is its centralizer resp. normalizer. Hence

\[ W_a = N_{sc}(L)/T_{sc}(L)_1 \]

is the Iwahori Weyl group of \( G_{sc} \). This group is called the affine Weyl group associated to \( S \), for the following reason. Let us fix a special vertex \( x \) in the apartment corresponding to \( S \). Then there exists a reduced root system \( \Sigma^x \) such that Proposition 2.5 (applied to \( G_{sc} \) instead of \( G \)) presents \( W_a \) as the affine Weyl group associated (in the sense of Bourbaki) to \( \Sigma^x \).

\[ W_a = W(\Sigma^x) \rtimes Q^\vee(\Sigma^x), \]

cf. [T], 1.7, compare also [HR]. In other words, we have an identification \( W_0 \cong W(\Sigma^x) \) and \( X_*(T_{sc})_I \cong Q^\vee(\Sigma^x) \) compatibly with the semidirect product decompositions (2.7) and Proposition 2.5. In particular, \( W_a \) is a Coxeter group.

There is a canonical injective homomorphism \( W_a \rightarrow \tilde{W} \) which induces an injection from \( X_*(T_{sc})_I \) into \( X_*(T)_I \). In fact, \( W_a \) is a normal subgroup of \( \tilde{W} \) and \( \tilde{W} \) is an extension,

\[ 1 \longrightarrow W_a \longrightarrow \tilde{W} \longrightarrow X_*(T)_I/X_*(T_{sc})_I \longrightarrow 1. \]

The affine Weyl group \( W_a \) acts simply transitively on the set of alcoves in the apartment of \( S \), cf. [T], 1.7. Since \( \tilde{W} \) acts transitively on the set of these alcoves and \( W_a \) acts simply transitively, \( \tilde{W} \) is the semidirect product of \( W_a \) with the normalizer \( \Omega \) of a base alcove, i.e. the subgroup of \( \tilde{W} \) which preserves the alcove as a set,

\[ \tilde{W} = W_a \rtimes \Omega. \]

In the sequel we will often identify \( \Omega \) with \( X_*(T)_I/X_*(T_{sc})_I \).
Remarks 2.6

(i) Let $K = K_a$ be a parahoric subgroup and $\mathcal{G} = \mathcal{G}_a$ the corresponding group scheme over $\text{Spec} \mathcal{O}_L$, cf. Remarks 2.2, (ii). Then $\tilde{W}^K$ can be identified with the Weyl group of the (maximal reductive quotient of the) special fiber $\mathcal{G} = \mathcal{G} \otimes_{\mathcal{O}_k} k$ of the group scheme $\mathcal{G}$.

(ii) Assume in Proposition 2.5 above that $x$ is a hyperspecial vertex. In this case $S = T$ and $W_0$ is the absolute Weyl group of $G$. In this case we have

$\tilde{W} = W_0 \rtimes X_*(S)$

and, since $\tilde{W}^K = W_0$,

$\tilde{W}^K \backslash \tilde{W} / \tilde{W}^K = X_*(S) / W_0$.

3. $\mu$-admissible and $\mu$-permissible set

We continue with the notation of the previous section. In particular, we let $N$ resp. $T$ be the normalizer resp. centralizer of a maximal split torus $S$ over $L$. Let

$W = N(T)/T(T)$

be the absolute Weyl group of $G$. Then $H^1(I,T) = (0)$ and $W_0 = W/I$ is the set of invariants.

Let $\{\mu\}$ be a conjugacy class of cocharacters of $G$ over $\overline{T}$. We denote by the same symbol the corresponding $W$-orbit in $X_*(T)$. We associate to $\{\mu\}$ a $W_0$-orbit $\Lambda = \Lambda(\{\mu\})$ in $X_*(T)_I$, as follows. Let $B$ be a Borel subgroup containing $T$, defined over $L$ (which exists since $G$ is automatically quasi-split). We denote the corresponding closed (absolute) Weyl chamber in $X_*(T)_R$ by $\overline{C}_B$. Let $\mu_B \in \{\mu\}$ be the unique element in $\overline{C}_B$. Then the $W_0$-orbit $\Lambda$ of the image $\lambda$ of $\mu_B$ in $X_*(T)_I$ is well-determined since any two choices of $B$ are conjugate under an element of $W_0$.

Lemma 3.1. — All elements in $\Lambda$ are congruent modulo $X_*(T_{sc})_I$.

Proof. — Let us fix a special vertex $x$ and let us identify $W_0$ with $W(\Sigma)$, cf. (2.7). We claim that for any $\lambda \in X_*(T)_I$ and any $w \in W_0$ we have

$w(\lambda) - \lambda \in Q^\vee(\Sigma)$.

By induction on the length of $w$ (w.r.t. some ordering of the roots) we may assume that $w = s_\alpha$ is a reflection about a simple root $\alpha \in \Sigma$. But

$s_\alpha(\lambda) - \lambda = -\langle \lambda, \alpha \rangle \cdot \alpha^\vee$.

The assertion will follow, once we show that $\langle \lambda, \alpha \rangle$ is an integer. This follows, since the image of $X_*(T)_I$ in $X_*(T_{ad})_I \otimes \mathbb{R} = X_*(S_{sc}) \otimes \mathbb{R}$ lies in the lattice of coweights $P^\vee$ for $\Sigma$ (this holds since $P^\vee$ acts simply transitively on the set of special vertices in the apartment and these are preserved under the subgroup $X_*(T)_I$ of $\tilde{W}$, compare [HR]).
We shall denote by \( \tau = \tau(\{\mu\}) \in \Omega \) the common image of all elements of \( \Lambda \). Let us now fix an alcove \( a \) in the apartment corresponding to \( S \). This defines a Bruhat order on the affine Weyl group \( W_a \) which we extend in the obvious way to the semidirect product \( \widetilde{W} = W_a \times \Omega \), cf. (2.9).

Using this Bruhat order we can now introduce the \( \mu \)-admissible subset of \( \widetilde{W} \) [KR1],

\[
(3.4) \quad \text{Adm}(\mu) = \{ w \in \widetilde{W}; \ w \leq \lambda \text{ for some } \lambda \in \Lambda \}
\]

Here we consider the elements \( \lambda \in X_*(T)_I \) as elements of \( \widetilde{W} \) (translation elements). Note that by definition all elements in \( \text{Adm}(\mu) \) have image \( \tau \) in \( \Omega \).

More generally, let \( a' \) be a facet of \( a \) and let \( K \) be the corresponding parahoric subgroup. Then the Bruhat order on \( \widetilde{W} \) induces a Bruhat order on the double coset space \( \widetilde{W}K \backslash \widetilde{W} / \widetilde{W}K \) characterized by

\[
(3.5) \quad \widetilde{W}Kw_1\widetilde{W}K \leq \widetilde{W}Kw_2\widetilde{W}K \iff \exists \ w'_1 \in \widetilde{W}Kw_1\widetilde{W}K \text{ and } \exists \ w'_2 \in \widetilde{W}Kw_2\widetilde{W}K \text{ such that } w'_1 \leq w'_2
\]

\[
\iff \text{the same holds for } w'_1 \text{ and } w'_2 \text{ the unique elements of minimal length in their respective double cosets.}
\]

We then define the \( \mu \)-admissible subset of \( \widetilde{W}K \backslash \widetilde{W} / \widetilde{W}K \),

\[
(3.6) \quad \text{Adm}_K(\mu) = \{ w \in \widetilde{W}K \backslash \widetilde{W} / \widetilde{W}K; \ w \leq \widetilde{W}K\lambda\widetilde{W}K \text{ for some } \lambda \in \Lambda \}
\]

Since the element of minimal length in a double coset is smaller than any element in it, the natural map

\[
(3.7) \quad \text{Adm}(\mu) \to \text{Adm}_K(\mu)
\]

is surjective. In other words,

\[
(3.8) \quad \text{Adm}_K(\mu) = \text{image of } \text{Adm}(\mu) \text{ under } \widetilde{W} \to \widetilde{W}K \backslash \widetilde{W} / \widetilde{W}K.
\]

We next introduce another subset of \( \widetilde{W} \). We first note that the apartment in \( B(G_{\text{ad}}, L) \) corresponding to \( S \) is a principal homogeneous space under \( X_*(S_{\text{ad}})_R \). Let \( \Lambda_{\text{ad}} \) be the image of \( \Lambda \) under the natural map

\[
X_*(T)_I \to X_*(T_{\text{ad}})_I \to X_*(S_{\text{ad}})_I \otimes R = X_*(S_{\text{ad}})_R.
\]

We denote by \( P_\mu = \text{Conv}(\Lambda_{\text{ad}}) \) the convex hull of \( \Lambda_{\text{ad}} \). Now we can define the \( \mu \)-permissible subset of \( \widetilde{W} \),

\[
(3.9) \quad \text{Perm}(\mu) = \{ w \in \widetilde{W}; \ w \equiv \tau \mod W_a \text{ and } w(a) - a \in P_\mu, \text{ for all } a \in a \}
\]

Note that by convexity it suffices to impose the second condition in (3.9) for the vertices \( a_i \) of \( a \). Again there is a variant for a parahoric subgroup \( K \) corresponding to a facet \( a' \) of \( a \). Since \( \widetilde{W}K \subset W_a \) [HR], the first condition in the next definition
makes sense,

\[(3.10) \quad \text{Perm}_K(\mu) = \left\{ w \in \widehat{W}^K \backslash \widehat{W} \middle| \text{ for all } a \in a' \right\}.
\]

Let us check that the second condition also depends only on the double coset of \( w \).

Let us write \( \widehat{W} = W_0 \rtimes X^* (T) \) (corresponding to the choice of a special vertex which defines the inclusion of \( W_a \) in \( \hat{W} \)). If \( x \in \widehat{W}^K \), let \( x = x_0 \cdot t_x \) with \( x_0 \in W_0 \) and \( \nu \in X^* (T) \). Then for \( a \in a' \) we have \( x(a) = a \) which implies \( x_0(a) + x_0(\nu) = a \).

Hence

\[
xw(a) - a = x_0w(a) + x_0(\nu) - a = x_0w(a) - x_0(a)
\]

\[
= x_0(w(a) - a) \in x_0P_\mu = P_\mu,
\]

which proves our claim.

There is a natural map

\[(3.11) \quad \text{Perm}(\mu) \to \text{Perm}_K(\mu).
\]

However, in contrast to (3.8) it is not clear whether this map is surjective, compare Proposition 3.10 below.

An important question is to understand the relation between the sets \( \text{Adm}(\mu) \) and \( \text{Perm}(\mu) \). In any case, the elements \( t_\lambda \) for \( \lambda \in \Lambda \) are contained in both of them. In fact, these elements are maximal in \( \text{Adm}(\mu) \). The following fact is proved in [KR1].

We repeat the proof.

**Proposition 3.2.** — Let \( G \) be split over \( L \). Then

\[
\text{Adm}(\mu) \subset \text{Perm}(\mu).
\]

In fact, \( \text{Perm}(\mu) \) is closed under the Bruhat order.

Note that, by the preliminary remarks above, the second claim implies the first.

The significance of the second claim becomes more transparent when we discuss local models in Section 6. Namely, in many cases \( \text{Perm}(\mu) \) is supposed to parametrize the set of Iwahori-orbits in the special fiber of the local model and, since the latter is in these cases a closed subvariety of an affine flag variety, it contains with an orbit also all orbits in its closure. The question of when \( \text{Adm}(\mu) \) coincides with \( \text{Perm}(\mu) \) is closely related to the flatness property of local models. In all cases known so far, this flatness property was established by proving that the special fiber is reduced and that the generic points of the irreducible components of the special fiber are in the closure of the generic fiber. The property \( \text{Adm}(\mu) = \text{Perm}(\mu) \) is supposed to mean that the only irreducible components of the special fiber are “the obvious ones” indexed by \( t_\lambda \) for \( \lambda \in \Lambda \), for which the liftability problem should be visibly true (cf. Görtz [G1]–[G3] for various cases).

Returning to Proposition 3.2, we note that when \( G \) is split over \( L \), we have \( S = T \) and the action of \( I \) is trivial. Furthermore \( W_0 = W \). The reflections in the affine...
Weyl group will be written as \( s_\beta - m \) where \( \beta \) is a root in the sense of the euclidean root system and \( m \in \mathbb{Z} \). The proposition is a consequence of the following lemma.

**Lemma 3.3.** — Let \( \mathcal{P} \) be a \( W_0 \)-stable convex polygon in \( X_*(\mathcal{M})_{\mathbb{R}} \). Let \( x, y \in \tilde{\mathcal{W}} \) with \( x \leq y \). Let \( v \in a \) and put \( v_x = x(v), \ v_y = y(v) \). Then

\[
\text{if } v_y \in v + \mathcal{P}, \ \text{then } v_x \in v + \mathcal{P}.
\]

**Proof.** — We may assume that \( x = s_\beta - m y \), with \( \ell(x) < \ell(y) \). Since \( \beta - m \) separates \( a \) from \( y(a) \), it weakly separates \( v \in a \) from \( v_y \).

Now \((3.12)\)

\[
(\beta - m)(v) = \beta(v) - m
\]

\((3.13)\)

\[
(\beta - m)(v_y) = \beta(v_y) - m.
\]

Hence we have 2 cases:

- First case: \( \beta(v) \leq m \leq \beta(v_y) \)
- Second case: \( \beta(v_y) \leq m \leq \beta(v) \).

Now \((3.14)\)

\[
v_x = s_\beta - m(v_y) = v_y - [\beta(v_y) - m] \beta'.
\]

Hence in either case, \( v_x \) lies on the segment joining \( v_y \) with \( v_y - [\beta(v_y) - \beta(v)] \cdot \beta' \).

Hence it suffices to show that \( s_\beta(v_y) + \beta(v) \beta' \in v + \mathcal{P} \).

But \( v_y = p + v \), with \( p \in \mathcal{P} \), hence

\[
(3.15) \quad s_\beta(v_y) = s_\beta(p) + s_\beta(v) = s_\beta(p) + v - \beta(v) \cdot \beta'.
\]

Hence \( s_\beta(v_y) + \beta(v) \beta' = s_\beta(p) + v \in v + \mathcal{P} \). \( \square \)

The converse inclusion is not true in general. We have the following result which generalizes [KR1] valid for minuscule \( \mu \).

**Theorem 3.4 (Haines, Ngo [HN1]).** — Let \( G \) be either \( GL_n \) or \( GSp_{2n} \). In the case of \( GSp_{2n} \) assume that the dominant representative of \( \{ \mu \} \) is a sum of minuscule dominant coweights. Then \( Adm(\mu) = Perm(\mu) \).

It may be conjectured that we have equality in general in Proposition 3.2, if \( \mu \) is a sum of minuscule dominant coweights. Note that, in the case of \( GL_n \), this condition on \( \mu \) is automatically satisfied. On the other hand, Haines and Ngo ([HN1]) have shown by example that for any split group \( G \) of rank \( \geq 4 \) not of type \( A_n \), there exists a dominant coweight \( \mu \) such that \( Adm(\mu) \neq Perm(\mu) \). In [HN1] the result for \( GSp_{2n} \) is obtained by relating the sets \( Adm(\mu) \) resp. \( Perm(\mu) \) with the corresponding sets for the “ambient” \( GL_{2n} \). It would be interesting to clarify this relation in other cases as well.
In the sequel, until Proposition 3.10, we investigate the intersections of $\text{Adm}(\mu)$ resp. $\text{Perm}(\mu)$ with the translation subgroup of $\tilde{W}$. These results are taken from unpublished notes of Kottwitz, as completed by Haines. They will not be used elsewhere.

**Proposition 3.5 (Kottwitz, Haines).** — Let $G$ be split over $L$. Then

$$X_+(T) \cap \text{Adm}(\mu) = X_+(T) \cap \text{Perm}(\mu).$$

To prove this we need a few more lemmas which will also be useful for other purposes. For the time being we assume $G$ split. We denote by $R$ the set of roots and by $R^+$ resp. $\Delta$ the set of positive resp. simple roots for a fixed ordering.

**Lemma 3.6.** — Let $\nu \in X_+(S)$. Then

$$\ell(t_\nu) = \langle \nu, 2w(\varrho) \rangle,$$

where $\varrho = \frac{1}{2} \Sigma_{\alpha > 0} \alpha$ and where $w \in W$ is such that $w^{-1}(\nu)$ is dominant. In other words, if $\nu$ is dominant then

$$\ell(t_\nu) = \langle \nu, 2\varrho \rangle \quad \text{and} \quad \ell(t_{w(\nu)}) = \ell(t_\nu), \; \forall w \in W.$$

**Proof.** — This is an immediate consequence of [IM], Prop. 1.23. □

**Lemma 3.7.** — For any $\beta \in R^+$ we have

$$\ell(s_\beta) < \langle \beta^\vee, 2\varrho \rangle.$$

**Proof.** — It suffices to prove the weak inequality since the left hand side is an odd integer and the right hand side an even integer (twice the height of the coroot $\beta^\vee$).

We use induction on $\ell(s_\beta)$, the case $\ell(s_\beta) = 1$ being trivial. So assume $\ell(s_\beta) \geq 3$. We first claim that

$$\exists \alpha \in \Delta \text{ such that } \ell(s_\alpha s_\beta s_\alpha) = \ell(s_\beta) - 2. \tag{3.16}$$

Indeed, let $\alpha \in \Delta$ such that $\ell(s_\alpha s_\beta) = \ell(s_\beta) - 1$. Then also $\ell(s_\beta s_\alpha) = \ell(s_\beta) - 1$. Hence there are 2 possible configurations of $s_\alpha s_\beta s_\alpha, s_\beta s_\alpha, s_\beta s_\beta s_\alpha, s_\beta s_\alpha$ in the Bruhat order,

1. $s_\alpha s_\beta s_\alpha < s_\alpha s_\beta < s_\beta$ and $s_\alpha s_\beta s_\alpha < s_\beta s_\alpha < s_\beta$  
2. $s_\alpha s_\beta < s_\beta$ and $s_\alpha s_\beta < s_\alpha s_\beta s_\alpha$ and $s_\beta s_\alpha < s_\beta$ and $s_\beta s_\alpha < s_\alpha s_\beta s_\alpha$.

It suffices to show that case (2) does not arise. In case 2, by Lemma 4.1 of [H3] we have $s_\alpha s_\beta s_\alpha = s_\beta$. Hence $\beta$ and $s_\alpha(\beta)$ are proportional, hence $s_\alpha(\beta) = \pm \beta$. The minus sign cannot occur since $\alpha$ is the only root in $R^+$ sent by $s_\alpha$ into $R^-$. Hence $s_\alpha(\beta) = \beta$, i.e. $\langle \alpha, \beta^\vee \rangle = \langle \alpha^\vee, \beta \rangle = 0$. Hence $s_\beta(\alpha) = \alpha \in R^+$. But for any $\gamma \in R^+$ and $w \in W_0$ we have $w^{-1}(\gamma) > 0 \Leftrightarrow w \leq s_\gamma$. Hence $s_\beta < s_\alpha s_\beta$, a contradiction.

Now start with $\alpha$ satisfying (3.16) and put $\beta' = s_\alpha(\beta)$. Then by induction hypothesis

$$\ell(s_\beta) - 2 = \ell(s_{\beta'}) \leq \langle \beta^\vee, 2\varrho \rangle = \langle \beta^\vee, 2\varrho \rangle - 2\langle \beta^\vee, \alpha \rangle.$$

Hence it suffices to show that $\langle \beta^\vee, \alpha \rangle \geq 1$. But if $\langle \beta^\vee, \alpha \rangle < 0$, then $s_\beta(\alpha) = \alpha - \langle \beta^\vee, \alpha \rangle \beta \in R^+$. Hence, arguing as before, $s_\beta < s_\alpha s_\beta$, a contradiction. □
Lemma 3.8. — Let \( \nu \in X_*(S) \) be dominant. Let \( \beta \in R^+ \) such that \( \nu - \beta^\vee \) is dominant. Then \( t_{\nu - \beta^\vee} \leq t_{\nu} \) in the Bruhat order on \( \tilde{W} \).

[Here the Bruhat order on \( \tilde{W} \) is defined by the alcove \( a \) in \( X_*(S) \) with apex 0 and bounded by hyperplanes \( \alpha = 0 \) for \( \alpha \in \Delta \).]

Proof. — We use the identity

\[
s_{\beta-1} \cdot s_\beta = t_{\beta^\vee}, \quad \beta \in R.
\]

Indeed, this follows from the expression

\[
s_{\beta+k}(x) = x - (\beta, x)\beta^\vee - k\beta^\vee, \quad x \in X_*(S) \subset R.
\]

This last identity also shows

\[
t_{\nu} \cdot s_\beta = s_{\beta-m} \cdot t_{\nu}, \quad \beta \in R, \ \nu \in X_*(S).
\]

Here \( m = \langle \beta, \nu \rangle \).

The assertion of the lemma follows from the following two statements.

\begin{align*}
(3.20) & \quad t_{\nu} \cdot s_\beta \leq t_{\nu} \\
(3.21) & \quad t_{\nu - \beta^\vee} \leq t_{\nu} \cdot s_\beta.
\end{align*}

Let us prove (3.20), i.e.

\[
s_{\beta-m} \cdot t_{\nu} \leq t_{\nu}, \quad m = \langle \beta, \nu \rangle.
\]

It is enough to show that \( \beta - m \) separates \( a \) from \( t_{\nu}(a) \). But

\[
(\beta - m)(a) = \beta(a) - m
\]

\[
(\beta - m)(t_{\nu}(a)) = \beta(a) \subset [0, 1].
\]

Hence it suffices to know that \( m \geq 1 \). But since \( \nu - \beta^\vee \) is dominant we have

\[
\langle \beta, \nu - \beta^\vee \rangle \geq 0, \quad \text{i.e.} \quad m = \langle \beta, \nu \rangle \geq \langle \beta, \beta^\vee \rangle = 2.
\]

Now let us prove (3.21). It follows with (3.17) that both sides of (3.21) differ by a reflection, since

\[
t_{\nu - \beta^\vee} = t_{\nu} \cdot s_{\beta^\vee}^{-1} = (t_{\nu} \cdot s_\beta) \cdot s_{\beta-1}.
\]

Hence it suffices to prove that \( \ell(t_{\nu - \beta^\vee}) < \ell(t_{\nu} \cdot s_\beta) \). But by Lemma 3.6 we have, since \( \nu \) and \( \nu - \beta^\vee \) are dominant,

\[
\ell(t_{\nu - \beta^\vee}) = \ell(t_{\nu}) - \langle 2\varrho, \beta^\vee \rangle < \ell(t_{\nu}) - \ell(s_\beta) \leq \ell(t_{\nu} s_\beta).
\]

For the first inequality we used Lemma 3.7.

Remark 3.9 (Haines). — In the course of the proof of Lemma 3.8 we proved the chain of inequalities

\[
t_{\nu - \beta^\vee} \leq t_{\nu} s_\beta \leq t_{\nu}
\]
which is stronger than the assertion of the Lemma. Here is a simpler argument for the assertion of Lemma 3.8 which does not make use of Lemma 3.7. To prove \( t_{\nu - \beta^\vee} \leq t_{\nu} \), it is enough to prove the inequalities

\[
t_{\nu - \beta^\vee} \leq s_\beta t_{\nu - \beta^\vee} = s_{\beta - 1} t_{\nu} \leq t_{\nu}.
\]

The middle equality follows from \( s_{\beta - 1} = t_{\beta^\vee} s_\beta \), cf. (3.17). The left inequality holds because \( \nu - \beta^\vee \) dominant implies \( \ell(s_\beta t_{\nu - \beta^\vee}) = \ell(s_\beta) + \ell(t_{\nu - \beta^\vee}) \). We used that for any dominant \( \lambda \) and any simple reflection \( s \) we have

\[
\ell(st\lambda) = \sum_{\alpha > 0 \atop s(\alpha) < 0} |\langle \alpha, \lambda \rangle| + 1 \sum_{\alpha > 0 \atop s(\alpha) > 0} |\langle \alpha, \lambda \rangle| \]

([IM], Prop. 1.23). For the right inequality, it is enough to show that \( a \) and \( t_{\nu}(a) \) are on opposite sides of the hyperplane fixed by \( s_{\beta - 1} \). But using \( m = \langle \beta, \nu \rangle = \langle \beta, \nu - \beta^\vee + \beta^\vee \rangle \geq 2 \), we have

\[
(\beta - 1)(a) \subset [-1, 0] \quad (\beta - 1)(t_{\nu}a) = m - 1 + \beta(a) \subset [m - 1, m] \subset [1, \infty).
\]

**Proof of Proposition 3.5.**—L et \( \nu \in X_*(T) \cap \text{Perm}(\mu) \) and let us prove that \( t_{\nu} \) is \( \mu \)-admissible. Since \( \nu \) is \( \mu \)-permissible we have \( \mu - \nu \in X_*(S_{\infty}) \). Let us first assume that \( \nu \) is dominant. Then \( \nu \leq \mu \), i.e. \( \mu - \nu \) is a non-negative sum with integer coefficients of simple coroots. Now if \( \lambda \) and \( \lambda' \) are dominant coweights, then

\[
\ell(t_{\lambda + \lambda'}) = \ell(t_{\lambda}) + \ell(t_{\lambda'}).
\]

Hence if \( \nu, \mu \) and \( \lambda \) are dominant coweights, then

\[
t_{\nu} \leq t_{\mu} \iff t_{\lambda + \nu} \leq t_{\lambda + \mu}.
\]

Returning to our \( \nu \leq \mu \), there exists a sequence of coweights \( \nu = \nu_0, \nu_1, \ldots, \nu_r = \mu \), such that consecutive terms look like \( \lambda - \beta^\vee, \lambda \) for some positive coroot \( \beta^\vee \). Adding a sufficiently dominant coweight to this sequence, we can assume that all terms are dominant and even regular dominant.\(^{(2)}\) This reduces us to the special case of Lemma 3.8 where both \( \nu \) and \( \nu - \beta^\vee \) are regular dominant. Applying Lemma 3.8 we conclude

\[
t_{\nu} \leq t_{m_1} \leq \cdots \leq t_{\mu},
\]

hence \( t_{\nu} \) is \( \mu \)-admissible.

If \( \nu \) is arbitrary there exists a conjugate under \( w \in W \) which is dominant, and any such conjugate by Lemma 3.6 has the same length. By a general lemma of Haines [H3], Lemma 4.5, elements of \( \tilde{W} \) which are conjugate under a simple reflection and of

\(^{(2)}\)This argument, due to the referee, allows us to avoid Stembridge’s Lemma, e.g. [R3], used in the original proof.
the same length are simultaneously \(\mu\)-admissible. The result follows by induction by writing \(w\) as a product of simple reflections and using Lemma 3.6 repeatedly.\(^{(3)}\) \(\square\)

In the preceding considerations we looked at the situation for an Iwahori subgroup. Let us now make some comments on the subsets \(\text{Adm}_K(\mu)\) and \(\text{Perm}_K(\mu)\) of \(\tilde{W}^K \setminus \tilde{W}/\tilde{W}^K\) for an arbitrary parahoric subgroup \(K\), cf. (3.6) and (3.10). First of all we note that as a consequence of Proposition 3.2 and the surjectivity of (3.8) we have the following statement.

**Proposition 3.10.** — Let \(G\) be split over \(L\). Then

\[
\text{Adm}_K(\mu) \subset \text{Perm}_K(\mu).
\]

If \(\text{Adm}(\mu) = \text{Perm}(\mu)\), and \(\text{Perm}(\mu) \to \text{Perm}_K(\mu)\) is surjective, then also \(\text{Adm}_K(\mu) = \text{Perm}_K(\mu)\).

We note that, as proved in [KR1], all these statements hold true if \(G\) is equal to \(GL_n\) or to \(GSp_{2n}\) and \(\mu\) is minuscule.\(^{(4)}\)

Let now \(G\) be split over \(L\), and let \(K\) be a special maximal parahoric subgroup. We may take the vertex fixed by \(K\) to be the origin in the apartment. This identifies

\[
(3.30) \quad \tilde{W}^K \setminus \tilde{W}/\tilde{W}^K = X_*(S)/W = X_*(S) \cap \overline{C},
\]

where \(X_*(S) \cap \overline{C}\) are the dominant elements for some ordering of the roots. Let us choose \(\mu \in W(\mu)\) dominant and introduce the partial order as before (3.28) (i.e., the difference is a sum of positive coroots).

**Proposition 3.11.** — Let \(G\) be split over \(L\) and let \(K\) be a special maximal parahoric subgroup. With the notations introduced we have

\[
\text{Adm}_K(\mu) = \text{Perm}_K(\mu) = \{\nu \in X_*(S) \cap \overline{C}; \nu \overset{1}{\leq} \mu\}.
\]

**Proof.** — Let \(\nu \in \text{Perm}_K(\mu)\). Then \(\nu\) and \(\mu\) have the same image in \(X_*(S)/X_*(S_{sc})\). Hence the condition on \(\nu\) to be \(\mu\)-permissible, which says that \(t_{\nu}(0) \in \mathcal{P}_\mu\), is equivalent to

\[
(3.31) \quad \nu \overset{1}{\leq} \mu.
\]

Hence it suffices to show that (3.31) implies \(t_\nu \leq t_\mu\). But this is shown by the proof of Proposition 3.5 above. \(\square\)

**Corollary 3.12.** — In the situation of the previous proposition assume that \(\mu\) is minuscule. Then \(\text{Adm}_K(\mu)\) consists of one element, namely \(\mu \in X_*(T)/W\).

**Proof.** — Indeed in this case the set appearing in the statement of the Proposition consists of \(\mu\) only, cf. [K1]. \(\square\)

\(^{(3)}\) See Notes at the end, n° 2
\(^{(4)}\) See Notes at the end, n° 3
4. Affine Deligne-Lusztig varieties

In this section we change notations. Let $F$ be a finite extension of $\mathbb{Q}_p$ and let $L$ be the completion of the maximal unramified extension of $F$ in a fixed algebraic closure $\overline{F}$ of $F$. We denote by $\sigma$ the relative Frobenius automorphism of $L/F$. Let $G$ be a connected reductive group over $F$ and let $\tilde{G}$ be the group over $L$ obtained by base change. Let $\mathcal{B} = \mathcal{B}(G_{\text{ad}}, L)$ be the Bruhat-Tits building of $G_{\text{ad}}$. The Bruhat-Tits building of $G_{\text{ad}}$ over $F$ can be identified with the set of $\langle \sigma \rangle$-invariants in $B$.

We fix a maximal split torus $\tilde{S}$ of $\tilde{G}$ which is defined over $F$ (such tori exist by [BT2], 5.1.12.) We also fix a facet $a'$ in the apartment corresponding to $\tilde{S}$ which is invariant under $\langle \sigma \rangle$. Let $\tilde{K} = \tilde{K}(a')$ be the corresponding parahoric subgroup of $\tilde{G}(L)$. The subgroup $K = \tilde{K} \cap G(L)$ is called the parahoric subgroup of $G(L)$ corresponding to $a'$. [The subgroup $K$ determines $a'$ uniquely, and hence we obtain a bijection between the set of parahoric subgroups of $G(L)$, the set of $\sigma$-invariant parahoric subgroups of $G(L)$, and the set of $\sigma$-invariant facets of $\mathcal{B}$.] From Prop. 2.4 we have a map (with obvious notation),

\[(4.1) \quad \text{inv} : \tilde{G}(L)/\tilde{K} \times G(L)/\tilde{K} \to \tilde{W}^K \backslash \tilde{W}/\tilde{W}^K,\]

in which the target space can be identified with the quotient of the source by the diagonal action of $\tilde{G}(L)$.

**Definition 4.1.** — Let $w \in \tilde{W}^K \backslash \tilde{W}/\tilde{W}^K$ and $b \in G(L)$. The generalized affine Deligne-Lusztig variety associated to $w$ and $b$ is the set

\[X_w(b) = \{ g \in \tilde{G}(L)/\tilde{K} ; \text{inv}(b \sigma(g), g) = w \}.\]

When $\tilde{K}$ is an Iwahori subgroup, in which case $\tilde{W}^K$ is trivial, i.e. $w \in \tilde{W}$, this set is called the affine Deligne-Lusztig variety associated to $w$ and $b$.

Let

\[J_b(F) = \{ h \in G(L) ; h^{-1} b \sigma(h) = b \}.\]

Then $J_b(F)$ acts on $X_w(b)$ via $g \mapsto hg$.

**Remarks 4.2**

(i) If $b' \in G(L)$ is $\sigma$-conjugate to $b$, i.e. $b' = h^{-1} b \sigma(h)$, then the map $g \mapsto g' = hg$ induces a bijection

\[X_w(b) \xrightarrow{\sim} X_w(b').\]

Sometimes it is useful to indicate the parahoric subgroup $K$ in the notation. If $K' \subset K$, then denoting by $w$ a representative of $w$ in $\tilde{W}^K \backslash \tilde{W}/\tilde{W}^K$, there is a natural map

\[X_w(b)_{K'} \to X_w(b)_K.\]

(ii) One could hope to equip $X_w(b)$ with the structure of an algebraic variety locally of finite type over the residue field $\mathbf{F}$ of $\mathcal{O}_L$. 

ASTÉRISQUE 298
Example 4.3. — Let $G = GL_2$ and $K = K_0$ = standard Iwahori subgroup. We associate to $b \sigma$ its slope vector $\lambda = (\lambda_1, \lambda_2) \in \mathbb{Q}^2$. Then $\lambda_1 \geq \lambda_2$ with $\lambda_1 + \lambda_2 \in \mathbb{Z}$ and $\lambda_2 \in \mathbb{Z}$ if $\lambda_1 \neq \lambda_2$. If $X_w(b) \neq \emptyset$ then the image of $w$ in

$$X_w(T)_I / X_w(S_{w_0}) = \mathbb{Z}$$

coincides with $\lambda_1 + \lambda_2$. Conversely, let us assume this and let us enumerate the $w \in \tilde{W}$ for which $X_w(b) \neq \emptyset$. We distinguish cases.

(a) $b$ basic, i.e. $\lambda_1 = \lambda_2$.

(a1) $\lambda_1 + \lambda_2$ odd.

In this case $X_w(b) \neq \emptyset$ $\iff$ $\ell(w)$ is even and then $\dim X_w(b) = \ell(w)/2$.

(a2) $\lambda_1 + \lambda_2$ even.

In this case $X_w(b) \neq \emptyset$ $\iff$ either the projection of $w$ to $W_a$ is trivial, in which case $\dim X_w(b) = 0$, or $\ell(w)$ is odd, in which case $\dim X_w(b) = (\ell(w) + 1)/2$.

(b) $b$ hyperbolic, i.e. $\lambda_1 \neq \lambda_2$.

In this case $X_w(b) \neq \emptyset$ $\iff$ either $w = t(\lambda_1, \lambda_2)$, in which case $\dim X_w(b) = 0$, or $\ell(w) = \ell(t(\lambda_1, \lambda_2))$, and $\ell(w) \equiv \lambda_1 - \lambda_2 + 1 \mod 2$, in which case $\dim X_w(b) = (\ell(w) - (\lambda_1 - \lambda_2 - 1))/2$.

Before going on, we recall some definitions of Kottwitz [K2], [K4]. Let $B(G)$ be the set of $\sigma$-conjugacy classes of elements of $G(L)$. The homomorphism $\kappa_G$, cf. (2.1), induces a map

$$\kappa_G : B(G) \longrightarrow X^*(\hat{Z}(G)^\Gamma).$$

Here $\Gamma = \text{Gal}(\overline{F}/F)$ denotes the absolute Galois group of $F$. We also recall the Newton map,

$$\varphi : B(G) \longrightarrow \mathfrak{A}^+.$$

Here the notation is as follows. Let $G^*$ be the quasisplit inner form of $G$. Let $B^*$ be a Borel subgroup of $G^*$ defined over $F$ and let $T^*$ be a maximal torus in $B^*$. Then

\footnote{See Notes at the end, no 4}
\[ \mathfrak{A} = X_*(T^*)_{R}^{\mathfrak{A}} \] and \( \mathfrak{A}^+ \) denotes the intersection of \( \mathfrak{A} \) with the positive Weyl chamber in \( X_*(T^*)_{R} \) corresponding to \( B^* \). For instance, if \( G = GL_n \), then the Newton map associates to \( b \in G(L) \) the slopes of the isocrystal \( (L^n, b \sigma) \) with multiplicity of each slope equal to the dimension of the corresponding isotypical component, in decreasing order. An element \( b \in B(G) \) is called basic if \( \mathfrak{p}_b \) is central, \( i.e. \) if \( \mathfrak{p}_b \in X_*(Z)_{R} \). This is the analogue for general \( G \) of an isocrinal isocrystal. At the opposite extreme of the basic elements of \( B(G) \) are the unramified elements. Namely, let \( G = G^* \) be the quasisplit, and let \( A \) be a maximal split torus contained in \( T^* \). Let \( b \in A(L) \). Then \( \mathfrak{p}_b \) is the unique dominant element in the conjugacy class of \( \text{ord} (b) \in X_*(A) \subset \mathfrak{A} \) (this follows from the functoriality of the Newton map).

Let now \( \{ \mu \} \) be a conjugacy class of one-parameter subgroups of \( G \). Then \( \{ \mu \} \) determines a well-defined element \( \mu^* \) in \( X_*(T^*)_{R} \) lying in the positive Weyl chamber (use an inner isomorphism of \( G \) with \( G^* \) over \( \mathcal{O} \)). Let

\[
(4.4) \quad \mathfrak{p}^* = \left[ \Gamma : \Gamma_{\mu^*} \right]^{-1} \cdot \sum_{\tau \in \Gamma / \Gamma_{\mu^*}} \tau(\mu^*).
\]

Then \( \mathfrak{p}^* \in \mathfrak{A}^+ \). On the other hand, \( \{ \mu \} \) determines a well-defined element \( \mu^\flat \) of \( X^*(\tilde{Z}(G)\Gamma) \).

We define a finite subset \( B(G, \mu) \) of \( B(G) \) as the set of \( b \in B(G) \) satisfying the following two conditions,

\[
(4.5) \quad \kappa_G(b) = \mu^\flat
\]
\[
(4.6) \quad \mathfrak{p}_b \leq \mathfrak{p}^*,
\]

cf. [K4], section 6. Here in (4.6) there appears the usual partial order on \( \mathfrak{A}^+ \), for which \( \nu \preceq \nu' \) if \( \nu' - \nu \) is a nonnegative linear combination of simple relative coroots.

The motivation for the definition of \( B(G, \mu) \) comes from the following fact. We return to the notation of the beginning of this section. Let us assume that \( G \) is quasisplit over \( F \) and \( \tilde{G} \) split over \( L \), \( i.e. \), \( G \) is unramified. Let \( K \) be a hyperspecial maximal parahoric subgroup. Then \( T = \tilde{S} \) and \( \tilde{W}^K \setminus \tilde{W} / \tilde{W}^K \) can be identified with \( X_*(\tilde{S}) / W_0 \).

**Proposition 4.4 ([RR]).** — Let \( \mu \in X_*(\tilde{S}) / W_0 \). For \( b \in G(L) \) let \( [b] \in B(G) \) be its \( \sigma \)-conjugacy class. Then

\[ X_\mu(b) \neq \emptyset \implies [b] \in B(G, \mu). \]

This is the group theoretic version of Mazur’s inequality between the Hodge polygon of an \( F \)-crystal and the Newton polygon of its underlying \( F \)-isocrystal.(6)

**Example 4.5.** — Let \( G = GL_n \) and let \( T = S \) be the group of diagonal matrices, and \( K \) the stabilizer of the standard lattice \( O^*_F \) in \( F^n \). Then, with the choice of the upper

---

(6) See Notes at the end, n° 5
triangular matrices for $B^*$,

$$\mathfrak{A}^+ = (\mathbb{R}^n)_+ = \{ \nu = (\nu_1, \ldots, \nu_n) \in \mathbb{R}^n; \nu_1 \geq \nu_2 \geq \cdots \geq \nu_n \}.$$ 

For the usual partial order on $\mathfrak{A}^+$ we have $\nu \preceq \nu'$ iff

$$\sum_{i=1}^{r} \nu_i \leq \sum_{i=1}^{r} \nu'_i \text{ for } r = 1, \ldots, n-1 \text{ and } \sum_{i=1}^{n} \nu_i = \sum_{i=1}^{n} \nu'_i.$$ 

Let $b \in G(L)$ and let $(N, \Phi) = (L^n, b \cdot \sigma)$ be the corresponding isocrystal of dimension $n$. If $M$ is an $O_L$-lattice in $N$ we have

$$\mu(M) = \text{inv}(M, \Phi(M)) \in (\mathbb{Z}^n)_+.$$ 

Here $(\mathbb{Z}^n)_+ = \mathbb{Z}^n \cap (\mathbb{R}^n)_+$ and $\mu(M) = (\mu_1, \ldots, \mu_n)$ iff there exists a $O_L$-basis $e_1, \ldots, e_n$ of $M$ such that $\pi^{\mu_1} e_1, \ldots, \pi^{\mu_n} e_n$ is a $O_L$-basis of $\Phi(M)$. By $\pi$ we denoted a uniformizer of $F$. Denoting by $\nu_b$ the Newton vector of the isocrystal $(N, \Phi)$, Mazur’s inequality states that

$$\nu_b \leq \mu(M),$$

i.e. $[b] \in B(G, \mu(M))$.

Conjecture 4.6. — The converse in the implication of Proposition 4.4 holds.

In this direction we have the following results.\(^{(7)}\)

Theorem 4.7. — The converse implication in Proposition 4.4 holds in either of the following cases.

(i) \([KR2]\) $G = GL_n$ or $G = GSp_{2n}$.

(ii) \([R3]\) The derived group of $G$ is simply connected, $\mu \in X_*(\tilde{S})/W_0$ is $\langle \sigma \rangle$-invariant and $b \in A(F)$, where $A$ denotes a maximal $F$-split torus in $G$.

Whereas the individual affine Deligne-Lusztig varieties are very difficult to understand, the situation seems to change radically when we form a suitable finite union of them. This is the subject of the next section.

5. The sets $X(\mu, b)_K$

We continue with the notation of the previous section. Let $\{\mu\}$ be a conjugacy class of one-parameter subgroups of $G$. Equivalently, $\{\mu\}$ is a $W$-orbit in $X_*(T)$ where $T$ denotes the centralizer of $\tilde{S}$. We again introduce the subsets $\text{Adm}_K(\mu)$ resp. $\text{Perm}_K(\mu)$ of $\tilde{W} \backslash W / \tilde{W}$ and again set $K = \tilde{K} \cap G(F)$.

Let $b \in G(L)$. Then we define the following set, a finite union of generalized affine Deligne-Lusztig varieties,

$$X(\mu, b)_K = \{ g \in \tilde{G}(L)/\tilde{K}; \text{ inv}(g, b\sigma(g)) \in \text{Adm}_K(\mu) \}.$$  

\(^{(7)}\)See Notes at the end, n° 6
Let $E \subset \overline{F}$ be the field of definition of $\{\mu\}$. Let $E_0 = E \cap L$ and $r = [E_0 : F]$. We note that $\sigma$ acts in compatible way on $\tilde{W}$ and its subgroup $X_*(T)_I$, and that the map (4.1) is compatible with this action.

**Lemma 5.1.** — The subset $\Lambda(\{\mu\})$ of $X_*(T)_I$ (defined shortly after (3.1)) is invariant under $\sigma^r$. Hence also the subsets $\operatorname{Perm}_K(\mu)$ and $\operatorname{Adm}_K(\mu)$ of $\tilde{W}/\tilde{W}^r$ are invariant under $\sigma^r$.

**Proof.** — Let $\nu \in X_*(T)$ with image $[\nu]_I$ in $X_*(T)_I$. Then

$$\sigma^r([\nu]_I) = [\tau(\nu)]_I,$$

for $\tau \in \operatorname{Gal}(\overline{L}/F)$ an arbitrary lifting of $\sigma^r$. We take for $\tau$ an an extension of the automorphism $\operatorname{id} \otimes \sigma^r$ of $E_L = E \otimes_{E_0} L$. Then $\tau \in \operatorname{Gal}(\overline{L}/E)$ and hence preserves the orbit $\{\mu\}$ in $X_*(T)$. Furthermore $\tau$ normalizes $\operatorname{Gal}(\overline{L}/L)$. Hence if $\mu \in \overline{C}_E \cap \{\mu\}$ for a Borel subgroup defined over $L$, then $\tau(\mu) \in \overline{C}_E \cap \{\mu\}$ for another Borel subgroup defined over $L$. Hence if $\lambda \in X_*(T)_I$ denotes the image of $\mu$, then $\sigma^r(\lambda) = w_0(\lambda)$ for some $w_0 \in W_0$ which implies the first assertion. The second assertion follows (for $\operatorname{Adm}_K(\mu)$ use that $w_1 \leq w_2$ implies $\sigma(w_1) \leq \sigma(w_2)$).

Using this lemma we can now define an operator $\Phi$ on $X(\mu, b)_K$ by

$$\Phi(g) = (b \sigma)^r \cdot g \cdot \sigma^{-r} = b \cdot \sigma(b) \cdots \sigma^{-1}(b) \cdot \sigma^r(g).$$

Let us check that $\Phi$ indeed preserves the set $X(\mu, b)_K$. We have

$$\operatorname{inv}(\Phi(g), b \sigma(\Phi(g))) = \operatorname{inv}(\sigma^r(g), \sigma^r(b \sigma(g))) = \sigma^r(\operatorname{inv}(g, b \sigma(g))).$$

The claim follows from Lemma 5.1.

In the context of Remark 4.2, (ii), the set $X(\mu, b)_K$ may be expected to be the set of $F$-points of an algebraic variety over $F$, and $\Phi$ would define a Weil descent datum over the residue field $\kappa_E$ of $E$ in the sense of [RZ2].

As mentioned at the end of the last section, whereas it seems difficult to understand when the individual affine Deligne-Lusztig varieties which make up $X(\mu, b)_K$ are non-empty, their union seems to behave better, at least in the cases when $\{\mu\}$ is minuscule.

**Conjecture 5.2.** — Let $\{\mu\}$ be minuscule.

(a) $X(\mu, b)_K \neq \emptyset$ if and only if the class $[b]$ of $b$ in $B(G)$ lies in the subset $B(G, \mu)$.

(b) For $K \subset K'$, the induced map $X(\mu, b)_K \to X(\mu, b)_{K'}$ is surjective.

**Remark 5.3**

Suppose that $K$ is hyperspecial. Then, if $X(\mu, b)_K \neq \emptyset$, it follows that $[b] \in B(G, \mu)$, cf. Prop. 4.4. This holds even when $\{\mu\}$ is not minuscule. In general, it
is not clear whether the hypothesis that \( \{ \mu \} \) be minuscule is indeed necessary in Conjecture 5.2.\(^{(8)}\)

In the direction of Conjecture 5.2 we first note the following easy observation.

**Lemma 5.4.** — If \( X(\mu, b)_K \neq \emptyset \), then

\[
\kappa(b) = \mu^2.
\]

**Proof.** — We consider the composition \( \tilde{\kappa} = \tilde{\kappa}_G \) of \( \tilde{\kappa}_G \) and the natural surjection,

\[
G(L) \rightarrow X^*(\tilde{Z}(G)^L) \rightarrow X^*(\tilde{Z}(G)^L).
\]

The map \( \tilde{\kappa}_G \) induces \( \kappa_G \) on \( B(G) \). If \( gK \in X(\mu, b)_K \), then \( g^{-1}b\sigma(g) = k_1w_1 \) with \( k_1, k_2 \in K \) and with \( w \in \text{Adm}_K(\mu) \). Since \( k_1, k_2 \in \text{Ker} \tilde{\kappa}_G \) we conclude that

\[
\tilde{\kappa}(b) = \tilde{\kappa}(g^{-1}b\sigma(g)) = \tilde{\kappa}(k_1w_1) = \tilde{\kappa}(w).
\]

But

\[
\text{W}K_\text{wW}K_\text{wW}K \leq \text{W}K_t\mu\text{W}K,
\]

for a conjugate \( \mu' \) of \( \mu \). Since \( \text{W}K \subset W_a \) and \( W_a \subset \text{Ker} \tilde{\kappa} \) (since \( \tilde{\kappa}(G_{sc}(L)) = 0 \)), we conclude that \( \tilde{\kappa}(w) = \tilde{\kappa}(t_\mu) \). If \( \mu' = w_0(\mu) \) for \( w_0 \in W_0 \) we have

\[
\tilde{\kappa}(t_\mu) = \tilde{\kappa}(w_0t_\muw_0^{-1}) = \tilde{\kappa}(t_\mu),
\]

hence \( \tilde{\kappa}(b) = \tilde{\kappa}(t_\mu) = \mu^2. \)

**Theorem 5.5 ([KR2]).** — Conjecture 5.2 holds for minuscule \( \mu \), in the cases \( G = R_{F'/F}(GL_n) \) and \( G = R_{F'/F}(GSp_{2n}) \), where \( F' \) is an unramified extension of \( F \).

In fact, in loc. cit. also the case when \( G \) is an inner form of \( GL_n \) is treated.

We also mention the following case when Conjecture 5.2a) holds.

**Proposition 5.6.** — Assume that \( G \) splits over \( L \) and that the center of \( G \) is connected. Let \( b \in G(L) \) be such that \( [b] \in B(G)_{\text{basic}} \). Let \( K_0 \) be an Iwahori subgroup defined over \( F \). Then \( X(\mu, b)_{K_0} \neq \emptyset \) \( \iff \) \( [b] \in B(G, \mu) \).

**Proof.** — It is obvious that if \( [b] \in B(G)_{\text{basic}} \), then \( [b] \in B(G, \mu) \) iff \( \kappa(b) = \mu^2 \). Hence one implication (\( \Rightarrow \)) follows from Lemma 5.4. Now let \( [b] \in B(G, \mu) \cap B(G)_{\text{basic}} \).

**Claim.** — Let \( N \) be the normalizer of a maximal torus \( S \) which splits over \( L \), and let \( \tilde{K} \subset G(L) \) be a \( \sigma \)-stable Iwahori subgroup corresponding to an alcove in the apartment of \( S \). Then there exists a representative \( b' \) of \( [b] \) in \( N(L) \) which normalizes \( \tilde{K} \).

\(^{(8)}\)See Notes at the end, n° 7.
Proof of Claim. — Since the center of $G$ is connected, the map $G(L) \to G_{\text{ad}}(L)$ is surjective. Hence we may replace $G$ by $G_{\text{ad}}$, in which case $B(G)_{\text{basic}} = H^1(F, G)$. Hence any representative of $b$ defines an inner form of $G$ which splits over $L$. Assume that there exists a representative $b'$ of $[b]$ in $N(L)$ as in the claim. Then in the corresponding inner form of $G$ there exists a maximal torus which splits over $F_{\text{un}}$ and an $F$-rational Iwahori subgroup fixing an alcove in the apartment for this torus. Conversely, if the inner form of $G$ corresponding to a representative of $[b]$ has this property, then this representative normalizes this maximal torus and the Iwahori subgroup. Now, since any inner form of $G$ contains a maximal torus which splits over $F_{\text{un}}$ and an $F$-rational Iwahori subgroup fixing an alcove in the apartment for this torus, such a representative must exist, which proves the claim.

Let $g \in G(L)$ be such that $\tilde{K} = gK_0g^{-1}$. Then $gb\sigma(g)^{-1} \in \tilde{K}_0w\tilde{K}_0$, where $w \in \tilde{W}$ normalizes $\tilde{K}_0$. It follows that the component of $w$ in $W_\alpha$ is trivial, hence by Lemma 5.4, $w \leq t_\mu$. If $b' = h\sigma(h)^{-1}$ then $gb\tilde{K}_0 \in X_w(b\sigma) \subset X(\mu, \tilde{b})_{\tilde{K}_0}$.

The following statement yields an inequality which goes in a sense in the opposite direction to that defining $B(G, \mu)$.

Proposition 5.7. — Let $G$ be split over $F$. Let $S$ be a maximal split torus over $F$. Let $b \in S(L)$. Let $K_0$ denote the Iwahori subgroup fixing an alcove in the apartment of $B$ corresponding to $S$. Let $w \in \tilde{W}$ such $X_w(b\sigma) \neq \emptyset$, i.e.,

$$\exists \ g \in G(L) : g^{-1}b\sigma(g) \in K_0wK_0.$$

Then for the translation by the Newton point $w_\nu \in X(S) \cap \mathfrak{A}^+$ of $b$ we have

$$t_{w_\nu} \leq w.$$

Proof. — Let $\mathcal{A}$ be the apartment in $B$ corresponding to $S = S \otimes F L$, and let $a_0 \subset \mathcal{A}$ be the base alcove fixed by $K_0$, and let $a = ga_0$, for $g \in X_w(b\sigma)$. Let $\alpha$ be the automorphism of $B$ induced by $b\sigma$. Then we have for the component $w_\alpha$ of $w$ in the affine Weyl group

$$w_\alpha = \text{inv}(a, \alpha(a)).$$

Let $\mathcal{C}$ be any quartier corresponding to the positive vector Weyl chamber, after a choice of a special vertex of $a_0$. Only the germ of $\mathcal{C}$ will be relevant to us, i.e. $\mathcal{C}$ up to translation by an element of $X_S(S_{\text{sc}})R$. Let $\varrho_{\mathcal{A}, \mathcal{C}}$ be the corresponding retraction, i.e. $\varrho_{\mathcal{A}, \mathcal{C}} = \varrho_{\mathcal{A}, a'}$ for some alcove $a'$ far into the quartier, compare [BT1], Prop. 2.9.1. Then we have the following two statements.

$$\alpha \circ \varrho_{\mathcal{A}, \mathcal{C}}(a) = \varrho_{\mathcal{A}, \mathcal{C}} \circ \alpha(a)$$

(5.6)

$$\text{For any two alcoves } a, a', \quad \text{inv}(\varrho_{\mathcal{A}, \mathcal{C}}(a), \varrho_{\mathcal{A}, \mathcal{C}}(a')) \leq \text{inv}(a, a').$$

To see (5.6), note that $\alpha$ preserves $\mathcal{A}$ and the germ of $\mathcal{C}$, hence

$$\alpha \circ \varrho_{\mathcal{A}, \mathcal{C}} \circ \alpha^{-1}(a) = \varrho_{\alpha(\mathcal{A}), \alpha(\mathcal{C})}(a) = \varrho_{\mathcal{A}, \alpha(\mathcal{C})}(a).$$

(5.8)
Since $α$ also preserves the germ of $C$, it follows that $θ_{A,α(c)}(a) = θ_{A,c}(a)$.

To see (5.7), let $a = a_0, a_1, \ldots, a_ℓ = a'$ be a minimal gallery $Γ$ between $a$ and $a'$. This corresponds to a minimal decomposition of $x = \text{inv}(a, a')$,

(5.9) \[ x = s_1 \cdots s_ℓ. \]

Here $s_1, \ldots, s_ℓ$ are the reflections around the walls of type $a_0 \cap a_1, \ldots, a_ℓ−1 \cap a_ℓ$ of the base simplex (|[BT1], 2.3.10). The image of $Γ$ under $ϕ = θ_{A,c}$ is a gallery $\Gamma'$ between $\varpi = ϕ(a)$ and $\varpi' = ϕ(a')$. Furthermore, the type $\varpi_1, \ldots, \varpi_ℓ$ of $Γ'$ is identical with that of $Γ$ ([BT1], 2.3.4.) Let us replace $Γ'$ by a minimal gallery. Then we can write $\varpi = \text{inv}(\varpi, \varpi')$ as

(5.10) \[ \varpi = s_i_1 \cdots s_i_k \]

([BT1], 2.1.9 and 2.1.11). Hence $\varpi ≤ x$, which proves (5.7).

We now apply this to $a$ and $α(a)$. But for any $\varpi$ contained in $A$ we have

(5.11) \[ \text{inv}(\varpi, α(\varpi)) = (t_ν)_α. \]

Here $ν = ν_b ∈ X_*(S)$. Hence using (5.6) and (5.7),

\[ (t_ν)_α = \text{inv}(ϕ(a), αϕ(a)) = \text{inv}(ϕ(a), ϕ(α(a))) \]

\[ ≤ \text{inv}(a, α(a)) = w_α. \]

Taking into account the definition of the Bruhat order on $\hat{W}$, the assertion follows. □

 Whereas Proposition 5.6 concerned the case of a basic element $b$, the following proposition treats the other extreme, namely, unramified elements $b$.

**Proposition 5.8.** — Let $G, S, \hat{K}_0$ and $b ∈ S(L)$ be as in the previous proposition. Assume that $G_{der}$ is simply connected. Then $X(μ, b)_{K_0} ≠ \emptyset ⇔ [b] ∈ B(G, μ)$.

**Proof.** — If $X(μ, b)_{K_0} ≠ \emptyset$, there exists $w ∈ \text{Adm}(μ)$ such that $X_w(bσ) ≠ \emptyset$. This implies $κ(w) = κ(b)$ and $t_ν ≤ w$, by Lemma 5.4 and Proposition 5.7. Since $w ≤ t_μ'$ for some conjugate $μ'$ of $μ$ it follows that $t_ν ≤ t_μ'$ which implies that $W_0t_ν$ $W_0 ≤ W_0t_μW_0$ and hence $ν_b ≤ μ$, i.e. $[b] ∈ B(G, μ)$.

Conversely, let $[b] ∈ B(G, μ)$. Hence $ν_b$ and $μ$ are both dominant elements in $X_*(S)$ with $ν_b ≤ μ$. However, for any alcove $a$ in the apartment $A$ corresponding to $S$, we have $\text{inv}(a, bσ(a)) = (t_ν)_α$, hence $X_{t_ν}(bσ) ≠ \emptyset$. Hence it suffices to see that $t_ν ∈ \text{Adm}(μ)$. But $ν_b ≤ μ$, hence since $G_{der}$ is simply connected, $ν_b ≤ μ$. Therefore by Proposition 3.11 $t_ν ≤ t_μ$, i.e. $t_ν ∈ \text{Adm}(μ)$. □

As mentioned above, the sets $X(μ, b)_K$ should have the structure of an algebraic variety over the residue field $F$ of $O_L$, at least when $μ$ is minuscule. For their dimension there is a conjectural formula when $b$ is basic. To state it we first mention the following result. The set $B(G, μ)$ is partially ordered (a finite poset) by $[b] ≤ [b']$.
iff $\nu_1 \leq \nu_2$ in $\mathfrak{A}^+$. That this is indeed a partial order follows from the fact that the map $(\nu, \kappa) : B(G) \to \mathfrak{A}^+ \times X^*(\widetilde{Z}(G)^\Gamma)$ is injective [K4], [RR].

**Theorem 5.9 (Chai [C2]).** — Assume $G$ quasisplit over $F$.

(i) Any subset of $B(G, \mu)$ has a join, i.e. a supremum.

(ii) The poset $B(G, \mu)$ is ranked, i.e. any two maximal chains between two comparable elements have the same length.

(iii) Let $[b], [b'] \in B(G, \mu)$ with $[b] \leq [b']$. Then the length of the maximal chain between $[b]$ and $[b']$ is given by

$$\text{length}([b], [b']) = \ell \sum_{i=1}^{\ell} (|\langle \omega_i, \nu_{[b]} \rangle - \langle \omega_i, \nu^{\star} \rangle | - |\langle \omega_i, \nu_{[b]} \rangle - \langle \omega_i, \nu^{\star} \rangle |).$$

Here $\omega_1, \ldots, \omega_\ell$ are the fundamental $F$-weights of the adjoint group $G_{ad}$, i.e. $\langle \omega_i, \alpha_j^\vee \rangle = \delta_{ij}$ for any simple relative coroot $\alpha_j^\vee$. Also $[x]$ denotes the greatest integer $\leq x$.

We note that $B(G, \mu)$ has a unique minimal element, namely the unique basic element $[b_0]$ in $B(G, \mu)$, and a unique maximal element, namely the $\mu$-ordinary element $[b_1] = [b_\mu]$ for which $\nu_{[b_1]} = \nu^\star$. Given (i) and (ii) of Theorem 5.9, the formula in (iii) is equivalent to

$$\text{length}([b], [b_1]) = -\ell \sum_{i=1}^{\ell} (|\langle \omega_i, \nu_{[b]} \rangle - \langle \omega_i, \nu^{\star} \rangle |).$$

The dimension formula for $X(\mu, b)_K$ may now be given as follows.

**Conjecture 5.10.** — Let $K$ be a hyperspecial maximal parahoric. Let $[b] = [b_0] \in B(G, \mu)$ be basic. Then $X(\mu, b)_K$ is equidimensional of dimension

$$\dim X(\mu, b)_K = \langle 2\varrho, \nu^{\star} \rangle - \text{length}([b], [b_\mu])$$

$$= \langle 2\varrho, \nu^{\star} \rangle + \sum_{i=1}^{\ell} (|\langle \omega_i, \nu^{\star} - \nu_{[b]} \rangle |).$$

Here $\varrho$ denotes the half-sum of all positive roots$^{(9)}$.

The motivation for this formula comes from global considerations connected with the Newton strata of Shimura varieties, compare Theorem 7.4 below. It would be interesting to extend this conjecture to the non-basic case$^{(10)}$ and also to the case when $K$ is no longer hyperspecial.

$^{(9)}$See Notes at the end, no 8

$^{(10)}$Added in October 2004: In the light of recent results of Chai and Oort, and of Mierendorff, the following formula seems reasonable,

$$\dim X(\mu, b)_K = \langle 2\varrho, \nu^{\star} - \nu_{[b]} \rangle + \sum_{i=1}^{\ell} (|\langle \omega_i, \nu^{\star} - \nu_{[b]} \rangle |)$$

ASTÉRISQUE 298
6. Relations to local models

We continue with the notation of the last two sections. In particular, $G$ denotes a connected reductive group over $F$ and $\{\mu\}$ is a conjugacy class of one-parameter subgroups of $G$. Again $E$ is the field of definition of $\{\mu\}$. Let $K$ be a parahoric subgroup of $G(F)$ and $\tilde{K}$ the corresponding parahoric subgroup of $G(L)$. We denote by $G = G_\mu$ the group scheme over $O_F$ corresponding to $K$, cf. Remark 2.2, (ii).

To these data one would like to associate the local model, a projective scheme $\mathcal{M}^{loc} = \mathcal{M}^{loc}(G, \mu)_K$ over $\text{Spec} \ O_E$, equipped with an action of $G_{O_E}$, at least if $\{\mu\}$ is minuscule. It is not clear at the moment how to characterize $\mathcal{M}^{loc}$ or how to construct it in general. It should have at least the following properties.

(i) $\mathcal{M}^{loc}$ is flat over $\text{Spec} \ O_E$ with generic fiber isomorphic to $G/P_\mu$. Here $P_\mu$ denotes the variety of parabolic subgroups in the conjugacy class of parabolic subgroups corresponding to $\{\mu\}$.

(ii) There is an identification of the geometric points of the special fiber,

$$\mathcal{M}^{loc}(\tilde{\pi}_E) = \{g \in G(L)/\tilde{K}; \tilde{K}g\tilde{K} \in \text{Adm}_{\tilde{K}}(\mu)e\}.$$

(iii) $\mathcal{M}^{loc}(G, \mu)_K$ is functorial in $K$ and in $G$.

Examples 6.1

(1) If $K$ is hyperspecial, then we set $\mathcal{M}^{loc}(G, \mu)_K = G_{O_E}/P_\mu$, where $P_\mu$ is in the conjugacy class of parabolic subgroups in $G_{O_E}$ corresponding to $\{\mu\}$. In this case $\mathcal{M}^{loc}(G, \mu)_K$ is smooth over $\text{Spec} \ O_E$. Property (ii) follows from Corollary 3.12.

(2) Let $V$ be an $F$-vector space of dimension $n$. Let $G = GL(V)$ and let $\{\mu\}$ be minuscule of weight $r$ for some $0 \leq r \leq n$, i.e. $\omega_r \in \{\mu\}$, where $\omega_r(t) = \text{diag}(t, \ldots, t, 1, \ldots, 1)$ with $r$ times $t$ and $n - r$ times $1$. Let $e_1, \ldots, e_n$ be a basis of $V$ and, for $0 \leq i \leq n - 1$, let

$$\Lambda_i = \text{span}_{O_F}\{\pi^{-1}e_1, \ldots, \pi^{-1}e_i, e_{i+1}, \ldots, e_n\}.$$

For a non-empty subset $I$ of $\{0, 1, \ldots, n - 1\}$, let $K = K_I$ be the parahoric subgroup of $G(F)$ which is the common stabilizer of the lattices $\Lambda_i$, for $i \in I$. The local model $\mathcal{M}^{loc} = \mathcal{M}^{loc}(G, \mu)_K$ for this triple $(G, \{\mu\}, K_I)$ represents the following moduli problem on $(\text{Sch} / \text{Spec} \ O_F)$ (in this case $E = F$). To $S$ the functor associates the set of commutative diagrams of $O_S$-modules,

$$\begin{array}{ccccccc}
\Lambda_{i_0,S} & \rightarrow & \Lambda_{i_1,S} & \rightarrow & \cdots & \rightarrow & \Lambda_{i_m,S} \\
\cup & \cup & \cup & \cup & \cup & \cup & \\
\mathcal{F}_{i_0} & \rightarrow & \mathcal{F}_{i_1} & \rightarrow & \cdots & \rightarrow & \mathcal{F}_{i_m} \\
\end{array}$$

Here $I = \{i_0 < i_1 < \cdots < i_m\}$ and we have set $\Lambda_i,S = \Lambda_i \otimes_{O_F} O_S$. It is required that $\mathcal{F}_{i_j}$ is a locally free $O_S$-module of rank $r$ which is locally a direct summand of $\Lambda_i,S$. The main result of the paper [G1] of Görtz is that $\mathcal{M}^{loc}$ satisfies the conditions (i) and (ii) above.
Let $V$ be a $F$-vector space of dimension $2n$ with a symplectic form $\langle \ , \ \rangle$. Let 
$G = \text{GSp}(V, \langle \ , \ \rangle)$ and let $\{\mu\}$ be minuscule of weight $n$. Let $e_1, \ldots, e_{2n}$ be a symplectic basis of $V$, i.e.
\[
\langle e_i, e_j \rangle = 0, \langle e_{i+n}, e_{j+n} \rangle = 0, \langle e_i, e_{2n-j+1} \rangle = \delta_{ij}
\]
for $i, j = 1, \ldots, n$. Let $I$ be a non-empty subset of $\{0, \ldots, 2n - 1\}$ which with $i \neq 0$ also contains $2n - i$. Let $K = K_I$ be the parahoric subgroup of $G(F)$ which is the common stabilizer of the lattices $\Lambda_i$ for $i \in I$. The local model $\mathcal{M}^{\text{loc}} = \mathcal{M}^{\text{loc}}(G, \mu)_K$ for the triple $(G, \{\mu\}, K_I)$ represents the moduli problem on $(\text{Sch}/\mathcal{O}_F)$ which to $S$ associates the objects $(\mathcal{F}_{i_1}, \ldots, \mathcal{F}_{i_m})$ of the local model for $(\text{GL}(V), \{\mu\}, K_I)$ as in Example (2) above which satisfy the following additional conditions. If $0 \in I$, then $\mathcal{F}_0$ is isotropic for the symplectic form on $\Lambda_{0,S}$, and for each $i \in I$ with $i \neq 0$, the composition
\[
\mathcal{F}_i \longrightarrow \Lambda_{i,S} \cong \hat{\Lambda}_{2n-i,S} \longrightarrow \mathcal{F}_{2n-i}
\]
is the zero map. Here “hat” denotes the dual $\mathcal{O}_S$-module.

By the main result of [G2], $\mathcal{M}^{\text{loc}}$ satisfies the conditions (i) and (ii) above.

(4) Let $G = R_{F'/F}(\text{GL}_n)$ or $G = R_{F'/F}(\text{GSp}_{2n})$, where $F'$ is a totally ramified extension. Let $\{\mu\}$ be a minuscule conjugacy class of one-parameter subgroups and let $K$ be a parahoric subgroup of $G(F)$. In [PR1] resp. [PR2] local models $\mathcal{M}^{\text{loc}}(G, \mu)_K$ are constructed which satisfy conditions (i) and (ii) above. But in these cases it seems difficult to describe the functors that these local models represent.

In all these examples, property (ii) for local models can be considerably strengthened by identifying the special fiber of $\mathcal{M}^{\text{loc}}(G, \mu)_K$ with a closed subscheme of the partial flag variety corresponding to $\check{K}$ of the loop group over $\kappa_E$ associated to $G$, [G1], [PR2]. Here, to be on the safe side, we are assuming $G$ split. Via this identification there is a link between the theory of local models and the geometric Langlands program of Beilinson, Drinfeld et al. [BD].

The true significance of the local models becomes more transparent when they appear in the global context of Shimura varieties, cf. (7.1). It is in this context that they were first introduced, originally by P. Deligne and G. Pappas [DP] for the Hilbert-Blumenthal varieties, then by J. de Jong [J] for the Siegel moduli space and, still later, in [RZ2]. Here we explain their relation with the sets $X(\mu, b)_K$. Let
\[
\check{K}_1 = \ker(G(O_L) \longrightarrow G(\mathfrak{F}_E)).
\]
Let
\[
X(\mu, b)_{K_1} = \{g \in G(L)/\check{K}_1; \text{inv}(g\check{K}, b\sigma(g)\check{K}) \in \text{Adm}_K(\mu)\}.
\]
In other words, $X(\mu, b)_{K_1}$ is the inverse image of $X(\mu, b)_K$ under $G(L)/\check{K}_1 \rightarrow G(L)/\check{K}$. We define a map
\[
(6.1) \quad \overline{\gamma} : X(\mu, b)_{K_1} \longrightarrow \mathcal{M}^{\text{loc}}(G, \mu)_K(\mathfrak{F}_E)
\]
by
\[ \tilde{\gamma}(g\tilde{K}_1) = g^{-1}b\sigma(g) \cdot \tilde{K}. \]
This is well-defined since \( \tilde{K}_1 \) acts trivially on \( \mathcal{M}^{\text{loc}}(\overline{\kappa}_E) \). Noting that \( \tilde{K}/\tilde{K}_1 \) is a principal homogeneous space under \( \mathcal{G}(\overline{\kappa}_E) \), we may write \( \tilde{\gamma} \) more suggestively as a map on geometric points of algebraic stacks,
\[ (6.2) \quad \gamma: X(\mu, b)_K \longrightarrow [\mathcal{M}^{\text{loc}}/\mathcal{G} \otimes_{\mathcal{O}_E} \mathcal{O}_E](\overline{\kappa}_E). \]
It should be possible, at least if \( \{\mu\} \) is minuscule, to equip \( X(\mu, b)_K \) with the structure of an algebraic variety over \( \kappa_E \) and the map \( \gamma \) should be induced by a morphism of algebraic stacks over \( \kappa_E \),
\[ X(\mu, b)_K \longrightarrow [\mathcal{M}^{\text{loc}} \otimes_{\mathcal{O}_E} \overline{\kappa}_E] / \mathcal{G} \otimes_{\mathcal{O}_E} \overline{\kappa}_E]. \]
Furthermore, this morphism should be compatible with Weil descent data over \( \kappa_E \) on source and target.

After Görtz’s theorems the most interesting question is the following conjecture. A variant was also proposed by G. Pappas [P], Conj. 2.12.

**Conjecture 6.2.** — Assume that \( G \) is unramified over \( F \). Let \( \mathcal{M}^{\text{loc}}(G, \mu)_K \) be the local model over \( \text{Spec}\, \mathcal{O}_E \), corresponding to a parahoric subgroup \( K \) of \( G(F) \) and a minuscule conjugacy class of cocharacters \( \{\mu\} \), with its action of \( \mathcal{G} \otimes_{\mathcal{O}_F} \mathcal{O}_E \). Then there exists a \( \mathcal{G} \otimes_{\mathcal{O}_F} \mathcal{O}_E \)-equivariant blowing up in the special fiber \( \overline{\mathcal{M}}^{\text{loc}}(G, \mu)_K \rightarrow \mathcal{M}^{\text{loc}}(G, \mu)_K \) which has semistable reduction.

In Example 6.1, (2) the conjecture above is trivial for \( r = 1 \) (in this case \( \mathcal{M}^{\text{loc}} \) has semistable reduction). For \( r = 2 \), Faltings [F2] has constructed an equivariant blowing-up with semistable reduction, i.e., the conjecture holds in this case, compare also [L]. In Example 6.1, (3) the existence of a semistable blowing-up is due to de Jong [J] for \( n = 2 \) and to Genestier [Ge] for \( n = 3 \).

**PART II**

**GLOBAL THEORY**

**7. Geometry of the reduction of a Shimura variety**

In the global part we use the following notation. Let \( (G, \{h\}) \) be a Shimura datum, i.e. \( G \) is a connected reductive group over \( \mathbb{Q} \) and \( \{h\} \) a \( G(\mathbb{R}) \)-conjugacy class of homomorphisms from \( R_{C/R}\mathbb{G}_m \) to \( G_{\mathbb{R}} \) satisfying the usual axioms. We fix a prime number \( p \). We let \( K \) be an open compact subgroup of \( G(\mathbb{A}_f) \) which is of the form \( K = K^p \cdot K_p \), where \( K^p \subset G(\mathbb{A}_f^p) \) and where \( K_p = K^p \) is a parahoric subgroup of \( G(\mathbb{Q}_p) \). We also assume that \( K^p \) is sufficiently small to exclude torsion phenomena.
The corresponding Shimura variety $\text{Sh}(G, \{h\})_K$ is a quasi-projective variety defined over the Shimura field $E$, a finite number field contained in the field $\overline{\mathbb{Q}}$ of algebraic numbers. It is the field of definition of the conjugacy class $\{\mu_h\}$, where $\mu_h$ is the cocharacter corresponding to $h \in \{h\}$. Let $G = G \otimes \mathbb{Q} \mathbb{Q}_p$ and $F = \mathbb{Q}_p$. After fixing an embedding of $\overline{\mathbb{Q}}$ into an algebraic closure of $\mathbb{Q}_p$, we obtain a conjugacy class $\{\mu\}$ of one-parameter subgroups of $G$. It is defined over the completion $E$ of $E$ in the $p$-adic place $p$ induced by this embedding. We have therefore obtained by localization a triple $(G, \{\mu\}, K)$ as in the local part relative to $F = \mathbb{Q}_p$.

We make the basic assumption that the Shimura variety has a good integral model over $\text{Spec} \mathcal{O}_{E,p}$. Although we do not know how to characterize it, or how to construct it in general, we know a good number of examples all related to moduli spaces of abelian varieties. The “facts” stated below all refer to these moduli spaces, and the conjectures also concern these moduli spaces or are extrapolations to the general case. We denote by $\text{Sh}(G, h)_K$ the model over $\text{Spec} \mathcal{O}_E$ which is obtained by base change $\mathcal{O}_{E,p} \to \mathcal{O}_E$ from this good integral model.

The significance of the local model $\mathcal{M}^{\text{loc}}(G, \mu)_K$ is given by the relatively representable morphism of algebraic stacks over $\text{Spec} \mathcal{O}_E$,

$$\lambda : \text{Sh}(G, h)_K \to [\mathcal{M}^{\text{loc}}(G, \mu)_K / \mathcal{G}_{\mathcal{O}_E}].$$

Here $G$ is the group scheme over $\text{Spec} \mathbb{Z}_p$ corresponding to $K_p$. The morphism $\lambda$ is smooth of relative dimension $\dim G$. This statement is proved in those cases where an integral model of the Shimura variety $\text{Sh}(G, h)_K$ exists $[\text{PR2}], [\text{RZ2}]$. The crux is, just as in (1.11), to find a normal form for the DeRham homology modules of a polarized $L$-set of abelian varieties up to prime-to-$p$ isogeny. Once this is established, the assertion is a consequence of the theorem of Grothendieck-Messing.

We now turn to the special fiber. By associating to each point of $\text{Sh}(G, h)_K(\kappa_E)$ the isomorphism class of its rational Dieudonné module (again this makes sense only for those Shimura varieties which are moduli spaces of abelian varieties), we obtain a map

$$\delta : \text{Sh}(G, h)_K(\kappa_E) \to B(G).$$

We note that by Proposition 4.4 the image of $\delta$ is contained in $B(G, \mu)$, provided that $G$ is unramified and $K$ is hyperspecial.

**Conjecture 7.1.** — $\text{Im}(\delta) = B(G, \mu)$.

In particular, we expect that the basic locus is non-empty. Let $\overline{\text{Sh}(G, h)}_K = \text{Sh}(G, h)_K \otimes \mathcal{O}_E \kappa_E$. The basic locus $\overline{\text{Sh}(G, h)}_K,_{\text{basic}}$ is the set of points whose image under $\delta$ is the unique basic element $[b_0]$ of $B(G, \mu)$. More generally, for $[b] \in B(G)$, let

$$S_{[b]} = \delta^{-1}([b]).$$

**Conjecture 7.1.** — $\text{Im}(\delta) = B(G, \mu)$.
Proposition 7.2 ([RR]). — Each $S_{[b]}$ is a locally closed subvariety of $\overline{\text{Sh}}(G, h)_K$. Furthermore, for $[b], [b'] \in B(G)$, we have

$$S_{[b]} \cap \text{closure}(S_{[b']}) \neq \emptyset \implies [b'] \leq [b']$$

This is the group-theoretic version of Grothendieck's semicontinuity theorem, according to which the Newton vector of an isocrystal decreases under specialization (in the natural partial order on $(\mathbb{R}^n)_+$, cf. (4.5)). The subvarieties $S_{[b]}$ are called the Newton strata of $\overline{\text{Sh}}(G, h)_K$. The Newton stratification of the special fiber is very mysterious.

Questions 7.3. — Assume that $K_p$ is hyperspecial. Let $[b], [b'] \in B(G, \mu)$ with $[b] \leq [b']$.

(i) Is $S_{[b]} \cap \text{closure}(S_{[b']}) \neq \emptyset$?

(ii) Is $S_{[b]} \cap \text{closure}(C) \neq \emptyset$, for every irreducible component $C$ of $S_{[b']}$?

(iii) Is $S_{[b]} \subset \text{closure}(S_{[b']})$?

Obviously, (i) is implied by (iii). Here part (iii) has become known as the strong Grothendieck conjecture and (i) as the weak Grothendieck conjecture, although this denomination is somewhat abusive.

Theorem 7.4 (Oort [O]). — For the Shimura variety associated to $GSp_{2n}$ (the Siegel moduli space), question 7.3 (iii) has an affirmative answer. Also in this case, each Newton stratum $S_{[b]}$ is equidimensional of codimension in $\overline{\text{Sh}}(G, h)_K$ equal to

$$\text{codim} S_{[b]} = \text{length}([b], [b_\mu])$$

compare Theorem 5.9. Here $[b_\mu] = [b_1]$ denotes the $\mu$-ordinary element of $B(G, \mu)$, cf. (5.12).\(^{(11)}\)

Conjecture 7.5 (Chai [C2]). — Assume $K_p$ hyperspecial. Each Newton stratum $S_{[b]}$ is equidimensional of codimension given by the above formula.

Since the basic element in $B(G, \mu)$ is minimal, the basic locus $\overline{\text{Sh}}(G, h)_{K, \text{basic}}$ is a closed subvariety of the special fiber. This variety has been studied in many cases ([LO], [Ka], [Ri]): it is conceivable that one can give a group-theoretical “synthetic” description of it in general. At the other extreme is the $\mu$-ordinary element $[b_\mu] \in B(G, \mu)$. It is the unique maximal element of $B(G, \mu)$, cf. (5.12).

Conjecture 7.6 (Chai). — Let $K_p$ be hyperspecial. The orbit of any point of $S_{[b_\mu]}$ under $G(A_f^p)$ is dense in $\overline{\text{Sh}}(G, h)_K$.

Here the action of $G(A_f^p)$ is via Hecke correspondences. In this direction we have the following results.

Theorem 7.7 (Chai [C1]). — The conjecture 7.6 is true for the Siegel moduli space.

\(^{(11)}\)See Notes at the end n° 9
**Theorem 7.8 (Wedhorn [W]).** — We assume that the Shimura variety $\text{Sh}(G, h)_K$ corresponds to a PEL-moduli problem of abelian varieties. Let $K_p$ be hyperspecial. The $\mu$-ordinary stratum $S_{[b_{\mu}]}$ is dense in $\text{Sh}(G, h)_K$.

The hypothesis that $K_p$ be hyperspecial in Wedhorn’s theorem is indeed necessary, as the examples of Stamm [S] relative to the Hilbert-Blumenthal surfaces with Iwahori level structure at $p$ and of Drinfeld [D] relative to a group which is ramified at $p$ show.

We finally relate the maps $\gamma$ and $\lambda$. Let $[b] \in B(G, \mu)$ and let $b \in G(L)$ be a representative of $[b]$. As in Definition 4.1, we let

$$J_b(Q_p) = \{ g \in G(L); g^{-1}b \sigma(g) = b \}.$$ 

The Newton stratum $S_{[b]}$ has a covering $\tilde{S}_b$, for which we fix an isomorphism of the isocrystal in the variable point $x \in S_{[b]}$ with the model isocrystal with $G$-structure determined by $b$. Then $\tilde{S}_b$ is a principal homogeneous space under $J_b(Q_p)$ over $S_{[b]}$. The relation between $\gamma$ and $\lambda$ is then given by a commutative diagram of morphisms of algebraic stacks (over $\kappa_E$, or even compatible with Weil descent data over $\kappa_E$),

$$\begin{array}{ccc}
\tilde{S}_b & \xrightarrow{\tilde{\lambda}} & X(\mu, b)_{K_p} \\
\downarrow & & \downarrow \gamma \\
S_{[b]} & \xrightarrow{\lambda} & \left[ \tilde{\mathcal{M}}_{\text{loc}}^{\text{G}}(G, \mu)_{K_p}/\mathcal{G}_{\kappa_E} \right].
\end{array}$$

Here $\tilde{\mathcal{M}}_{\text{loc}}$ resp. $\mathcal{G}$ denote the special fibers of the local model resp. the group scheme corresponding to $K_p$, and by $\lambda$ we denoted the restriction of (7.1) to $S_{[b]}$. The commutativity of the diagram above reflects the well-known compatibility between DeRham homology and Dieudonné theory.

### 8. Pseudomotivic and quasi-pseudomotivic Galois gerbs

In this section and the next one we wish to give a conjectural description of the point set of $\text{Sh}(G, h)_K(\kappa_E)$ with its action of the Frobenius automorphism. This description is modeled on the one given in [LR], but differs from it in an important detail, compare Remark 9.3. The idea is to partition the point set into “isogeny classes”, as was done in the case of the elliptic modular curve in (1.1), and then to describe the point set of the individual isogeny classes in a manner reminiscent of (1.1) and (1.2) resp. (1.9) in the elliptic modular case. According to an idea of Grothendieck, the set of isogeny classes will be described in terms of representations of certain Galois gerbs. In this section we introduce these Galois gerbs, following Reimann’s book [Re1].

We first explain our terminology concerning Galois gerbs. Let $k$ be a field of characteristic zero and let $\Gamma = \text{Gal}(\overline{k}/k)$ be the Galois group of a chosen algebraic...
closure. A *Galois gerb over* \( k \) is an extension of topological groups
\[
1 \longrightarrow G(\overline{k}) \longrightarrow \mathcal{G} \overset{q}{\longrightarrow} \Gamma \longrightarrow 1.
\]

Here \( G \) denotes a linear algebraic group over \( \overline{k} \) and is called the *kernel* of \( \mathcal{G} \). The topology on \( \Gamma \) is the Krull topology and the topology on \( G(\overline{k}) \) is the discrete topology. The extension \( \mathcal{G} \) is required to satisfy the following two conditions:

(i) For any representative \( g_\sigma \in \mathcal{G} \) of \( \sigma \in \Gamma \), the automorphism \( g \mapsto g_\sigma g g_\sigma^{-1} \) of \( G(\overline{k}) \) is a \( \sigma \)-linear algebraic automorphism.

(ii) Let \( K/k \) be a finite extension over which \( G \) is defined. Let \( \Gamma_K = \text{Gal}(k/K) \) be the corresponding subgroup of \( \Gamma \). We choose a section of \( G \to \Gamma \) over \( \Gamma_K \) such that the automorphism \( g \mapsto g_\sigma g g_\sigma^{-1}, g \in G(\overline{k}) \) defines the \( K \)-structure on \( G \). Then the resulting bijection
\[
q^{-1}(\Gamma_K) = G(\overline{k}) \times \Gamma_K
\]
is a homomorphism.

A *morphism between Galois gerbs* \( \varphi : \mathcal{G} \to \mathcal{G}' \) is a continuous map of extensions which induces the identity map on \( \Gamma \) and an algebraic homomorphism on the kernel groups. Two homomorphisms \( \varphi_1 \) and \( \varphi_2 \) are called *equivalent* if there exists \( g' \in G'(\overline{k}) \) with \( \varphi_2 = \text{Int}(g') \circ \varphi_1 \). A *neutral gerb* is one isomorphic to the semi-direct product
\[
\mathcal{G}_G = G(\overline{k}) \rtimes \Gamma
\]
associated to an algebraic group \( G \) over \( k \). Sometimes we will have to consider a slightly more general notion. Let \( k' \) be a Galois extension of \( k \) contained in \( \overline{k} \) with Galois group \( \Gamma' = \text{Gal}(k'/k) \). Then one defines in the obvious way the notion of a \( k'/k \)-Galois gerb, which is an extension
\[
1 \longrightarrow G(k') \longrightarrow \mathcal{G} \longrightarrow \Gamma' \longrightarrow 1,
\]
where \( G \) is an algebraic group defined over \( k' \). A \( k'/k \)-Galois gerb defines in the obvious way a \( \overline{k}/k \)-Galois gerb (i.e. a Galois gerb): one first pulls back the extension by the surjection \( \Gamma \to \Gamma' \) and then pushes out via \( G(k') \to G(k) = (G \otimes_{k'} \overline{k})(\overline{k}) \).

In the sequel we will have to deal with projective limits of Galois gerbs of the previous kind. We transpose the above terminology to them. In particular two morphisms of pro-Galois gerbs will be called equivalent if they are projective limits of equivalent morphisms of Galois gerbs (in [Re1], B.1.1, this is called *algebraically equivalent*).

An important example is given by the Dieudonné gerb over \( \mathbf{Q}_p \), cf. [Re1], B.1.2. For every \( n \in \mathbf{Z}, n \geq 1 \), there is an explicitly defined \( \mathbf{Q}_p^{un}/\mathbf{Q}_p \)-gerb \( \mathcal{D}_n \) with kernel group \( \mathbf{G}_m \). For \( n' \) divisible by \( n \) there is a natural homomorphism \( \mathcal{D}_{n'} \to \mathcal{D}_n \) inducing the map \( x \mapsto x^{n'/n} \) on the kernel groups. Let
\[
\mathcal{D}_0 = \lim_{\rightarrow} \mathcal{D}_n
\]
be the pro-$\mathbb{Q}_p^\text{un}/\mathbb{Q}_p$-Galois gerb defined by this projective system. Then in $\mathcal{D}_0$ there is an explicit representative $d_\sigma$ of the Frobenius element. The Diedernd gerb $\mathcal{D}$ is the pro-$\mathbb{Q}_p^\text{un}/\mathbb{Q}_p$-Galois gerb defined by $\mathcal{D}_0$.

Another Galois gerb of relevance to us is the weight gerb $\mathcal{W}$. This is the Galois gerb over $\mathbb{R}$ with kernel $\mathcal{G}_m$, which is defined by the fundamental cocycle of $\text{Gal}(\mathbb{C}/\mathbb{R})$

\[
(w_{q,\sigma} = -1 \text{ if } q = \sigma \text{ = complex conjugation; otherwise } w_{q,\sigma} = 1).
\]

We now recall some pertinent facts about the pro-Galois gerbs appearing in the title of this section. We fix an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ and for every place $\ell$ of $\mathbb{Q}$ an embedding $\mathbb{Q} \subset \overline{\mathbb{Q}}$. Let $L/\mathbb{Q}$ be a finite Galois extension contained in $\overline{\mathbb{Q}}$.

There is an initial object $(\mathbb{Q}_L, \nu(\infty), \nu(p))$ in the category of all triples $(T, \nu(\infty), \nu(p))$ where $T$ is a $\mathbb{Q}$-torus which splits over $L$ and such that $\nu(\infty), \nu(p) \in X_*(T)$ satisfy

\[
[\mathbb{L}_: \mathbb{R}]^{-1} \cdot \text{Tr}_{L/\mathbb{Q}}(\nu(\infty)) + [\mathbb{L}_p : \mathbb{Q}_p]^{-1} \cdot \text{Tr}_{L/\mathbb{Q}}(\nu(p)) = 0,
\]

cf. [Re1], B.2.2.

Similarly, assume that $L$ is a CM-field and denote by $L_0$ its maximal totally real subfield. Then there is an initial object $(P^L, \nu(\infty)^L, \nu(p)L)$ in the category of all triples $(T, \nu(\infty), \nu(p))$ where $T$ is a $\mathbb{Q}$-torus which splits over $L$ and such that $\nu(\infty), \nu(p) \in X_*(T)$ are defined over $\mathbb{Q}$ and $\mathbb{Q}_p$ respectively and such that

\[
\nu(\infty) + [\mathbb{L}_p : \mathbb{Q}_p]^{-1} \cdot \text{Tr}_{L/L_0}(\nu(p)) = 0,
\]

cf. [Re1], B.2.3. Since obviously condition (8.5) implies condition (8.4), there is a canonical morphism

\[
(Q^L, \nu(\infty)^L, \nu(p)L) \rightarrow (P^L, \nu(\infty)^L, \nu(p)L).
\]

If $L \subset L'$ then we obtain morphisms of tori in the opposite direction,

\[
Q^L \rightarrow Q'^L, \quad P^L \rightarrow P'^L.
\]

Let $Q$ resp. $P$ denote the pro-torus defined by this projective system. Then there are homomorphisms of pro-tori over $\mathbb{Q}$,

\[
\nu(\infty) : \mathcal{G}_{m,\mathbb{R}} \rightarrow \mathbb{Q}_R, \quad \text{resp. } \nu(p) : \mathcal{D} \rightarrow \mathbb{Q}_p,
\]

whose composite with $Q \rightarrow Q^L$ is $\nu(\infty)^L$ if $L$ is totally imaginary resp. is $[L_p : \mathbb{Q}_p] \cdot \nu(p)L$. Here $\mathcal{D}$ denotes the pro-torus with character group equal to $\mathbb{Q}$.

Similarly, we obtain

\[
\nu(\infty) : \mathcal{G}_{m,\mathbb{R}} \rightarrow \mathbb{P}_R, \quad \nu(p) : \mathcal{D} \rightarrow \mathbb{P}_Q.
\]

We can now introduce the pro-Galois gerbs which will be relevant for the theory of Shimura varieties.
A quasi-pseudomotivic Galois gerb is a pro-Galois gerb $\mathfrak{Q}$ over $\mathbb{Q}$ with kernel $Q$ together with morphisms
\begin{align*}
\zeta_\infty & : W \longrightarrow \mathfrak{Q}_R \\
\zeta_p & : D \longrightarrow \mathfrak{Q}_p \\
\zeta_\ell & : \Gamma_\ell \longrightarrow \mathfrak{Q}_\ell, \quad \ell \neq \infty, p
\end{align*}
(8.10)
such that $\nu(\infty)$ is induced by $\zeta_\infty$ resp. $\nu(p)$ is induced by $\zeta_p$ on the kernel group $G_{mR}$ of the weight gerb $W$ resp. the kernel group $D$ of the Dieudonné gerb $D$. In addition, a coherence condition on the family $\{\zeta_\ell; \ell \neq \infty, p\}$ is imposed, cf. [Re1], B.2.7. Similarly one defines a pseudomotivic Galois gerb $(\mathfrak{P}, \zeta_p^{\prime})$. These pro-Galois gerbs are uniquely defined up to an isomorphism preserving the morphisms $\zeta_\ell$ up to equivalence for $\ell = \infty, p$ and $\ell \neq p$. Furthermore, these isomorphisms are unique up to equivalence. There is a morphism
\begin{equation}
\mathfrak{Q} \longrightarrow \mathfrak{P}
\end{equation}
(8.11)
compatible with the morphisms $\zeta_\ell$ resp. $\zeta_p^{\prime}$ and inducing the homomorphism (8.6) above on the kernel groups.

For a pair $(T, \mu)$ consisting of a $\mathbb{Q}$-torus $T$ and an element $\mu \in X_*(T)$, there is associated a morphism of Galois gerbs
\begin{equation}
\psi_\mu : \mathfrak{Q} \longrightarrow \mathfrak{G}_T,
\end{equation}
(8.12)
cf. [Re1], B.2.10. This morphism factors through $\mathfrak{P}$ if and only if the following two conditions are satisfied,

(i) the image of $\nu(\infty)$ in $X_*(T)$ is defined over $\mathbb{Q}$

(ii) the image of $\nu(p)$ in $X_*(T) \otimes \mathbb{Q}$ satisfies the Serre condition, i.e., it is defined over a CM-field and its weight is defined over $\mathbb{Q}$, cf. [Re1], B.2.11. This is the case if $T$ itself satisfies the Serre condition, i.e., $(\text{id} + i)(\tau - \text{id}) = (\tau - \text{id})(\text{id} + i)$ in $\text{End}(X_*(T))$, where $i$ denotes the complex conjugation and $\tau \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is arbitrary.

**Remark 8.1.** — The pseudomotivic Galois gerb was introduced in [LR] with the aim of describing the points in the reduction of a Shimura variety when $(G, \{h\})$ satisfies the Serre condition. When this last condition is dropped, the pseudomotivic Galois gerb cannot suffice for this purpose. However, the quasi-pseudomotivic Galois gerb in [LR], introduced there to cover the cases when the Serre condition fails, does not exist (there is a fatal error in the construction of loc.cit.). Two replacements have been suggested, one by Pfau [Pf] and one by Reimann [Re1]. We follow here the latter.

9. Description of the point set in the reduction

In this section we return to the notation used in section 7. Therefore $\text{Sh}(G, h)_K$ is an integral model over $\text{Spec} \mathcal{O}_E$ of the Shimura variety associated to $(G, \{h\}, K =
$K^p.K_p$), and $\{\mu_h\}$ is the associated conjugacy class of cocharacters of the reductive group $G$ over $\text{Spec } Q$. Our purpose is to describe the set $\text{Sh}(G, h)_K(\mathcal{O}_E)$ of our model over $\text{Spec } \mathcal{O}_E$ of the Shimura variety $\text{Sh}(G, h)_K$. We make the blanket assumption that the derived group of $G$ is simply connected.

The description of the points in the reduction will be in terms of admissible morphisms of pro-Galois gerbs

$$\varphi : \Omega \rightarrow \mathcal{G}_G.$$  

**Definition 9.1.** — A morphism (9.1) of pro-Galois gerbs over $Q$ is called admissible if it satisfies the four conditions a)–d) below.

Let $D = G/G_{\text{der}}$ and let $\mu_D$ be the image of $\{\mu_h\}$ in $X_*(D)$. The first condition is global:

(a) The composition $\Omega \rightarrow \mathcal{G}_G \rightarrow \mathcal{G}_D$ is equivalent to $\psi_{\mu_D}$, cf. (8.12).

The next three conditions will be local, one for each place of $Q$. To formulate the next condition we remark that for $h \in \{h\}$ with corresponding weight homomorphism $w_h : G_{m, \mathbb{R}} \rightarrow G_{\mathbb{R}}$ the pair $(w_h, \mu_h(-1))$ corresponds to a morphism of Galois gerbs over $\mathbb{R}$,

$$\xi_{\infty} : \mathcal{W} \rightarrow \mathcal{G}_{\mathbb{R}}.$$

(b) The composition $\varphi \circ \zeta_{\infty}$ is equivalent to $\xi_{\infty}$.

(c) For any $\ell \neq \infty, p$ the composition $\varphi \circ \zeta_{\ell}$ is equivalent to the canonical section $\xi_{\ell}$ of $\mathcal{G}_{\mathcal{A}_{\ell}}$.

For the final condition we remark that (the equivalence class of) the composition $\varphi \circ \zeta_p : D \rightarrow \mathcal{G}_{\mathcal{Q}_p^\text{un}}$ defines an element $[b] = [b(\varphi_p)] = [b(\varphi)]$ of $B(G)$. More precisely, let $\mathcal{D}_0$ be the explicit unramified version of the Dieudonné gerb as in [Re1], B.2, compare (8.3). Then there exists a morphism

$$\theta_0 : \mathcal{D}_0 \rightarrow G(\mathcal{Q}_p^\text{un}) \times \hat{\mathbb{Z}}$$

such that $\varphi \circ \zeta_p$ is equivalent to the pullback $\overline{\theta}_0$ of $\theta_0$ to $\Gamma$. Then $[b]$ is the class of $b = \theta_0(d_\sigma)$, where $d_\sigma \in \mathcal{D}_0$ is the explicit representative of the Frobenius $\sigma$.

(d) The element $[b]$ lies in $B(G, \mu)$.

We note that, whereas the local components $\varphi \circ \zeta_{\infty}$ and $\varphi \circ \zeta_{\ell}$ ($\ell \neq p$) are uniquely determined up to equivalence by the Shimura data, the $p$-component $\varphi \circ \zeta_p$ is allowed to vary over a finite set of equivalence classes.

To every admissible morphism $\varphi$ we shall associate a set $S(\varphi)$ with an action from the right of $Z(\mathcal{Q}_p) \times G(A_f^p)$ and a commuting action of an automorphism $\Phi$. Here $Z$ denotes the center of $G$. For $\ell \neq \infty, p$ let

$$X_{\ell} = \{g \in G(\overline{\mathbb{Q}}_\ell) ; \text{ Int}(g) \circ \xi_{\ell} = \varphi \circ \zeta_{\ell}\}.$$
By condition (c) this set is non-empty. We put

\[(9.5) \quad X^p = \prod_{\ell \neq \infty, p} X_\ell,\]

where the restricted product is explained in [Re1], B.3.6. The group \(G(A_f^p)\) acts simply transitively on \(X^p\). We also put \(X_p = X(\mu, b)_{K_p}\) in the notation of (5.1), where \(b \in G(L)\) is as above. It is equipped with commuting actions of \(Z(Q_p)\) and an operator \(\Phi\) (cf. (5.2)). Finally we introduce the group of automorphisms \(I_\varphi = \text{Aut}(\varphi)\). The group \(I_\varphi(Q)\) obviously operates on \(X^p\). Let \(g_p \in G(\overline{Q}_p)\) be such that

\[(9.6) \quad \varphi_p \circ \zeta_p = \text{Int} g_p \circ \overline{\theta}_0\]

where the notation is as in the formulation of condition (d) above. Then we obtain an embedding

\[(9.7) \quad I_\varphi(Q) \to J_b(Q_p), \quad h \mapsto g_p h g_p^{-1}\]

where \(J_b(Q_p)\) is the group associated to the element \(b\) which acts on \(X_p = X(\mu, b)_{K_p}\), cf. Definition 4.1. We now define

\[(9.8) \quad S(\varphi)_{K_p} = \lim_{K_p} I_\varphi(Q) \backslash X_p \times X^p / K_p, \quad \text{resp.} \quad S(\varphi)_K = I_\varphi(Q) \backslash X_p \times X^p / K^p,\]

where the limit is over all open compact subgroups \(K^p \subset G(A_f^p)\). On \(S(\varphi)_{K_p}\) we have commuting actions of the automorphism \(\Phi\) and of \(Z(Q_p) \times G(A_f^p)\) from the right.

**Conjecture 9.2.** — Assume that the derived group of \(G\) is simply connected. Then for every sufficiently small \(K^p\) there is a model \(\text{Sh}(G, \{h\})_K\) of \(\text{Sh}(G, \{h\})_K\) over \(\text{Spec} \mathcal{O}_{E(p)}\), such that the point set of its special fiber is a disjoint sum of subsets invariant under the actions of the Frobenius automorphism over \(\kappa_E\) and of \(Z(Q_p)\) and \(G(A_f^p)\),

\[\text{Sh}(G, h)_K(\kappa_E) = \coprod_{\varphi} \text{Sh}(G, h)_{K, \varphi},\]

and for each \(\varphi\) a bijection

\[\text{Sh}(G, h)_{K, \varphi} = S(\varphi)_K,\]

which carries the action of the Frobenius automorphism over \(\kappa_E\) on the left into the action of \(\Phi\) on the right and which commutes with the actions of \(Z(Q_p)\) and of \(G(A_f^p)\) (for variable \(K^p\)) on both sides. Here the disjoint union is taken over a set of representatives of equivalence classes of admissible morphisms \(\varphi : \Omega \to \mathcal{G}_G\).

We remark that if \(D\) splits over a CM-field and the weight homomorphism \(w_h\) is defined over \(Q\), every admissible morphism \(\varphi : \Omega \to \mathcal{G}_G\) factors through \(\mathfrak{P}\), cf. [Re1], B.3.9.

Note that we are not proposing a characterization of the model \(\text{Sh}(G, h)_K\). In the case where \(K_p\) is hyperspecial, such a characterization was suggested by Milne [M2]. In this case we expect \(\text{Sh}(G, h)_K\) to be smooth over \(\text{Spec} \mathcal{O}_{E(p)}\). In [Re3],

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
Reimann gives a wider class of parahoric subgroups $K_p$ for which one should expect the smoothness of this model, and he conjectures that this class is exhaustive.

Conjecture 9.2 has been proved by Reimann [Re1], Prop. 6.10 (and Remark 4.9) and Prop. 7.7, in the following case: $G$ is the multiplicative group of a quaternion algebra over a totally real field in which $p$ is unramified, which is either totally indefinite or which is split at all primes above $p$. Furthermore, $K_p$ is a maximal compact subgroup of $G(\mathbb{Q}_p)$. It has been proved for Shimura varieties of PEL-type by Milne [M1], when $K_p$ is hyperspecial. In [LR] it is shown how the conjecture is related to a hypothetical good theory of motives, compare also [M3].

Remark 9.3. — We note that if Conjecture 5.2 holds, then each summand in Conjecture 9.2 is non-empty. In [LR] (apart from a very special case of bad reduction) it was assumed that $K_p$ is hyperspecial, and the admissibility condition (d) was replaced by the condition that $X(\mu, b)_{K_p}$ be non-empty. From Remark 5.3 it follows that then $[b] \in B(G, \mu)$, i.e. condition d) above holds.

Remark 9.4. — Assume Conjecture 9.2. In [RZ2] it was shown that in certain very rare cases the Shimura variety $\text{Sh}(G, \{h\})_{K}$ admits a $p$-adic uniformization by (products of) Drinfeld upper half spaces. The proof in loc. cit. is a generalization of Drinfeld’s proof [D] of Cherednik’s uniformization theorem in dimension one. From the proof in [RZ2] it is clear that this can occur only when all admissible morphisms are locally equivalent, provided that all summands in Conjecture 9.2 are non-empty. It comes to the same to ask that $\varphi_p \circ \zeta_p$ is basic for any admissible morphism $\varphi$. In [K4] it is shown that when $G$ is adjoint simple such that $B(G, \mu)$ consists of a single element (which is then basic), then $(G, \mu)$ is the adjoint pair associated to $(D_{1/n}^{\times}, (1, 0, \ldots, 0))$ or $(D_{-1/n}^{\times}, (1, \ldots, 1, 0))$, where $D_{1/n}^{\times}$ resp. $D_{-1/n}^{\times}$ denotes the inner form of $GL_n$ associated to the central division algebra of invariant $1/n$ resp. $-1/n$. In other words, this result of Kottwitz implies in conjunction with Conjecture 9.2 and Conjecture 5.2 that there is no hope of finding cases of $p$-adic uniformization essentially different from those in [RZ2]. In particular, in all cases of $p$-adic uniformization the uniformizing space will be a product of Drinfeld upper half spaces.

10. The semi-simple zeta function

We continue with the notation of the previous section. One ultimate goal of the considerations of the previous section is to determine the local factor of the zeta function of $\text{Sh}(G, h)_{K}$ at $p$. Our present approach is through the determination of the local semi-simple zeta function [R2]. We refer to [HN2], §3.1 for a systematic exposition of the concepts of the semi-simple zeta function and semi-simple trace of Frobenius. The decisive property of the semi-simple trace of Frobenius on representations of the local Galois group is that it factors through the Grothendieck group. In the case of good reduction, the semi-simple zeta function coincides with the usual zeta function.
To calculate the semi-simple zeta function we may use the Lefschetz fixed point formula. Let $\kappa_E^n$ be the extension of degree $n$ of $\kappa_E$ contained in $\pi_E$. For $x \in \text{Sh}(G, h)_K(\kappa_E^n)$ we introduce the semisimple trace
\begin{equation}
\text{Contr}_n(x) = \text{tr}^s(Fr_n; R\Psi_x(\mathcal{Q}_\ell)).
\end{equation}
Here $R\Psi(\mathcal{Q}_\ell)$ denotes the complex of nearby cycles. By $Fr_n$ we denote the geometric Frobenius in $\text{Gal}(\kappa_E^n/\kappa_E)$. This is the contribution of $x$ to the Lefschetz fixed point formula over $\kappa_E^n$. In the case of good reduction, or more generally if $x$ is a smooth point of $\text{Sh}(G, h)_K$, then $\text{Contr}_n(x) = 1$.

For an admissible homomorphism $\varphi : \mathfrak{Q} \to \mathcal{G}_G$ as in Conjecture 9.2, we introduce the contribution of $\varphi$ (or its equivalence class) to the Lefschetz fixed point formula over $\kappa_E^n$:
\begin{equation}
\text{Contr}_n(\varphi) = \sum_{x \in \text{Sh}(G, h)_K(\kappa_E^n)} \text{Contr}_n(x).
\end{equation}

**Definition 10.1.** — A morphism $\varphi : \mathfrak{Q} \to \mathcal{G}_G$ is called special if there exists a maximal torus $T \subset G$ and an element $\mu \in X^*(T)$ which defines a one-parameter subgroup of $G$ in the conjugacy class $\{\mu_t\}$ such that $\varphi$ is equivalent to $i \circ \psi_\mu$, cf. (8.12). Here $i : G_T \to G_G$ denotes the canonical morphism defined by the inclusion of $T$ in $G$.

If $K_p$ is hyperspecial (and $G_{\text{der}}$ is simply connected, as is assumed throughout this section), then every admissible morphism is special ([LR], Th. 5.3), at least if it factors through $\Psi$.

**Conjecture 10.2.** — We have $\text{Contr}_n(\varphi) = 0$ unless $\varphi$ is special.

For some cases of this conjecture related to $GL_2$, compare [R1] and [Re1]. Note that this is really a conjecture about bad reduction. In the case of good reduction the cancellation phenomenon predicted by Conjecture 10.2 cannot occur since each point $x$ in $\text{Sh}(G, h)_K(\kappa_E^n)$ contributes 1 in this case. This is compatible with the remark immediately preceding the statement of the conjecture, which says that the conjecture is empty if $K_p$ is hyperspecial.

Let us explain how one would like to give a group-theoretic expression for $\text{Contr}_n(x)$. Let $x$ be represented by $(x_p, x^p) \in X_p \times X^p/K^p$ under the bijection (9.8). Let $n' = n \cdot r$, where as shortly after (5.1) $r = [\kappa_E : \mathbb{F}_p]$. Since $x$ is fixed under the $n$-th power of the Frobenius over $\kappa_E$, it is fixed under the $n'$-th power of the absolute Frobenius and we obtain an equation of the form
\begin{equation}
(\Phi^{n'}x_p, x^p) = h \cdot (x_p, x^p),
\end{equation}
for some $h \in I_\varphi(\mathbb{Q})$. By [K1], Lemma 1.4.9, it follows that there exists $c \in G(L)$ such that
\begin{equation}
c \cdot h^{-1} \cdot \Phi^{n'} \cdot c^{-1} = \sigma^{n'}.
\end{equation}
This is an identity in the semi-direct product $G(L) \rtimes \langle \sigma \rangle$, where $L$ is the completion of the maximal unramified extension $Q_p^{un}$ of $Q_p$. Let $Q_{p'}$, be the fixed field of $\sigma^{n'}$. Then the element $\delta \in G(L)$ which is defined by the equation

\[(10.5) \quad c \cdot (b\sigma) \cdot c^{-1} = \delta \sigma\]

lies in $G(Q_{p'})$. Here $b \in G(L)$ is the element defined before (9.3). Also, always by [K1], we have that $x' = c \cdot x \in G(Q_{p'})$. To simplify the notation put $\delta g' = \tilde{K}_p^{(\sigma^{n'})}$. Let

\[(10.6) \quad H = H(G(Q_{p'})\|K_{p'})\]

be the Hecke algebra corresponding to the parahoric subgroup $K_{p'}$. It may be conjectured that there exists an element $\phi_{p'} \in H$ with the following property. Let $g' \in G(Q_{p'})$ be a representative of $x'$. Then

\[(10.7) \quad \text{Constr}_n(x) = \phi_{p'}^{n'}(g'^{-1}\delta\sigma(g'))\]

Appealing to [K1], 1.5, we therefore obtain that the contribution of the admissible homomorphism $\varphi : \Omega \to G\sigma$ to the Lefschetz fixed point formula over $\kappa_E$ is a sum of terms of the form

\[(10.8) \quad v \cdot O_h(\phi^p) \cdot TO_\delta(\phi^p)\]

Here $O_h(\phi^p)$ is the orbital integral over $h \in I_p(Q)$ of the characteristic function of $K^p$ and $TO_\delta(\phi^p)$ the twisted orbital integral of $\phi^p$ over the twisted conjugacy class of $\delta \in G(Q_{p'})$. Furthermore, $v$ is a certain volume factor.

For the function $\phi^p_{p'}$ there is the following conjecture.

**Conjecture 10.3 (Kottwitz).** — Assume that $G$ splits over $Q_p$. Let $K_{p'}^{\circ}$ be an Iwahori subgroup of $G(Q_{p'})$ contained in $K_{p'}$. Then $\phi^p_{p'}$ is the image of $p^{n'(\sigma^{n'})} \cdot z_\lambda$ under the homomorphism of Hecke algebras

\[\mathcal{H}(G(Q_{p'})\|K_{p'}^{\circ}) \longrightarrow \mathcal{H}(G(Q_{p'})\|K_{p'})\]

Here $z_\lambda$ denotes the Bernstein function in the center of the Iwahori Hecke algebra associated to $\mu$, compare [H2], 2.3.

Recall that the center of $\mathcal{H}(G(Q_{p'})\|K_{p'}^{\circ})$ has a basis as a $\mathbb{C}$-vector space formed by the Bernstein functions $z_\lambda$, where $\lambda$ runs through the conjugacy classes of one-parameter subgroups of $G$.

In this direction we have the following facts.

**Theorem 10.4.** — Conjecture 10.3 holds in the following cases.

1. (Haines, Ngo [HN2]): $G = GL_n$ or $G = GSp_{2n}$.
2. (Haines [H2]): $G$ is an inner form of $GL_n$ and $\{\mu\} \ni \omega_1$ (the Drinfeld case).
It would be interesting to extend the statement of Kottwitz’ conjecture to the general case. Once this is done (and the corresponding conjecture proved!) it remains to calculate the sum over all equivalence classes of admissible homomorphisms $\varphi$ of the expressions (10.8). In fact, one would like to replace the twisted orbital integrals in (10.8) by an ordinary orbital integral of a suitable function on $G(\mathbb{Q}_p)$ and compare the resulting expression with the trace of a suitable function on $G(\mathbb{A})$ in the automorphic spectrum. When the Shimura variety is not projective, one also has to deal with the contribution of the points on the boundary. Even when the Shimura variety is projective, the phenomenon of $L$-indistinguishability complicates the picture. But, at least these complications are of a different nature from the ones addressed in this report. They are of a group-theoretic nature, not of a geometric nature.

References


[M2] The points on a Shimura variety modulo a prime of good reduction, in
The zeta functions of Picard modular surfaces, Univ. Montréal, Montreal, QC, 1992,

Sympos. Pure Math., vol. 55, American Mathematical Society, Providence, RI, 1994,
p. 401–459.


[O] F. Oort – Newton polygon strata in the moduli space of abelian varieties, in Moduli
of abelian varieties (Texel Island, 1999), Progress in Math., vol. 195, Birkhäuser,


[PR1] G. Pappas & M. Rapoport – Local models in the ramified case I. The EL-case, J.

[PR2] , Local models in the ramified case II. Splitting models, math.AG/0205021,
2002.

[Pf] M. Pfau – The conjecture of Langlands and Rapoport for certain Shimura varieties

[R1] M. Rapoport – On the local zeta function of quaternionic Shimura varieties with

[R2] , On the bad reduction of Shimura varieties, in Automorphic forms, Shimura

[R3] , A positivity property of the Satake isomorphism, Manuscripta Math. 101

[RR] M. Rapoport & M. Richartz – On the classification and specialization of F-

Monodromiefiltration und verschwindende Zyklen in ungleicher Charakteristik, Invent.


[Re3] , Reduction of Shimura varieties at parahoric levels, Manuscripta Math. 107

[Ri] M. Richartz – Klassifikation von selbstdualen Dieudonné-gittern in einem dreidi-

[S] H. Stamm – On the reduction of the Hilbert-Blumenthal-moduli scheme with \( \Gamma_0(p) \)-

American Mathematical Society, Providence, RI, 1979, p. 29–69.

[W] T. Wedhorn – Ordinariness in good reductions of Shimura varieties of PEL-type,
Notes added June 2003

We give here some complements and mention some recent developments. A list of supplementary references can be found at the end.

(1) The Iwahori subgroup of a torus, compare Remarks 2.2, (iii). At the request of the referee we give more details. Let $L$ be a complete discretely valued field with algebraically closed residue field. Let $T$ be a torus over $L$. Then $T$ has a lift Néron model over Spec $O_L$ (cf. [BLR] = S. Bosch, W. Lütkebohmert, M. Raynaud, Néron Models, Springer Verlag; 10.1, Prop. 6). Let $T^I$ be the maximal subgroup scheme of finite type over Spec $O_L$ and let $T^o$ be its identity component.

**Proposition**

(a) $T^I(O_L)$ is the unique maximal bounded subgroup $\tilde{K}$ of $T(L)$ and $K := T^o(O_L)$ is a subgroup of finite index in $\tilde{K}$.
(b) $K$ is the Iwahori subgroup of $T(L)$.
(c) For the group of connected components of the special fiber $\mathcal{T}$ of $T$, we have

$$\pi_0(\mathcal{T}) = X_*(T)_I.$$ 

**Proof**

(a) By loc. cit., $T^I(O_L)$ is a bounded subgroup, and hence is contained in $\tilde{K}$. On the other hand, $\tilde{K}$ has a Néron model in the rigid-analytic sense which is contained in $T$ (cf. S. Bosch & K. Schlöter, Néron models in the setting of formal and rigid geometry, Math. Ann. 301 (1995), p. 339–362), and therefore contained in $T^I$ by maximality. Hence $\tilde{K} \subset T^I(O_L)$. This proves the first claim, and the second is immediate.

(b) If $T$ is induced, then $X_*(T)_I = \text{Hom}(X^*(T)^I, \mathbb{Z})$, and hence

$$\text{Ker} \tilde{\kappa}_T = \tilde{K} = \{ x \in T(L); |\chi(x)| = 1, \forall \chi \in X^*(T)^I \}.$$ 

On the other hand, $T$ is a product of tori of the form $R^I_{F'/F}(G_m)$ for a finite extension $L'$ of $L$ and then $T^I$ is a product of group schemes of the form $R_{O^I_{F'/F}}(G_m)$ and hence is connected, whence $T^I(O_L) = T^o(O_L) = K$.

In the general case we choose induced tori $R, S$ over $L$ and an exact sequence of $I$-modules,

$$X_*(S) \to X_*(R) \to X_*(T) \to 0.$$ 

We obtain a commutative diagram with surjective vertical arrows,

$$\begin{array}{ccc}
S(L) & \to & R(L) \to T(L) \to 0 \\
\downarrow \tilde{\kappa}_S & & \downarrow \tilde{\kappa}_R \downarrow \tilde{\kappa}_T \\
X_*(S)_I & \to & X_*(R)_I \to X_*(T)_I \to 0.
\end{array}$$

Hence the map $\text{Ker} \tilde{\kappa}_T \to \text{Ker} \tilde{\kappa}_R$ is surjective. On the other hand, the map on lift Néron models $R(O_L) \to T(O_L)$ is surjective and this implies that $R^o(O_L) \to T^o(O_L)$
is also surjective, cf. [BLR], 9.6, Lemma 2. From $K_R = \text{Ker } \tilde{\kappa}_R$ we conclude $K_T = \text{Ker } \tilde{\kappa}_T$.

(c) This now follows easily from the surjectivity of the Kottwitz map $\tilde{\kappa}_T$.

If we replace $L$ by an arbitrary complete discretely valued field $F$ with perfect residue field $k$, then (a) and (b) are still true (with $L$ replaced by $F$). In (c) the LHS has to be interpreted as the group of connected components of $\mathcal{F} \times_{\text{Spec } k} \text{Spec } \overline{k}$.

(2) The equality $X_*(T) \cap \text{Adm}(\mu) = X_*(T) \cap \text{Perm}(\mu)$ and the closure relations in affine Grassmannians, compare Prop. 3.5. As T. Haines pointed out, Lemma 3.8 can be used to give a simple proof of the closure relations in the affine Grassmannian (for the case of $SL_n$, compare A. Beauville, Y. Laszlo: Conformal blocks and generalized theta functions, Comm. Math. Phys. 164 (1994), p. 385–419).

Let $LG$ denote the loop group associated to the split group $G$ over a field $k$. We denote by $I \subset K$ the loop group versions of the Iwahori and special maximal parahoric subgroup. We obtain the natural morphism between the affine flag variety and the affine Grassmannian,

$$\pi : LG/I \longrightarrow LG/K.$$ 

We choose the convention of viewing elements of the extended affine Weyl group $\tilde{W}$ as elements in $LG$ by sending $\lambda \in X_*(T)$ to $t^{-\lambda} := \lambda(t^{-1})$, and $w \in W_0$ to any lift in the normalizer $N_G(T) \subset G \subset LG$. With this convention, we have the BN-pair axioms

$$IwIsI = IwsI, \quad \text{if } w < ws$$

$$IwIsI = IwI \cup IwsI, \quad \text{if } ws < w,$$

where the Bruhat order $\leq$ is defined using the simple reflections $s$ through the walls of the alcove fixed by $I$. (It is a subtle point, but it is important to use the element $t^{-1}$ rather than $t$ for embedding $X_*(T) \hookrightarrow LG$ here.)

Then we have the usual Demazure resolution

$$T \pi I/I \times_t I s_1 I/I \times_t \cdots \times_t I s_r I/I \longrightarrow TwI/I,$$

where $w = \tau s_1 \cdots s_r$ is any reduced expression. This being a proper morphism, and an isomorphism over the inverse image of $IwI/I$, is all we need to conclude that the closure relations are given by the Bruhat order,

$$TwI/I = \bigsqcup_{w' \leq w} Iw'I/I.$$

Now we want to use this information to deduce the analogous fact for affine Grassmannians:

$$Kt_{\mu}K/K = \bigsqcup_{\lambda \leq \mu} Kt_{\lambda}K/K.$$ 

(Note that here $t_{\lambda} = \lambda(t^{-1})$ as above.)

Now Lemma 3.8 states that for dominant coweights $\lambda, \mu$, we have

$$\lambda \leq \mu \iff w_0 t_{\lambda} \leq w_0 t_{\mu}.$$
This follows since \( w_0t_\lambda \leq w_0t_\mu \Leftrightarrow t_\lambda \leq t_\mu \). Now it follows easily that the set of elements \( w' \leq w_0t_\mu \) is the set of all elements \( w_1t_\lambda w_2 \), where \( w_1,w_2 \in W_0 \), and \( \lambda \) is dominant with \( \lambda \leq \mu \). So we have from the closure relations in the affine flag variety

\[
\overline{Iw_0t_\mu I/I} = \bigcup_{w' \leq w_0t_\mu} \overline{Iw'I/I} = \bigcup_{\lambda \leq \mu} ( \bigcup_{w_1,w_2 \in W_0} \overline{Iw_1t_\lambda w_2I/I} ).
\]

Applying \( \pi \) to both sides, we deduce the analogous fact for affine Grassmannians

\[
\overline{Kt_\mu K/K} = \bigcup_{\lambda \leq \mu} \overline{Kt_\lambda K/K},
\]

by using the following lemma.

**Lemma**

1. \( \pi(\bigcup_{w_1,w_2 \in W_0} \overline{Iw_1t_\lambda w_2I/I}) = \overline{Kt_\lambda K/K} \).

2. \( \pi(\overline{Iw_0t_\mu I/I}) = \overline{Kt_\mu K/K} \).

**Proof.** — The first statement is easy. As for the second statement, the continuity of \( \pi \) gives

\[
\pi(\overline{Iw_0t_\mu I/I}) \subset \pi(\overline{Iw_0t_\mu I/I}) \subset \overline{Iw_0t_\mu K/K} \subset \overline{Kt_\mu K/K}.
\]

On the other hand, since \( \pi \) is proper, \( \pi(\overline{Iw_0t_\mu I/I}) \) is a closed set containing each \( \pi(\overline{Iw_\mu I/I}) = \overline{Iw_\mu K/K} \) (for \( w \in W_0 \)), hence also the closure of their union \( \overline{Kt_\mu K/K} \).

\( \square \)

(3) The surjectivity of \( \text{Perm}(\mu) \to \text{Perm}_K(\mu) \) and the equality \( \text{Adm}_K(\mu) = \text{Perm}_K(\mu) \), compare Prop. 3.10. This has now been proved by U. Görtz [G3] for \( G = \text{GL}_n \) and arbitrary \( \mu \) (loc. cit., Cor. 7.6) and for \( G = \text{GSp}_{2n} \) and \( \mu \) a sum of minuscule coweights (loc. cit., Cor. 7.10). Also, the referee points out that it might be useful in this context to generalize Deodhar’s Lemma (Lemma 6.3 in [HN1]). This might help to prove in general the inclusion \( \text{Perm}_K^\ast(\mu) \subset \text{Adm}_K(\mu) \), compare [HN1]; here \( \text{Perm}_K^\ast(\mu) \) is the obvious generalization to the parahoric case of the strongly permissible set \( \text{Perm}_K^\ast(\mu) \) of loc. cit. As T. Haines points out, this together with the equality \( \text{Perm}_K^\ast(\mu) = \text{Perm}_K(\mu) \) for \( G = \text{GL}_n \) could be used to show that \( \text{Adm}_K^{\text{GSp}}(\mu) = \text{Perm}_K^{\text{GL}}(\mu) \cap \overline{W}(\text{GSp}) \), compare [HN1].

(4) Structure of the affine Deligne-Lusztig varieties, compare Example 4.3. The question for which pairs \( (w,b) \) the affine Deligne-Lusztig variety \( X_w(b) \) is non-empty has been investigated by D. Reuman in his Chicago PhD thesis [Reu1], for \( G = \text{SL}_3 \) and \( G = \text{Sp}_4 \), and for various \( b \) (mostly \( b = 1 \), but also for others). For the question on the dimension of affine Deligne-Lusztig varieties compare [Reu2]; see also note (7) below.

(5) Group-theoretic version of Mazur’s inequality, compare Prop. 4.4. The proof in [RR] is by reduction to the usual Mazur inequality between the Hodge polygon and the Newton polygon of an F-crystal. Recently R. Kottwitz [K5] has given
a purely group-theoretical proof by reduction to the inequality of Arthur-Harish-Chandra. That paper also contains a group-theoretical version of the Hodge-Newton decomposition of an $F$-crystal.

(6) **Converse to Mazur’s inequality, compare Conjecture 4.6.** In [KR2] there is an abstract criterion (Prop. 4.6) for the non-emptiness of $X_{\mu}(b)$, valid for any split group. In [K5] this is generalized to the case of an unramified group. As to the actual converse to Mazur’s inequality, this has been proved by C. Leigh [Le] in the case of an arbitrary classical split group. For $GL_n$, there is a proof, different from [KR2], by reduction to the converse to Fontaine’s inequality, using Laffaille’s theorem, cf. [FR]. For a general unramified group J.-P. Wintenberger [Wi] deduces the converse, in case $\mu$ is minuscule, from his theorem mentioned in (7) below.

(7) **On the sets $X(\mu, b)_K$, compare Conjecture 5.2 and Remark 5.3.** In the light of recent developments, it seems that the hypothesis that $\mu$ be minuscule is not necessary for part a) of Conjecture 5.2. Indeed, Wintenberger [Wi], Th. 2 proves that if $[b] \in B(G, \mu)$, then $X(b, \mu)_K \neq \emptyset$ for any parahoric subgroup, in case $G$ is quasi-split and $\{\mu\}$ defined over $L$ (in a slightly stronger form, replacing the usual Bruhat order in the definition of $\text{Adm}(\mu)$ by the weak Bruhat order).

Now assume that $G$ is unramified and that $X(b, \mu)_K \neq \emptyset$ for some Iwahori subgroup $K$. Then, as pointed out by the referee, it follows conversely that $[b] \in B(G, \mu)$. Indeed, let $K' = K_a$ be the hyperspecial parahoric subgroup over $F$ fixing a vertex $a$ of the alcove $a$ corresponding to $K$. Let $w \in \text{Adm}(\mu)$ with $X_w(b)_K \neq \emptyset$. Put $\nu = w(a) - a$. Then $X_{\nu}(b)_{K'} \neq \emptyset$. Since $w \in \text{Adm}(\mu) \subset \text{Perm}(\mu)$ (cf. Prop. 3.2), $\nu$ lies in the convex hull of $W_0\mu$. By Prop. 4.4 it follows that $[b] \in B(G, \nu) \subset B(G, \mu)$.

(8) **The dimension formula for $X(\mu, b)_K$, compare Conjecture 5.10.** Note that $\dim X(\mu, b)_K = \dim X_\mu(b)_K$ (recall that $K$ is hyperspecial). The dimension formula is easy to verify for $G = GL_2$. It is compatible with the dimension formulae in Example 4.3 for $X_w(b)$ for $G = GL_2$ and the Iwahori subgroup $K_0$ in the following sense. Denoting by $\pi : G(L)/K_0 \to G(L)/K$ the natural projection, we have for any dominant $\lambda$

$$\pi^{-1}X_\lambda(b)_K = \bigcup_{w \in W_0W_{0}^{s}} X_w(b)_{K_0},$$

and $\dim \pi^{-1}(X_\lambda(b)_K) = \dim X_\lambda(b)_K + \dim \tilde{K}/\tilde{K}_0$, compare [Reu2].

Let $G$ be split, and assume that $\mu^2 = 0$. Then $\tilde{\mu} = \mu$ and $\langle \omega_i, \mu \rangle \in \mathbb{Z}$ for $i = 1, \ldots, \ell$. The conjecture in this case then states

$$\dim X_\mu(b) = \langle \varrho, \mu \rangle.$$

This formula was rediscovered by D. Reuman [Reu2] and proved by him for $G = SL_2, SL_3, Sp_4$. He mentions work in progress with R. Kottwitz which seems to confirm this formula in general (under the assumptions spelled out above).

(9) **On the generalized Newton stratification, compare Questions 7.3 and Theorem 7.4.** The strong stratification property of the generalized Newton stratification ((iii)
of Question 7.3) has also been proved for Shimura varieties associated to certain unitary similitude groups, as well as the codimension formula analogous to Theorem 7.4, compare [O] and [BW]. These questions are also addressed in the survey [R4].

**Additional references**


L\textit{-}MODULES AND THE CONJECTURE OF RAPOPORT AND GORESKY-MACPHERSON

by

Leslie Saper

Abstract. — Consider the middle perversity intersection cohomology groups of various compactifications of a Hermitian locally symmetric space. Rapoport and independently Goresky and MacPherson have conjectured that these groups coincide for the reductive Borel-Serre compactification and the Baily-Borel-Satake compactification. This paper describes the theory of L-modules and how it is used to solve the conjecture. More generally we consider a Satake compactification for which all real boundary components are equal-rank. Details will be given elsewhere \[26\]. As another application of L-modules, we prove a vanishing theorem for the ordinary cohomology of a locally symmetric space. This answers a question raised by Tilouine.


2000 Mathematics Subject Classification. — 11F75, 14G35, 22E40, 22E45, 32S60, 55N33.
Key words and phrases. — Intersection cohomology, Shimura varieties, locally symmetric spaces, compactifications.

This research was supported in part by the National Science Foundation under grants DMS-8957216, DMS-9100383, and DMS-9870162, and by a grant from the Institut des Hautes Études Scientifiques. Early stages of this research were supported by The Duke Endowment, an Alfred P. Sloan Research Fellowship, and the Maximilian-Bickhoff-Stiftung during a visit to the Katholischen Universität Eichstätt. The author wishes to thank all these organizations for their hospitality and support. The manuscript was prepared with the \texttt{AMS-LATeX} macro system; diagrams were generated with the \texttt{XY-pic} package.
1. Introduction

In a letter to Borel in 1986 Rapoport made a conjecture (independently rediscovered by Goresky and MacPherson in 1988) regarding the equality of the intersection cohomology of two compactifications of a locally symmetric variety, the reductive Borel-Serre compactification and the Baily-Borel compactification. In this paper I describe the conjecture, introduce the theory of $L$-modules which was developed to attack the conjecture, and explain the solution of the conjecture. The theory of $L$-modules actually applies to the study of many other types of cohomology. As a simple illustration, I will answer at the end of this paper a question raised during the semester by Tilouine regarding the vanishing of the ordinary cohomology of a locally symmetric variety below the middle degree. Except in this final section, proofs are omitted; the details will appear in [26].

This paper is an expanded version of lectures I gave during the Automorphic Forms Semester (Spring 2000) at the Centre Émile Borel in Paris; I would like to thank the organizers for inviting me and providing a stimulating environment. During this research I benefited from discussions with numerous people whom I would like thank, in particular A. Borel, R. Bryant, M. Goresky, R. Hain, G. Harder, J.-P. Labesse, J. Tilouine, M. Rapoport, J. Rohlfs, J. Schwermer, and N. Wallach.

2. Compactifications

We consider a connected reductive algebraic group $G$ defined over $\mathbb{Q}$ and its associated symmetric space $D = G(\mathbb{R})/KA_G$, where $K$ is a maximal compact subgroup of $G(\mathbb{R})$ and $A_G$ is the identity component of the $\mathbb{R}$-points of a maximal $\mathbb{Q}$-split torus in the center of $G$. Let $\Gamma \subset G(\mathbb{Q})$ be an arithmetic subgroup which for simplicity here we assume to be neat. (Any arithmetic subgroup has a neat subgroup of finite index; the neatness hypothesis ensures that all arithmetic quotients in what follows will be smooth as opposed to $V$-manifolds or orbifolds.) The locally symmetric space $X = \Gamma \backslash D$ is in general not compact and we are interested in three compactifications (see Figure 1), belonging respectively to the topological, differential geometric, and (if $D$ is Hermitian symmetric) complex analytic categories.

Let $\mathcal{P}$ (resp. $\mathcal{P}_1$) denote the partially ordered set of $\Gamma$-conjugacy classes of parabolic (resp. maximal parabolic) $\mathbb{Q}$-subgroups of $G$. For $P \in \mathcal{P}$, let $L_P$ denote the Levi quotient $P/N_P$, where $N_P$ is the unipotent radical of $P$. (When it is convenient we will identify $L_P$ with a subgroup of $P$ via an appropriate lift.) The Borel-Serre compactification [4] has strata $Y_P = \Gamma_P \backslash P(\mathbb{R})/K_PA_P$ indexed by $P \in \mathcal{P}$ (for $P = G$ we simply have $Y_G = X$). Here $\Gamma_P = \Gamma \cap P$, $K_P = K \cap P$, and $A_P$ is the identity component of the $\mathbb{R}$-points of a maximal $\mathbb{Q}$-split torus in the center of $L_P$. The Borel-Serre compactification $\overline{X}$ is a manifold with corners, homotopically equivalent with $X$ itself.
The arithmetic subgroup $\Gamma$ induces arithmetic subgroups $\Gamma_{N_P} = \Gamma \cap N_P$ in $N_P$ and $\Gamma_{L_P} = \Gamma_P / N_P$ in $L_P$. Let $D_P = L_P(\mathbb{R}) / K_P A_P$ be the symmetric space associated to $L_P$ and let $X_P = \Gamma_{L_P} \backslash D_P$ be its arithmetic quotient. Each stratum of $\overline{X}$ admits a fibration $Y_P \to X_P$ with fibers being compact nilmanifolds $\Gamma_{N_P} \backslash N_P(\mathbb{R})$. The union $\hat{X} = \bigsqcup P \in \mathcal{P} X_P$ (with the quotient topology from the natural map $\overline{X} \to \hat{X}$) is the reductive Borel-Serre compactification; it was introduced by Zucker [34]. The reductive Borel-Serre compactification is natural from a differential geometric standpoint since the locally symmetric metric on $X$ degenerates precisely along these nilmanifolds near the boundary of $\overline{X}$.

Finally assume now that $D$ is Hermitian symmetric. Then each $D_P$ factors into a product $D_{P,\ell} \times D_{P,h}$, where $D_{P,h}$ is again Hermitian symmetric (see Figure 2). This induces a factorization (modulo a finite quotient) $X_P = X_{P,\ell} \times X_{P,h}$ of the arithmetic quotients and we consider the projection $X_P \to X_{P,h}$ onto the second factor. Now among the different $P \in \mathcal{P}$ that yield the same $X_{P,h}$, let $P^\dagger \in \mathcal{P}_1$ be the maximal one and set $F_{P^\dagger} = X_{P,h}$. Thus each stratum of $\hat{X}$ has a projection $X_P \to F_{P^\dagger}$. The union $X^* = \bigsqcup R \in \mathcal{P}_1 F_R$ (with the quotient topology from the map $\hat{X} \to X^*$) is the Baily-Borel-Satake compactification $X^*$. Topologically $X^*$ was constructed by Satake [29], [30] (though the description we have given is due to Zucker [35]); if $\Gamma$ is contained in the group of biholomorphisms of $D$, the compactification $X^*$ was given the structure of a normal projective algebraic variety by Baily and Borel [2].

The simplest example where all three compactifications are distinct is the Hilbert modular surface case. Here $G = R_{k/\mathbb{Q}} \operatorname{SL}(2)$ where $k$ is a real quadratic extension. There is only one proper parabolic $\mathbb{Q}$-subgroup $P$ up to $G(\mathbb{Q})$-conjugacy; $Y_P$ is a torus bundle over $X_P = S^1$ and $F_P$ is a point.
3. The Conjecture

Assume that $D$ is Hermitian symmetric. Let $E \in \mathfrak{Mod}(G)$, the category of finite dimensional regular representations of $G$ and let $E$ denote the corresponding local system on $X$. Let $\mathcal{IC}(\hat{X}; E)$ and $\mathcal{IC}(X^*; E)$ denote middle perversity intersection cohomology sheaves on $\hat{X}$ and $X^*$ respectively [10].

For example, $\mathcal{IC}(\hat{X}; E) = \tau^{\leq p(\text{codim } X_P)} j_{\ast} E$ if $\hat{X}$ has only one singular stratum $X_P$; here $j_{\ast}$ denotes the derived direct image functor of the inclusion $j : \hat{X} \setminus X_P \hookrightarrow \hat{X}$, codim $X_P$ denotes the topological codimension, $p(k)$ is one of the middle perversities $\lfloor (k-1)/2 \rfloor$ or $\lfloor (k-2)/2 \rfloor$, and $\tau^{\leq p(k)}$ truncates link cohomology in degrees $> p(k)$. In general the pattern of pushforward/truncate is repeated over each singular stratum. Note that since $\hat{X}$ may have odd codimension strata, $\mathcal{IC}(\hat{X}; E)$ depends on the choice of the middle perversity $p$; on the other hand, since $X^*$ only has even codimension strata, $\mathcal{IC}(X^*; E)$ is independent of $p$.

**Main Theorem (Rapoport’s Conjecture).** — Let $X$ be an arithmetic quotient of a Hermitian symmetric space. Then $\pi_{\ast} \mathcal{IC}(\hat{X}; E) \cong \mathcal{IC}(X^*; E)$. (That is, they are isomorphic in the derived category.)

Following discussions with Kottwitz, Rapoport conjectured the theorem in a letter to Borel [22] and later provided motivation for it in an unpublished note [23]. Previously Zucker had noticed that the conjecture held for $G = \text{Sp}(4)$, $E = \mathbb{C}$. The conjecture was later rediscovered by Goresky and MacPherson and described in an unpublished preprint [11] in which they also announced the theorem for $G = \text{Sp}(4)$, $\text{Sp}(6)$, and (for $E = \mathbb{C}$) $\text{Sp}(8)$. The first published appearance of the conjecture was in a revised version of Rapoport’s note [24] and included an appendix by Saper and Stern giving a proof of the theorem when $\mathbb{Q}$-rank $G = 1$.

---

(1) By a “sheaf” we will always mean a complex of sheaves representing an element of the derived category. A derived functor will be denoted by the same symbol as the original functor, thus we will write $\pi_{\ast}$ instead of $R\pi_{\ast}$. 

ASTÉRISQUE 298
To see one reason why the conjecture might be useful in the theory of automorphic forms, note that the right hand side $\mathcal{I}C(X^*; E)$ is isomorphic to the $L^2$-cohomology sheaf $\mathcal{L}_{(2)}(X^*; E)$ by (the proof of) Zucker’s conjecture [17], [28]. The trace of a Hecke operator on $L^2$-cohomology could then be studied topologically via the Lefschetz fixed point formula for $\mathcal{I}C(X^*; E)$. However, the singularities of $\hat{X}$ are simpler than those of $X^*$ so a Lefschetz fixed point formula for $\mathcal{I}C(\hat{X}; E)$ should be easier to calculate. The conjecture says that this should give the same result. Also note that a Lefschetz fixed point formula for $\mathcal{I}C(\hat{X}; E)$ involves a sum over $\mathcal{P}$, while a Lefschetz fixed point formula for $\mathcal{I}C(X^*; E)$ involves a sum over $\mathcal{P}_1$. Thus it is more likely that the former can be directly related to the Arthur-Selberg trace formula for a Hecke operator on $L^2$-cohomology [1].

This program has been pursued by Goresky and MacPherson, but instead of $\mathcal{I}C(\hat{X}; E)$ they use the “middle weighted cohomology” WC($\hat{X}; E$) in which cohomology classes in the link are truncated according to their weight as opposed to their degree. Thus weighted cohomology is an algebraic analogue of $L^2$-cohomology. Goresky and MacPherson prove (in joint work with Harder [8]) the analogue of the above theorem, $\pi_* WC(\hat{X}; E) \cong \mathcal{I}C(X^*; E)$, calculate the Lefschetz fixed point formula [12], and (in joint work with Kottwitz) show that it agrees with Arthur’s trace formula for $L^2$-cohomology [9].

Nonetheless the original conjecture remains interesting for a number of reasons. First of all, intersection cohomology is a true topological invariant and the local cohomology of $\mathcal{I}C(\hat{X}; E)$ behaves better than that of WC($\hat{X}; E$) when $E$ varies. Secondly, the local property (“micro-purity”) one needs to prove is much deeper for $\mathcal{I}C(\hat{X}; E)$ than for WC($\hat{X}; E$) and should have applications elsewhere. And finally the method used to attack the conjecture, the theory of $L$-modules, has application to other cohomology, in particular, weighted cohomology, $L^2$-cohomology, and ordinary cohomology.

In §§5–10 we will indicate how the Main Theorem follows from three theorems in the theory of $L$-modules.

4. A generalization

This section is optional; we will indicate a more general context in which the Main Theorem holds. First we sketchily recall the general theory of Satake compactifications [29], [30], [35], [6]. By embedding $D$ into a real projective space via a finite-dimensional representation $\sigma$ of $G$ and then taking the closure, Satake constructed a finite family of Satake compactifications $sD^*$ of $D$. Each of these is equipped with an action of $G(\mathbb{R})$ and is formed by adjoining to $D$ certain real boundary components. Let $D^*$ denote the union of $D$ together with those real boundary components whose normalizer is defined over $\mathbb{Q}$; call these the rational boundary components. In the
**geometrically rational** case (a condition satisfied for example if $\sigma$ is $\mathbb{Q}$-rational\(^{(2)}\)) one may equip $D^*$ with a suitable topology so that $X^* = \Gamma \backslash D^*$ is a Hausdorff compactification of $X$; this is also called a Satake compactification. For $D$ Hermitian symmetric, one of the Satake compactifications is (topologically equivalent to) the closure of the realization of $D$ as a bounded symmetric domain and it is geometrically rational; the corresponding compactification of $X$ is the Baily-Borel-Satake compactification.

Let $0G = \bigcap_{\chi \in X_0(G)} \ker \chi^2$ so that $G(\mathbb{R}) = 0G(\mathbb{R}) A_G$ [4]. Suppose that $\text{rank } 0G = \text{rank } K$, that is, $0G(\mathbb{R})$ has discrete series representations. This is equivalent to the assumption that the maximal $\mathbb{R}$-split torus in the center of $G$ is also $\mathbb{Q}$-split and that the real points of $G^{\text{der}}$ (the semisimple derived group) has discrete series representations. (We may also substitute here the adjoint group $G^{\text{ad}}$ for $G^{\text{der}}$.) We say in this case that $D$ is an equal-rank symmetric space. A Satake compactification $\mathbb{R}^* D^*$ of $D$ will be called a real equal-rank Satake compactification if all the real boundary components of $\mathbb{R}^* D^*$ are also equal-rank symmetric spaces. The possible $D$ that admit real equal-rank Satake compactifications are listed in [36]; they include the Hermitian symmetric cases but there are other infinite families as well. If such a $\mathbb{R}^* D^*$ is geometrically rational\(^{(3)}\) then the corresponding compactification $X^*$ of $X$ is also called a real equal-rank Satake compactification; note that we impose the equal-rank condition on all real boundary components even though only the rational boundary components contribute to $X^*$.

The generalization we alluded to above is that the Main Theorem holds for real equal-rank Satake compactifications. (Note that Borel conjectured that the analogue of the Zucker conjecture should remain true for such $X^*$ and Saper and Stern (unpublished) observed that their proof could be adapted to this case.)

**5. $L$-modules**

Now again let $G$ be any connected reductive group over $\mathbb{Q}$ (with no Hermitian hypothesis). The “sheaf” $\mathcal{IC}(\hat{X}; \mathbb{E})$ is actually an object of $D_X(\hat{X})$, the derived category of complexes of sheaves $S$ on $\hat{X}$ that are constructible. Here the constructibility of $S$ means that if for all $P \in \mathcal{P}$ we let $i_P : X_P \hookrightarrow \hat{X}$ denote the inclusion, then the local cohomology sheaf $H(i_P^* S) = H(S|_{X_P})$ is locally constant, or equivalently the cohomology sheaf $E_P = H(i_P^* S)$ is locally constant on $X_P$. Thus by the correspondence between local systems and representations of the fundamental group one obtains a family of objects $E_P \in \text{Gr}(\Gamma_{L_P})$, the category of graded $\Gamma_{L_P}$-modules, one for each $P \in \mathcal{P}$.

\(^{(2)}\)Borel points out that in his 1962 Bruxelles conference paper “Ensembles fondamentaux pour les groupes arithmétiques” he proves geometric rationality only when $\sigma$ is strongly $\mathbb{Q}$-rational. In [27] we prove geometric rationality for the general $\mathbb{Q}$-rational case.

\(^{(3)}\)We show in [27] that this always holds except for certain explicitly described situations in $\mathbb{Q}$-rank 1 and 2 involving restriction of scalars.
Instead of $\mathcal{S}$ we wish to work with a combinatorial analogue in which $\text{Gr}(\Gamma_{L_{\mathcal{P}}})$ is replaced by $\text{Gr}(L_{\mathcal{P}})$, the category of graded regular $L_{\mathcal{P}}$-modules. This analogue is what we will call an $\mathcal{L}$-module on $\tilde{X}$. We will describe just what an $\mathcal{L}$-module is more precisely later, but first let us give some of the properties of the categories $\text{Mod}(\mathcal{L}_W)$ of $\mathcal{L}$-modules on $W$, where $W$ is any locally closed union of strata of $\tilde{X}$:

(i) if $W = X_{\mathcal{P}}$, then $\text{Mod}(\mathcal{L}_{X_{\mathcal{P}}}) = \text{C}(L_{\mathcal{P}})$, the category of complexes of regular $L_{\mathcal{P}}$-modules;

(ii) for any inclusion $j : W \hookrightarrow W'$, there exist functors $j^*, j^! : \text{Mod}(\mathcal{L}_{W'}) \rightarrow \text{Mod}(\mathcal{L}_W)$ and $j_* : \text{Mod}(\mathcal{L}_W) \rightarrow \text{Mod}(\mathcal{L}_{W'})$, as well as a degree truncation functor $\tau_{\leq \mathcal{P}} : \text{Mod}(\mathcal{L}_W) \rightarrow \text{Mod}(\mathcal{L}_W)$;

(iii) there is a realization functor $\mathcal{S}_W : \text{Mod}(\mathcal{L}_W) \rightarrow D_X(W)$ which commutes with the functors in (ii) and for which the following diagram commutes:

$$
\begin{array}{ccc}
\text{Mod}(\mathcal{L}_{X_{\mathcal{P}}}) & \longrightarrow & D_X(X_{\mathcal{P}}) \\
\uparrow \quad \downarrow H(\cdot) & & \downarrow H(\cdot) \\
\text{Gr}(L_{\mathcal{P}}) & \xrightarrow{\text{Res}} & \text{Gr}(\Gamma_{L_{\mathcal{P}}})
\end{array}
$$

Note that one advantage of $\mathcal{L}$-modules over sheaves is that the left hand vertical arrows in (iii) are equivalences of categories, unlike those on the right; this is because $\text{Mod}(L_{\mathcal{P}})$ is a semisimple category.

So roughly speaking an $\mathcal{L}$-module is like a sheaf $\mathcal{S}$ with the “extra structure” that $E_{\mathcal{P}} = H(i_{\mathcal{P}}^*\mathcal{S})$ is associated to a regular $L_{\mathcal{P}}$-module, as opposed to merely a $\Gamma_{L_{\mathcal{P}}}$-module. Condition (ii) implies that the usual operations on sheaves preserve this “extra structure”. The following example shows this is reasonable. Let $E$ be a local system on $X$ associated to a regular representation $E$ of $G$. The smooth part of the link bundle of a real codimension $k$ stratum $X_{\mathcal{P}} \subset \tilde{X}$ is the flat bundle with fiber $|\Delta_{\mathcal{P}}|^* \times \Gamma_{N_{\mathcal{P}}} \backslash N_{\mathcal{P}}(\mathbb{R})$, where $|\Delta_{\mathcal{P}}|^*$ is an open $(k - 1)$-simplex and $\Gamma_{L_{\mathcal{P}}}$ acts via conjugation on the second factor [8, §8]. Thus $H(i_{\mathcal{P}}^*g_\mathcal{S}, E) \cong H(\Gamma_{N_{\mathcal{P}}} \backslash N_{\mathcal{P}}(\mathbb{R}); E)$, the local system associated to the $\Gamma_{L_{\mathcal{P}}}$-module $H(\Gamma_{N_{\mathcal{P}}} \backslash N_{\mathcal{P}}(\mathbb{R}); E)$. However by van Est’s theorem [7], $H(\Gamma_{N_{\mathcal{P}}} \backslash N_{\mathcal{P}}(\mathbb{R}); E)$ is isomorphic to the restriction of the regular $L_{\mathcal{P}}$-module $H(\mathfrak{n}_\mathcal{P}; E)$, where $\mathfrak{n}_{\mathcal{P}}$ is the Lie algebra of $N_{\mathcal{P}}(\mathbb{R})$.

In fact this also suggests how to precisely define $\mathcal{L}$-modules. Let $\mathcal{P}(W) \subseteq \mathcal{P}$ correspond to the strata of $W$. For $P \leq Q$ let $\mathfrak{n}_P^Q$ be the Lie algebra of $N_P(\mathbb{R})/N_Q(\mathbb{R})$. An $\mathcal{L}$-module $\mathcal{M} \in \text{Mod}(\mathcal{L}_W)$ is a family $(E_{\cdot} \cdot f_{\cdot})$ consisting of objects $E_P \in \text{Gr}(L_{\mathcal{P}})$ for every $P \in \mathcal{P}(W)$ and degree 1 morphisms $f_{PQ} : H(\mathfrak{n}_P^Q; E_Q) \xrightarrow{[1]} E_P$ for every $P \leq Q \in \mathcal{P}(W)$ such that

$$
\sum_{P \leq Q \leq R} f_{PQ} \circ H(\mathfrak{n}_P^Q; f_{QR}) = 0
$$
for all $P \leq R \in \mathcal{P}(W)$. The functors $i^*_P$ and $i^!_P$ are given by

$$
i^*_PM = (E_P, f_{PP}),
$$

$$
i^!_PM = \left( \bigoplus_{P \leq R} H(n_R^P; E_R), \sum_{P \leq R \leq S} H(n_R^P; f_{RS}) \right).
$$

We define the global cohomology $H(\hat{X}; \mathcal{M})$ of an $\mathcal{L}$-module $\mathcal{M}$ to be the hypercohomology of its realization, $H(\hat{X}; S_{\hat{X}}(\mathcal{M}))$. In general we will often write simply $\mathcal{M}$ for both the $\mathcal{L}$-module and its realization $S_{\hat{X}}(\mathcal{M})$; it should be clear what is meant from the context.

6. Examples of $\mathcal{L}$-modules

(i) Let $E \in \text{Mod}(G)$. Then the $\mathcal{L}$-module $i_{G!*}E$ defined by $E_G = E$ and $E_P = 0$ for $P \neq G$ corresponds via $S_{\hat{X}}$ to $i_{G!*}E$ and its cohomology is the ordinary cohomology $H(X; E) = H(\Gamma; E)$.

(ii) It follows immediately from the properties of $\mathcal{L}$-modules in the previous section that given $E \in \text{Mod}(G)$ there exists an $\mathcal{L}$-module $\mathcal{IC}(\hat{X}; E)$ which maps under $S_{\hat{X}}$ to the intersection cohomology sheaf $\mathcal{IC}(\hat{X}; E)$. For example, if $\mathcal{P} = \{G, P\}$ (that is, $\hat{X}$ has only one singular stratum) and $p = p(\text{codim } X_P)$, then

$$\mathcal{IC}(\hat{X}; E) = \left( E_G = E, E_P = (\tau^{>p}H(n_P^P; E))[-1], f_{PG} : H(n_P^P; E) \to \tau^{>p}H(n_P^P; E) \right)$$

where $\tau^{>p}H(n_P^P; E) = \bigoplus_{i > p} H^i(n_P^P; E)[-i]$ and $f_{PG}$ is the projection. Note that the truncation $\tau^{\leq p}$ of local cohomology at $X_P$ has been implemented externally via a mapping cone; this is valid in view of the quasi-isomorphism $\tau^{\leq p}C \sim \text{Cone}(C \to \tau^{>p}C)[-1]$ for any complex $C$.

(iii) The weighted cohomology sheaf and the $L^2$-cohomology sheaf may also be lifted to $\mathcal{L}$-modules $WC(\hat{X}; E)$ and $L(2)(\hat{X}; E)$; for the latter we must replace $\text{Mod}(L_P)$ by the category of locally regular $L_P$-modules to handle the potentially infinite dimensional local cohomologies.

7. Micro-support of $\mathcal{L}$-modules

The support of a sheaf $\mathcal{S}$ is the set of points $x$ such that $H(\mathcal{S})_x \neq 0$. As is well-known the global cohomology of $\mathcal{S}$ vanishes if the support is empty (that is, the sheaf is quasi-isomorphic to $0$). For an $\mathcal{L}$-module $\mathcal{M}$ we will state in the next section a more subtle vanishing result based on the micro-support of $\mathcal{M}$ which we now define; this is a rough analogue of the corresponding notion for sheaves [13].

Let $P \in \mathcal{P}$ and let $\mathfrak{Irr}(L_P)$ denote the set of irreducible regular $L_P$-modules. For $V \in \mathfrak{Irr}(L_P)$ let $\xi_V$ be the character by which $A_P$ acts on $V$. Let $\Delta_P$ be the simple
roots of the adjoint action of $A_P$ on $\mathfrak{p}_P$; the parabolic $\mathbb{Q}$-subgroups $Q \supseteq P$ are indexed by subsets $\Delta^Q_P$ of $\Delta_P$. Define $P \leq Q_V \leq Q'_V \in \mathcal{P}$ by

$$\Delta^Q_P = \{ \alpha \in \Delta_P \mid (\xi_V + \rho, \alpha) < 0 \},$$

$$\Delta^{Q'}_P = \{ \alpha \in \Delta_P \mid (\xi_V + \rho, \alpha) \leq 0 \},$$

where $\rho$ denotes one-half the sum of the positive roots of $G$ and the inner product is induced by the Killing form of $G$. Let $M_P = 0_{L_P}$ so that $L_P(\mathbb{R}) = M_P(\mathbb{R})A_P$. Let $V|_{M_P}$ denote the restriction of the representation $V$ to $M_P$.

The micro-support $SS(M)$ of $\mathcal{M}$ is the subset of $\prod_{P \in \mathcal{P}} \mathfrak{rr}(L_P)$ consisting of those $V \in \mathfrak{rr}(L_P)$ satisfying

(i) $(V|_{M_P})^* \cong V|_{M_P}$, and

(ii) there exists $Q_V \leq Q \leq Q'_V$ such that

$$(7.1) \quad H(i^*_P i^{1*}_Q(M))_V \neq 0.$$  

Here $i_Q : \hat{X}_Q \hookrightarrow \hat{X}$ is the inclusion of the closure of the stratum $X_Q$ and the subscript $V$ indicates the $V$-isotypical component. A simple example of the computation of micro-support will be given in §11.

Condition (i) is equivalent to the existence of a nondegenerate sesquilinear form on $V$ which is invariant under the action of $M_P$.

As for condition (ii), let $j_Q : \hat{X} \setminus \hat{X}_Q \hookrightarrow \hat{X}$ be the open inclusion. Note that we have a short exact sequence

$$0 \longrightarrow i^*_P i^{1*}_Q M \longrightarrow i^*_P M \longrightarrow i^*_P j_Q j^{1*}_Q M \longrightarrow 0$$

and a corresponding long exact sequence. Topologically, this is the long exact sequence of the pair $(\hat{U}, \hat{U} \setminus (\hat{U} \cap \hat{X}_Q))$ where $\hat{U}$ is a small neighborhood of a point of $X_P$. Thus condition (ii) means that

$$(7.1) \quad H(i^*_P i^{1*}_Q(M))_V \neq 0.$$  

is not an isomorphism for some degree and for some $Q$ between $Q_V$ and $Q'_V$.

It is convenient to define the essential micro-support $SS_{\text{ess}}(\mathcal{M})$ of $\mathcal{M}$ to be the subset consisting of those $V \in SS(\mathcal{M})$ for which

Type$_V(\mathcal{M}) = \text{Image}(H(i^*_P i^{1*}_Q(M))_V \longrightarrow H(i^*_P i^{1*}_Q'(M))_V)$

is nonzero. The essential micro-support of $\mathcal{M}$ determines the micro-support (though not the actual parabolics $Q$ that arise in condition (ii)). In fact the relation between $SS(\mathcal{M})$ and $SS_{\text{ess}}(\mathcal{M})$ is analogous to the relation between the strata of a nonreduced variety (possibly with embedded components) and the smooth open strata of the
irreducible components: there exists a partial order $\preceq$ on $\coprod_{P \in \mathcal{P}} \mathfrak{H}(L_P)$ such that if $V \in \mathcal{S}(\mathcal{M})$ then there exists $\tilde{V} \in \mathcal{S}_{\text{ess}}(\mathcal{M})$ with $V \preceq \tilde{V}$, and if $\tilde{V} \in \mathcal{S}_{\text{ess}}(\mathcal{M})$ and $V \not\preceq \tilde{V}$ then $V \in \mathcal{S}(\mathcal{M})$.

8. A vanishing theorem for $\mathcal{L}$-modules

The justification for the definition of $\mathcal{S}(\mathcal{M})$ is that it is an ingredient for a vanishing theorem for $H(\tilde{X}; \mathcal{M})$. To state the theorem we need some more notation.

Let $V \in \mathfrak{H}(L_P)$ have highest weight $\mu \in \mathfrak{h}_c^*$ where $\mathfrak{h}$ is a fundamental (maximally compact) Cartan subalgebra for the Lie algebra $\mathfrak{l}_P$ of $L_P(\mathbb{R})$ equipped with a compatible ordering. Assume $(V|_{M_P})^* \cong \tilde{V}|_{M_P}$ and define

$$L_P(\mu) = \text{the centralizer of } \mu \in \mathfrak{h}_C^* \subset \mathfrak{t}_P,$$

$$= \text{the reductive subgroup of } L_P \text{ with roots } \{ \gamma \in \mathfrak{P}(\mathfrak{t}_P, \mathfrak{h}_C) \mid (\gamma, \mu) = 0 \},$$

$$D_P(\mu) = \text{the associated symmetric space } L_P(\mu)/(K_P \cap L_P(\mu))A_P.$$ Choose a compatible ordering for which $\dim D_P(\mu)$ is maximized and let $D_P(V) = D_P(\mu)$. Suppose now that $V \in \mathcal{S}_{\text{ess}}(\mathcal{M})$. Let $c(V; \mathcal{M}) \leq d(V; \mathcal{M})$ be the least and greatest degrees in which Type$_V(\mathcal{M})$ is nonzero, and define

$$\bar{c}(V; \mathcal{M}) = \frac{1}{2}(\dim D_P - \dim D_P(V)) + c(V; \mathcal{M}),$$

$$\bar{d}(V; \mathcal{M}) = \frac{1}{2}(\dim D_P + \dim D_P(V)) + d(V; \mathcal{M}).$$

Set

$$c(\mathcal{M}) = \inf_{V \in \mathcal{S}_{\text{ess}}(\mathcal{M})} \bar{c}(V; \mathcal{M}), \quad d(\mathcal{M}) = \sup_{V \in \mathcal{S}_{\text{ess}}(\mathcal{M})} \bar{d}(V; \mathcal{M}).$$

(One can show that the same values are obtained if instead we consider all $V \in \mathcal{S}(\mathcal{M})$ and let $c(V; \mathcal{M}) \leq d(V; \mathcal{M})$ be the least and greatest degrees in which (7.1) is nonzero (for any $Q$).)

**Theorem 1.** — $H^i(\tilde{X}; \mathcal{M}) = 0$ for $i \notin [c(\mathcal{M}), d(\mathcal{M})]$.

Let us comment briefly on the proof which uses combinatorial Hodge-de Rham theory. The sheaf $\mathcal{S}(\mathcal{M})$ has an incarnation as a complex of fine sheaves whose global sections are “combinatorial” differential forms. That is, an element of $\Gamma(\tilde{X}; \mathcal{S}(\tilde{X})(\mathcal{M}))$ is a family $(\omega_P)_{P \in \mathcal{P}}$, where each $\omega_P$ is a special differential form on $X_P$ with coefficients in $\mathbb{E}_P$. (For $P = G$, the special differential forms [8, (13.2)] on $X = X_G$ are those which near each boundary stratum $Y_Q$ of the Borel-Serre compactification $\overline{X}$ are the pullback of an $N_Q(\mathbb{R})$-invariant form on $Y_Q$; they form a resolution of $\mathbb{E}_G$.) The differential is a sum of the usual de Rham exterior derivative (on each $\omega_P$) together with operators based on the $f_{PQ}$.

To do harmonic theory we need a metric; unfortunately the locally symmetric metric on each $X_P$ is not appropriate since it would introduce unwanted $L^2$-growth conditions on the differential forms. Instead the theory of tilings from [25] gives a
natural piecewise analytic diffeomorphism of $\overline{X}$ onto a closed subdomain $\overline{X}_0$ of the interior $X$; the pullback of the locally symmetric metric under this map yields metrics on all $X_P$ which extend to nondegenerate metrics on their boundary strata. Now a spectral analogue of the Mayer-Vietoris sequence as in [28] reduces the problem to a vanishing theorem for combinatorial $L^2$-cohomology near each stratum $X_P$. After unraveling the combinatorics one obtains contributions to the cohomology of the form $H_{(2)}(X_P; V) \otimes \text{Type}_V(\mathcal{M})$ for $V \in \text{SS}_\text{ess}(\mathcal{M})$; by Raghunathan’s vanishing theorem [20], [21], [28] this is zero outside the degree range $[\overline{c}(V; \mathcal{M}), \overline{d}(V; \mathcal{M})]$. (The proof is actually more complicated since there are infinite dimensional contributions from $\text{SS}(\mathcal{M}) \setminus \text{SS}_\text{ess}(\mathcal{M})$ as well.)

9. Micro-purity of intersection cohomology

We will say an $\mathcal{L}$-module $\mathcal{M}$ on $\widehat{X}$ is $V$-micro-pure if $\text{SS}_\text{ess}(\mathcal{M}) = \{V\}$ with $\text{Type}_V(\mathcal{M})$ concentrated in degree 0.

**Theorem 2.** — Assume the irreducible components of the $\mathbb{Q}$-root system of $G$ are of type $A_n$, $B_n$, $C_n$, $BC_n$, or $G_2$. Let $E \in \mathcal{Irr}(G)$ satisfy $(E|_{\mathcal{U}G})^* \cong E|_{\partial G}$. Then $\mathcal{IC}(\widehat{X}; E)$ is $E$-micro-pure.

If $D$ is a Hermitian symmetric space (or an equal-rank symmetric space admitting a real equal-rank Satake compactification as in §4) $G$ will have a $\mathbb{Q}$-root system of the indicated type and thus the theorem applies in the context of Rapoport and Goresky-MacPherson’s conjecture. In fact it is quite possible that this restriction in the theorem may be removed; it is only required at one crucial stage in the proof.

What the theorem is asserting is that $V \notin \text{SS}_\text{ess}(\mathcal{IC}(\widehat{X}; E))$ for $V \in \mathcal{Irr}(L_P)$ with $P \neq G$. When $P$ is a maximal parabolic we can give a brief indication of how this is proven; for definiteness we assume $p$ is the upper middle perversity. In this case

\begin{equation}
H(i_P^! \mathcal{IC}(\widehat{X}; E)) = \begin{cases}
\tau^{\leq p} H(n_P; E) & \text{for } Q = G, \\
(\tau^{> p} H(n_P; E))[1] & \text{for } Q = P,
\end{cases}
\end{equation}

where $p = \lfloor \frac{1}{2} \dim n_P \rfloor$. Let $\lambda$ be the highest weight of $E$. By Kostant’s theorem [15] an irreducible component $V$ of $H(n; E)$ has highest weight $w(\lambda + \rho) - \rho$ where $w \in W_P = \{w \in W \mid w^{-1} \gamma > 0 \text{ for all positive roots } \gamma \text{ of } I_P\}$,

the set of minimal length representatives of the Weyl group quotient $W_{L_P}\backslash W$. Furthermore $V$ occurs in degree $\ell(w)$, the length of $w$, with multiplicity 1. Assume now that $V \in \text{SS}_\text{ess}(\mathcal{IC}(\widehat{X}; E))$. Since the two cases in (9.1) above do not share a common component we must have $Q_V = Q'_V$, that is, $(\xi_V + \rho, \alpha) \neq 0$ for the unique $\alpha \in \Delta_P$.

Furthermore (9.1) also shows that the possibilities $(\xi_V + \rho, \alpha) < 0$ and $(\xi_V + \rho, \alpha) > 0$ correspond respectively to $\ell(w) \leq \frac{1}{2} \dim n_P$ and $\ell(w) > \frac{3}{2} \dim n_P$. However the following lemma from [26] shows that in fact the opposite relation between weight and
Lemma 3. — Let $V \in \mathcal{H}(L_\nu)$ have highest weight $w(\lambda + \rho) - \rho$ where $w \in W_\nu$ and $\lambda \in \mathfrak{h}_C^* \rightarrow \mathfrak{h}_C^*$ is dominant. Assume that $(V|_{\mathfrak{M}_\nu})^* \cong \overline{V|_{\mathfrak{M}_\nu}}$.

(i) If $(\xi_V + \rho, \alpha) \leq 0$ for all $\alpha \in \Delta_\nu$, then $\ell(w) \geq \frac{1}{2}(\dim \mathfrak{m}_\nu + \dim \mathfrak{m}_\nu(V))$.

(ii) If $(\xi_V + \rho, \alpha) \geq 0$ for all $\alpha \in \Delta_\nu$, then $\ell(w) \leq \frac{1}{2}(\dim \mathfrak{m}_\nu - \dim \mathfrak{m}_\nu(V))$.

The only remaining possibility is that $\ell(w) = \frac{1}{2} \dim \mathfrak{m}_\nu$, but since $(\xi_V + \rho, \alpha) \neq 0$ and $(E|_{\mathfrak{g}})^* \cong E_{|\mathfrak{g}}$, this is impossible by an argument based on [3]. By the way, Lemma 3 is basic to the proofs of Theorems 1, 4, and 5 as well and has its origin in a result of Casselman for $\mathbb{R}$-rank one [5].

When $P$ is not a maximal parabolic the situation is far more complicated. The irreducible components of $H(i_P^* \mathcal{I}(X; E))$ are among those of $H(n_P; E)$, but they may occur in various degrees and with multiplicity. Since we do not know a nonrecursive formula for $H(i_P^* \mathcal{I}(X; E))$ we must rely on the inductive definition. However condition (i) in the definition of micro-support is not preserved upon passing to a larger stratum. Specifically, let $P < R$ and suppose $V$ is an irreducible component of $H(n_P; E) = H(n_P^R; E)$, but they may occur in various degrees and with multiplicity. Since we do not know a nonrecursive formula for $H(i_P^* \mathcal{I}(X; E))$ we must rely on the inductive definition. However condition (i) in the definition of micro-support is not preserved upon passing to a larger stratum. Specifically, let $P < R$ and suppose $V$ is an irreducible component of $H(n_P; E) = H(n_P^R; E)$. It must lie within $H(n_P^R; V_R)$ for some irreducible component $V_R$ of $H(n_R; E)$. The difficulty in using induction is that $(V|_{\mathfrak{M}_\nu})^* \cong \overline{V|_{\mathfrak{M}_\nu}}$ does not imply $(V_R|_{\mathfrak{M}_R})^* \cong \overline{V_R|_{\mathfrak{M}_R}}$.

These difficulties do not apply to $\mathcal{W}(X; E)$ and in fact a fairly simple argument shows that Theorem 2 holds for $\mathcal{W}(X; E)$ without any hypothesis on the $\mathbb{Q}$-root system and for either middle weight profile. Indeed since $\mathcal{W}(X; E)$ is defined directly in terms of weight the relationship between weight and degree provided by Lemma 3 is not needed and hence the condition $(V|_{\mathfrak{M}_\nu})^* \cong \overline{V|_{\mathfrak{M}_\nu}}$ plays no role in the proof.

10. Functoriality of micro-support and proof of the Main Theorem

Let $\mathcal{M}$ be an $\mathcal{L}$-module which is $E$-micro-pure (for example, $\mathcal{M} = \mathcal{I}(X; E)$ by Theorem 2) and assume we are in the context of Rapoport and Goresky-MacPherson’s conjecture, that is, $D$ is Hermitian symmetric and $\pi: \hat{X} \rightarrow X^*$ is the projection onto the Baily-Borel-Satake compactification. The desired equality $\pi_* \mathcal{M} = \mathcal{I}(X^*; E)$ is equivalent to certain local vanishing and covanishing conditions on $\pi_* \mathcal{M}$ [10]. To state them, let $i_x : \{x\} \hookrightarrow X^*$ denote the inclusion of a point in a stratum $F_R \subset X^*$. Since every stratum of $X^*$ has even codimension, $p(\text{codim } F_R) = \frac{1}{2} \text{codim } F_R - 1$. The local conditions that characterize intersection cohomology now can be expressed as

\begin{align}
H^i(i_x^* \pi_* \mathcal{M}) &= 0 \quad \text{for } x \in F_R, \ i \geq \frac{1}{2} \text{codim } F_R, \text{ and} \\
H^i(i_x^* \pi_* \mathcal{M}) &= 0 \quad \text{for } x \in F_R, \ i \leq \frac{1}{2} \text{codim } F_R
\end{align}

for every stratum $F_R \subset X^*$.\n
ASTÉRISQUE 298
Recall that for every \( P \in \mathcal{P} \) with \( P^1 = R \) there is a factorization \( X_P = X_{P, \ell} \times F_R \) and that \( \pi|_{X_P} \) is simply projection onto the second factor. Thus \( \pi^{-1}(x) = \coprod_{\ell = R} X_{P, \ell} \times \{ x \} = \tilde{X}_{R, \ell} \times \{ x \} \) and we let \( \iota_{R, \ell} : \tilde{X}_{R, \ell} \cong \pi^{-1}(x) \hookrightarrow \tilde{X} \) be the inclusion. Since \( H^i(i^{\bar{\pi}}_{x, \ast} \mathcal{M}) = H^i(\tilde{X}_{R, \ell}; \tilde{i}^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) and \( H^i(i^{\bar{\pi}}_{x, \ast} \mathcal{M}) = H^i(\tilde{X}_{R, \ell}; \tilde{i}^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) we can use Theorem 1 to see these vanish for \( i > d(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) and \( i < c(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) respectively. Thus the following theorem implies that (10.1) holds (and hence completes the proof of the Main Theorem):

**Theorem 4.** — Let \( \mathcal{M} \) be an \( E \)-micro-pure \( \mathcal{L} \)-module and let \( F_R \) be a stratum of the Baily-Borel-Satake compactification \( X^* \). Then

\[
d(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \leq \frac{1}{2} \, \text{codim} \, F_R - 1 \quad \text{and} \quad c(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \geq \frac{1}{2} \, \text{codim} \, F_R + 1.
\]

The same result holds if \( D \) is an equal-rank symmetric space and \( \mathcal{S} \) is a real equal-rank Satake compactification as in §4. This theorem is actually a special case of a more general result on the functoriality of micro-support: for \( \mathcal{M} \) an arbitrary \( \mathcal{L} \)-module and \( X^* \) a real equal-rank Satake compactification as in §4, the theorem gives a bound on \( SS(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) and \( SS(i^{\bar{\pi}}_{R, \ell} \mathcal{M}) \) in terms of \( SS(\mathcal{M}) \).

Since \( WC(\tilde{X}; E) \) is also \( E \)-micro-pure, the same argument yields a new proof of the main result of [8] (and in fact a generalization to real equal-rank Satake compactifications).

### 11. Example/application: ordinary cohomology

As another application of \( \mathcal{L} \)-modules we consider the ordinary cohomology \( H(X; \mathcal{E}) \) or \( H(\Gamma; E) \) with coefficients in \( E \in \mathfrak{c}(G) \). This is the cohomology \( H(\tilde{X}; \mathcal{M}) \) for the \( \mathcal{L} \)-module \( \mathcal{M} = i_G \mathcal{E} \) which has \( E_G = E \) and \( E_P = 0 \) for \( P \neq G \) (see §6(i)).

We calculate the micro-support of \( i_G \mathcal{E} \). Since \( i^1_Q i_G \mathcal{E} = E_Q \) we see that

\[
H(i^1_P i^1_Q i_G \mathcal{E}) = \begin{cases} H(n_P; E) & \text{for } Q = G , \\ 0 & \text{for } Q \neq G . \end{cases}
\]

Thus for \( V \in \mathfrak{c}(L_P) \) to be in \( SS(i_G \mathcal{E}) \) it must be an irreducible component of \( H(n_P; E) \) satisfying \( (V|_{M_P})^* \cong \overline{V}|_{M_P} \) and \( (\xi_V + \rho, \alpha) \leq 0 \) for all \( \alpha \in \Delta_P \) (since \( Q = G \) implies \( Q'_V = G \)). The essential micro-support will consist of such \( V \) satisfying in addition the strict inequalities \( (\xi_V + \rho, \alpha) < 0 \).

Let \( \lambda \) be the highest weight of \( E \). As in §9, the irreducible components of \( H(n_P; E) \) are the modules \( V_{w(\lambda + \rho) - \rho} \in \mathfrak{c}(L_P) \) with highest weight \( w(\lambda + \rho) - \rho \) for \( w \in W_P \). Let \( \tau_P : \mathfrak{h}_P \to \mathfrak{h}_P \) transform the highest weight of a representation of \( L_P \) into the highest weight of its complex conjugate contragredient; we assume that \( \mathfrak{h} = \mathfrak{h}_P + a_P = \cdots \)

SOCIÉTÉ MATHÉMATIQUE DE FRANCE 2005
\[ \mathfrak{b}_{P,\mathfrak{l}} + \mathfrak{b}_{P,\mathfrak{p}} + \mathfrak{a}_{P} \] is a fundamental Cartan subalgebra of \( \mathfrak{I}_P \) equipped with a compatible order so that \( \tau_P \) is simply the Cartan involution [3]. We can now reexpress our calculation as

\[
\text{SS}_{\text{end}}(i_{G^*}E) = \prod_P \left\{ V_{\tau_P} - \mu \ni w \in W_P, \ (w(\lambda + \rho), \alpha) < 0 \text{ for all } \alpha \in \Delta_P, \right. \\
\left. \text{and } \tau_P(\alpha + \rho)|_{b_P} = \alpha + \rho \right\}.
\]

(In the last equation we have used the fact that \( \tau_P|_{b_P} = \rho|_{b_P} \).) Furthermore since \( V = V_{\tau_P} \) occurs in \( H(i_P^* i_Q^* i_G^* E) \) in degree \( \ell(w) \) we see that

\[
(11.1) \quad \tilde{c}(V; i_G^* E) = \frac{1}{2}(\dim D_P - D_P(V)) + \ell(w).
\]

We use Lemma 3 to estimate \( \ell(w) \), however now we need the term \( \dim \mathfrak{n}_P(V) \). To define it, recall we have defined \( L_P(\mu) \subseteq L_P \) in \S 8 to have roots \( \gamma \perp \mu = w(\lambda + \rho) - \rho. \) Since \( (w(\lambda + \rho) - \rho)|_{b_P} \) is invariant under \( \tau_P \), the roots of \( L_P(\mu) \) are stable under \( \tau_P \).

Thus given an \( L_P(\mu) \)-irreducible submodule of \( \mathfrak{n}_P \mathcal{C} \), the transform by \( -\tau_P \) of its weights are the weights of another \( L_P(\mu) \)-irreducible submodule of \( \mathfrak{n}_P \mathcal{C} \). Define \( \mathfrak{n}_P(\mu) \) to be the sum of the \( L_P(\mu) \)-irreducible submodules of \( \mathfrak{n}_P \) whose weights are stable under \( -\tau_P \). Choose a compatible ordering for which \( \dim \mathfrak{n}_P(\mu) \) is maximized and let \( \mathfrak{n}_P(V) = \mathfrak{n}_P(\mu) \). Note that \( \mathfrak{n}_P(V) \) contains the root spaces of the positive \( -\tau_P \)-invariant roots, that is, the real roots.

We now make two assumptions: that \( D \) is Hermitian symmetric, or more generally equal-rank, and that \( E \) has regular highest weight \( \lambda. \) By the first assumption the Lie algebra of \( \hat{G}(\mathbb{R}) \) also possesses a compact Cartan subalgebra and therefore by the Kostant-Sugiura theory of conjugacy classes of Cartan subalgebras [14], [31], [32] there must exist at least \( \dim \mathfrak{b}_{P,\mathfrak{l}} + \dim \mathfrak{a}_P = \dim \mathfrak{a}_G \) orthogonal real roots.

Thus

\[
(11.2) \quad \dim \mathfrak{n}_P(V) \geq \dim \mathfrak{b}_{P,\mathfrak{l}} + \dim \mathfrak{a}_P - \dim \mathfrak{a}_G.
\]

On the other hand, note that if \( \gamma^\vee = 2\gamma/(\gamma, \gamma) \) then \( (\rho, \gamma^\vee) = 1 \) if and only if \( \gamma \) is simple. Consequently for \( \gamma \) a simple root of \( L_P \) in any compatible ordering we have

\[
\gamma \text{ is a root of } L_P(\mu) \iff (w(\lambda + \rho), \gamma^\vee) = (\rho, \gamma^\vee) \iff (\lambda + \rho, w^{-1}\gamma^\vee) = 1 \iff (\lambda, w^{-1}\gamma) = 0 \text{ and } w^{-1}\gamma \text{ is simple}.
\]

Thus the second assumption implies that \( L_P(\mu) = H, \) the Cartan subgroup, and hence

\[
(11.3) \quad \dim D_P(V) = \dim \mathfrak{b}_{P,\mathfrak{l}}.
\]

There is a significant improvement here since we have \( \dim \mathfrak{a}_P = \dim \mathfrak{a}_G \). Thus Lemma 3(i) and equations (11.1)–(11.3) yield the estimate

\[
\tilde{c}(V; i_G^* E) \geq \frac{1}{2}(\dim D_P + \dim \mathfrak{a}_P + \dim \mathfrak{n}_P - \dim \mathfrak{a}_G) = \frac{1}{2} \dim X.\]

Thus Theorem 1 implies

**Theorem 5.** — If \( X \) is an arithmetic quotient of a Hermitian or equal-rank symmetric space and \( E \) has regular highest weight then \( H^i(X; \mathcal{E}) = 0 \) for \( i < \frac{1}{2} \dim X. \)
This resolves a question posed by Tilouine during the Automorphic Forms Semester. For the case \( G = \mathbb{R}_k/\mathbb{Q} \text{GSp}(4) \) where \( k \) is a totally real number field the theorem is proven in [33] using results of Franke. For applications of the theorem see [18], [19]. While this paper was being prepared we heard that Li and Schwermer also had a proof of the theorem.\(^{(4)}\)

A vanishing range for the case where \( E \) does not have regular highest weight may be obtained by replacing (11.2) and (11.3) by the more subtle estimate on \( \dim \mathfrak{n}_P(V) \) given in [26].

References


\(^{(4)}\) Added Oct. 2003: See [16]. The methods are completely different. They show vanishing in the range \( i < \frac{1}{2}(\dim X - (\text{rank} \, G - \text{rank} \, K)) \) without assuming \( D \) is equal-rank. This strengthened theorem also follows from the methods of the present paper: if \( D \) is not equal-rank, equation (11.2) remains true provided we subtract \( \text{rank} \, G - \text{rank} \, K \) from the right-hand side.

L. SAPER, Department of Mathematics, Duke University, Box 90320, Durham, NC 27708, U.S.A.
E-mail : saper@math.duke.edu • Url : http://www.math.duke.edu/faculty/saper

ASTÉRISQUE 298
ON LANGLANDS Functoriality
FROM CLASSICAL GROUPS TO $GL_n$

by

David Soudry

Abstract. — This article is a survey of the descent method of Ginzburg, Rallis and Soudry. This method constructs, for an irreducible, automorphic, cuspidal, self-conjugate representation $\tau$ on $GL_n(\mathbb{A})$, an irreducible, automorphic, cuspidal, generic representation $\sigma(\tau)$, on a corresponding quasi-split classical group $G$, which lifts weakly to $\tau$. This construction works well also for all representations of $GL_n(\mathbb{A})$, which are in the so-called “tempered” part of the expected image of Langlands functorial lift from $G$ to $GL_n$.

Résumé (Sur la fonctorialité de Langlands des groupes classiques à $GL_n$). — Cet article est une exposition de la méthode de descente de Ginzburg, Rallis et Soudry. Cette méthode construit, pour une représentation irréductible, automorphe et cuspidale $\tau$ telle que $\tau = \tau^\ast$, une représentation irréductible, automorphe, cuspidale et générique $\sigma(\tau)$ d’un groupe classique quasi-deployé $G$ (qui dépend de $GL_n$ et $\tau$), telle que $\tau$ corresponde à $\sigma(\tau)$ par la correspondance fonctorielle faible (« weak lifting »). Cette construction est valable aussi pour toutes les représentations de $GL_n(\mathbb{A})$ qui appartiennent à la partie dite « tempérée » de l’image de la correspondance fonctorielle de Langlands de $G$ à $GL_n$.

Introduction

In these notes, I survey a long term work, joint with D. Ginzburg and S. Rallis, where we develop a descent method, which associates to a given irreducible automorphic representation $\tau$ of $GL_n(\mathbb{A})$, an irreducible, automorphic, cuspidal, generic representation $\sigma_\tau$ on a given appropriate split classical group $G$, such that $\sigma_\nu$ lifts to $\tau_\nu$, for almost all places $\nu$, where $\tau_\nu$ is unramified. Of course, not every $\tau$ is obtained in such a way. We have to restrict ourselves to $\tau$ which lies in the expected

2000 Mathematics Subject Classification. — 11F70, 11R39.

Key words and phrases. — Descent method, Eisenstein series, Whittaker model, $L$-functions, functorial lift.

Partially supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.
(conjectural) image of the functorial lift from $G$ to $GL_n$, restricted to cuspidal representations $\sigma$ of $G(\mathbb{A})$. We restrict ourselves even more and consider only generic $\sigma$. This also applies to quasi-split unitary groups $G$. Here $\mathbb{A}$ denotes the adele ring of a number field $F$. Thus, for example, let $E$ be a quadratic extension of $F$, and let $\tau$ be an irreducible, automorphic, cuspidal representation of $GL_{2n+1}(\mathbb{A}_E)$, such that its partial Asai $L$-function $L^S(\tau, Asai, s)$ has a pole at $s = 1$. Then we construct an irreducible, automorphic, cuspidal, generic representation $\sigma_\tau$ of $U_{2n+1}(\mathbb{A})$, which lifts weakly (i.e. lifts at all places, where $\tau$ is unramified) to $\tau$. Here, $U_{2n+1}$ is the quasi-split unitary group in $2n + 1$ variables, which corresponds to $E$. We regard it as an algebraic group over $F$. Note that $\sigma_\tau$ would probably be a generic member of “an $L$-packet which lifts to $\tau$”. Of course, $\sigma_\tau$ is a generic member of the near equivalence class which lifts to $\tau$.

The basic ideas of our descent method (backward lift) can be found in [GRS7, GRS8]. A more detailed account appears in [GRS1], where we also start focusing on the descent from cuspidal $\tau$ on $GL_{2n}(\mathbb{A})$, such that $L^S(\tau, \Lambda^2, s)$ has a pole at $s = 1$, and $L(\tau, 1/2) \neq 0$, to $\psi$-generic cuspidal representations $\sigma$ on the metaplectic cover of $Sp_{2n}$. We complete the study of this case (for non-cuspidal $\tau$ as well) in [GRS2, GRS3, GRS4, GRS6]. In [GRS9], we consider the lift from (split) $SO_{2n+1}$ to $GL_{2n}$. I review this last case in Chapter 1 of these notes. Here we can prove more; namely, that the generic cuspidal representation $\sigma_\tau$ is unique up to isomorphism. This is achieved due to a “local converse theorem” for generic representations of $SO_{2n+1}(k)$, over a $p$-adic field $k$, proved in [Ji.So.1]. In Chapter 2, I review integral representations for standard $L$-functions for $G \times GL_m$ (valid only for generic representations). The integrals are of Rankin-Selberg or Shimura type. They are certain Gelfand-Graev, or Fourier-Jacobi coefficients applied to Eisenstein series or cusp forms. In Chapter 3, I review the descent from $GL_n$ to $G$ in general, and in Chapter 4, I illustrate various proofs through low rank examples.

This survey is the content of a minicourse that I gave at Centre Émile Borel, IHP, Paris, when I took part in the special semester in automorphic forms (Spring 2000). I thank the organizers H. Carayol, M. Harris, J. Tilouine, and M.-F. Vignéras for their invitation, and I thank my audience for their attention.

Frequently used notation

- $F$ - a number field.
- $\mathbb{A} = \mathbb{A}_F$ - the adele ring of $F$.
- $F_\nu$ - the completion of $F$ at a place $\nu$.
- $\mathcal{O}_\nu$ - the ring of integers of $F_\nu$, in case $\nu < \infty$.
- $\mathcal{P}_\nu$ - the prime ideal of $\mathcal{O}_\nu$.
- $q_\nu = |\mathcal{O}_\nu/\mathcal{P}_\nu|$.
- $SO_m(F) = \{g \in GL_m(F) | t^\nu J g = J\}$, where $J = \begin{pmatrix} 1 & \cdots \\ & \ddots & 1 \\ & & 1 \end{pmatrix}$. 
Let $\mathbb{R}^+$ denote the group of positive real numbers. Let $i : \mathbb{R}^+ \to \mathbb{A}^*$ be defined by $i(r) = \{x_\nu\}$, where for all finite places $\nu$, $x_\nu = 1$, and for each archimedean place $\nu$, $x_\nu = r$. We denote $i(\mathbb{R}^+) = \mathbb{A}^+_\nu$. For an irreducible representation $\tau$, $\omega_\nu$ denotes its central character. Sometimes we denote by $V_\tau$ a vector space realization of $\tau$. When $\tau$ is an automorphic cuspidal representation, we assume that $\tau$ comes together with a specific vector space realization of cusp forms, which we sometimes denote by $\tau$ as well. Finally, given representations $\tau_1, \ldots, \tau_r$ of $GL_n(F_\nu), \ldots, GL_n(F_\nu)$ respectively, we denote by $\tau_1 \otimes \cdots \otimes \tau_r$ the representation of $GL_n(F_\nu), n = n_1 + \cdots + n_r$, induced from the standard parabolic subgroup, whose Levi part is isomorphic to $GL_{n_1}(F_\nu) \times \cdots \times GL_{n_r}(F_\nu)$, and the representation $\tau_1 \otimes \cdots \otimes \tau_r$.

1. The weak lift from $SO_{2n+1}$ to $GL_{2n}$

In this chapter we survey the results on the weak lift from $SO_{2n+1}$ to $GL_{2n}$, obtained after applying our descent method (backward lift). Together with the existence of this weak lift for generic representations $[C.K.P.S.S.]$, we obtain a fairly nice description of this weak lift, which turns out to be not weak at all.

1.1. Some preliminaries. — Let $\sigma \cong \otimes \sigma_\nu$ be an irreducible, automorphic, cuspidal representation of $SO_{2n+1}(\mathbb{A})$. For almost all $\nu$, $\sigma_\nu$ is unramified and is completely determined by a semisimple conjugacy class $[a_\nu]$ in $SO_{2n+1}^0 = Sp_{2n}(\mathbb{C})$, so that $L(\sigma_\nu, s) = \det(I_{2n} - g_\nu^{-s} a_\nu)^{-1}$. Let $i$ be the embedding $Sp_{2n}(\mathbb{C}) \subset GL_{2n}(\mathbb{C})$. Then the conjugacy class $[i(a_\nu)]$ in $GL_{2n}(\mathbb{C})$ determines an unramified representation $\tau_\nu$ of $GL_{2n}(F_\nu)$, such that $L(\tau_\nu, s) = L(\sigma_\nu, s)$. The unramified representation $\tau_\nu$ is called the local Langlands lift of $\sigma_\nu$. This notion (of local Langlands lift) is conjecturally defined at all finite places and is well defined at archimedean places. For an archimedean place $\nu$, $\sigma_\nu$ is determined by its Langlands parameter, which is an admissible homomorphism $\varphi_\nu : W_\nu \to Sp_{2n}(\mathbb{C})$ from the Weil group of $F_\nu$. The local lift of $\sigma_\nu$ is the representation $\tau_\nu$ of $GL_{2n}(F_\nu)$, whose Langlands parameter is $i \circ \varphi_\nu : W_\nu \to GL_{2n}(\mathbb{C})$. (For finite places $\nu$, where $\sigma_\nu$ is not unramified, $\sigma_\nu$ is conjecturally parameterized by an admissible homomorphism from the Weil-Deligne group $\varphi_\nu : W_\nu \times SL_2(\mathbb{C}) \to Sp_{2n}(\mathbb{C})$, and an irreducible representation $\tau_\nu$ of $GL_{2n}(F_\nu)$ would be a local lift of $\sigma_\nu$, if $\tau_\nu$ corresponds to the homomorphism $i \circ \varphi_\nu$, under the local Langlands reciprocity law for $GL_{2n}$, now proved by Harris-Taylor $[H.T.]$ and by Henniart $[H.]$.) An irreducible, automorphic representation $\tau \cong \otimes \tau_\nu$ is a weak lift of $\sigma$, if for every archimedean place $\nu$ and for almost all finite places $\nu$ where $\sigma_\nu$ is unramified, $\tau_\nu$ is the local lift of $\sigma_\nu$. Using the converse theorem for $GL_m$ $[C.P.S.]$ and $L$-functions for $SO_{2n+1} \times GL_k$ constructed and studied by Shahidi $[Sistle]$, the existence of a weak lift from $SO_{2n+1}$ to $GL_{2n}$ was established for \textit{globally generic} $\sigma$, by J. Cogdell, H. Kim, I. Piatetski-Shapiro and F. Shahidi.
Theorem ([C.K.P.S.S.]). — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $SO_{2n+1}(\mathbb{A})$. Then $\sigma$ has a weak lift to $GL_{2n}(\mathbb{A})$.

Here we remark that a weak lift of $\sigma$ is realized as an irreducible subquotient of the space of automorphic forms on $GL_{2n}(\mathbb{A})$. Moreover, by the strong multiplicity one property for $GL_{2n}$ [J.S.], all weak lifts of $\sigma$ are constituents of one representation of $GL_{2n}(\mathbb{A})$ of the form $\tau_1 \times \cdots \times \tau_r$, where $\tau_i$ are (irreducible, automorphic) cuspidal representations of $GL_{m_i}(\mathbb{A})$, $m_1 + \cdots + m_r = 2n$ and the set $\{\tau_1, \ldots, \tau_r\}$ is uniquely determined. In particular, if $\sigma$ has a cuspidal weak lift, then it is unique. We are going to describe the image of the above weak lift, starting with its cuspidal part.

1.2. The cuspidal part of the image. — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $SO_{2n+1}(\mathbb{A})$. Assume that $\sigma$ has a cuspidal weak lift $\tau$ on $GL_{2n}(\mathbb{A})$. As we just remarked, $\tau$ is uniquely determined (even with multiplicity one). Clearly $\tau_\nu \cong \hat{\tau}_\nu$ (and $\omega_\nu = 1$), for almost all $\nu$. By the strong multiplicity one and multiplicity one properties for $GL_{2n}$, [J.S.], [Sk], we have $\tau = \hat{\tau}$, i.e. $\tau$ is self-dual. (Similarly, $\omega_\nu = 1$). Let $S$ be a finite set of places, including those at infinity, outside which $\sigma$ and $\tau$ are unramified. We have

$$L^S(\sigma \times \tau) = L^S(\tau \times s\tau) = L^S(\hat{\tau} \times \tau),$$

and hence $L^S(\sigma \times \tau)$ has a pole at $s = 1$. Recall that

$$L^S(\tau \times s\tau) = L^S(\tau,\text{sym}^2) L^S(\tau,\Lambda^2, s).$$

By Langlands’ conjectures, one expects $\tau$ to be “symplectic”, and so the pole of $L^S(\tau \times s\tau)$ at $s = 1$ should come from $L^S(\tau,\Lambda^2, s)$.

Theorem 1. — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $SO_{2n+1}(\mathbb{A})$. Assume that $\sigma$ has a cuspidal weak lift $\tau$ on $GL_{2n}(\mathbb{A})$. Then $L^S(\tau,\Lambda^2, s)$ has a pole at $s = 1$.

Proof. — Let us express the pole at $s = 1$ of $L^S(\sigma \times \tau)$ through a Rankin-Selberg type integral which represents this $L$-function [So1], [G.P.S.R.]. It has the form

$$L(\varphi_\sigma,f_{\tau,s}) = \int_{SO_{2n+1}(F) \backslash SO_{2n+1}(\mathbb{A})} \varphi_\sigma(g) E^\psi(f_{\tau,s},g) dg,$$

where $\varphi_\sigma$ is a cusp form in the space of $\sigma$, $E(f_{\tau,s},\cdot)$ is an Eisenstein series on split $SO_{4n}(\mathbb{A})$ corresponding to a $K$-finite holomorphic section $f_{\tau,s}$ in $\text{Ind}_{P_{2n}(\mathbb{A})} SO_{4n}(\mathbb{A}) | \det |^{s-1/2}$, where $P_{2n}$ is the Siegel parabolic subgroup of $SO_{4n}$. $E^\psi$ denotes a Fourier coefficient along the subgroup

$$N_n = \left\{ u = \left( \begin{array}{c} z \\ I_{2n+2} \\ y' \end{array} \right) \in SO_{4n} \mid z \in Z_{n-1} = \left( \begin{array}{c} 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ z \\ 1 \end{array} \right) \right\},$$

ASTÉRISQUE 298
with respect to the character
\[ \chi_\psi : u \mapsto \psi(2z_1 + 2z_3 + \cdots + z_{2n-2,n-1} + y_{n-1,n+1} - y_{n-1,n+2}). \]
Here \( \psi \) is a fixed nontrivial character of \( F \setminus \mathbb{A} \). The stabilizer of \( \chi_\psi \) inside
\[
\begin{pmatrix}
I_{n-1} & 0 \\
SO_{2n+2} & I_{n-1}
\end{pmatrix}
\]
is the subgroup of all \( \begin{pmatrix} I_{n-1} & g \\ I_{n-1} & \end{pmatrix} \), where \( g \) fixes the vector
\[
\begin{pmatrix}
0 \\
\vdots \\
1 \\
\vdots \\
0
\end{pmatrix}
\]
(inside \( F^{2n+2} \)). This defines (split) \( SO_{2n+1} \) and its embedding (over \( F \)) inside \( SO_{4n} \),
all implicit in the definition of \( L(\varphi_\sigma, f_{\tau,s}) \). For a suitable choice of data,
\[
L(\varphi_\sigma, f_{\tau,s}) = \frac{L^S(\sigma \times \tau, s)}{L^S(\tau, \Lambda^2, 2s)} R(s),
\]
where \( R(s) \) is a meromorphic function, which can be made holomorphic and nonzero
at a neighbourhood of a given point \( s_0 \). We consider \( s_0 = 1 \). Since \( \tau \) is unitary,
\( L^S(\tau, \Lambda^2, 2s) \) is holomorphic at \( s = 1 \). We conclude from the last equation that
\( L(\varphi_\sigma, f_{\tau,s}) \), and hence \( E(f_{\tau,s}, \cdot) \), has a pole at \( s = 1 \) (for some choice of data).
This implies that the constant term of \( E(f_{\tau,s}, I) \), along the radical of \( P_{2n} \), has a pole at
\( s = 1 \), for some decomposable section, and this has the form
\[
(1.3) \quad f_{\tau,s}(I) + \prod_{\nu \in S'} M(f^{(\nu)}_{\tau,s}) \frac{L^S(\tau, \Lambda^2, 2s - 1)}{L^S(\tau, \Lambda^2, 2s)},
\]
for some finite set of places \( S' \) containing \( S \). By \([K, \text{Lemma 2.4}]\), \( M(f^{(\nu)}_{\tau,s}) \) (the
corresponding local intertwining operator at \( I \)) is holomorphic for \( \text{Re}(s) \geq 1 \). We
conclude that \( L^S(\tau, \Lambda^2, s) \) has a pole at \( s = 1 \). Since \( L(\tau, \Lambda^2, s) \) is nonzero for
(each \( s \) and) each \( \nu \), \( L^S(\tau, \Lambda^2, s) \) has a pole at \( s = 1 \).

Remarks

1. For each place \( \nu \), \( L(\tau, \Lambda^2, s) \) is holomorphic at \( s = 1 \). We thus may replace
\( L^S(\tau, \Lambda^2, s) \) by \( L^S(\tau, \Lambda^2, s) \), for any \( S' \) and even by \( L(\tau, \Lambda^2, s) \).

2. If \( \sigma \) is not (globally) generic, \( L(\varphi_\sigma, f_{\tau,s}) \) is identically zero.

The argument in the last proof proves the second direction of the following proposition.
(The first direction is easy and appears in \([GRS1, p. 814]\).)

**Proposition 2.** — Let \( \tau \) be an irreducible, automorphic, cuspidal representation of
\( \text{GL}_k(\mathbb{A}) \), \( k \geq 2 \). Assume that the central character of \( \tau \) is trivial on \( \mathbb{A}^*_\infty \). Let \( s_0 \in \mathbb{C} \)
be such that \( \text{Re}(s_0) \geq 1 \). Then \( E(f_{\tau,s}, \cdot) \) (similarly constructed on \( SO_{2k}(\mathbb{A}) \)) has
a pole at \( s_0 \) (as \( f_{\tau,s} \) varies), if and only if \( k \) is even, \( s_0 = 1 \), and \( L(\tau, \Lambda^2, s) \) has a pole
at \( s = 1 \).
From this proposition we conclude

**Theorem 3.** — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $\text{SO}_{2n+1}(\mathbb{A})$, and let $\tau$ be an irreducible, automorphic, cuspidal representation of $\text{GL}_k(\mathbb{A})$, $k \geq 2$, such that $\omega_\tau|_{\mathbb{A}_\infty^\times} = 1$. Then $L^S(\sigma \times \tau, s)$ is holomorphic for $\text{Re}(s) > 1$, and if $L^S(\sigma \times \tau, s)$ has a pole at $s_0$, such that $\text{Re}(s_0) = 1$, then $k$ is even, $s_0 = 1$ and $L^S(\tau, \Lambda^2, s)$ has a pole at $s = 1$. $(S$, as usual, is a finite set of places, outside of which both $\sigma$ and $\tau$ are unramified.) Finally, if $\tau$ is an automorphic character of $k^*$, then $L^S(\sigma \times \tau, s)$ is entire.

**Proof.** — As in the proof of Theorem 1, we can express $L^S(\sigma \times \tau, s)$ using global integrals (see [G], [So1], [G.PS.R.]). We will review them in more detail later. They involve the Eisenstein series $E(f_{\tau,s}, \cdot)$ on $\text{SO}_{2k}(\mathbb{A})$ when $k \geq 2$, so that, as in Theorem 1, if $L^S(\sigma \times \tau, s)$ has a pole at $s_0$, $\text{Re}(s_0) > 1$, then $E(f_{\tau,s}, \cdot)$ has a pole at $s_0$, and by Proposition 2, we get what we want. In case $k = 1$, the global integrals turn out to be entire, and then it is easy to conclude that $L^S(\sigma \times \tau, s)$ is entire as well.

Let us start now with an irreducible, automorphic, cuspidal representation $\tau$ of $\text{GL}_{2n}(\mathbb{A})$, such that $L(\tau, \Lambda^2, s)$ has a pole at $s = 1$. As we have seen in Theorem 1, this is a necessary condition for (a cuspidal) $\tau$ to lie in the image of the weak lift from $\text{SO}_{2n+1}(\mathbb{A})$. If $\tau$ is a weak lift of a generic $\sigma$, then by (1.2) $L(\varphi_\sigma, f_{\tau,s})$ has a pole at $s = 1$ (for suitable choice of data), and hence (see (1.1)) there is a non-trivial $L^2$-pairing between (the space of) $\sigma$ and

\[
(1.4) \quad \sigma_\psi(\tau) = \text{Span} \{ \text{Res}_{s=1} E^{v^{-1}}(f_{\tau,s}, \cdot)|_{\text{SO}_{2n+1}(\mathbb{A})} \}. 
\]

Now we note that $\sigma_\psi(\tau)$ can be defined as in (1.4) for any cuspidal $\tau$, such that $L(\tau, \Lambda^2, s)$ has a pole at $s = 1$. $\sigma_\psi(\tau)$ is a space of automorphic functions on $\text{SO}_{2n+1}(\mathbb{A})$. The descent map $\tau \mapsto \sigma_\psi(\tau)$ is the main vehicle, which will lead us to the description of the functorial lift from $\text{SO}_{2n+1}$ to $\text{GL}_{2n}$. One of the main theorems is

**Theorem 4.** — Let $\tau$ be an irreducible, automorphic, cuspidal representation of $\text{GL}_{2n}(\mathbb{A})$. Assume that $L(\tau, \Lambda^2, s)$ has a pole at $s = 1$. Then $\sigma_\psi(\tau)$ is a nonzero, irreducible, automorphic, cuspidal, generic representation of $\text{SO}_{2n+1}(\mathbb{A})$, which weakly lifts to $\tau$. Every other such representation has a non-trivial $L^2$-pairing with $\sigma_\psi(\tau)$.

**Guidelines to the proof**

1. $\sigma_\psi(\tau)$ is cuspidal: put, for short $e_\tau(h) = \text{Res}_{s=1} E(f_{\tau,s}, h)$. We have to show that all constant terms of $e_\tau^{v^{-1}}$, along unipotent radicals (of parabolic subgroups) in $\text{SO}_{2n+1}$, vanish. Consider then the constant term of $e_\tau$ along the unipotent radical of the standard parabolic subgroup of $\text{SO}_{2n+1}$, which preserves a $p$-dimensional isotropic...
subspace, \(1 \leq p \leq n\). This constant term (evaluated at \(h = I\)) equals [GRS1, Chapter 2]

\[
(1.5) \quad \sum_{\gamma \in Z_p(F) \setminus GL_p(F)} \int_{\mathcal{Z}_p(\mathbb{A})} e^{(N_{n-p},\psi^{-1})_\gamma(x)} dx,
\]

where \(Z_p\) is the standard maximal unipotent subgroup of \(GL_p\), \(\mathcal{Z}_p\) is a certain unipotent subgroup inside the Levi part of \(P_{2n}\), \(\beta\) is a certain Weyl element of \(SO_{4n}\), and 
\(\hat{\gamma} = \left( \gamma, I_{2(n-p)} \right)\), \(e^{(N_{n-p},\psi^{-1})_\gamma}\) is the Fourier coefficient of \(e_\gamma\) along

\[
N_{n-p} = \left\{ u = \left( \begin{array}{ccc} z & y & e \\ y & z^* & \psi \end{array} \right) \in SO_{4n} \mid z \in Z_{n+p-1} \right\},
\]

with respect to the character

\[
\chi^{(n-p)}_\psi : u \mapsto \psi^{-1} \left( \sum_{i=1}^{n+p-2} z_{i,i+1} \right) \psi^{-1} (y_{n+p-1,n-p+1} - y_{n+p-1,n-p+2}).
\]

As for the case \(p = 0\), \(\chi^{(n-p)}_\psi\) is fixed by \(SO_{2(n-p)+1}\), appropriately embedded in \(SO_{4n}\), and we may consider

\[
\sigma^{(n-p)}_\psi(\tau) = \text{Span}\{ e^{(N_{n-p},\psi^{-1})_\gamma} | \gamma \in SO_{2(n-p)+1}(\mathbb{A}) \}.
\]

The cuspidality of \(\sigma^{(n-p)}_\psi(\tau)\) is implied by

\[
(1.6) \quad \sigma^{(k)}_\psi(\tau) = 0, \quad \forall 0 \leq k < n.
\]

This is proved using just one place. First, note that the residues \(e_\gamma\) are square integrable. Next, take an irreducible summand \(\pi\) of the space of the residues \(e_\gamma\). At a place \(\nu\), where \(\pi_\nu\) is unramified, \(\pi_\nu\) is the spherical constituent of \(\text{Ind}_{P_{2n}(\mathbb{F}_\nu)}^{SO_{4n}(\mathbb{F}_\nu)} \tau_\nu \mid \det^{-1/2}\). One shows, using Bruhat theory, that the corresponding Jacquet modules vanish

\[
(1.7) \quad J_{N_h(\mathbb{F}_\nu),\chi^{(n)}_\psi}(\pi_\nu) = 0, \quad \forall 0 \leq k < n.
\]

This depends only on the fact that (unramified) \(\tau_\nu\) is self-dual and \(\omega_{\tau_\nu} = 1\).

(2) \(\sigma_\psi(\tau)\) is nontrivial: this depends only on the fact that \(\tau\) is (globally) generic. We can relate the \(\psi\)-Whittaker coefficient of \(\sigma_\psi(\tau)\) to that of \(\tau\).

(3) Write \(\sigma_\psi(\tau) = \bigoplus \sigma_i\) - a direct sum of irreducible (cuspidal) representations. Each summand \(\sigma_i\) weakly lifts to \(\tau\). This follows from the fact that at a place \(\nu\), where \(\pi_\nu\) (as in (1.7)) and \(\tau_\nu\) are unramified, \(J_{N_h(\mathbb{F}_\nu),\chi^{(n)}_\psi}(\pi_\nu)\), which surjects on \(\sigma_{i,\nu}\), shares its unramified constituent with that of \(\text{Ind}_{B(\mathbb{F}_\nu)}^{SO_{2n+1}(\mathbb{F}_\nu)} \mu_{1,\nu} \otimes \cdots \otimes \mu_{n,\nu}\), where \(B\) is the Borel subgroup of \(SO_{2n+1}\), and \(\tau_\nu\) is the unramified constituent of \(\mu_{1,\nu} \times \cdots \times \mu_{n,\nu} \times \mu_{n,\nu}^{-1} \times \mu_{1,\nu}^{-1}\) on \(GL_{2n}(\mathbb{F}_\nu)\) (\(\mu_{i\nu}\) are unramified characters of \(\mathbb{F}_\nu^*\)).
(4) Decompose $\sigma_\psi(\tau)$ into a direct sum $\oplus \sigma_i$ of irreducible cuspidal representations. Each summand $\sigma_i$ has a non-trivial $L^2$-pairing with $\sigma_\psi(\tau)$, and so by definition ((1.4)), $L(\varphi_{\tau_i}, f_{\tau_i}) \neq 0$ (see (1.1)). By Remark (2), after the proof of Theorem 1, $\sigma_i$ must be generic for all $i$.

Note that since $\sigma_i$ is generic, its weak lift $\tau'$ on $GL_{2n}(\mathbb{A})$ [C.K.P.S.S.]. By the strong multiplicity one and multiplicity one properties for $GL_{2n}$, we must have $\tau' = \tau$. In particular, $\tau_\nu$ is the local lift of $\sigma_{i,\nu}$ at infinite places as well.

(5) $\sigma_\psi(\tau)$ is multiplicity free: if $\sigma_i$ and $\sigma_j$ acting in subspaces $V_{\sigma_i}, V_{\sigma_j}$ are isomorphic summands, choose an isomorphism (of representations) $T : V_{\sigma_i} \rightarrow V_{\sigma_j}$, such that $T(\varphi) - \varphi$ has a zero $\psi$-Whittaker coefficient for all cusp forms $\varphi \in V_{\sigma_i}$. This follows from the uniqueness up to scalars of a Whittaker functional. The argument of (4) applied to $\sigma_i'$ acting in $\{T(\varphi) - \varphi \mid \varphi \in V_{\sigma_i}\}$ shows that $\sigma_i$ must be globally generic. This is a contradiction, unless $T = id$.

(6) $\sigma_\psi(\tau)$ is irreducible: it follows from Cor. 4 in Sec. 6 of [C.K.P.S.S.], that for any two summands $\sigma_i, \sigma_j$, and any place $\nu$, we have an equality of local gamma factors:

$$\gamma(\sigma_{i,\nu} \times \eta, s, \psi_\nu) = \gamma(\sigma_{j,\nu} \times \eta, s, \psi_\nu),$$

for any irreducible representation $\eta$ of $GL_k(F_\nu)$, $k = 1, 2, \ldots$. By the local converse theorem (for generic representation of $SO_{2n+1}(F_\nu)$ of [Ji.So.1], we conclude that $\sigma_{i,\nu} \cong \sigma_{j,\nu}$, for all finite places $\nu$. For archimedean $\nu$, we already know that $\sigma_{i,\nu} \cong \sigma_{j,\nu}$ (both representations have the same Langlands parameter as $\tau_\nu$, for $\nu$ archimedean). We conclude that $\sigma_i \cong \sigma_j$, and by (5) $\sigma_i = \sigma_j$, and so $\sigma_\psi(\tau)$ has only one irreducible summand (appearing with multiplicity one) i.e $\sigma_\psi(\tau)$ is irreducible.

1.3. Description of the image in general, and endoscopy. — In general, an irreducible, automorphic, cuspidal, generic representation $\sigma$ of $SO_{2n+1}(\mathbb{A})$ weakly lifts an irreducible automorphic representation $\tau$ of $GL_{2n}(\mathbb{A})$, which is a constituent of an induced representation of the form

$$\delta_1 |\det z_1| \cdots \delta_j |\det z_j| \tau_1 \cdots \tau_{k} \times \delta_j |\det z_j| \cdots \delta_1 |\det z_1|,$$

where $\text{Re}(z_1) \leq \cdots \leq \text{Re}(z_j) \leq 0$, and each of the representations $\delta_i, \tau_k$ is irreducible, automorphic, unitary, cuspidal, or an automorphic character of the idele group, so that their central characters are trivial on $\mathbb{A}_{\infty}^+$, and also $\tau_i = \tilde{\tau}_i$, for $i = 1, \ldots, \ell$. We have (for appropriate $S$)

$$L^S(\sigma \times \delta_1, s) = \prod_{i=1}^j L^S(\delta_i \times \delta_1, s + z_i) L^S(\delta_1 \times \delta_1, s - z_1) \prod_{i=1}^\ell L^S(\tau_i \times \delta_1, s).$$

This product has a pole at $s = 1 - z_1$. (It comes from $L^S(\delta_1 \times \delta_1, s + z_1)$. Note that $\text{Re}(1 - z_1), \text{Re}(1 - z_1 \pm z_i) \geq 1$, so that the other factors in the product do not cancel this pole.) From Theorem 3, we conclude, in particular, that $\delta_1$ is not a character of the idele group, $z_1 = 0$ and $\delta_1 = \tilde{\delta}_1$, but then $L^S(\sigma \times \delta_1, s)$ has a double pole at
s = 1, which is impossible. (The global integral which represents \( L^S(\sigma \times \delta_1, s) \) involves the Eisenstein series on \( \text{SO}_{2k_1}(\mathbb{A}) \), induced from \( \delta_1 \) and the Siegel parabolic subgroup. This Eisenstein series can have at most simple poles for \( \text{Re}(s) \geq 1/2 \).) We conclude that “there are no \( \delta_i \)-s”, and

\[
\tau \cong \tau_1 \times \tau_2 \times \cdots \times \tau_\ell,
\]

where \( \tau_i \) are irreducible, self-dual, automorphic, cuspidal, such that (again by Theorem 3) \( L^S(\tau_i, \Lambda^2, s) \) has a pole at \( s = 1 \), and also \( \tau_i \neq \tau_j \), for \( 1 \leq i \neq j \leq \ell \). (We just need to repeat the last argument.) Note that for any irreducible, automorphic, unitary representations \( \tau_1, \ldots, \tau_\ell \) (on \( \text{GL}_{k_1}(\mathbb{A}), \ldots, \text{GL}_{k_\ell}(\mathbb{A}) \) respectively) the representation \( \tau_1 \times \cdots \times \tau_\ell \) is irreducible. This proves

**Theorem 5.** — Let \( \sigma \) be an irreducible, automorphic, cuspidal, generic representation of \( \text{SO}_{2n+1}(\mathbb{A}) \). Then \( \sigma \) weakly lifts to a representation (on \( \text{GL}_{2n}(\mathbb{A}) \)) of the form \( \tau = \tau_1 \times \cdots \times \tau_\ell \), where \( \tau_1, \ldots, \tau_\ell \) are pairwise different irreducible, automorphic, cuspidal representations of \( \text{GL}_{2n_1}(\mathbb{A}), \ldots, \text{GL}_{2n_\ell}(\mathbb{A}) \), \( n_1 + \cdots + n_\ell = n \), respectively, such that \( L^S(\tau_i, \Lambda^2, s) \) has a pole at \( s = 1 \), for \( 1 \leq i \leq \ell \).

Conversely, let \( \tau \) be an irreducible representation of \( \text{GL}_{2n}(\mathbb{A}) \) of the form just described in Theorem 5. We can apply the same procedure as in Sec.1.2 (case \( \ell = 1 \)) and construct \( \sigma_\psi(\tau) \) — an irreducible, automorphic, cuspidal, generic representation of \( \text{SO}_{2n+1}(\mathbb{A}) \), which lifts weakly to \( \tau \). For this, we consider the Eisenstein series on \( \text{SO}_{4n}(\mathbb{A}) \) corresponding to a \( K \)-finite, holomorphic section \( f_{\tau, \underline{s}} \) in \( \text{Ind}_{\mathbb{Q}_k}^{\text{SO}_{4n}(\mathbb{A})} \chi_1(\det (\cdot)^{s_1-1/2} \cdots \tau_\ell(\cdot)^{s_\ell-1/2}) \), where \( \underline{s} = (s_1, \ldots, s_\ell) \) and \( Q \) is the standard parabolic subgroup of \( \text{SO}_{4n} \), whose Levi part is isomorphic to \( \text{GL}_{2n_1} \times \cdots \times \text{GL}_{2n_\ell} \). Denote this Eisenstein series by \( E(f_{\tau, \underline{s}}, h) \). As in [GRS4, Theorem 2.1], we can prove that the function

\[
(s_1 - 1)(s_2 - 1) \cdots (s_\ell - 1)E(f_{\tau, \underline{s}}, h)
\]

is holomorphic at \( \underline{s} = (1, 1, \ldots, 1) \) and is not identically zero, as the section varies. Consider

\[
\text{Res}_{\underline{s} = 1} E(f_{\tau, \underline{s}}, h) = \lim_{\underline{s} \rightarrow 1} (s_1 - 1) \cdots (s_\ell - 1)E(f_{\tau, \underline{s}}, h),
\]

where \( 1 = (1, \ldots, 1) \). These residues generate a square integrable automorphic representation of \( \text{SO}_{4n}(\mathbb{A}) \). Consider, as in (1.4)

\[
\sigma_\psi(\tau) = \text{Span}\{\text{Res}_{\underline{s} = 1} E^{s-1}(f_{\tau, \underline{s}}, \cdot)\mid \text{SO}_{2n+1}(\mathbb{A})\}.
\]

**Theorem 6.** — Let \( \tau = \tau_1 \times \tau_2 \times \cdots \times \tau_\ell \) be the irreducible representation of \( \text{GL}_{2n}(\mathbb{A}) \), induced from \( \tau_1 \otimes \cdots \otimes \tau_\ell \), where \( \tau_1, \ldots, \tau_\ell \) are pairwise inequivalent irreducible, automorphic, cuspidal representations on \( \text{GL}_{2n_1}(\mathbb{A}), \ldots, \text{GL}_{2n_\ell}(\mathbb{A}) \) respectively, \( n_1 + \cdots + n_\ell = n \), such that for each \( 1 \leq i \leq \ell \), \( L^S(\tau_i, \Lambda^2, s) \) has a pole at \( s = 1 \). Then \( \sigma_\psi(\tau) \) is a nonzero, irreducible, automorphic, cuspidal, generic representation of \( \text{SO}_{2n+1}(\mathbb{A}) \).
which weakly lifts to \( \tau \). Any other such representation has a non-trivial \( L^2 \)-pairing with \( \sigma_\psi(\tau) \).

**Proof.** — The nontriviality of \( \sigma_\psi(\tau) \) is shown exactly as in case \( \ell = 1 \). As we mentioned in the proof of Theorem 4, only the fact that \( \tau \) is generic is important here. The cuspidality of \( \sigma_\psi(\tau) \) is shown as in case \( \ell = 1 \), only we need also to use induction on \( \ell \). Let \( \sigma \) be an irreducible summand of \( \sigma_\psi(\tau) \). Then

\[
\int_{\SO_{2n+1}(F) \setminus \SO_{2n+1}(A)} \varphi_\sigma(g) \Res_{\mathfrak{r}_g} E^\psi(f_{\tau,\mathfrak{r}}, g) dg \neq 0,
\]
as the data \( \varphi_\sigma \) and \( f_{\tau,\mathfrak{r}} \) vary. In particular

\[
\mathcal{L}(\varphi_\sigma, f_{\tau,\mathfrak{r}}) = \int_{\SO_{2n+1}(F) \setminus \SO_{2n+1}(A)} \varphi_\sigma(g) E^\psi(f_{\tau,\mathfrak{r}}, g) dg \neq 0.
\]

As in (1.4), also in this case the integrals \( \mathcal{L}(\varphi_\sigma, f_{\tau,\mathfrak{r}}) \) represent

\[
\prod_{i=1}^\ell L^S(\sigma \times \tau_i, s_i) \prod_{1 \leq i < j \leq \ell} L^S(\tau_i \times \tau_j, s_i + s_j) \prod_{i=1}^\ell L^S(\tau, \Lambda^2, 2s_i),
\]
for generic \( \sigma \). Moreover, as in case \( \ell = 1 \), if \( \sigma \) is not (globally) generic, then the last two integrals above are identically zero. The rest of the proof is now exactly as in Theorem 4. In particular, the irreducibility of \( \sigma_\psi(\tau) \) follows from the local converse theorem in [Ji.So.1].

As a corollary, we obtain that generic cuspidal representations of \( \SO_{2n+1}(A) \) satisfy the strong multiplicity one property.

**Theorem 7.** — Let \( \sigma_1 \) and \( \sigma_2 \) be two irreducible, automorphic, cuspidal, generic representations of \( \SO_{2n+1}(A) \). Assume that \( \sigma_{1,\nu} \cong \sigma_{2,\nu} \), for almost all places \( \nu \). Then \( \sigma_1 \cong \sigma_2 \).

**Proof.** — Both \( \sigma_1 \) and \( \sigma_2 \) weakly lift to the same representation \( \tau \) on \( \GL_{2n}(A) \). \( \tau \) has the form as in Theorem 6. By Theorem 6, \( \sigma_1 \) and \( \sigma_2 \) have non-trivial \( L^2 \)-pairings with \( \sigma_\psi(\tau) \). In particular \( \sigma_1 \cong \sigma_\psi(\tau) \cong \sigma_2 \).

**Example.** — Consider the group \( \SO_2(A) \cong \PGSp_4(A) \). Every irreducible, automorphic, cuspidal, generic representation of \( \PGSp_4(A) \) has a unique weak lift to \( \GL_4(A) \). The image of this lift consists of all irreducible, automorphic, cuspidal representations \( \tau \) of \( \GL_4(A) \), such that \( L^S(\tau, \Lambda^2, s) \) has a pole at \( s = 1 \), and of all representations of the form \( \tau_1 \times \tau_2 \), where \( \tau_1 \) and \( \tau_2 \) are different, irreducible, automorphic, cuspidal representations of \( \GL_2(A) \), each one having a trivial central character.

**Remark.** — In [Ji.So.1, Ji.So.2] a Langlands reciprocity law is established for generic representations of \( \SO_{2n+1}(F_\nu) \) (\( \nu \) finite). Theorem 6.3 of [Ji.So.2] says (in above notation) that if \( \sigma \) weakly lifts to \( \tau \), then at all places \( \nu \), \( \sigma_\nu \) locally lifts to \( \tau_\nu \) in the
sense that both $\sigma_\nu$ and $\tau_\nu$ correspond to the same Langlands parameter (which is symplectic).

Finally, if $\sigma$ (as before) does not lift to a cuspidal representation of $\GL_{2n}(\A)$ then, as in Theorems 5,6, it lifts to a representation $\tau = \tau_1 \times \cdots \times \tau_\ell$, as in Theorem 6. By Theorem 4, each $\tau_i$ is the lift of $\sigma_i = \sigma_\psi(\tau_i)$ on $\SO_{2n_i+1}(\A)$. Thus $\sigma$ is the (generalized) endoscopic lift of $\sigma_1 \otimes \cdots \otimes \sigma_\ell$ on $\SO_{2n_1+1}(\A) \times \cdots \times \SO_{2n_\ell+1}(\A)$. This lift is compatible with the $L$-group map

$$\Sp_{2n_1}(\C) \times \cdots \times \Sp_{2n_\ell}(\C) \rightarrow \Sp_{2n}(\C).$$

Conversely, let $\sigma_1, \ldots, \sigma_\ell$ be irreducible, automorphic, cuspidal, generic representations of $\SO_{2n_1+1}(\A), \ldots, \SO_{2n_\ell+1}(\A)$ respectively. Consider the lifts $\tau_i$ of $\sigma_i$ to $\GL_{2n_i}(\A)$, $\tau_i = \tau_{i1} \times \cdots \times \tau_{i\ell}$, $i = 1, \ldots, \ell$. Denote $C_i = \{\tau_{ij}\}_{j=1}^{\ell_i}$. Clearly, if $C_i \cap C_{i'} = \emptyset$ for all $1 \leq i \neq i' \leq \ell$, then $\tau = \times_{i=1}^{\ell} \tau_i = \times_{i=1}^{\ell} \times_{j=1}^{\ell_i} \tau_{ij}$ lies in the image of the lift from $\SO_{2n_1+1}(\A)$, and hence $\sigma_\psi(\tau)$ is an irreducible, automorphic, cuspidal, general representation of $\SO_{2n+1}(\A)$, which is the lift of $\sigma_1 \otimes \cdots \otimes \sigma_\ell$. Summarizing

**Theorem 8.** — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $\SO_{2n+1}(\A)$. Assume that the lift of $\sigma$ to $\GL_{2n}(\A)$ is not cuspidal. Then there exist irreducible, automorphic, cuspidal, generic representations $\sigma_1, \sigma_2, \ldots, \sigma_\ell$ on $\SO_{2n_1+1}(\A), \SO_{2n_2+1}(\A), \ldots, \SO_{2n_\ell+1}(\A)$ respectively, $n_1 + \cdots + n_\ell = n$ such that $\sigma$ is the lift of $\sigma_1 \otimes \cdots \otimes \sigma_\ell$. The set $\{\sigma_1, \sigma_2, \ldots, \sigma_\ell\}$ is unique up to permutation and up to isomorphism.

Conversely, let $\sigma_1, \sigma_2, \ldots, \sigma_\ell$ be irreducible, automorphic, cuspidal, generic representations of $\SO_{2n_1+1}(\A), \ldots, \SO_{2n_\ell+1}(\A)$ respectively, $n_1 + \cdots + n_\ell = n$. Consider the sets $\{C_i\}_{i=1}^{\ell}$ as above. If $C_i \cap C_j = \emptyset$ for all $1 \leq i \neq j \leq \ell$, then there is a unique up to isomorphism, irreducible, automorphic, cuspidal, general representation $\sigma$ of $\SO_{2n+1}(\A)$, which is a lift of $\sigma_1 \otimes \cdots \otimes \sigma_\ell$. Otherwise, cuspidal data on $\SO_{2n+1}(\A)$ can be specified, so that $\sigma_1 \otimes \cdots \otimes \sigma_\ell$ lifts to a constituent of the corresponding induced representation.

**Example.** — Let $\sigma_1, \ldots, \sigma_n$ be pairwise different irreducible, automorphic, cuspidal representations of $\PGL_2(\A)$. Then, up to isomorphism, there is a unique irreducible, automorphic, cuspidal, generic representation $\sigma$ of $\SO_{2n+1}(\A)$, which is the lift of $\sigma_1 \otimes \cdots \otimes \sigma_n$.

**1.4. Base change.** — Let us compose our descent map $\tau \mapsto \sigma_\psi(\tau)$ (“backward lift”) with the base change lift for $\GL_{2n}$. Let $E/F$ be a cyclic extension of odd prime degree $p$. Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of
Let $\tau$ be the lift of $\sigma$ on $\text{GL}_{2n}(A)$. We would like to follow the diagram

$$
\begin{array}{c}
\sigma' = \sigma_{\psi}(\tau') \\
\text{SO}_{2n+1}(A_{E}) \\
\text{GL}_{2n}(A_{E}) \\
\tau' = bc(\tau) \\
\sigma = \sigma_{\psi}(\tau) \\
\text{SO}_{2n+1}(A_{F}) \\
\text{GL}_{2n}(A_{F}) \\
\end{array}
$$

Here $\tau' = bc(\tau)$ is the base change lift of $\tau$ [A.C.]. The top arrow of the diagram exists if we show that $\tau'$ lies in the image of the lift (restricted to generic representations) from $\text{SO}_{2n+1}(A_{E})$. The image is described in Theorems 5,6. This is indeed the case. For this, choose a nontrivial character $\eta$ of $A_{E}^*/F^*N_{E/F}A_{F}^*$, and a generator $\varepsilon$ of $\text{Gal}(E/F)$. Starting with a generic $\sigma$ on $\text{SO}_{2n+1}(A_{F})$, we know that its lift $\tau$ on $\text{GL}_{2n}(A_{F})$ has the form $\tau_1 \times \cdots \times \tau_{\ell}$ as in Theorem 5. Since $bc(\tau) = bc(\tau_1) \times \cdots \times bc(\tau_{\ell})$, we have to analyze each representation $bc(\tau_i)$. There are two cases according to whether $\tau_i$ is isomorphic or not isomorphic to $\tau_i \otimes \eta$. If $\tau_i \neq \tau_i \otimes \eta$, then $bc(\tau_i) = \theta_i$ is cuspidal and $\varepsilon$-invariant. We have

$$L^S(\theta_i, \Lambda^2, s) = \prod_{k=0}^{p-1} L^S(\tau_i, \Lambda^2 \otimes \eta^k, s).$$

It is a theorem of Shahidi [Sh2] that each factor in the last product is nonzero at $s = 1$, and since $L^S(\tau_i, \Lambda^2, s)$ has a pole at $s = 1$, we conclude that $L^S(\theta_i, \Lambda^2, s)$ has a pole at $s = 1$. If $\tau_i = \tau_i \otimes \eta$, then $p\mid 2n_i$, and

$$bc(\tau_i) = \theta_i \times \theta_i^e \times \cdots \times \theta_i^{e^{p-1}},$$

where $\theta_i$ is cuspidal, such that $\theta_i \neq \theta_i^e$. We have

$$[L^S(\tau_i, \Lambda^2, s)]^p = \prod_{k=0}^{p-1} L^S(\tau_i, \Lambda^2 \otimes \eta^k, s) = L^S(bc(\tau), \Lambda^2, s)$$

$$= \prod_{\theta_j \neq \theta_i \otimes \eta} L^S(\theta_i^e \times \theta_i^{e^j}, \Lambda^2, s) \prod_{j=0}^{p-1} L^S(\theta_i^e, \Lambda^2, s).$$

We conclude that the product has a pole of order $p$ at $s = 1$. It is easy to see that $\theta_i$ is self-dual. (This follows from the self-duality of $\tau_i$ and the fact that $p$ is odd.) In particular, $\theta_i^{e^j} \neq \theta_i^{e^k}$, for $0 \leq j < k \leq p$. We conclude that $\prod_{j=0}^{p-1} L^S(\theta_i^{e^j}, \Lambda^2, s)$ has a pole of order $p$ at $s = 1$, and hence $L^S(\theta_i^e, \Lambda^2, s)$ has a pole at $s = 1$, for $0 \leq j \leq p - 1$. Finally, it is easy to see that in

$$bc(\tau) = bc(\tau_1) \times \cdots \times bc(\tau_\ell) = \prod_{\tau_i \neq \tau_i \otimes \eta} \theta_i \times \prod_{\tau_i = \tau_i \otimes \eta} \left( \prod_{j=0}^{p-1} \theta_i^{e^j} \right),$$

ASTÉRISQUE 298
all factors are different. This shows (by Theorem 6) that \( \tau' = bc(\tau) \) is in the image of the lift from \( \text{SO}_{2n+1}(\mathbb{A}_E) \). The representation \( \sigma' = \sigma_\psi(\tau') \) is an irreducible, automorphic, cuspidal and generic, and it is a base change lift of \( \sigma \). Summarizing

**Theorem 9.** — Let \( E/F \) be a cyclic extension of odd prime degree. Then there is a base change lift from irreducible, automorphic, cuspidal, generic representations of \( \text{SO}_{2n+1}(\mathbb{A}_E) \) to irreducible, automorphic, cuspidal, generic representations of \( \text{SO}_{2n+1}(\mathbb{A}_F) \).

**Conclusion.** — The descent map (backward lift) \( \tau \mapsto \sigma_\psi(\tau) \) is a very powerful tool. This chapter demonstrated the nice results obtained for \( \text{SO}_{2n+1} \) using the descent map. The ideas and methods are general and apply to other quasi-split classical groups \( G \). The definition of \( \sigma_\psi(\tau) \) (for appropriate \( \tau \)) is intimately related to global integrals (of Rankin-Selberg type, or of Shimura type) representing the standard \( L \)-function for \( G \times \text{GL}_n \). These integrals are available, and we will survey them in the next chapter. These integrals suggest the construction of \( \sigma_\psi(\tau) \), which arises as a natural object; it is constructed so that \( L^S(\sigma_\psi(\tau) \times \tau, s) \) has a pole at \( s = 1 \). The representation \( \sigma_\psi(\tau) \) is defined by taking certain Gelfand-Graev, or Fourier-Jacobi coefficients of the residue at 1 of a certain Eisenstein series induced from \( \tau \). The study of \( \sigma_\psi(\tau) \) is now the study of these Gelfand-Graev, or Fourier-Jacobi coefficients of the residual Eisenstein series induced from \( \tau \). The three main problems concerning \( \sigma_\psi(\tau) \) are the following (for appropriate \( \tau \), i.e. in the expected image of the lift from \( G \) to \( \text{GL}_N \), for appropriate \( N \)).

1. Show that \( \sigma_\psi(\tau) \neq 0 \).
2. Show that \( \sigma_\psi(\tau) \) is cuspidal.
3. Show that each summand of \( \sigma_\psi(\tau) \) weakly lifts to \( \tau \).

In Chapters 4-6, we will indicate how to prove these properties through low rank examples. In this way we construct examples of generic cuspidal representations \( \sigma \) on \( G \), which weakly lift to a given \( \tau \) in the expected image. Similarly, we get examples of (generalized) endoscopy and base change. Once the existence of the weak lift from \( G \) to \( \text{GL}_N \) is established (and not much is missing for the proof by converse theorem to be completed) then our examples above give the general case.

*Note added in proof.* — Recently, the existence of the weak lift of cuspidal generic representations on \( G \) to \( \text{GL}_N \) has indeed been established. See [C.K.PS.S.1].

2. \( L \)-functions for \( G \times \text{GL}_k \), where \( G \) is a quasi-split classical group

(generic representations)

In this chapter, we survey the global integrals (of Rankin-Selberg type, or of Shimura type) which represent the standard \( L \)-functions for generic representations on \( G \times \text{GL}_k \). Note that these \( L \)-functions were obtained by Shahidi [Sh1] using the
Langlands-Shahidi method. However, the integrals we present here relate the fact that \( L^S(\sigma \times \tau, s) \) has a pole at \( s = 1 \), and the fact that \( \sigma \) has a nontrivial \( L^2 \)-pairing with the descent applied to \( \tau \).

We'll first present the notions of certain Gelfand-Graev models and Fourier-Jacobi models, which enter in the definitions of the global integrals.

2.1. Gelfand-Graev models. — Let \( F \) be a field of characteristic different than 2. (Eventually we'll be interested in a number field \( F \) or in its completion in one of its places.) Let \( E \) be either \( F \) or a quadratic extension of \( F \). Denote by \( x \mapsto x \) the nontrivial element of \( \text{Gal}(E/F) \) in case \( [E:F] = 2 \). If \( E = F \), we agree that \( x = x \) on \( F \). Let \( V \) be a finite dimensional vector space over \( E \), equipped with a non-degenerate bilinear form \((\cdot,\cdot)\), which is either symmetric, or anti-symmetric in case \( E = F \), and is Hermitian in case \([E:F] = 2 \). Let \( H = H(V) \) be the connected component of the isometry group of \((V,(\cdot,\cdot))\). We assume that \( H \) acts on \( V \) from the left.

Assume that

\[
V = V^\ell_+ + W + V^\ell_-,
\]

where \( V^\ell_{\pm} \) are isotropic subspaces of dimension \( \ell \), which are in duality under \((\cdot,\cdot)\) (i.e. \((\cdot,\cdot)\) restricted to \( V^\ell_+ \times V^\ell_- \) is non-degenerate), and \( W = (V^\ell_+ + V^\ell_-)^\perp \). Let \( P_\ell \) be the parabolic subgroup of \( H \), which preserves \( V^\ell_+ \). Write its Levi decomposition

\[
P_\ell = M_\ell \ltimes U_\ell.\]

Let us write the elements of \( H \) in matrix form, following the decomposition (2.1). Then (with evident notation)

\[
M_\ell = \left\{ \begin{pmatrix} g & h \\ g^* & g \end{pmatrix} \mid g \in \text{GL}(V^\ell_+), h \in H(W) \right\},
\]

\[
U_\ell = \left\{ u = \begin{pmatrix} I_{V^\ell_+} & g & y \\ I_W & x & y' \\ I_{V^\ell_-} & g^* & v_0 \end{pmatrix} \in H \right\}.
\]

Fix nonzero vectors \( w_0, v_0^- \in W, v_0^- \in V^\ell_- \). Define for \( u \in U_\ell \) (written as in (2.3)) the following rational character

\[
\chi_{w_0,v_0^-}(u) = (u \cdot w_0, v_0^-).
\]

We have

\[
\text{Stab}_{M_\ell}(\chi_{w_0,v_0^-}) = \left\{ \begin{pmatrix} g & h \\ g^* & g \end{pmatrix} \in H \mid h \cdot w_0 = w_0, \quad g^* \cdot v_0^- = v_0^- \right\}.
\]

Thus, if \( w_0 \) is anisotropic, then \( h \cdot w_0 = w_0 \) means that \( h \in H(w_0^+ \cap W) \), and if \( w_0 \) is isotropic, then \( h \cdot w_0 = w_0 \) means that \( h \) lies in the parabolic subgroup \( P_{W,w_0} \) of \( H(W) \), which fixes the isotropic subspace \( E \cdot w_0 \) (and also \( h \cdot w_0 = w_0 \)). Put in this case (i.e. \( (w_0,w_0) = 0 \))

\[
P_{W,w_0}^1 = \{ h \in P_{W,w_0} \mid h \cdot w_0 = w_0 \}.\]
The condition $g^* v^\circ_0 = v^\circ_0$ in (2.4) means that $g$ lies in the so called “mirabolic” subgroup of $GL(V^+_\ell)$. Let us insert more coordinates. Choose a basis $\{v_1, \ldots, v_\ell\}$ of $V^+_\ell$ and a dual basis $\{v_{-\ell}, \ldots, v_{-1}\}$ of $V^-_{\ell}$ (i.e. $(v_i, v_{-j}) = \delta_{ij}$, for $1 \leq i, j \leq \ell$). We assume that $v^\circ_0 = v_{-\ell}$. We identify $GL(V^\pm \ell)$ with $GL\ell(E)$ using these bases. Note that for $g \in GL\ell(E)$, $g^* = w_\ell \overline{g}^{-1} w_\ell$, where $w_\ell = \begin{pmatrix} \vdots \mid 1 \end{pmatrix}$, and $g^* v_{-\ell} = v_{-\ell}$ means that $g \in \mathbb{A}^{-\infty}$. Let $Z_\ell$ be the standard maximal unipotent subgroup of $GL\ell(E)$. Put

$$Z_\ell = \left\{ z = (z_1, \ldots, z_{\ell}) \mid z \in Z_\ell \right\},$$

$$L_{W, w_0} = \begin{cases} H(w_0^+ \cap W), & (w_0, w_0) \neq 0 \\ P_{W, w_0}, & (w_0, w_0) = 0 \end{cases},$$

$$N_\ell = \tilde{Z}_\ell U_\ell,$$

$$R_{\ell, w_0} = N_\ell L_{W, w_0}.$$ Fix a nontrivial character $\psi$ of $F$. Put $\psi_E = \psi \circ tr_{E/F}$. Let $\psi_{\ell, w_0}$ be the following character of $N_\ell$

$$\psi_{\ell, w_0}(z \cdot u) = \psi_{\ell}(z) \psi_E(\chi_{w_0, v^\circ_\ell}(u)) = \psi_E \left( \sum_{i=1}^{\ell-1} z_{i, i+1} \right) \psi_E \left( (u \cdot w_0, v^\circ_\ell) \right).$$

Assume now that $w_0$ is anisotropic. (This precludes symplectic groups $H$.) Let $F$ be a local field, and let $\sigma$ be an irreducible (smooth) representation of $H(w_0^+ \cap W)$. We say that an irreducible (smooth) representation $\pi$ of $H$ has a Gelfand-Graev model with respect to $(R_{\ell, w_0}; \sigma, \psi)$ if

$$\Hom_{R_{\ell, w_0}}(\pi, \psi_{\ell, w_0} \otimes \overline{\sigma}) \neq 0.$$ ($\psi_{\ell, w_0}$ may be viewed as a character of $R_{\ell, w_0}$ by trivial extension.)

Now assume that $F$ is a global field, that $\psi$ is a non-trivial character of $F \setminus \mathbb{A}$ ($\mathbb{A} = \mathbb{A}_F$), and that $\pi$ is an automorphic representation of $H_{\mathbb{A}}$, acting in a space of automorphic forms $V_{\pi}$. Put, for $\varphi_{\pi} \in V_{\pi}$

$$\varphi_{\psi_{\ell, w_0}}(h) = \int_{N_{\ell}(F) \backslash N_{\ell}(\mathbb{A})} \varphi_{\pi}(vh^{-1}) \psi_{\ell, w_0}^{-1}(v) dv.$$ Note that $\varphi_{\psi_{\ell, w_0}}(\gamma h) = \varphi_{\psi_{\ell, w_0}}(h)$, for $\gamma \in H(w_0^+ \cap W)_F$. We call the Fourier coefficient (2.6) the Gelfand-Graev coefficient of $\varphi_{\pi}$ with respect $\psi_{\ell, w_0}$.

Let $\sigma$ be an automorphic representation of $H(w_0^+ \cap W)_\mathbb{A}$ (acting in a space of automorphic forms $V_{\sigma}$). We say that $\pi$ has a global Gelfand-Graev coefficient with respect to $(R_{\ell, w_0}; \sigma, \psi)$ if (the following integral converges absolutely and)

$$b(\varphi_{\pi}, \varphi_{\sigma}) = \int_{H(w_0^+ \cap W)_F \setminus H(w_0^+ \cap W)_\mathbb{A}} \varphi_{\psi_{\ell, w_0}}(g) \varphi_{\sigma}(g) dg \neq 0,$$
as \( \varphi_\pi \) varies in \( V_\pi \) and \( \varphi_\sigma \) varies in \( V_\sigma \). The corresponding Gelfand-Graev model of \( \pi \) is the space of functions on \( H_K \) spanned by the functions \( h \to b(\pi(h)\varphi_\pi, \varphi_\sigma) \), as \( \varphi_\pi \) varies in \( V_\pi \) and \( \varphi_\sigma \) varies in \( V_\sigma \). In practice, one of \((\pi, \sigma)\) will be cuspidal and the other will be “Eisensteinian”.

2.2. Fourier-Jacobi models. — We continue with the previous notations. Assume that \( w_0 \) is isotropic and that \((,\,\,\,\,\,\,)\) is not symmetric (i.e. \( H \) is either symplectic or unitary). Write

\[
W = Ew_0 + W' + Ew_{-0},
\]

where \( w_{-0} \) is isotropic, \((w_0, w_{-0}) = 1 \) and \( W' = (Ew_0 + Ew_{-0})^1 \cap W \). Put \( v_{\ell+1} = w_0, v_{-(\ell+1)} = w_{-0}, V_{\ell+1}^+ = \text{Span}\{v_1, \ldots, v_{\ell}, v_{\ell+1}\}, V_{\ell+1}^- = \text{Span}\{v_{-(\ell+1)}, v_{-\ell}, \ldots, v_{-1}\} \)

and identify, as before, \( \text{GL}(V_{\ell+1}^+) \) with \( \text{GL}_{\ell+1}(E) \). Using these coordinates, an element of \( \mathcal{U}_E \) has the form

\[
u = \begin{pmatrix}
I_{\ell} & y & * & * & * \\
1 & 0 & * & * & * \\
0 & * & I_{w'} & 0 & * \\
1 & y' & * & * & 1 \\
\end{pmatrix}, \quad g \in H(W').
\]

The unipotent radical of \( L_{W,w_0} \) is isomorphic to the Heisenberg group of \( W', \mathcal{H}_{W'} = W' \oplus F \). Note that \( N_\ell \setminus N_{\ell+1} \cong \mathcal{H}_{W'} \). Fix an isomorphism \( j : N_\ell \setminus N_{\ell+1} \to \mathcal{H}_{W'} \). Let \( F \) be a local field. Let \( \omega_\psi \) be the Weil representation of \( \mathcal{H}_{W'} \rtimes \tilde{\text{Sp}}(W') \). If \( H \) is a symplectic group, then \( H(W') = \text{Sp}(W') \). If \( H \) is a unitary group, then so is \( H(W') \), and we embed \( H(W') \) inside \( \text{Sp}(W') \) \((W' \text{ viewed over } F)\). This requires a choice of a character \( \gamma \) of \( E^* \), such that \( \gamma|_{E^*} = \omega_{E/F} \) — the non-trivial quadratic character of \( E^* \), associated to \( E \). See [Ge.Ro.]. Denote, in this case, by \( \omega_{\psi, \gamma} \) the restriction of \( \omega_\psi \) to the image of \( H(W') \). Put \( \omega_{\psi, 1} = \omega_\psi \) in case \( H \) is symplectic (thus denoting here \( \gamma = 1 \)).

Let \( \sigma \) be an irreducible representation of \( H(W') \), in case \( H \) is unitary, and of \( H(W')^\varepsilon, \varepsilon = 0, 1 \), in case \( H \) is symplectic, where

\[
H(W')^\varepsilon = \begin{cases}
\text{Sp}(W'), & \varepsilon = 0 \\
\tilde{\text{Sp}}(W'), & \varepsilon = 1
\end{cases}
\]

Then \( \omega_{\psi, \gamma} \otimes \tilde{\sigma} \) is a representation of \( \mathcal{H}_{W'} \rtimes H(W') \) in case \( H \) is unitary, and of \( \mathcal{H}_{W'} \rtimes H(W')^{1-\varepsilon} \) in case \( H \) is symplectic. Let \( R_{\ell,w_0} \) denote \( R_{\ell,w_0} \) in case \( H \) is unitary, or \( \varepsilon = 1 \), and \( N_{\ell+1} \cdot \tilde{\text{Sp}}(W') \) in case \( \varepsilon = 0 \). We view \( \psi_{\ell,w_0} \) as a character of \( R_{\ell,w_0} \) by trivial extension.
Let \( \pi \) be an irreducible representation of \( H \) in case \( H \) is unitary and of \( H^{1-\varepsilon} = H(V)^{1-\varepsilon} \) in case \( H \) is symplectic. We say that \( \pi \) has a Fourier-Jacobi model with respect to \( (R_{\ell, w_0}; \psi, \gamma, \sigma) \) if

\[
\text{Hom}_{\tilde{\mathcal{R}}_{\ell}}(\pi, \psi_{\ell} \otimes (\omega_{\psi, \gamma} \otimes \delta)) \neq 0,
\]

where we shorten the notation in this case: \( \psi_{\ell} = \psi_{\ell, w_0}, \ R_{\ell} = R_{\ell, w_0} \). Here is a short table which summarizes the above cases.

<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( H(V) )-unitary</th>
<th>( \text{Sp}(V) )</th>
<th>( \tilde{\text{Sp}}(V) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
<td>( H(W') )-unitary</td>
<td>( \tilde{\text{Sp}}(W') )</td>
<td>( \text{Sp}(W') )</td>
</tr>
<tr>
<td>( R_{\ell}^* )</td>
<td>( R_{\ell} )</td>
<td>( R_{\ell} )</td>
<td>( N_{\ell + 1} \cdot \tilde{\text{Sp}}(W') )</td>
</tr>
</tbody>
</table>

Note that \( R_{\ell} \cong N_{\ell} \times (\mathcal{H}_{W'} \times H(W')) \) (using the isomorphism \( j : N_{\ell} \backslash N_{\ell + 1} \xrightarrow{\sim} \mathcal{H}_{W'} \)).

Assume now that \( F \) is a global field and that \( \psi \) is a non-trivial character of \( F \backslash \mathbb{A} \).

Let \( \omega_{\psi} \) be the Weil representation of \( \tilde{\text{Sp}}(W')_{\mathbb{A}} \), and in case \( H \) is a unitary group, fix a character \( \gamma \) of \( E^1 \backslash E \), such that \( \gamma|_{\mathbb{A}_{F}} = \theta_{E/F} \), and denote by \( \omega_{\psi, \gamma} \) the restriction of \( \omega_{\psi} \) to the image of \( H(W')_{\mathbb{A}} \) determined by \( (\gamma, \psi) \). Denote, as before, \( \omega_{\psi, 1} = \omega_{\psi} \) in the symplectic case. Denote, for a Schwartz function \( \phi \) in a Schrodinger model of \( \omega_{\psi} \), by \( \theta_{E/F}^{\psi, \gamma} \) the corresponding theta series.

Let \( \pi \) be an automorphic representation of \( H_{\mathbb{A}} \), in case \( H \) is a unitary group, or of \( H^{1-\varepsilon}_{\mathbb{A}} \), in case \( H \) is symplectic.

Put, for \( \varphi_{\pi} \in V_{\pi} \)

\[
\varphi_{\pi}^{\psi, \gamma, \phi}(h) = \int_{N_{\ell + 1}(F) \backslash N_{\ell + 1}(\mathbb{A})} \varphi_{\pi}(v h) \psi_{\ell}^{-1}(v) \theta_{E/F}^{\psi, \gamma, \phi}(j(v) h) dv
\]

Recall that \( \psi_{\ell} \) is extended trivially to \( N_{\ell + 1} \) and \( j \) is the isomorphism \( N_{\ell} \backslash N_{\ell + 1} \xrightarrow{\sim} \mathcal{H}_{W'} \). (We keep denoting by \( j \) its composition with \( N_{\ell + 1} \rightarrow N_{\ell} \backslash N_{\ell + 1} \).) Note that \( \varphi_{\pi}^{\psi, \gamma, \phi} \) is called a Fourier-Jacobi coefficient of \( \varphi_{\pi} \) with respect to \( \omega_{\psi, \gamma} \) (and \( \phi \)). Let \( \sigma \) be an automorphic representation of \( H(W')_{\mathbb{A}} \) in case \( H \) is a unitary group, and of \( H(W')^{\varepsilon}_{\mathbb{A}} \) in case \( H \) is symplectic. We say that \( \pi \) has a global Fourier-Jacobi model with respect to \( (R_{\ell}; \psi, \gamma, \sigma) \) if (the following integral is absolutely convergent and)

\[
\int_{H(W')^\varepsilon_{\mathbb{A}} \backslash H(W')_{\mathbb{A}}} \varphi_{\pi}^{\psi, \gamma, \phi}(g) \varphi_{\sigma}(g) dg \neq 0,
\]

as \( \varphi_{\pi} \) and \( \varphi_{\sigma} \) vary in \( V_{\pi} \) and \( V_{\sigma} \) respectively. (In both cases, local, or global, representations of metaplectic covers are assumed to be genuine.) In practice, we will take one of \( (\pi, \sigma) \) to be cuspidal and the other to be “Eisensteinian”.

---

SOCIETÉ MATHÉMATIQUE DE FRANCE 2005
In the following remark, we relate the above models to degenerate Whittaker models, as formulated in [M.W.]. It is meant just for completeness sake, and may be skipped at a first reading.

**Remark.** — The equivariance properties with respect to $N_\ell$ or $N_\ell + 1$ of the models just introduced are special cases of the general set-up of degenerate Whittaker models. To relate to the terminology [M.W.], we have to choose a nilpotent element $f$ in Lie ($H$), and a one parameter subgroup $\varphi$ of $H$, such that

$$\text{Ad}(\varphi(t)) \cdot f = t^{-2} \cdot f, \quad \forall t \in F^*.$$  

We realize

$$\text{Lie}(H) = \{ A \in End_E(V) \mid (Av_1, v_2) + (v_1, Av_2) = 0, \quad \forall v_1, v_2 \in V \},$$

and write its elements in matrix form following (2.1). Consider again the rational character $\chi_{w_0,v_0}$ of $U_\ell$. Clearly, there is a unique element $f_1(w_0) \in \text{Hom}(V_\ell^+, W)$, such that

$$\chi_{w_0,v_0}(\exp(S)) = \text{tr}(f_1(w_0) \cdot S).$$

From (2.12) and (2.13) we have, for $S \in \text{Lie}(N_\ell)$,

$$\psi_{\ell,w_0}(\exp(S)) = \psi_E(\text{tr}(f_\ell,w_0 \cdot S)).$$

Next, we have to explain what was our choice of a one parameter subgroup $\varphi$ of $H$. Let

$$\varphi_\ell(t) = \begin{pmatrix} a_\ell(t) & I_W \\ I_{a_\ell(t)}^* & a_\ell(t) \end{pmatrix} \in H,$$

where

$$a_\ell(t) = \text{diag}(t^{2\ell}, t^{2\ell-2}, t^{2\ell-4}, \ldots, t^2).$$
If \( w_0 \) is anisotropic, we choose \( \varphi = \varphi_t \), and if \( w_0 \) is isotropic, we choose
\[
\varphi(t) = \left( \begin{array}{c} t a_i(t) \\ t I_{W'}^{-1} t^{-1}_i a_i(t)^* \end{array} \right).
\]
Note that (2.11) is satisfied. Now decompose
\[
\text{Lie}(H) = \bigoplus_i \text{Lie}(H)_i,
\]
where
\[
\text{Lie}(H)_i = \{ S \in \text{Lie}(H) \mid \text{Ad} \varphi(t) \cdot S = t^i S, \ \forall t \in F^* \}.
\]
Clearly, if \( w_0 \) is anisotropic, then
\[
\text{Lie}(N_t) = \bigoplus_{i \geq 2} \text{Lie}(H)_i = \bigoplus_{i \geq 1} \text{Lie}(H)_i.
\]
If \( w_0 \) is isotropic, then
\[
\text{Lie}(N_t \cdot \text{Center}(L_{W_0})) = \bigoplus_{i \geq 2} \text{Lie}(H)_i,
\]
and
\[
\text{Lie}(N_{t+1}) = \bigoplus_{i \geq 1} \text{Lie}(H)_i.
\]
(Note that \( N_t \cdot \text{Center}(L_{W_0}) = j^{-1}(\text{Center}(\mathcal{H}_{W'})) \), where \( j \) is the composition of \( N_{t+1} \to N_t \).)

### 2.3. The global integrals: overview

The general form of the global integrals is just an application of a global Gelfand-Graev model, or a global Fourier-Jacobi model to an Eisenstein series on \( H_\kappa \), or on \( H_\kappa^\lambda \), in case \( H \) is symplectic, induced from a cuspidal representation on a maximal parabolic subgroup of \( H \). The global model is taken against a cuspidal representation \( \sigma \) on \( H(w_0^+ \cap W)_\kappa \), in case \( w_0 \) is anisotropic, on \( H(W')_\kappa \), in case \( H \) is unitary, or on \( H(W')_\kappa \), in case \( H \) is symplectic. Thus, in (2.7) and in (2.9), \( \pi \) is an Eisenstein series induced from a cuspidal representation \( \tau \otimes \sigma_0 \) on a parabolic subgroup, whose Levi part is isomorphic to \( \text{GL}_k \times H(W_k) \), where \( V = V_1^+ + W_k + V_1^- \), as in (2.1). With normalized Eisenstein series, these integrals represent \( L^S(\sigma \times \tau, s) \), the partial standard \( L \)-function for \( H(w_0^+ \cap W) \times \text{GL}_k \), (resp. \( H(W') \times \text{Res}_{E/F} \text{GL}_k \), resp. \( H(W')^e \times \text{GL}_k \)) provided \( \sigma \) and \( \sigma_0 \) are related through an appropriate global Gelfand-Graev model (resp. Fourier-Jacobi model).

For example, if \( W_k \) is a subspace of \( w_0^+ \cap W \), in case \( w_0 \) is anisotropic, or a subspace of \( W' \), in case \( w_0 \) is isotropic, then \( \sigma \) should have a global model with respect to a subgroup \( R_{e', w_0'} \subset H(w_0^+ \cap W) \) (resp. \( H(W') \)), whose reductive part is isomorphic to \( H(W_k) \), on which we take \( \sigma_0 \). In this generality, the global integrals were studied in [G.P.S.R.] for orthogonal groups \( H \). Special cases were treated in [Ge.Ro.] (Fourier-Jacobi model for \( \pi \) cuspidal on \( U_{2,1} \)) and in [N] (Gelfand-Graev model for \( \pi \) cuspidal on \( \text{SO}_{3,2} \) (actually on \( \text{GSp}_4 \)). We will be interested here in the case where the bilinear form has maximal Witt index (i.e. \( \frac{1}{2} \dim_E V \)). Thus, if \( E = F \), \( H \) is split and if
$E : F = 2$, $H$ is the quasi-split unitary group in $\dim_E V$ variables. In this case, we will apply the above global models to $\pi$ – an Eisenstein series induced from the Siegel parabolic subgroup and a cuspidal representation $\tau$ on $\Res_{E/F} \GL_k$. We will choose $w_0$, (when anisotropic), such that $H(w_0^+ \cap W)$ is quasi-split or split. Again, with normalized Eisenstein series, the integrals (2.7) and (2.9) represent $L^2(\sigma \times \tau, s)$. These cases were studied in \cite{So1, So2, So3} ($H$ – even orthogonal) and in \cite{GRS3} ($H$ symplectic or metaplectic). The case where $H(w_0^+ \cap W)$ is of rank one less than $k$ was studied in \cite{G.PS} ($H$ orthogonal). The remaining cases are treated similarly and will appear in detail in future works. We will summarize them here. In these cases, except for $H = U_{2k+1}$, these integrals are identically zero, unless $\sigma$ has a global Whittaker model (with respect to an appropriate character). Finally, we also consider the cases where $\pi$ is cuspidal on $H$ (resp. on $H^{1-\varepsilon}$, when $H$ is symplectic) and $\sigma$ is an Eisenstein series on $H(w_0^+ \cap W)$, when $(w_0, w_0) \neq 0$ (resp. on $H(W')^{1-\varepsilon}$, when $H$ is symplectic, or $H(W')$, when $H$ is unitary and $w_0$ is isotropic). This Eisenstein series is induced from the Siegel parabolic subgroup and a cuspidal representation $\tau$ on $\Res_{E/F} \GL_n$, $n = \frac{1}{2}(\dim(w_0^+ \cap W))$, or $\frac{1}{2}\dim(W')$. Again, in these cases, except when $H(w_0^+ \cap W)$ or $H(W')$ are $U_{2n+1}$, the integrals (2.7), (2.9) are identically zero, unless $\pi$ has a global Whittaker model (with respect to an appropriate character) and then they represent $L^2(\pi \times \tau, s)$ once the Eisenstein series is normalized. These cases were studied in \cite{G} ($H$ – split orthogonal) and in \cite{GRS3} ($H$ symplectic or metaplectic). The cases where $k = n$ were studied in \cite{G.PS}, \cite{T}, \cite{W}. The remaining cases ($H$ unitary, $k > n$) are treated similarly and will appear in detail in future works. We will summarize them here.

From now on, we assume that $(,)$ has Witt index $\left\lfloor \frac{1}{2}\dim_E V \right\rfloor$. We will denote $r = \dim_E V$, and $H = H_r$. We realize $V$ as the column space $E^r$ and represent $(,)$ in terms of the matrix

$$
\begin{pmatrix}
1 & 1 \\
\varepsilon & . \\
& .
\end{pmatrix}
$$

where $\varepsilon = \pm 1$; $\varepsilon = -1$ is reserved just for symplectic groups. $H_r$ is realized as a matrix group. We denote by $P_r$ the Siegel parabolic subgroup of $H_r$. Its Levi part is isomorphic to $\Res_{E/F} \GL_{[r/2]}$.

### 2.4. The global integrals: Gelfand-Graev models

It remains to specify $w_0$. We do this in the following table. Write $^{t}w_0 = (0, ^{t}w_0', 0)$, where $0$ denotes a zero row vector in $\ell$ coordinates. Recall that for

$$
v = \begin{pmatrix} z & y & x \end{pmatrix}_{\ell \omega} \in N_{\ell} \quad (z \in Z_{\ell}),
$$

(2.14) \begin{equation}
\psi_{^{t}w_0}(v) = \psi_E(z)\psi_E(y_{\ell} \cdot w_0') = \psi_E\left(\sum_{i=1}^{\ell-1} z_{i,i+1} + y_{\ell} \cdot w_0'\right),
\end{equation}

where $y_{\ell}$ denotes the last row of $y$. In the following table (2.15) we indicate the choice of $^{t}w_0'$. We also write $\ell$ in terms of $m + 1 = \dim_E W$ and $r = \dim_E V$. 

---

*ASTÉRISQUE 298*
<table>
<thead>
<tr>
<th>$H = H_r$</th>
<th>$\dim_E W = m + 1$</th>
<th>$\ell$</th>
<th>$\ell w'_0$</th>
<th>$\psi_E(y_{\ell} \cdot w'_0)$</th>
<th>$H(w'_0 \cap W) \cong H^{(\alpha)}_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) $SO_{2k}$</td>
<td>$2n + 2$</td>
<td>$k - n - 1$</td>
<td>$(0, \ldots, 0, 1, -1, 0, \ldots, 0)$</td>
<td>$\psi(y_{\ell,n+1} - y_{\ell,n+2})$</td>
<td>$SO_{2n+1}$</td>
</tr>
<tr>
<td>(2) $U_{2k}$</td>
<td>$2n + 2$</td>
<td>$k - n - 1$</td>
<td>$(0, \ldots, 0, 1 - 1, 0, \ldots, 0)$</td>
<td>$\psi_E(y_{\ell,n+1} - y_{\ell,n+2})$</td>
<td>$U_{2n+1}$</td>
</tr>
<tr>
<td>(3) $SO_{2k+1}$</td>
<td>$2n + 1$</td>
<td>$k - n$</td>
<td>$(0, \ldots, 0, 1, 0, \alpha, 0, \ldots, 0)$</td>
<td>$\psi(y_{\ell,n} + \alpha y_{\ell,n+2})$</td>
<td>$SO_{2n}$</td>
</tr>
<tr>
<td>(3$'$) $SO_{2k+1}$</td>
<td>$2n + 1$</td>
<td>$k - n$</td>
<td>$(0, \ldots, 0, 1, 0, \ldots, 0)$</td>
<td>$\psi(y_{\ell,n+1})$</td>
<td>$SO_{2n}$</td>
</tr>
<tr>
<td>(4) $U_{2k+1}$</td>
<td>$2n + 1$</td>
<td>$k - n$</td>
<td>$(0, \ldots, 0, 1, 0, \ldots, 0)$</td>
<td>$\psi_E(y_{\ell,n+1})$</td>
<td>$U_{2n}$</td>
</tr>
</tbody>
</table>

Table (2.15)
Here, we also denote \( H_m^{(-1)} = H_m \), so that in all cases except (3), \( \alpha = -1 \). In case (3), \( \text{SO}^{(n)}_{2n} = H_{2n}^{(\alpha)} \), \( \alpha \in F^* \), denotes the quasi-split orthogonal group with respect to the symmetric form, whose matrix is

\[
\begin{pmatrix}
1 & 0 & w_{n-1} \\
0 & -2\alpha & \vdots \\
w_{n-1} & \vdots & 1
\end{pmatrix}
\]

Note that \( \text{SO}^{(n)}_{2n} \cong \text{SO}_{2n} \), if and only if \( 2\alpha \in (F^*)^2 \). (In this case we may replace Case (3) by Case (3').) We denote by \( \psi_{\ell,\alpha} \) the character (2.14). Let \( \tau \) be an irreducible, automorphic, cuspidal representation of \( \text{GL}_k(A_F) \) \( (k = \lfloor r/2 \rfloor) \). We consider now all cases except case (4). Denote

\[
\rho_{\tau,s}^H = \text{Ind}_{\mathcal{H}}^{H_r(K_F)} \tau |_{K_F} \det \|t\|^{-1/2}.
\]

Let \( \xi_{\tau,s} \) be a holomorphic \( K \)-finite section for \( \rho_{\tau,s} \) and denote by \( E_{H_r}(\xi_{\tau,s}, h) \) the corresponding Eisenstein series on \( H_r(K_F) \). Let \( \sigma \) be an irreducible, automorphic, cuspidal representation of \( H_m^{(\alpha)}(A_F) \) \( (\alpha = -1 \) for all cases except (3)). Fix an \( F \)-isomorphism \( H_m^{(\alpha)} \sim H(w_0^1 \cap W) \), and denote by \( i_{m,r} \) its composition with the inclusion \( H(w_0^1 \cap W) \hookrightarrow H_r \). Define, for a cusp form \( \varphi_\sigma \) in the space of \( \sigma \),

\[
\mathcal{L}(\varphi_\sigma, \xi_{\tau,s}) = \int_{H_m^{(\alpha)}(F) \setminus H_m^{(\alpha)}(A_F)} \varphi_\sigma(g) E_{H_r}^{\psi_{\ell,s}}(\xi_{\tau,s}, i_{m,r}(g)) dg.
\]

These integrals converge absolutely and are meromorphic in \( s \). For \( \text{Re}(s) \) large enough, the integral (2.16) equals an Eulerian integral which depends on the \( \psi \)-Whittaker coefficient of \( \varphi_\sigma \). [For example, for \( H_r = U_{2k} \) \( (H_m^{(\alpha)} = U_{2n+1}) \) and \( \text{Re}(s) \gg 0 \),

\[
\mathcal{L}(\varphi_\sigma, \xi_{\tau,s}) = \int_{N_F \setminus U_{2n+1}(A_F)} \int_{M_{E,(n+1)}(A_F) \times h_{E}(A_F)} \xi_{\tau,s}^{-1} \beta_{k,n} \left( \begin{array}{ccc}
I_x & 0 & y \\
I_{n+1} & 0 & 0 \\
I_{n+1} & 0 & I_x
\end{array} \right) \psi_E(x_{\ell,n+1}) d(x,e) dg
\]

where \( N \) is the standard maximal unipotent subgroup of \( U_{2n+1} \),

\[
W_{\varphi_\sigma}^\psi(g) = \int_{N_F \setminus N_{A_F}} \varphi_\sigma(ug) \psi_N^{-1}(u) du
\]

is the \( \psi \)-Whittaker function of \( \varphi_\sigma \) \( (\psi_N(u) = \psi_E(\sum_{i=1}^m u_{i,i+1})) \); \( \xi_{\tau,s}^{-1}(h) \) is the composition of \( \xi_{\tau,s} \) with the \( \psi^{-1} \)-Whittaker coefficient on \( \tau \), i.e.

\[
\xi_{\tau,s}^{-1}(h) = \int_{Z_k(E) \setminus Z_k(A_E)} \xi_{\tau,s}(\left( \begin{array}{ccc}
z & 0 \\
\tilde{u} & z
\end{array} \right) h) \psi_E(z) dz;
\]
\[ \beta_{k,n} \text{ is the Weyl element} \]
\[
\begin{pmatrix}
0 & I_{n+1} & 0 & 0 \\
0 & 0 & 0 & I_{l} \\
I_{l} & 0 & 0 & 0 \\
0 & 0 & I_{n+1} & 0
\end{pmatrix}
\]

and \( h_{\ell} = \{ A \in M_{\ell}(E) \mid \left( \overline{Aw_{\ell}} \right) + (Aw_{\ell}) = 0 \} \). The integrals (2.16) are identically zero, unless the \( \psi \)-Whittaker coefficient of \( \varphi_{\sigma} \) is nontrivial as a function on \( H_{m}^{(\alpha)}(\mathbb{A}_F) \). Thus \( \sigma \) has to be globally \( \psi \)-generic. (This is not the condition one gets in case (4). This is why we exclude it now.) Assume then that \( \sigma \) is \( \psi \)-generic. For decomposable data, the Eulerian integral of (2.16) has the form

\[ L(\varphi_{\sigma}, \xi_{\tau,s}) = R(s) \frac{L^{S}(\sigma \times \tau, s)}{L^{S}(\tau, \delta, 2s)}. \]  

Here \( S \) is a finite set of places of \( F \), including the ones at infinity, outside which \( \sigma, \tau \) and the components of \( \varphi_{\sigma}, \xi_{\tau,s} \) are unramified. \( R(s) \) is a finite product of “local integrals” (over \( S \)), where data can be chosen so that \( R(s) \) is holomorphic and nonzero at a neighborhood of a given point \( s_0 \). \( L^{S}(\tau, \delta, z) \) is the partial \( L \)-function which enters in the normalizing factor of \( E_{H_r}(\xi_{\tau,s}, h) \). Let us summarize this in the following table

<table>
<thead>
<tr>
<th>( L^{S}(\sigma \times \tau, s) ) for the group</th>
<th>( L^{S}(\tau, \delta, 2s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{SO}_{2n+1} \times \text{GL}_k )</td>
<td>( k &gt; n )</td>
</tr>
<tr>
<td>( \text{U}<em>{2n+1} \times \text{Res}</em>{E/F}(\text{GL}_k) )</td>
<td>( k &gt; n )</td>
</tr>
<tr>
<td>( \text{SO}_{2n}^{(\alpha)} \times \text{GL}_k )</td>
<td>( k \geq n )</td>
</tr>
</tbody>
</table>

(2.19)

Next, we may take a cusp form on \( H \) and an Eisenstein series on \( H_{m}^{(\alpha)} \). We go back to table (2.15) and assume now that case (2) is excluded, and also in case (3) we consider \( \alpha = -1 \) (and so we may replace (3) by (3')). Let \( \sigma \) be an irreducible, automorphic, cuspidal representation of \( H_r(\mathbb{A}) \). Let \( \tau \) be an irreducible, automorphic, cuspidal representation of \( \text{GL}_n(\mathbb{A}_E) \), and consider the Eisenstein series \( E_{H_m}(\xi_{\tau,s}, g) \) corresponding to a \( K \)-finite holomorphic section \( \xi_{\tau,s} \) for \( \rho_{H_{m}^{(\alpha)}}^{H_r} \). Define, for a cusp form \( \varphi_{\sigma} \) in the space of \( \sigma \)

\[ L(\varphi_{\sigma}, \xi_{\tau,s}) = \int_{H_m(F) \backslash H_m(\mathbb{A}_F)} \varphi_{\sigma}^{\psi_{H_{m}^{(\alpha)}}^{-1}}(i_{m,r}(g)) E_{H_m}(\xi_{\tau,s}, g) dg. \]  

As before, for \( \text{Re}(s) \) large enough, the integral (2.20) equals an Eulerian integral which depends on the \( \psi \)-Whittaker coefficient of \( \varphi_{\sigma} \). [For example, for \( H_r = U_{2k+1} \)
\((H_m = U_{2n})\) and \(\text{Re}(s) \gg 0\)

\[
(2.21) \quad \mathcal{L}(\varphi_\sigma, \xi_{\tau,s}) = \int_{N_{k,F} \setminus U_{2n}(A_F)} \int_{M_{(k-n) \times n}(k,F)} W^\psi_{\varphi_\sigma} \left( \left( \begin{array}{c|c} \mathbb{I}_n & I_{k-n} \\ \hline & \mathbb{I}_n \end{array} \right) \right) \cdot \hat{w}_{n,k+2n,2k+1}(g) \xi^{\psi^{-1}}_{\tau,s}(g) dg,
\]

where, for \(g \in \text{GL}_k(A_E)\), we denote \(\hat{g} = \left( \begin{array}{c} 1 \\ g^* \end{array} \right)\), \(w_{n,k} = \left( \begin{array}{c} I_{n} \\ I_{k-n} \end{array} \right)\). \(N\) denotes the standard maximal unipotent subgroup of \(U_{2n}\). The remaining notation is as in (2.17).

As before, \(\mathcal{L}(\varphi_\sigma, \xi_{\tau,s})\) is identically zero, unless the \(\psi\)-Whittaker coefficient of \(\varphi_\sigma\) is non-trivial (as a function on \(H_r(A_F)\)). Thus, assume that \(\sigma\) is globally \(\psi\)-generic, and then for decomposable data the Eulerian integral of (2.10) has the form

\[
(2.22) \quad \mathcal{L}(\varphi_\sigma, \xi_{\tau,s}) = R(s)^{L^S(\sigma \times \tau, s)} L^S(\tau, \delta, 2s),
\]

as in (2.18). \(L^S(\tau, \delta, 2s)\) is given by (2.19), switching roles of \(k\) and \(n\). More precisely

<table>
<thead>
<tr>
<th>(L^S(\sigma \times \tau, s)) for the group</th>
<th>(L^S(\tau, \delta, 2s))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{SO}_{2k+1} \times \text{GL}_n,) (k \geq n)</td>
<td>(L^S(\tau, \Lambda^2, 2s))</td>
</tr>
<tr>
<td>(U_{2k+1} \times \text{Res}_{E/F}(\text{GL}_n),) (k \geq 2n)</td>
<td>(L^S(\tau, \text{Asai}, 2s))</td>
</tr>
<tr>
<td>(\text{SO}_2 \times \text{GL}_n,) (k &gt; n)</td>
<td>(L^S(\tau, \text{sym}^2, 2s))</td>
</tr>
</tbody>
</table>

\[
(2.23)
\]

### 2.5. The global integrals: Fourier-Jacobi models.

We use the notation of Sec. 2.2, where \(w_0\) was already chosen. We will denote by \(H(W')\) the group \(H(W')\) in case \(H\) is a unitary group, and in case \(H\) is symplectic \(H(W') = H(W')^{\xi}, \xi = 0, 1\).

The cases we consider appear in the following table (\(r = \dim_E V\))

<table>
<thead>
<tr>
<th>(H = H_r)</th>
<th>(\dim_E W = m + 2)</th>
<th>(\ell)</th>
<th>(H(W') \simeq H_m)</th>
<th>(H_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) (\text{Sp}_{2k})</td>
<td>(2n + 2)</td>
<td>(k - n - 1)</td>
<td>(\text{Sp}_{2n})</td>
<td>(\text{Sp}_{2n})</td>
</tr>
<tr>
<td>(2) (\text{Sp}_{2k})</td>
<td>(2n + 2)</td>
<td>(k - n - 1)</td>
<td>(\text{Sp}_{2n})</td>
<td>(\text{Sp}_{2n})</td>
</tr>
<tr>
<td>(3) (U_{2k})</td>
<td>(2n + 2)</td>
<td>(k - n - 1)</td>
<td>(U_{2n})</td>
<td>(U_{2n})</td>
</tr>
</tbody>
</table>

\[
(2.24)
\]

Let \(\tau\) be an irreducible, automorphic, cuspidal representation of \(\text{GL}_k(A_E)\). Denote by \(\rho_{\tau, r, 2k}\) the representation of \(H_{2k}(\text{A}_E)^{\sim}\) induced from \(\tau|\det|^{s-1/2}\) on the Siegel parabolic subgroup in cases (1), (3). In case (2) we replace \(\tau\) by \(\gamma_{\psi} \cdot \tau\), where \(\gamma_{\psi}\) is the Weil factor. Let \(\xi_{\tau,s}\) be a \(K\)-finite holomorphic section for \(\rho_{\tau,s}\), and let \(E_{H_{2k}}(\xi_{\tau,s}, h)\) be the corresponding Eisenstein series. Let \(\sigma\) be an irreducible, automorphic, cuspidal representation of \(H_{2n}(\text{A}_F)\). Fix an \(F\)-embedding \(j_{2n,2k} : H_{2n} \to H_{2k}\), so that the
image of \( j_{2n,2k} \) is \( H(W) \). Define for a cusp form \( \varphi_\sigma \) in the space of \( \sigma \)
\[
(2.25) \quad \mathcal{L}(\varphi_\sigma, \phi, \xi) = \int_{H_{2n}(F)/H_{2n}(\mathbb{A}_F)} \varphi_\sigma(g) E_{H_{2n}}^{\psi, \gamma, \varphi}(\xi, s, j_{2n,2k}(g)) \, dg.
\]

As before, this integral equals an Eulerian integral, for \( \text{Re}(s) \gg 0 \), and it depends on the \( \psi \)-Whittaker coefficient of \( \varphi_\sigma \). [For example, for \( H_{2k} = U_{2k} \) \( (H_{2n} = U_{2n}) \), we get, for \( \text{Re}(s) \gg 0 \),
\[
(2.26) \quad \mathcal{L}(\varphi_\sigma, \phi, \xi) = \int_{N_{\mathbb{A}_F} \setminus U_{2n}(\mathbb{A}_F)} \int_{M_{(k-n) \times n}(\mathbb{A}_E) \times h_{k-n}(\mathbb{A}_F)} W^\psi_{\varphi_\sigma}(g) \xi^{-1}_{\tau, s} \left( \alpha_{k,n} \begin{pmatrix} I_{k-n} & x & 0 & y \\ I_n & 0 & 0 \\ I_n & 0 \\ l_{k-n} \end{pmatrix} \right) j_{2n,2k}(g) \, \omega_{\psi, \gamma, \varepsilon}(x_{k-n}, 0; \text{Im}(y_{k-n})g) \phi(\varepsilon_n) \, dx \, dy \, dg.
\]

Here we assume that \( E = F[\sqrt{b}] \), and \( W^\psi_{\varphi_\sigma}(g) \) is the Whittaker coefficient of \( \varphi_\sigma \) (at \( g \)) with respect to the non-degenerate character of \( N_{\mathbb{A}_F} \setminus \mathbb{A}_F \) given by
\[
u \mapsto \psi_E \left( u_{12} + u_{23} + \cdots + u_{n-1,n} + \frac{u_{n,n+1}}{2\sqrt{b}} \right); \quad \alpha_{k,n} = \begin{pmatrix} 0 & \frac{1}{\sqrt{n}} I_n & 0 & 0 \\ 0 & 0 & 0 & I_{k-n} \\ 0 & 0 \\ 0 & 0 & -\sqrt{n} & 0 \end{pmatrix}.
\]

We regard \( E^{2n} \) as a symplectic space over \( F \) with respect to the form \( (v_1, v_2) = -2 \text{Im}(v_1, v_2) \). (Im denotes the “imaginary” part: \( \text{Im}(a + \sqrt{b}) = b \).) This defines \( \omega_{\psi, \gamma, \varepsilon} \), realized in \( S(\mathbb{A}_E) \). Finally, \( \phi \in S(\mathbb{A}_E) \) and \( \varepsilon_n = (0, \ldots, 0, 1) \). The rest of the notation is as in (2.17).]

Thus, assume that \( \sigma \) is globally \( \psi \)-generic. For decomposable data the Eulerian integral of (2.25) has the form

\[
(2.27) \quad \mathcal{L}(\varphi_\sigma, \phi, \xi) = R(s) L^S(\sigma, \tau, s)
\]

where \( L^S(\sigma, \tau, s) \) is given by the following table

<table>
<thead>
<tr>
<th>( \sigma ) on ( H_{2n}^{\infty} )</th>
<th>( \tau ) on ( H_{2k}^{\infty} )</th>
<th>Eisenstein series ( E(\xi, s) ) on ( H_{2n}^{\infty} )</th>
<th>( L^S(\sigma, \tau, s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \widetilde{Sp}_{2n} )</td>
<td>( GL_k )</td>
<td>( \text{Sp}_{2k} )</td>
<td>( L^S(\sigma \times \tau, s) )</td>
</tr>
<tr>
<td>( \text{Sp}_{2n} )</td>
<td>( GL_k )</td>
<td>( \text{Sp}_{2k} )</td>
<td>( L^S(\sigma \times \tau, s) )</td>
</tr>
<tr>
<td>( U_{2n} )</td>
<td>( \text{Res}_{E/F} GL_k )</td>
<td>( U_{2k} )</td>
<td>( L^S(\sigma \times \gamma^{-1}, s) )</td>
</tr>
</tbody>
</table>

(2.27) \( (k > n) \)

In case (1) there is no canonical way to attach an \( L \)-function to \( \sigma \times \tau \). At places \( \nu \) where \( \sigma \) is unramified (and \( \psi \) normalized) we write the unramified characters corresponding to \( \sigma_\nu \) in the form \( \gamma_\psi \cdot \mu_\nu \), where \( \mu_\nu \) is an unramified character of \( F_\nu^* \). We write the parameter of \( \sigma_\nu \) as a conjugacy class in \( \text{Sp}_{2n}(\mathbb{C}) \) (constructed from the \( \mu_\nu(p_\nu)\pm 1 \)). Another choice \( \gamma_\psi \cdot \mu_\nu \) would yield a different conjugacy class. This explains the dependence on \( \psi \) in \( L^S(\sigma \times \tau, s) \). The function \( R(s) \) in (2.26) can be chosen to have the same properties as in (2.18), (2.22).
Finally, as in the previous case (Gelfand-Graev models) we may reverse the roles of \( H_{r}^{+} \) and \( H_{r}^{-} \). We go back to table (2.24) and consider now an irreducible, automorphic, cuspidal representation \( \sigma \) of \( H_{r}^{-}(\mathbb{A}_{F}) \) and an irreducible, automorphic, cuspidal representation \( \tau \) of \( \text{GL}_{n}(\mathbb{A}_{F}) \). Consider the Eisenstein series \( E_{H_{2n}^{+}}(\xi_{r,s}, g) \) on \( H_{2n}^{+}(\mathbb{A}_{F}) \) corresponding to a holomorphic \( K \)-finite section \( \xi_{r,s} \) for \( \rho_{r,s}^{+} \). Define for a cuspidal form \( \varphi_{\sigma} \) in the space of \( \rho_{r,s}^{+} \),

\[
(2.28) \quad \mathcal{L}(\varphi_{\sigma}, \phi, \xi_{r,s}) = \int_{H_{2n}(F) \backslash H_{2n}(\mathbb{A}_{F})} \varphi_{\sigma}^{\psi_{r}, \gamma_{r}}(j_{2n,r}(g)) E_{H_{2n}^{+}}(\xi_{r,s}, g) dg.
\]

Again, for \( \text{Re}(s) \) large, the integral (2.28) equals an Eulerian integral which depends on the \( \psi \)-Whittaker function of \( \varphi_{\sigma} \). [For example, for \( H_{r}^{+} = U_{2k} \) \( (H_{2n}^{+} = U_{2n}) \) and \( \sigma \) on \( U_{2k}(\mathbb{A}_{F}) \), and \( \tau \) on \( \text{GL}_{n}(\mathbb{A}_{F}) \), we get for \( \text{Re}(s) \gg 0 \)

\[
(2.29) \quad \mathcal{L}(\varphi_{\sigma}, \phi, \xi_{r,s}) = \int_{N_{k} \backslash U_{2n}(\mathbb{A}_{F})} \int_{M(\mathbb{A})} W_{\varphi_{\sigma}}^{\psi}(\begin{pmatrix} I_{n} \xi_{k-n} \\ x \end{pmatrix}) \hat{w}_{n,k} j_{2n,r}(g) \omega_{\psi^{-1}, \gamma^{-1}}(g) \phi(x_{k-n}) \xi_{r,s}^{-1}(g) dx dg
\]

Here \( W_{\varphi_{\sigma}}^{\psi} \) is as in (2.16). For \( g \in \text{Res}_{E/F} \text{GL}_{k} \), we denote \( \hat{g} = \begin{pmatrix} g & g_{*} \\ g & g_{*} \end{pmatrix} \in U_{2k} \). For \( x \in M(\mathbb{A}) \times n \), \( x_{k-n} \) denotes the last row of \( x \); \( w_{n,k} = \begin{pmatrix} I_{k-n} \\ I_{n} \end{pmatrix} \). The rest of the notation is as before.]

Assume that \( \sigma \) is globally \( \psi \)-generic. Then for decomposable data

\[
(2.30) \quad \mathcal{L}(\varphi_{\sigma}, \phi, \xi_{r,s}) = R(s) \mathcal{L}^{S}(\sigma, \tau, s).
\]

\( \mathcal{L}^{S}(\sigma, \tau, s) \) is given by the last column of table (2.27) (where we switch the roles of \( E(\xi_{r,s}, \cdot) \) and \( \sigma \)).

In (2.25), (2.28) the case \( k = n \) is missing. Here, for a \( \psi \)-generic cuspidal representation \( \sigma \) on \( H_{2n}^{+}(\mathbb{A}_{F}) \) and a cuspidal representation \( \tau \) on \( \text{GL}_{n}(\mathbb{A}_{F}) \), we consider

\[
(2.31) \quad \mathcal{L}(\varphi_{\sigma}, \phi, \xi_{r,s}) = \int_{H_{2n}(F) \backslash H_{2n}(\mathbb{A}_{F})} \varphi_{\sigma}(g) \theta_{\psi_{r}, \gamma_{r}}^{\phi_{\sigma}}(g) E_{H_{2n}^{+}}(\xi_{r,s}, g) dg,
\]

where, as before, for \( H_{2n} = U_{2n}, H_{2n}^{+} = U_{2n} \), and for \( H_{2n} = \text{Sp}_{2n} \), if \( \sigma \) is on \( H_{2n}^{+} \), then the Eisenstein series is on \( H_{2n}^{1-\varepsilon} \), \( \varepsilon = 0, 1 \). For \( \text{Re}(s) \gg 0 \), we obtain as in the previous cases (for decomposable data)

\[
\mathcal{L}(\varphi_{\sigma}, \phi, \xi_{r,s}) = R(s) \mathcal{L}^{S}(\sigma, \tau, s),
\]

as in the last two cases.
3. On the weak lift from a quasi-split classical group to GL$_N$.

We construct examples of cuspidal generic representations on a given quasi-split classical group $G$, which weakly lift to automorphic representations on GL$_N$ (appropriate $N$) in the expected image of this lift. The methods are those of Chapter 1, constructing a descent map (backward lift), as suggested by the global integrals reviewed in Chapter 2. We use the notation of Chapter 2.

3.1. The cuspidal part of the image of the weak lift from $G$ to GL$_N$

Let $G$ be a group of the form $H(w_0^\perp \cap W)$ or $H(W')^\perp$, as in table (2.15) (without case (4)), or table (2.24). (For the moment dim$_E \mathcal V$ is not so important.) Let $N$ be the degree of the standard representation of $^L G^0$. The Langlands conjectures predict the existence of a functorial lift from irreducible, automorphic, cuspidal representations of $G_{k,F}$ to irreducible automorphic representations of GL$_N(\mathbb A_F)$. Let $\sigma \cong \otimes \nu$, be such a representation, and assume that $\sigma$ has a weak lift to an irreducible automorphic representation $\tau$ of GL$_N(\mathbb A_F)$, where the notion of a weak lift is similar to the one explained in Sec. 1.1. It is clear that $\tau^*_\nu \cong \nu$, and $\omega_{\tau^*}|_{F^*_\nu} = 1$, for almost all $\nu$, except in case (3) of (2.24), when $2\alpha$ is not a square in $F^*_\nu$, in which case, $\omega_{\tau^*}$ is the quadratic character associated to $2\alpha$. Here $\tau^*_\nu = \tilde{\tau}^*_\nu$ and $\tau^*_\nu'$ is the composition of $\tau^*_\nu$ with the automorphism $x \mapsto \mathfrak F$ of $E\nu$ over $F\nu$. (If $E = F$, then $\mathfrak F = x$, and $\tilde{\tau}^*_\nu = \tilde{\tau}_\nu$.) We conclude that $\omega_{\tau}|_{\mathbb A_F^*} = 1$, except in case (3) of (2.24), when $2\alpha$ is not a square, in which case $\omega_{\tau}$ is the quadratic character, associated to $2\alpha$. Let us assume that $\tau$ is cuspidal. Then by the strong multiplicity one and multiplicity one properties for GL$_N$, we conclude that $\tau^* = \tau$, and we also have that $L^S(\sigma \times \tilde{\tau},s) = L^S(\tilde{\tau} \times \tilde{\tau},s)$ has a simple pole at $s = 1$, for an appropriate finite set of places $S$. (In case $G$ is metaplectic, we have to fix $\psi$, a nontrivial character of $F^\perp \mathbb A_F$ and consider $L^S_\psi(\sigma \times \tilde{\tau},s)$ instead.) Assume further that $\sigma$ is globally $\psi$-generic. Then we can use the global integrals of Sections 2.4, 2.5 to represent the partial $L$-function of $\sigma$ twisted by $\tilde{\tau}$, and consider its pole at $s = 1$. Let $H$ be the group in the first column of (2.15) or (2.24), which has a Siegel parabolic subgroup whose Levi part is isomorphic to GL$_N$. Now consider the integrals (2.16) or (2.25) which represent the above $L$-function. Note that if $G$ is not a unitary group, then $\tilde{\tau} = \tau$, and we take the Eisenstein series on $H_{k,F}$ corresponding to $\rho^H_{\tau,s}$. If $G = U_{2n+1}$, $\tilde{\tau} = \tau'$ and we take $\rho^{U_{2n+1}}_{\tau',s}$. If $G = U_{2n}$, we take $\rho^{U_{2n}}_{\tau\otimes\gamma,s}$. For decomposable data the integrals above are of the forms (2.18) or (2.27) respectively, and we can choose $R(s)$ to be holomorphic and nonzero at $s = 1$. Looking at the quotients (2.18) in table (2.19) and in table (2.27), we see that the denominators are holomorphic and nonzero at $s = 1$. Since $L^S(\sigma \times \tilde{\tau},s)$ (resp. $L^S_\nu(\sigma \times \tau,s)$ if $G$ is metaplectic) has a pole at $s = 1$, we conclude that the global integral $L(\phi_{\tau,s})$ in (2.18), $L(\phi_{\tau,s})$ in (2.27), cases (1), (2), and $L(\phi_{\tau,s})$ in (2.27), case 3 has a pole at $s = 1$. This pole then comes from the Eisenstein series which appears in $L(\phi_{\tau,s},...)$ Therefore, we expect that the (partial) $L$-function $L^S(\tau,\beta,s)$ which
appears in the normalizing factor of this Eisenstein series to have a pole at \( s = 1 \).

The following table summarizes the various cases, when we take \( N = N_{2n} \). (In table (3.1), \( N_k = k \) in cases (1), (2), (4), (5), and \( N_k = k + 1 \) in cases (3), (6).)

<table>
<thead>
<tr>
<th>( G )</th>
<th>( \operatorname{Res}<em>{E/F} ) ( \operatorname{GL}</em>{N_k} )</th>
<th>( H = H_{G,k} )</th>
<th>( L^S(\tau, \beta, s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) ( \text{SO}_{2n+1} )</td>
<td>( \operatorname{GL}_k )</td>
<td>( \text{SO}_{2k} )</td>
<td>( L^S(\tau, \Lambda^2, 2s - 1) )</td>
</tr>
<tr>
<td>(2) ( \text{SO}_{2n} )</td>
<td>( \operatorname{GL}_k )</td>
<td>( \text{SO}_{2k+1} )</td>
<td>( L^S(\tau, \text{sym}^2, 2s - 1) )</td>
</tr>
<tr>
<td>(3) ( U_{2n+1} )</td>
<td>( \operatorname{Res}<em>{E/F} ) ( \text{GL}</em>{k+1} )</td>
<td>( U_{2k+2} )</td>
<td>( L^S(\tau', \text{Asai}, 2s - 1) )</td>
</tr>
<tr>
<td>(4) ( \text{Sp}_{2n} )</td>
<td>( \operatorname{GL}_k )</td>
<td>( \text{Sp}_{2k} )</td>
<td>( L^S(\tau, s - \frac{1}{2}) L^S(\tau, \Lambda^2, 2s - 1) )</td>
</tr>
<tr>
<td>(5) ( U_{2n} )</td>
<td>( \operatorname{Res}_{E/F} ) ( \text{GL}_k )</td>
<td>( U_{2k} )</td>
<td>( L^S(\tau \otimes \gamma', \text{Asai}, 2s - 1) )</td>
</tr>
<tr>
<td>(6) ( \text{Sp}_{2n} )</td>
<td>( \text{GL}_{k+1} )</td>
<td>( \text{Sp}_{2k+2} )</td>
<td>( L^S(\tau, \text{sym}^2, 2s - 1) )</td>
</tr>
</tbody>
</table>

Case (2) in table (3.1) includes both split and quasi-split even orthogonal groups.

We now proceed exactly as in case (1), which was proved in Theorem 1. The constant term of the Eisenstein series mentioned before, evaluated at \( I \), is the sum of the section evaluated at \( I \) and the corresponding intertwining operator, applied to the section, and evaluated at \( I \). The first summand is holomorphic, and hence the pole at \( s = 1 \) occurs for the second summand, which for decomposable data, equals as in (1.3) to a finite product, over a finite set of places \( S \) of local intertwining operators times a quotient of of the form \( \frac{L^S(\tau, \beta, s)}{L^S(\tau, \beta, s + 1)} \), except in case (4) of table (3.1) (\( \beta = \beta_{H,G,2n} \)), where it is \( \frac{L^S(\tau, s - \frac{1}{2}) L^S(\tau, \Lambda^2, 2s - 1)}{L^S(\tau, \Lambda^2, 2s)} \). In all cases, it is easy to see that the denominator of the last quotient is holomorphic and nonzero at \( s = 1 \). By [K, Lemma 2.4], the local intertwining operators above are holomorphic and nonzero for \( \text{Re}(s) \geq 1 \). (Note that the standard module conjecture needed in loc. cit. is needed here just for \( (\operatorname{Res}_{E/F} \text{GL}_{2n})(F_v) \) or \( (\operatorname{Res}_{E/F} \text{GL}_{2n+1})(F_v) \), and hence is valid.) We conclude that \( L^S(\tau, \beta, s) \) has a pole at \( s = 1 \). Summarizing

**Theorem 10.** — Let \( \sigma \) be an irreducible, automorphic, cuspidal representation of \( G_{k_E} \). Assume that \( \sigma \) is globally \( \psi \)-generic, and that \( \sigma \) has a weak lift to an irreducible, automorphic, cuspidal representation \( \tau \) of \( \text{GL}_{N_{2n}}(\mathbb{A}_E) \), as in table (3.1). Then \( \tau^* = \tau \), the partial L-function \( L^S(\tau, \beta_{H,G,2n}, s) \) has a pole at \( s = 1 \), and \( \omega_{\tau} \mid_{\mathbb{A}_E^*} = 1 \), except in case \( G = \text{SO}^\circ_{2n} \), when \( 2\alpha \) is not a square, in which case \( \omega_{\tau} \) is the quadratic character associated to \( 2\alpha \).

We conclude in exactly the same way, using the global integrals of Sec. 2.4, 2.5, the analogs of Proposition 2 and Theorem 3.

**Theorem 11.** — Let \( \sigma \) be an irreducible, automorphic, cuspidal representation of \( G_{k_E} \). Assume that \( \sigma \) is globally \( \psi \)-generic. Let \( \tau \) be an irreducible, automorphic, cuspidal representation of \( \text{GL}_k(\mathbb{A}_E) \), \( k \geq 2 \), such that \( \omega_{\tau} \mid_{(\mathbb{A}_E)^*} = 1 \). Then \( L^S(\sigma \times \tau, s) \)
(resp. $L^S_{\psi}(\sigma \times \tau, s)$ if $G$ is metaplectic) is holomorphic for $\text{Re}(s) > 1$ and if it has a pole at $s_0$, such that $\text{Re}(s_0) = 1$, then $s_0 = 1$ and $L^S(\tau, \beta_H, s)$ (table 3.1) has a pole at $s = 1$. The same assertions hold true, if $\tau$ is an automorphic unitary character of the idele group, which is trivial on $(\mathbb{A}_F)^+_\infty$, except in cases (1), (4). In case (1), we know that $L^S(\sigma \times \tau, s)$ is entire, and in case (4), the $L$ function (with respect to $\psi$, where $\sigma$ is globally $\psi$-generic) may have a pole for $\text{Re}(s) > 1$, and then it must be at $s = 3/2$, and $\tau$ must be trivial.

We remark that the last case of Theorem 11 occurs when $\sigma$ is a theta lift with respect to $\psi$ from a generic cuspidal representation of $\text{SO}_{2n-1}(\mathbb{A}_E)$.

Start now with an irreducible, automorphic, cuspidal representation $\tau$ of $\text{GL}_N(\mathbb{A}_E)$, $(N = N_{2n})$ such that $\omega_\tau|_{\mathbb{A}_F^*} = 1$, except in case $\tau$ is on $\text{GL}_{2n}(\mathbb{A}_F)$, where we allow $\omega_\tau$ to be either trivial or quadratic. If the quadratic character is associated to $2\tau$, then in the following, the Gelfand-Graev coefficient is taken with respect to $\text{SO}_{2n}$.

Assume that $L^S(\tau, \beta, s)(\beta = \beta_{H, 2n})$ has a pole at $s = 1$ in case (5) (as data vary). Thus, the Gelfand-Graev coefficient (resp. the Fourier-Jacobi coefficient) of the residue at $s = 1$ must be either trivial or quadratic. If the quadratic character is associated to $2\tau$, then we use the notation of table 3.1, with $\sigma = \text{SO}_{2n-1}(\mathbb{A}_E)$.

By Theorem 10, these are necessary conditions that (cuspidal) $\tau$ needs to satisfy in order to be in the image of the weak lift from generic cuspidal representations on $G_{\mathbb{A}_F}$. The second condition implies $\tau^* = \tau$. If $\tau$ is a weak lift of $\sigma$ (generic, cuspidal) on $G_{\mathbb{A}_F}$, then by (2.18), (2.27), $\mathcal{L}(\varphi, \xi, \xi_\tau, s)$ has a pole at $s = 1$ in cases (1)–(3) of Table (3.1), $\mathcal{L}(\varphi, \phi, \xi, s)$ has a pole at $s = 1$ in cases (4),(6), and $\mathcal{L}(\varphi, \phi, \xi_{\tau} \otimes \gamma, s)$ has a pole at $s = 1$ in case (5) (as data vary). Thus, the Gelfand-Graev coefficient (resp. the Fourier-Jacobi coefficient) of the residue at $s = 1$ of the Eisenstein series which appear in the global integrals has a non-trivial $L^2(G_F \setminus G_{\mathbb{A}_F})$-pairing against $\sigma$. This leads us to define

$$
\sigma_{\psi}(\tau) = \begin{cases} 
\text{Span}\{\text{Res}_{s=1} E_{H_{n+1}}^{v_{n+1}}(\xi, s, \cdot)|_{G_{\mathbb{A}_F}}\}, & G = \text{SO}_{2n+1} \\
\text{Span}\{\text{Res}_{s=1} E_{H_{2n+1}}^{v_{2n+1}}(\xi, s, \cdot)|_{G_{\mathbb{A}_F}}\}, & G = \text{SO}_{2n}, U_{2n+1} \\
\text{Span}\{\text{Res}_{s=1} E_{H_{2n}}^{v_{2n}}(\xi, \gamma, \phi)|_{G_{\mathbb{A}_F}}\}, & G = \text{Sp}_{2n}, U_{2n} \\
\text{Span}\{\text{Res}_{s=1} E_{H_{2n}}^{v_{2n}}(\xi, \cdot)|_{G_{\mathbb{A}_F}}\}, & G = \text{Sp}_{2n} 
\end{cases}
$$

(3.2)

Our main theorem is

**Theorem 12.** — Let $\tau$ be an irreducible, automorphic, cuspidal representation of $\text{GL}_N(\mathbb{A}_E)$, with central character, as above. Assume that $L^S(\tau, \beta, s)$ has a pole at $s = 1$. (We use the notation of table 3.1, with $N = N_{2n}, \beta = \beta_{H, 2n}$). Assume also that $n \geq 2$, in case $G = \text{SO}_{n,n}$. Then

1. $\sigma_{\psi}(\tau) \neq 0$
2. The representation $\sigma_{\psi}(\tau)$ of $G_{\mathbb{A}_F}$ is cuspidal.
Let $\sigma$ be an irreducible summand of $\sigma_\psi(\tau)$. Then $\sigma$ is globally $\psi$-generic, and $\sigma_\tau$ lifts to $\tau_\nu$, for almost all finite places $\nu$. (If $G = Sp_{2n}$, $\sigma_\nu$ lifts to $\tau_\nu$ with respect to $\psi_\nu$).

(4) Every irreducible, automorphic, cuspidal, $\psi$-generic representation $\sigma$ of $G_{K,F}$, which lifts weakly to $\tau$ has a nontrivial $L^2$-pairing with $\sigma_\psi(\tau)$.

(5) $\sigma_\psi(\tau)$ is a multiplicity free representation.

Remark. — The guidelines to the proof are similar to those of Theorem 4, except that the proof of (1) in case $G$ is even orthogonal or symplectic is not direct. In these cases, we show, once we fix $\psi$, that there is $\beta \in F^*$, such that $\sigma_{\psi,\beta}(\tau) \neq 0$, where $\sigma_{\psi,\beta}(\tau)$ is defined as in (3.2) only that the coefficient (Gelfand-Graev, or Fourier-Jacobi) of the residual Eisenstein series induced from $\tau$ is taken with respect $\psi_n^{-1}$, in case $G = SO_{2n}$, and in case $G = Sp_{2n}$, we take in (2.9) a residual Eisenstein series, induced from $\tau$, on $Sp_{4n}(A)$, instead of $\varphi$, and $\theta^\beta_{\nu,\beta}$, instead of $\theta^\beta_{\psi,-1}$ ($\gamma = 1$). In the first case we obtain a non-trivial cuspidal representation $\sigma_{\psi,\beta}(\tau)$ of $H^{(\beta)}_{2n}(A)$ (see table (2.15)), for which the following Whittaker coefficient is nontrivial

$$\begin{pmatrix} z & x & y \\ 1 & z & \end{pmatrix} \mapsto \psi(z_{12} + z_{23} + \cdots + z_{n-2,n-1} + x_{n-1,2}).$$

Here $z \in Z_{n-1}(A)$, and we write the elements of $H^{(\beta)}_{2n}$, with respect to $\begin{pmatrix} 1 & -2n^{-1} \\ w_{n-1} & 2n^{-1} \end{pmatrix}$.

Let $\sigma$ be an irreducible summand of $\sigma_{\psi,\beta}(\tau)$, which is globally generic with respect to the character (3.3). Then it has a weak lift to $\tau$, and hence, $\omega_\tau(t) = (2\beta, t)$ (Hilbert symbol). This implies that $\sigma_\psi(\tau)$ is non-trivial. In the second case, $(G = Sp_{2n})$, $\sigma_{\psi,\beta}(\tau)$ is a (nontrivial) automorphic cuspidal representation of $Sp_{2n}(A)$, which is globally $\psi^\beta$-generic. Let $\sigma$ be such an irreducible summand of $\sigma_{\psi,\beta}(\tau)$. Examining the unramified parameters of $\sigma$, we show that

$$L^S(\sigma, s) = \frac{L^S(\tau \times \chi_\beta, s)}{L^S(\chi_\beta, s)} L^S(1, s).$$

Here, $\chi_\beta = (\beta, t)$ (Hilbert symbol). If $\chi_\beta \neq 1$, this implies that $L^S(\sigma, s)$ has a pole at $s = 1$. By [GRS5], we conclude that $\sigma$ is a theta lift (with respect to an appropriate character) of a generic cuspidal representation $\pi$ on split $SO_{2n}(A)$. We have

$$L^S(\tau, s) = L^S(\pi \times \chi_\beta, s) L^S(1, s),$$

and hence $L^S(\tau, s)$ has a pole at $s = 1$. This is impossible, and so $\chi_\beta = 1$, i.e. $\sigma_\psi(\tau)$ is nontrivial.

3.2. The image (in general) of the weak lift from $G$ to $GL_N$. — Let $\sigma$ be an irreducible, automorphic, cuspidal generic representation of $G_{K,F}$. Assume that $\sigma$ has a weak lift to $GL_N$, and that it lifts to an irreducible, automorphic representation $\tau$, which as in Sec. 1.3, is a constituent of

$$\delta_1|\det|^{-121} \times \cdots \times \delta_j|\det|^{-21} \times \tau_1 \times \cdots \times \tau_\ell \times \delta^*_j|\det|^{-21} \cdots \times \delta^*_1|\det|^{-21}$$
where $\text{Re}(z_1) \leq \cdots \leq \text{Re}(z_j) \leq 0$, the representations $\delta_i, \tau_k$ are irreducible, automorphic and unitary, with central characters which are trivial on $(\mathbb{A}_E)^{\times}$, and $\tau_i = \tau_i^*=1$, for $1 \leq i \leq \ell$. If $\delta_i$ (resp. $\tau_k$) is on $\text{GL}_r(\mathbb{A})$, $r > 1$, we assume it is cuspidal.

Consider $L^S(\sigma \times \delta_1, s)$. As in Sec. 1.3, we see that $L^S(\sigma \times \delta_1, s)$ has a pole at $s = 1 - z_1$. (If $G$ is metaplectic, consider $L^S(\sigma \times \hat{\delta}_1, s)$.) By Theorem 11, except in case $G$ is metaplectic, and $\delta_1 = 1$, we have $z_1 = 0$ and $L^S(\hat{\delta}_1, \beta_{G,r,\gamma_1}, s)$ has a pole at $s = 1$. Here $\delta_1$ is on $\text{GL}_r(\mathbb{A}_E)$, and $r' = r$ in all cases of Table (3.1), except cases (3) and (6), where $r' = r - 1$. Note that since $L^S(\hat{\delta}_1, \beta_{G,r,\gamma}, s)$ has a pole at $s = 1$, we must have $\delta_1 = \delta_1^*$. (For example, in case of a unitary group, and $\eta = \hat{\delta}_1$, (3.5)

$$L^S(\eta \otimes \eta', s) = L^S(\eta, \text{Asai}, s)L^S(\eta \otimes \gamma, \text{Asai}, s),$$

and since one of the factors on the r.h.s. of (3.5) has a pole at $s = 1$, $L^S(\eta \otimes \eta', s)$ has a pole at $s = 1$, which implies that $\hat{\eta}' = \eta$, i.e. $\eta^* = \eta$. We conclude that $L^S(\sigma \times \hat{\delta}_1, s)$ has a double pole at $s = 1$. This is impossible, and we conclude that (3.4) has the form

$$\tau_1 \times \cdots \times \tau_\ell,$$

and repeating the last argument, we conclude that $L^S(\tau_i, \beta_{G,r,\gamma}, s)$ has a pole at $s = 1$, for $1 \leq i \leq \ell$, and also that $\tau_i \neq \tau_j$, for $1 \leq i \neq j \leq \ell$. Here $\tau_i$ is on $\text{GL}_r(\mathbb{A}_E)$.

Finally, in case $G$ is metaplectic, we see from Theorem 11, that it is possible to have $\delta_1 = 1$, and $z_1 = -1$, and as we remarked before, in this case $\sigma$ is a (ψ) theta lift from a cuspidal generic representation of $\text{SO}_{2n-1}(\mathbb{A})$, so that by Section 1.5, the lift of $\sigma$ to $\text{GL}_{2n}(\mathbb{A})$ has the form $\{ | \cdot |^{-1/2} \times \tau_1 \times \cdots \times \tau_\ell \times | \cdot |^{1/2}$, where $\tau_i$ are as before, one with its exterior square L-function having a pole at $s = 1$. This proves

**Theorem 13.** — Let $\sigma$ be an irreducible, automorphic, cuspidal, generic representation of $G_{\mathbb{A}_E}$. Assume that $\sigma$ lifts weakly to an irreducible automorphic representation $\tau$ of $\text{GL}_{N_{2n}}(\mathbb{A}_E)$ as in Table (3.1). Then except in case (4), $\tau$ has the form $\tau_1 \times \cdots \times \tau_\ell$, where for $1 \leq i \leq \ell$, $\tau_i$ is an irreducible, automorphic, unitary representation of $\text{GL}_{r_i}(\mathbb{A}_E)$, cuspidal in case $r_i > 1$, such that $\tau_i = \tau_i^*, \omega_\tau|_{\mathbb{A}_E^\times} = 1$, except in case $G = \text{SO}_{2n}^\circ$, in which case $\omega_\tau = \chi_{2n}$. The partial L-function $L^S(\tau_i, \beta_{G,r,\gamma}, s)$ has a pole at $s = 1$, and $\tau_i \neq \tau_j$, for all $1 \leq i \neq j \leq \ell$. In case (4), either $\tau$ has the form above, or it has the form $\{ | \cdot |^{-1/2} \times \tau_1 \times \cdots \times \tau_\ell \times | \cdot |^{1/2}$, where the product of the $\tau_i$ is in the image of the lift from generic cuspidal representations from (split) $\text{SO}_{2n-1}(\mathbb{A})$ to $\text{GL}_{2n-2}(\mathbb{A})$.

We consider the converse to Theorem 13, except the last case mentioned there. Let $\tau_1, \ldots, \tau_\ell$ be different irreducible, automorphic, unitary representations of $\text{GL}_{r_1}(\mathbb{A}_E), \ldots, \text{GL}_{r_\ell}(\mathbb{A}_E)$ respectively, and $\tau_i$ is cuspidal, if $r_i > 1$, and such that $r_1 + \cdots + r_\ell = N = N_{2n}$ (as in table (3.1)), $\tau_i = \tau_i^*$, and $L^S(\tau_i, \beta_{G,r,\gamma}, s)$ has a pole at $s = 1$, for $i = 1, \ldots, \ell$. Let $\tau = \tau_1 \times \cdots \times \tau_\ell$. Assume also that $\omega_\tau|_{\mathbb{A}_E^\times} = 1$, except in case $N$ is even, and $E = F$, where we also allow $\omega_\tau$ to be quadratic, and if it is,
say, \( \chi_{2\alpha} \), then, in the sequel, we’ll take the Gelfand-Graev coefficient with respect to \( \text{SO}_{2n}^{\text{\ast}} \). If \( \tau \) is a lift at almost all finite places of an irreducible, automorphic, cuspidal, \( \psi \)-generic representation \( \sigma \) on \( G_{\mathbb{A}_F} \), then by (2.18), (2.27), \( L(\varphi_\sigma, \xi_{\mathbb{A}_F}, s) \) has a pole at \( s = 1 \) in cases (1)–(3) of Table (3.1), \( L(\varphi_\sigma, \psi, \xi_{\mathbb{A}_F}, s) \) has a pole at \( s = 1 \) in cases (4), (6), and \( L(\varphi_\sigma, \phi, \xi_{\mathbb{A}_F}^{1\gamma}, s) \) has a pole at \( s = 1 \) in case (5), as data vary, and \( i = 1, \ldots, \ell \). Consider the Eisenstein series on \( H = H_{G,2n} \) (Table (3.1)) induced from \( \tau_1 | | s_1 - 1/2 \times \cdots \times \tau_\ell | | s_\ell - 1/2 \) and the standard parabolic subgroup of \( H \), whose Levi part is isomorphic to \( \text{Res}_{E/F} \text{GL}_{r_1} \times \cdots \times \text{Res}_{E/F} \text{GL}_{r_\ell} \). Denote it, for a \( K \)-finite holomorphic section \( \xi_{\tau, \mathfrak{p}} \) by \( E_H(\xi_{\tau, \mathfrak{p}}) \) where \( \mathfrak{p} = (s_1, \ldots, s_\ell) \). We can show that \( (s_1 - 1) \cdots (s_\ell - 1)E_H(\xi_{\tau, \mathfrak{p}}) \) is holomorphic and nontrivial at \( \mathfrak{p} = (1, \ldots, 1) \). Denote the value at \( (1, \ldots, 1) \) by \( \text{Res}_{(1, \ldots, 1)} E_H(\xi_{\tau, \mathfrak{p}}) \), and now define \( \sigma_\psi(\tau) \) on \( G_{\mathbb{A}_F} \) exactly as in (3.2).

Our main theorem in its most general form is

**Theorem 14.** — Fix the group \( G \). Let \( N = N_{2n} \) as in Table (3.1). Let \( \tau = \tau_1 \times \cdots \times \tau_\ell \) be the irreducible representation of \( \text{GL}_N(\mathbb{A}_E) \) induced from \( \tau_1 \otimes \cdots \otimes \tau_\ell \), where \( \tau_1, \ldots, \tau_\ell \) are pairwise inequivalent, irreducible, automorphic, unitary representations of \( \text{GL}_{r_1}(\mathbb{A}_E), \ldots, \text{GL}_{r_\ell}(\mathbb{A}_E) \) respectively, \( \tau_1 \) is cuspidal in case \( r_1 > 1 \), such that \( r_1 + \cdots + r_\ell = N \), \( \tau_1^* \tau_\ell = \tau_\ell \), and \( L^2(\tau_1, \beta_{H_{G,r_i}}, s) \) has a pole at \( s = 1 \), for \( i = 1, \ldots, \ell \). Assume that the central character of \( \tau \) is as above. Then

1. \( \sigma_\psi(\tau) \neq 0 \).
2. The representation \( \sigma_\psi(\tau) \) of \( G_{\mathbb{A}_F} \) is cuspidal.
3. Let \( \sigma \) be an irreducible summand of \( \sigma_\psi(\tau) \). Then \( \sigma \) is globally \( \psi \)-generic, and \( \sigma \) lifts to \( \tau_\nu \), for almost all finite places \( \nu \). (If \( G = \text{Sp}_{2n} \), \( \sigma_\nu \) lifts to \( \tau_\nu \) with respect to \( \psi_\nu \).)
4. Every irreducible, automorphic, cuspidal, \( \psi \)-generic representation \( \sigma \) of \( G_{\mathbb{A}_F} \), which lifts to \( \tau \) at almost all finite places, has a nontrivial \( L^2 \)-pairing with \( \sigma_\psi(\tau) \).
5. \( \sigma_\psi(\tau) \) is a multiplicity free representation.

Assume, for simplicity that \( \omega_{\tau_i} |_{\mathbb{A}_{\mathbb{F}}} |_{\mathbb{A}_{\mathbb{F}}} = 1 \), for each \( i \) in the last theorem. Then for each \( \tau_i \), we may apply Theorem 12 and consider the cuspidal \( \psi \)-generic representation \( \sigma_\psi(\tau_i) \) on a corresponding group \( G_i(\mathbb{A}_F) \). Let \( \sigma_i \) be an irreducible summand of \( \sigma_\psi(\tau_i) \), \( i = 1, \ldots, \ell \), and let \( \sigma \) be an irreducible summand of \( \sigma_\psi(\tau) \) (\( \sigma_1, \ldots, \sigma_\ell \) are all \( \psi \)-generic). Then \( \sigma_1 \otimes \cdots \otimes \sigma_\ell \) (on \( G_1(\mathbb{A}_F) \times \cdots \times G_\ell(\mathbb{A}_F) \)) lifts at almost all finite places to \( \sigma \). Both representations lift at almost all places to \( \tau \) on \( \text{GL}_N(\mathbb{A}_E) \). These are examples of (generalized) endoscopy. The following table summarizes the various cases, where we stay a little vague in specifying central characters, and in specifying even orthogonal groups and base change lifts to even unitary groups. (So far, for simplicity, we constructed only lifts from \( \text{U}_{2n} \) to \( \text{Res}_{E/F} \text{GL}_{2n} \), with central character, whose restriction to \( A_F \) is trivial.) Here, as above, \( \sigma_i \) is an irreducible summand of \( \sigma_\psi(\tau_i) \).
<table>
<thead>
<tr>
<th>$\tau_1 \otimes \cdots \otimes \tau_\ell$ on $\GL_{r_1}(A_F) \times \cdots \times \GL_{r_\ell}(A_F)$</th>
<th>pole condition for $\tau_i$</th>
<th>$\sigma_1 \otimes \cdots \otimes \sigma_\ell$ on $G_1(A_F) \times \cdots \times G_1(A_F)$</th>
<th>$\sigma$ on $G(A_F)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\GL_{2r_1}(A_F) \times \cdots \times \GL_{2r_\ell}(A_F)$</td>
<td>$\Res_{s=1} L^5(\tau_i, \chi, s) \neq 0$</td>
<td>$\SO_{2r_1+1}(A_F) \times \cdots \times \SO_{2r_\ell+1}(A_F)$</td>
<td></td>
</tr>
<tr>
<td>$\GL_{2r_1}(A_F) \times \cdots \times \GL_{2r_\ell}(A_F) \times \GL_{2m_1+1}(A_F) \times \cdots \times \GL_{2m_\ell+1}(A_F)$</td>
<td>$\Res_{s=1} L^5(\tau_i, \chi, \Sym^2, s) \neq 0$</td>
<td>$\Sp_{2(n_1+\cdots+n_\ell+1)}(A_F)$</td>
<td>$\SO_{2(n_1+\cdots+n_\ell+1)}(A_F)$</td>
</tr>
<tr>
<td>$\GL_{2r_1}(A_F) \times \cdots \times \GL_{2r_\ell}(A_F) \times \GL_{2m_1+1}(A_F) \times \cdots \times \GL_{2m_\ell+1}(A_F)$</td>
<td>$\Res_{s=1} L^5(\tau_i, \chi, \Sym^2, s) \neq 0$</td>
<td>$\Sp_{2(n_1+\cdots+n_\ell+1)}(A_F)$</td>
<td>$\SO_{2(n_1+\cdots+n_\ell+1)}(A_F)$</td>
</tr>
<tr>
<td>$\GL_{n_1}(A_F) \times \cdots \times \GL_{n_\ell}(A_F)$</td>
<td>$\Res_{s=1} L^5(\tau_i^\ast, \Asai, s) \neq 0$, if $n_i \equiv 1 \pmod{2}$</td>
<td>$U_{n_1}(A_F) \times \cdots \times U_{n_\ell}(A_F)$</td>
<td>$U_{n_1+\cdots+n_\ell}(A_F)$</td>
</tr>
<tr>
<td></td>
<td>$\Res_{s=1} L^5(\tau_i^\ast \otimes \gamma, \Asai, s) \neq 0$, if $n_i \equiv 0 \pmod{2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\GL_{2n_1}(A_F) \times \cdots \times \GL_{2n_\ell}(A_F)$</td>
<td>$\Res_{s=1} L^5(\tau_i, \chi, s - \frac{1}{2}) L^5(\tau_i, \chi, 2s - 1) \neq 0$</td>
<td>$\Sp_{2n_1}(A_F) \times \cdots \times \Sp_{2n_\ell}(A_F)$</td>
<td>$\Sp_{2(n_1+\cdots+n_\ell)}(A_F)$</td>
</tr>
</tbody>
</table>

(Table 3.6)
Example. — The functorial lift \( U_3 \to \text{Res}_{E/F}\text{GL}_3 \) is completely known from the work of Rogawski [R]. The cuspidal part of the image is the set of all irreducible, automorphic, cuspidal representations \( \tau \) of \( \text{GL}_3(\mathbb{A}_E) \), such that \( \tau^* = \tau \) and \( \omega_\tau|_{\mathbb{A}_E^*} = 1 \).

In this case, this is equivalent to \( L^S(\tau', \text{Asai}, s) \) having a pole at \( s = 1 \). In this case, using the multiplicity one property for cuspidal representations on \( U_3(\mathbb{A}_E) \) it follows that \( \sigma_\psi(\tau) \) is an irreducible, automorphic, cuspidal, generic representation of \( U_3(\mathbb{A}_E) \), which lifts to \( \tau \). \( \sigma_\psi(\tau) \) is the generic member of the \( L \)-packet on \( U_3(\mathbb{A}_E) \), parametrized by \( \tau \). The following representations occur in the non-cuspidal part of the image of the lift above, restricted to generic representations.

\( (1) \) \( \mu_\eta \times \pi \), where \( \eta \) is an automorphic character of \( U_3(\mathbb{A}_E) \) and \( \mu_\eta \) is the character of \( \mathbb{A}_E^* \) defined by \( \mu_\eta(x) = \eta(x/|x|) \). The representation \( \pi \) is on \( \text{GL}_2(\mathbb{A}_E) \), and it is irreducible, automorphic, and cuspidal such that \( \pi^* = \pi \), \( \omega_\pi|_{\mathbb{A}_E^*} = 1 \) and \( L^3(\pi' \otimes \gamma, \text{Asai}, s) \) has a pole at \( s = 1 \). The representation \( \sigma_\psi(\mu_\eta \times \pi) \) is an irreducible, automorphic, cuspidal, generic representation of \( U_3(\mathbb{A}_E) \), which lifts to \( \mu_\eta \times \pi \).

\( (2) \) \( \mu_{\eta_1} \times \mu_{\eta_2} \times \mu_{\eta_3} \), where \( \{\eta_1, \eta_2, \eta_3\} \) are three different automorphic characters of \( U_3(\mathbb{A}_E) \). The representation \( \sigma_\psi(\mu_{\eta_1} \times \mu_{\eta_2} \times \mu_{\eta_3}) \) is an irreducible, automorphic, cuspidal, generic representation of \( U_3(\mathbb{A}_E) \), which lifts to \( \mu_{\eta_1} \times \mu_{\eta_2} \times \mu_{\eta_3} \). See [G.J.R.], [Ge.Ro.So1, Ge.Ro.So2, Ge.Ro.So3].

In the remaining part of this paper, we will illustrate the proof of Theorem 12 through (low rank) examples.

4. Illustrations of Proofs in Low Rank Examples

4.1. An observation on unramified factors of residual Eisenstein series

Fix the group \( G \). Let \( N = N_{2n} \) as in Table (3.1). Let \( \tau \) be an irreducible, automorphic, cuspidal representation of \( \text{GL}_N(\mathbb{A}_E) \), such that \( \tau^* = \tau \), \( \omega_\tau|_{\mathbb{A}_E^*} = 1 \), and \( L^3(\tau, \beta_{H,2n}, s) \) has a pole at \( s = 1 \). Consider the residue at \( s = 1 \) of the Eisenstein series on \( H_{G,2n}(\mathbb{A}_E) \) induced from \( \tau' \otimes \gamma \cdot |\det \cdot|^s \). Denote this residual representation by \( E_\tau \). (In all cases, except case (5) in Table (3.1), \( \gamma = 1 \). Also \( \tau' = \tau \) in all cases except cases (3), (5).) We abuse notation and think of \( E_\tau \) also as the space of automorphic forms spanned by the residues. So, for example, when we refer to a constant term of \( E_\tau \), we mean that we consider this constant term applied to all automorphic forms in (the space of) \( E_\tau \). It is easy to check that \( E_\tau \) consists of square integrable automorphic forms. Indeed, \( E_\tau \) is concentrated along the Siegel parabolic subgroup (i.e. all constant terms, with respect to unipotent radicals of standard parabolic subgroups, other than the Siegel parabolic subgroup, vanish on \( E_\tau \)). The constant term of \( E_\tau \) along the Siegel radical has one exponent, which is negative. Now use Jacquet’s criterion for square integrability [J]. Consider an unramified factor \( \pi_\nu \) at a place \( \nu \) of (an irreducible summand of) \( E_\tau \). By our assumption on \( \tau \), we have \( \pi_\nu \cong \pi_\nu \) and
\( \omega_{\tau_\nu}|_{F^*_\nu} = 1 \). Since \( \tau_\nu \) is unramified, we see that \( \tau_\nu \) is the unramified constituent of a representation of \( \text{GL}_N(E_\nu) \) induced from the Borel subgroup and an unramified character of the torus of the form

\[
\text{diag}(t_1, \ldots, t_{2n}) \mapsto \mu_1 \left( \frac{t_1}{t_{2n}} \right) \cdots \mu_n \left( \frac{t_n}{t_{n+1}} \right), \quad \text{if } N = 2n
\]

\[
\text{diag}(t_1, \ldots, t_{2n+1}) \mapsto \mu_1 \left( \frac{t_n}{t_{2n+1}} \right) \cdots \mu_n \left( \frac{t_n}{t_{n+2}} \right), \quad \text{if } N = 2n + 1.
\]

Recall that if \( E = F, \mathfrak{T} = t, \) for \( t \in E \). If \( [E : F] = 2 \) and \( \nu \) is a place which splits in \( F \), then \( E_\nu = F_\nu \oplus F_\nu \), \( (a, b) = (b, a) \) and the characters \( \mu_i \) are given by pairs of characters of \( F_\nu^* \). Let \( Q \) be the standard parabolic subgroup of \( H = H_{G,2n} \), whose Levi part is isomorphic to \( (\text{Res}_{E/F} \text{GL}_2)^n \) in cases (1),(2),(4),(5) of Table (3.1), or to \( (\text{Res}_{E/F} \text{GL}_2)^n \times H_0 \) where \( H_0 = U_2 \) in case (3) and \( H_0 = \text{SL}_2 \) in case (6). (In case (6) we should really take the inverse image in \( \widetilde{Sp}_{4n+2} \), at each place \( \nu \) : \( \text{GL}_2(F_\nu)^n \times \text{SL}_2(F_\nu) \)). Denote by \( \pi_{\mu_1, \ldots, \mu_n} \) the unramified constituent of the representation \( \rho_{\mu_1, \ldots, \mu_n} \) of \( H(F_\nu) \) induced from \( Q(F_\nu) \) and the character \( (\mu_1 \cdot \det) \otimes \cdots \otimes (\mu_n \cdot \det) \). (In cases (3) and (6) of Table (3.1), it is trivial on \( H_0(F_\nu) \). In case (6) we also have to multiply by \( \gamma_\nu \).) Denote \( \mu'_j(t) = \mu_j(\mathfrak{T}) \). Denote by \( \omega \) the simple Weyl reflection in \( O_{4n} \), which flips the two middle coordinates in the diagonal subgroup.

**Proposition 15.** — Using the notation above, let \( \tau_\nu \) be the unramified representation of \( \text{GL}_N(E_\nu) \), corresponding to the unramified character (4.1). Then \( \pi_{\nu} \cong \pi_{\mu_1, \ldots, \mu_n} \), except in case 1 of Table 3.1, with \( n \) odd, where we have \( \pi_{\nu} \cong \pi_{\nu}^{\omega} \mu_1, \ldots, \mu_n \) (outer conjugation).

**Proof.** — Denote by \( \rho_{\tau_\nu} \otimes \gamma_\nu \) the representation of \( H(F_\nu) \) induced from the Siegel parabolic subgroup and \( \tau_\nu' \otimes \gamma_\nu \) \( |1/2\). (We have to modify by \( \gamma_\nu \) in case (6).) Consider first cases (1),(4),(5) in Table (3.1). In case (1), assume for simplicity that \( n \) is even. Here \( \rho_{\tau_\nu} \otimes \gamma_\nu \) is induced from the following character of the Borel subgroup

\[
(4.2) \quad \text{diag}(t_1, \ldots, t_{2n}, \mathfrak{T}_{2n}^{-1}, \ldots, \mathfrak{T}_1^{-1}) \mapsto \mu'_1 \gamma_\nu \left( \frac{t_1}{t_{2n}} \right) |t_1 t_{2n}|^{1/2} \cdots \mu'_n \gamma_\nu \left( \frac{t_n}{t_{n+1}} \right) |t_n t_{n+1}|^{1/2}
\]

This character is conjugate, under a suitable Weyl element of \( H \), to the character

\[
(4.3) \quad \text{diag}(t_1, \ldots, t_{2n}, \mathfrak{T}_{2n}^{-1}, \ldots, \mathfrak{T}_1^{-1}) \mapsto \mu'_1 \gamma_\nu (t_1 t_{2n}) \left( \frac{t_1}{t_{2n}} \right)^{1/2} \cdots \mu'_n \gamma_\nu (t_n t_{n+1}) \left( \frac{t_n}{t_{n+1}} \right)^{1/2},
\]

and this character is conjugate, under a suitable Weyl element of \( \text{GL}_N \), to the character

\[
(4.4) \quad \text{diag}(t_1, \ldots, t_{2n}, \mathfrak{T}_{2n}^{-1}, \ldots, \mathfrak{T}_1^{-1}) \mapsto \mu'_1 \gamma_\nu (t_1 t_2) \left( \frac{t_1}{t_2} \right)^{1/2} \cdots \mu'_n \gamma_\nu (t_{2n-1} t_{2n}) \left( \frac{t_{2n-1}}{t_{2n}} \right)^{1/2}.
\]
Thus $\pi_\varphi$ is the unramified constituent of the representation $\eta_{\mu_1, \gamma_1, \ldots, \mu_n, \gamma_n}$ induced from the character of the Borel subgroup defined by (4.4). Clearly $\eta_{\mu_1, \gamma_1, \ldots, \mu_n, \gamma_n}$ maps onto $\rho_{\mu_1, \gamma_1, \ldots, \mu_n, \gamma_n}$. Since the last representation is still unramified, we conclude that $\pi_\varphi$ is the unramified constituent of $\rho_{\mu_1, \gamma_1, \ldots, \mu_n, \gamma_n}$. (If $n$ is odd in case (1), we get that $\pi_\varphi \cong \pi_{\mu_1, \gamma_1, \ldots, \mu_n, \gamma_n}$, where $\omega$ is as above.) In case (2) the proof is the same, only that in (4.2)–(4.4), the left hand side is diag$(t_1, \ldots, t_{2n+1}, t_{2n+1}^{-1}, \ldots, t_1^{-1})$ and in the right hand side there is no change except that $\mu'_i = \mu_i$, $\gamma'_\varphi = 1$. In case (4) the proof is the same, only that in (4.2)–(4.4) the l.h.s. is diag$(t_1, \ldots, t_{2n+1}, t_{2n+1}^{-1}, \ldots, t_1^{-1})$. The r.h.s. of (4.2)–(4.4) remains the same. In case (6), the l.h.s of (4.2)–(4.4) is diag$(t_1, \ldots, t_{2n+1}, t_{2n+1}^{-1}, \ldots, t_1^{-1})$, and in the r.h.s. we have to multiply by $\gamma_\varphi(t_1 \cdots t_{2n+1})$ (and take $\mu'_i = \mu_i$, $\gamma_\varphi = 1$). \hfill \Box

4.2. Nonvanishing of $\sigma_\varphi(\tau)$: Case $G = U_3$, $H = U_6$, $\tau$ on $GL_3(\mathbb{A}_E)$

Let $\tau$ be a irreducible, automorphic, cuspidal representation of $GL_3(\mathbb{A}_E)$, such that $\tau^* = \tau$, $\omega_\varphi|_{\mathbb{A}_F}^* = 1$, and $L^2(\tau^*, \text{Asai}, s)$ has a pole at $s = 1$. (Actually, the last condition is equivalent to the first two conditions). The proof that $\sigma_\varphi(\tau) \neq 0$ consists of two steps. First, we introduce (in (4.8)) a unipotent group $V$ of $U_6$, and a certain character $\psi_V$ of $V|_F \backslash V|_F$, and prove that the Fourier coefficient along $V$, with respect to $\psi_V$, is nontrivial on (the space of) $E_\tau$ (Proposition 16). To do so, we prove that this nontriviality is equivalent to the nontriviality of another Fourier coefficient on $E_\tau$. This last Fourier coefficient is along a unipotent subgroup $U$, and with respect to a character $\psi_U$ of $U|_F \backslash U|_{\mathbb{A}_F}$. The group $U$ is almost the maximal unipotent subgroup of $U_6$. It "misses" just one root subgroup, namely the simple root which lies in the Siegel radical. The character $\psi_U$ is the restriction to $U|_{\mathbb{A}_F}$ of the standard nondegenerate character determined by $\psi$. Thus, the nontriviality of the $(U, \psi_U)$ coefficient on $E_\tau$ follows from the fact that $\tau$ is (globally) generic. In the second step we show that the nontriviality of the $(V, \psi_V)$ coefficient on $E_\tau$ is equivalent to the nonvanishing of $\sigma_\varphi(\tau)$. We develop for these proofs (and for the sequel) a tool that we call, for lack of a better name, "exchanging roots". In practice, it enables us to conclude that an automorphic representation, realized in a given space of automorphic forms, has a nontrivial $(V_1, \psi_{V_1})$ Fourier coefficient, if and only if it has a nontrivial $(V_2, \psi_{V_2})$ Fourier coefficient, where the unipotent groups $V_1, V_2$ are generated by root subgroups, and the passage from $V_1$ to $V_2$ is by "deleting" a certain root subgroup, and "replacing it, in exchange", with another certain root subgroup (outside $V_1$). The characters $\psi_{V_i}$ are equal on the subgroup generated by the roots common to $V_1$ and $V_2$, and extend trivially to "the rest of" $V_i$.

Let $H = U_6$, and let $P$ be the Siegel parabolic subgroup. Let $\rho_{\tau^*, s} = \text{Ind}_{U_6(\mathbb{A}_F)}^{U_6(\mathbb{A}_F)}(\tau^*|_{\mathbb{A}_F})$ det$|s-1/2$, and consider for a holomorphic, $K$-finite section $\xi_{\tau^*, s}$ of $\rho_{\tau^*, s}$, the corresponding Eisenstein series $E(\xi_{\tau^*, s}, h)$ on $U_6(\mathbb{A}_F)$. We know that $E(\xi_{\tau^*, s}, h)$ has a simple pole at $s = 1$, as data vary. Recall that the space of $\sigma_\varphi(\tau)$
is spanned by the $\psi_{1,1}^{-1}$ – Fourier coefficients of $\text{Res}_{s=1} E(\xi_{r,s}, \cdot)$ along $N_1$. Let us repeat the definitions in this case

\[(4.5) \quad N_1 = \{ u = \left( \begin{array}{c} 1 \\ u \\ 1 \\ v' \\ 1 \end{array} \right) \in U_6 \} \].

For $u \in N_1(\mathbb{A}_F)$ as in (4.5),

\[(4.6) \quad \psi_{1,1}(u) = \psi_E(y_2 - y_3). \]

The stabilizer of $\psi_{1,1}$ inside $\left( \begin{array}{c} 1 \\ u_4 \\ 1 \end{array} \right)$ is

\[ L = \left\{ \left( \begin{array}{c} 1 \\ h \\ 1 \end{array} \right) \in U_6 \mid h \left( \begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) = \left( \begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) \right\} \].

We fix an $F$-isomorphism $i : U_3 \simeq L$. The representation $\sigma_\psi(\tau)$ of $U_3(\mathbb{A}_F)$ acts in the space of automorphic functions spanned by

\[(4.7) \quad g \mapsto \int_{N_1(F) \backslash N_1(\mathbb{A}_F)} \text{Res}_{s=1} E(\xi_{r,s}, u \psi(g)) \psi_{1,1}^{-1}(u) du. \]

In this section we show that (4.7) is not identically zero. Consider the following subgroup of $U_6$

\[(4.8) \quad V = \{ v = \left( \begin{array}{c} I_2 \\ a \\ b \\ I_2 \\ a' \\ I_2 \end{array} \right) \in U_6 \}, \]

and the following character of $V_F \backslash V_{\mathbb{A}_F}$

\[ \psi_V(v) = \psi_E(a_{11} - a_{22}). \]

Let us denote by $E_\tau$ the residual representation of $U_6(\mathbb{A}_F)$ acting in $\text{Span}\{ \text{Res}_{s=1} E(\xi_{r,s}, \cdot) \}$.

**Proposition 16.** — The Fourier coefficient of $E_\tau$ with respect $\psi_V$ along $V_F \backslash V_{\mathbb{A}_F}$ is nontrivial, i.e.

\[ \int_{V_F \backslash V_{\mathbb{A}_F}} \text{Res}_{s=1} E(\xi_{r,s}, v) \psi_{V}^{-1}(v) dv \neq 0. \]

**Proof.** — Let

\[ w = \left( \begin{array}{cccccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right). \]

Write $v$ in (4.8) in the form

\[(4.9) \quad v = \left( \begin{array}{cccccc} 1 & 0 & a & b & * & * \\ 1 & c & d & * & * \\ 1 & 0 & -\frac{d}{a} & -\frac{b}{a} \\ 1 & -\frac{c}{a} & -\frac{d}{a} \\ 1 & 0 & 0 & 1 \end{array} \right). \]
Then
\[
(4.10) \quad wvw^{-1} = \begin{pmatrix}
  1 & a & \ast \\
  1 & 0 & -\tau \\
  0 & \ast & 1
\end{pmatrix}
\]
(zeroes elsewhere). Let \( V' = wVw^{-1} \). Then by (4.10), the elements of \( V' \) have the form
\[
(4.11) \quad v' = \begin{pmatrix}
  z & y \\
  y & z
\end{pmatrix} \in U_6,
\]
where \( z \) is upper unipotent, \( x, y \) are upper nilpotent (such that \( x_{23} = \bar{x}_{12}, y_{23} = -\bar{y}_{12} \)).

The conjugation (4.10) takes the character \( \psi_V \) to the character \( \psi_{V'} \) of \( V'_F \setminus V'_k \), defined by
\[
\psi_{V'}(v') = \psi_E(z_{12} + z_{23})
\]
(\( v' \) is of the form (4.11)). Since \( \text{Res}_{s=1} E(\xi_{r,s}, w \cdot v) = \text{Res}_{s=1} E(\xi_{r,s}, v) \) and
\[
\text{Res}_{s=1} E(\xi_{r,s}, whw^{-1}) = E_r(w^{-1}) \left( \text{Res}_{s=1} E(\xi_{r,s}, \cdot) \right)(h),
\]
what we have to prove is equivalent to
\[
(4.12) \quad \int_{V'_F \setminus V'_{k}} \text{Res}_{s=1} E(\xi_{r,s}, v') \psi^{-1}_{V'}(v')dv' \neq 0.
\]

We will now “exchange roots” in \( V' \) in (4.12), in the sense that (4.12) is equivalent to
\[
(4.13) \quad \int_{D_F \setminus D_k} \text{Res}_{s=1} E(\xi_{r,s}, r) \psi^{-1}_D(r)dr \neq 0,
\]
where
\[
(4.14) \quad D = \left\{ r = \begin{pmatrix}
  1 & a & \gamma \\
  1 & \delta & \beta \\
  1 & \ast & \ast
\end{pmatrix} \in U_6 \right\}, \quad \psi_D(r) = \psi_E(\alpha + \delta).
\]

Note that \( D \) is obtained from \( V' \) by exchanging \( c \) and \( -\bar{c} \) with the zeroes in coordinates (1.4),(3.6) in (4.10). This is done as follows. Let
\[
Z = \begin{pmatrix}
  1 & 0 & \ast \\
  1 & 1 & \ast \\
  1 & \ast & 1
\end{pmatrix} \in \text{Res}_{E/F} \text{GL}_3, \quad m(Z) = \left\{ (z, \bar{z}) \in U_6 \mid z \in Z \right\}
\]
\[
X_0 = \left\{ \begin{pmatrix}
  0 & t & c \\
  0 & 0 & -\tau \\
  0 & 0 & 0
\end{pmatrix} \mid e + \bar{e} = 0 \right\}, \quad X = \left\{ x \in \text{Res}_{E/F} M_{3 \times 3} \mid w_3 x + t(w_3 x) = 0 \right\}
\]
\[
\ell(X) = \left\{ \begin{pmatrix}
  1 & x \\
  1 & \ell_1
\end{pmatrix} \mid x \in X \right\}, \quad \bar{\ell}(X) = \left\{ \begin{pmatrix}
  \ell \bar{1} & x \\
  \ell_1 & \ell_1
\end{pmatrix} \mid x \in X \right\}.
\]

Denote
\[
\bar{Y}^{12} = \left\{ \ell \begin{pmatrix}
  0 & c & 0 \\
  0 & 0 & \bar{c} \\
  0 & 0 & 0
\end{pmatrix} \right\}, \quad \bar{Y}^{13} = \left\{ \ell \begin{pmatrix}
  0 & 0 & c \\
  0 & 0 & 0 \\
  0 & 0 & 0
\end{pmatrix} \mid e + \bar{e} = 0 \right\}, \quad X^{11} = \left\{ \ell \begin{pmatrix}
  t & 0 & -\tau \\
  0 & t & 0 \\
  0 & 0 & 0
\end{pmatrix} \right\}
\]
\[
C = m(Z)\ell(X_0)\bar{Y}^{13}.
\]
Then it is easy to check that $C$ is a group, (it is a subgroup of $V'$) and that the following properties are satisfied.

(i) Let $\psi_C = \psi_{\gamma'}|_{C_{\mathbb{A}_F}}$. Then $Y^{12}$ and $X^{11}$ normalize $C$ and (their adele points) preserve $\psi_C$.

(ii) $[X^{11}, Y^{12}] \subset C$

(iii) The characters $\psi_C(xy^{-1}y^{-1})$ on $X_{F_0}^{11} \setminus X_{A_F}^{11}$ (resp. on $X^{12}_F \setminus X_{A_F}^{12}$) as $y$ (resp. $x$) varies in $Y_F$ (resp. $X_F^{11}$) are all characters of $X^{11}_{F_0} \setminus X_{A_F}^{11}$ (resp. $X^{12}_F \setminus X_{A_F}^{12}$).

\[
\begin{array}{c}
Y^{12} & & Y^{12} \\
\downarrow & & \downarrow \\
V' & & X^{11} \\
\downarrow & & \downarrow \\
C & & C_X^{11} = D \\
\end{array}
\]

Let us check (iii), for example. We have

\[
(4.16) \quad \left( \begin{array}{cc}
I_s & x \\
y & I_s
\end{array} \right) \left( \begin{array}{cc}
I_s & -x \\
y & I_s
\end{array} \right) = \left( \begin{array}{cc}
I_s & y \\
y & -x
\end{array} \right) = \left( \begin{array}{cc}
I_s & yx \\
y & -xy
\end{array} \right)
\]

Now, for $y = \left( \begin{array}{cc} 0 & 0 \\
0 & 0 \end{array} \right)$, $x = \left( \begin{array}{c} 1 \\
0 \end{array} \right)$, $yxy = 0$, and hence (4.16) equals (note that $T(y) \in Y^{12}$, $\ell(x) \in X^{11}$) $(z, z^*)$, where $z = \left( \begin{array}{cc} 1 & -et \\
0 & 1 \end{array} \right)$. Hence $\psi_C$ applied to the l.h.s. of (4.16) equals $\psi_E^{-1}(ct)$, which represents a general character of $t$ (resp. $c$), as $c$ (resp. $t$) varies.

Let us explain now the equivalence of (4.12) and (4.13). Put $e_\xi(h) = \text{Res}_{s=1} E(\xi_{x, s}, h)$. We have

\[
\int_{V_F \setminus V'_{F_0}} e_\xi(v') \psi_{v'}^{-1}(v') dv' = \int_{Y^{12}_F \setminus Y^{12}_{A_F}} \int_{C_F \setminus C_{A_F}} e_\xi(cy) \psi_C^{-1}(c) dcdy
\]

\[
= \int_{Y^{12}_F \setminus Y^{12}_{A_F}} \sum_{\lambda \in E} \int_{E \setminus A_F} \int_{C_F \setminus C_{A_F}} e_\xi(c \ell(\begin{array}{cc} 1 & 0 \\
0 & \tau \end{array})) y) \psi_E^{-1}(\lambda t) \psi_C^{-1}(c) dtdcdy
\]

\[
= \int_{D \subset C^{11}_X} \sum_{\lambda \in E} \int_{D_F \setminus D_{A_F}} e_\xi(ry) \psi_{C, \lambda}^{-1}(r) drdy.
\]

Here, for $r = c \ell(\begin{array}{cc} 1 & 0 \\
0 & \tau \end{array}) \in C^{11}_X = D$, $\psi_{C, \lambda}(r) = \psi_E(\lambda t) \psi_C(c)$. Let $y_0 \in Y^{12}_F$. Then $e_\xi(ry) = e_\xi(y_0ry) = e_\xi(y_0y^{-1}y_0y)$. Recall that $y_0$ normalizes $D_{A_F}$, and it preserves $D_F$. Also, for $r = c \cdot x$, $x \in X^{11}_{A_F}$, $c \in C_{A_F}$, $y_0^{-1}xy = [y_0^{-1}, x] \in C_{A_F} X^{11}_{A_F}$, and...
In the one before last integral, we changed variable $r \mapsto r^{-1}y_{0}$. Thus, for each $y_{0} \in \mathbb{F}_{\ell}$, we have

\[
\int_{D_{F} \setminus D_{K}} \frac{e_{\xi}(y_{0}ry) \psi^{-1}_{C,\lambda}(r)dr}{\gamma} = \int_{D_{F} \setminus D_{K}} \frac{e_{\xi}(ry_{0}y) \psi^{-1}_{C,\lambda}(y^{-1}_{0}ry_{0})dr}{\gamma} = \int_{X^{11}_{K} \setminus X^{11}_{K}} \frac{e_{\xi}(cyry_{0}y) \psi^{-1}_{C,\lambda}([y^{-1}, x] \psi^{-1}_{C}(c)c)dc}{\gamma}. \]

We could take even a variable $y_{\lambda} \in \mathbb{F}_{\ell}$, $\lambda \in E$, and get the same results. Take $y_{\lambda} = \ell \left( \begin{smallmatrix} 0 & -\lambda \\ \lambda & 0 \end{smallmatrix} \right)$. Then for $x = \ell \left( \begin{smallmatrix} t \\ 0 \end{smallmatrix} \right)$, we have seen in (4.16) that $\psi_{C,\lambda}([y^{-1}_{\lambda}, x]) = \psi_{E}(-\lambda t)\psi_{E}(\lambda t) = 1$. Put $\psi_{D}(xc) = \psi_{C}(c)$. We get that the l.h.s. of (4.12) equals

\[
\int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} \sum_{y_{\lambda} \in \mathbb{F}_{\ell}} \int_{D_{F} \setminus D_{K}} \frac{e_{\xi}(ry_{0}y) \psi^{-1}_{D}(r)dr}{\gamma} dy_{\lambda} = \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} \int_{D_{F} \setminus D_{K}} \frac{e_{\xi}(ry) \psi^{-1}_{D}(r)dr}{\gamma}. \]

Thus, we have shown that

\[
(4.17) \quad \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} \frac{e_{\xi}(v') \psi_{V}^{-1}(v')dv'}{\gamma} = \int_{D_{F} \setminus D_{K}} \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} \frac{e_{\xi}(ry) \psi^{-1}_{D}(r)dr}{\gamma}. \]

We claim that the r.h.s. of (4.17) is not identically zero, if and only if

\[
\int_{D_{F} \setminus D_{K}} \frac{e_{\xi}(r) \psi^{-1}_{D}(r)dr}{\gamma} \neq 0, \]

which is (4.13). Indeed, assume that the r.h.s. of (4.17) is identically zero. Apply the convolution operator $\int_{\mathbb{A}_{E}} \phi(t)E_{\tau}(\xi \left( \begin{smallmatrix} t \\ 0 \end{smallmatrix} \right))dt$, for $\phi \in S(\mathbb{A}_{E})$. We get (denoting $\phi(\ell \left( \begin{smallmatrix} t \\ 0 \end{smallmatrix} \right)) = \phi(t)$)

\[
0 \equiv \int_{X^{11}_{K} \setminus X^{11}_{K}} \int_{D_{F} \setminus D_{K}} \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} \phi(x)e_{\xi}(r[y, x]xy)\psi^{-1}_{D}(r)drdydx \]

\[
= \int_{Y^{12}_{K} \setminus Y^{12}_{K}} \int_{X^{11}_{K} \setminus X^{11}_{K}} \phi(x)\psi_{D}([y, x])dx \int_{D_{F} \setminus D_{K}} \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} e_{\xi}(ry)\psi^{-1}_{D}(r)drdy \]

\[
= \int_{Y^{12}_{K} \setminus Y^{12}_{K}} \int_{D_{F} \setminus D_{K}} \hat{\phi}(y) \int_{\mathbb{F}_{\ell} \setminus \mathbb{F}_{\ell}} e_{\xi}(ry)\psi^{-1}_{D}(r)drdy. \]

In the one before last integral, we changed variable $r \mapsto r[y, x]^{-1}x^{-1}$. Recall that $\psi_{D}|_{X^{11}_{K}} = 1$. In the last integral, $\hat{\phi}(y) = \int_{X^{11}_{K}} \phi(x)\psi_{D}([y, x])dx$. This is a Fourier
transform of $\phi$, since $x \mapsto \psi_D([y, x])$ is a general character of $x$, as $y$ varies. Thus, (for all $\xi$)
\[
\int_{Y_{kF}^{12}} \hat{\phi}(y) \int_{D_F \backslash D_{kF}} e_{\xi}(ry) \psi_D^{-1}(r) dr dy \equiv 0,
\]
for all $\phi \in S(X_{kF}^{11})$. This is equivalent to (4.13). In the passage from (4.12) to (4.13) we “exchanged” $Y^{12} \equiv X^{11}$. (see (4.15)).

We have to prove (4.13). Let
\[
X^{22} = \left\{ \ell \begin{pmatrix} 0 & t & 0 \\ t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bigg| t + \overline{t} = 0 \right\}.
\]
Then $X^{22}$ normalizes $D$ and preserves $\psi_D$. Put $\widetilde{D} = D \cdot X^{22}$, and extend $\psi_D$ to $\widetilde{D}$, by making it trivial on $X^{22}$. Denote this extension by $\psi_{\widetilde{D}}$. Let
\[
X^{21} = \left\{ \ell \begin{pmatrix} 0 & 0 & 0 \\ t & 0 & 0 \\ 0 & \overline{t} & 0 \end{pmatrix} \bigg| \overline{t} = t \right\}
\]
Then one can check that $X^{21}$ normalizes $\widetilde{D}$ and preserves $\psi_{\widetilde{D}}$. Let $D^+ = \widetilde{D} \cdot X^{21}$, and extend $\psi_{\widetilde{D}}$ to a character $\psi_{D^+}$ of $D^+$, by making it trivial on $X^{21}$. In order to prove (4.13), it is enough to prove
\[
(4.18) \quad \int_{D_F \backslash D_{kF}^+} \operatorname{Res}_{s=1} E(\xi', s, r) \psi^{-1}_{D^+}(r) dr \not\equiv 0
\]
Let
\[
X^{21} = \left\{ \ell \begin{pmatrix} 0 & 0 & 0 \\ t & 0 & 0 \\ 0 & -\overline{t} & 0 \end{pmatrix} \bigg| \overline{t} = -t \right\}
\]
We can “exchange” in (4.18) $Y^{13}$ by $X^{21}$. More precisely, this is done as follows. Let $C^+ = m(Z)\ell(X)X^{21}$. This is a subgroup of $D^+$. Put $\psi_{C^+} = \psi_{D^+}|_{C^+}$. Then

(i) $Y^{13}$ and $X^{21}$ normalize $C^+$ and preserve $\psi_{C^+}$.
(ii) $[X^{21}, Y^{13}] \subset C^+$
(iii) The characters $\psi_{C^+}(xy^{-1}y^{-1})$ on $X^{21}_F \backslash X^{21}_{kF}$ (resp. on $Y^{13}_F \backslash Y^{13}_{kF}$) as $y$ (resp. $x$) varies in $Y^{13}_F$ (resp. $X^{21}_F$) are all characters of $X^{21}_F \backslash X^{21}_{kF}$ (resp. $Y^{13}_F \backslash Y^{13}_{kF}$).
Extend $\psi_{C^+}$ to a character $\psi_U$ of $U$ by making it trivial on $X_{21}$. As before, (4.18) is equivalent to

$$\int_{U_F \backslash U_{A_F}} \text{Res}_{s=1} E(\xi_{r'}, s, r) \psi_U^{-1}(r)dr \neq 0.$$  

Note that $r \in U_{A_F}$ has the form

$$r = \begin{pmatrix} 1 & a & \ast & \ast & \ast & \ast \\ b & 0 & \ast & \ast & \ast & \ast \\ 1 & -\tau & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \end{pmatrix} \in U_6(A_{F})$$

and

$$\psi_U(r) = \psi_E(a + b).$$

$U$ is a subgroup of the standard maximal unipotent subgroup $N$ of $U_6$. Extend $\psi_U$ to $\psi_N$ on $N_{A_F}$ by making it trivial on the Siegel radical $S$. Clearly (4.19) will follow from the nonvanishing of the Fourier coefficient of $\text{Res}_{s=1} E(\xi_{r'}, s, \cdot)$ with respect to $\psi_N$ along $N_F \backslash N_{A_F}$. This last Fourier coefficient is just the constant term of $\text{Res}_{s=1} E(\xi_{r'}, s, \cdot)$ along $S$, followed by the Whittaker coefficient for the Levi part of the Siegel parabolic subgroup. Writing the constant term of $\text{Res}_{s=1} E(\xi_{r'}, s, \cdot)$ in terms of the intertwining operator, we see that the last Fourier coefficient is just a Whittaker coefficient applied to $\tau'$ with respect to the standard nondegenerate character defined by $\psi_E$, which is, of course, not identically zero. This completes the proof of Proposition 16.

We now conclude that $\sigma_\psi(\tau) \neq 0$. For this, let

$$\gamma = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$ 

Then, by Proposition 16,

$$\int_{V_F \backslash V_{A_F}} \text{Res}_{s=1} E(\xi_{r'}, s, \gamma^{-1}v\gamma) \psi_{V'}^{-1}(v)dv \neq 0.$$ 

Note that for $v \in V_{A_F}$ of the form

$$v = \begin{pmatrix} 1 & 0 & a & b & \ast & \ast \\ 1 & c & d & \ast & \ast & \ast \\ 1 & -\tau & -\pi & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \end{pmatrix},$$

$$\gamma^{-1}v\gamma = \begin{pmatrix} 1 & 0 & a & b & \ast & \ast \\ 1 & c & \ast & \ast & \ast & \ast \\ 1 & -\tau & -\pi & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \\ 1 & \ast & \ast & \ast & \ast & \ast \end{pmatrix}.$$
Let \( \tilde{\psi} \) be the character, which takes \( v \) in \( V_{h,F} \) of the form (4.21) to \( \psi(a - b - d) \). Thus

\[
\int_{V_F \backslash V_{h,F}} \text{Res}_{s=1} E(\xi_{\tau}, v)\tilde{\psi}^{-1}(v)dv \neq 0.
\]

Change variable in (4.22), \( c \mapsto c + a, \, d \mapsto d + b \). Consider the following subgroups.

\[
J = \left\{ \left( \begin{array}{cc} I_2 & x \\ y & I_2 \end{array} \right) \in V \mid x = \left( \begin{array}{cc} a & b \\ d & d \end{array} \right) \right\}
\]

\[
K = \left\{ \left( \begin{array}{cc} 1 & c \\ 1 & -\tau \\ d & d \end{array} \right) \right\}
\]

\[
L = \left\{ \left( \begin{array}{cc} 1 & I_2 \\ 1 & 1 \end{array} \right) \right\}.
\]

Put \( \psi_J = \tilde{\psi} \mid J \). Then

(i) The subgroups \( K, L \) normalize \( J \) and preserve \( \psi_J \).

(ii) \([K, L] \subset J\)

(iii) The characters \( \psi_J(xy x^{-1}y^{-1}) \) describe general characters of \( x \) in \( K_F \backslash K_{h,F} \) (resp. \( y \in L_F \backslash L_{h,F} \)) as \( y \) varies in \( L_F \) (resp. as \( x \) varies in \( K_F \)).

Note that \( V = J \cdot K \). Denote \( U' = JL \), and extend \( \psi_J \) to a character of \( U' \), by making it trivial on \( L \). Now “exchange” \( K \) and \( L \) in (4.22). We get that

\[
\int_{U_{h,F}' \backslash U_{h,F}} \text{Res}_{s=1} E(\xi_{\tau'}, \tau)\psi_{U'}^{-1}(r)dr \neq 0.
\]

Note that \( r \in U_{h,F}' \) has the form

\[
r = \left( \begin{array}{cccc} 1 & I & a & b \\ 1 & d & * & * \\ 1 & 0 & -d & -b \\ 1 & \tau & -\tau & 1 \end{array} \right)
\]

and

\[
\psi_{U'}(r) = \psi_E(a - b) \cdot \psi_E(d).
\]

This means that the l.h.s. of (4.23) is the integration (4.7), which defines \( \sigma_\psi(\tau) \), followed by the Whittaker coefficient with respect to \( \psi_E \) along \( i(N) \), where \( N \) is the standard maximal unipotent subgroup of \( G = U_3 \). In particular \( \sigma_\psi(\tau) \neq 0 \), and we also showed that the \( \psi_E \)-Whittaker coefficient of \( \sigma_\psi(\tau) \), as a representation of \( U_3(h,F) \) is nontrivial.
4.3. The tower property: Case $H = \text{Sp}_k$, $\tau-$ on $\text{GL}_4(\mathbb{A}_F)$, $G = \widetilde{\text{Sp}}_4$

Let $\tau$ be an irreducible, automorphic, cuspidal representation of $\text{GL}_4(\mathbb{A}_F)$, such that $L^S(\tau, A^2, s)$ has a pole at $s = 1$, and $L(\tau, \frac{1}{2}) \neq 0$. (This implies in particular that $\tilde{\tau} = \tau$ and $\omega_\tau = 1$). Let $H = \text{Sp}_k$, and let $P$ be the Siegel parabolic subgroup of $H$. Let $\rho_{\tau,s} = \text{Ind}_{\tilde{P}_{\mathbb{A}_F}}^{H_{\mathbb{A}_F}} \tau \mid \det(\cdot)^{s-1/2}$, and consider the corresponding Eisenstein series $E(\xi_{\tau,s}, h)$ on $\text{Sp}_k(\mathbb{A}_F)$, for a holomorphic, $K$-finite section $\xi_{\tau,s}$. $E(\xi_{\tau,s}, h)$ has a simple pole at $s = 1$, as data vary. Recall that the space of $\sigma_\psi(\tau)$ is spanned by the Fourier-Jacobi coefficients of type $(\psi_1, 1, \phi)$ of $\text{Res}_{s=1} E(\xi_{\tau,s}, \cdot)$ along $N_2$. We repeat the definitions in this case

\begin{equation}
N_2 = \left\{ v = \begin{pmatrix}
1 & x & * & * \\
1 & y & t & * \\
l_4 & y' & * & * \\
1 & -x & 1 & 1
\end{pmatrix} \in \text{Sp}_k \right\}.
\end{equation}

For $v \in N_2(\mathbb{A}_F)$ as in (4.24),

$$\psi_1(v) = \psi(x).$$

The group $N_2$ surjects onto the Heisenberg group $H$ in five variables by

$$j(v) = (y; t),$$

for $v \in N_2$, as in (4.24). Let $\omega_{\psi^{-1}}$ be the Weil representation of $\widetilde{\text{Sp}}_4(\mathbb{A}_F) \ltimes H_{\mathbb{A}_F}$, acting on $S(\mathbb{A}_F^2)$, corresponding to the character $\psi^{-1}$. Denote, for $\phi \in S(\mathbb{A}_F^2)$, the corresponding theta series by $\theta_{\omega_{\psi^{-1}}}^\phi(\cdot)$. The representation $\sigma_{\psi}(\tau)$ of $\text{Sp}_4(\mathbb{A}_F)$ acts in the space of automorphic functions spanned by

\begin{equation}
\begin{aligned}
\tilde{g} &\mapsto \int_{N_2(\mathbb{F}) \backslash N_2(\mathbb{A}_F)} \text{Res}_{s=1} E(\xi_{\tau,s}, vj(g))\theta_{\omega_{\psi^{-1}}}^\phi(j(v)\tilde{g})\psi_1^{-1}(v)dv.
\end{aligned}
\end{equation}

Here $\tilde{g}$ is the projection of $\tilde{g}$ in $\widetilde{\text{Sp}}_4(\mathbb{A}_F)$ onto $\text{Sp}_4(\mathbb{A}_F)$, and we extend $j$ to an embedding of $\text{Sp}_4(\mathbb{A}_F) \cdot H_{\mathbb{A}_F}$ inside $\text{Sp}_8(\mathbb{A}_F)$ by $j(g) = \begin{pmatrix} I_2 & g \\
I_2 & 0 \end{pmatrix}$.

In order to prove that $\sigma_\psi(\tau)$ is cuspidal, we have to show that the constant terms along unipotent radicals (of parabolic subgroups of $\text{Sp}_4$) vanish on $\sigma_\psi(\tau)$. The tower property that we reveal when we compute these constant terms is that they are expressed in terms of “deeper descents” $\sigma_\psi^{(k)}(\tau)$ ($k < n = 2$), which in our case means $k = 0, 1$. Here $\sigma_\psi^{(0)}(\tau)$ is simply the “space” of $\psi$-Whittaker coefficients on the group “$\widetilde{\text{Sp}}_4(\mathbb{A}_F)$” which by definition is $\{1\}$, of the residue representation $E_\tau$ (acting on $\text{Span}\{\text{Res}_{s=1} E(\xi_{\tau,s}, \cdot)\}$). Since the $\psi$-Whittaker coefficient of $E(\xi_{\tau,s}, \cdot)$ is holomorphic at $s = 1$, the last space is zero dimensional, i.e. $\sigma_\psi^{(0)}(\tau) = 0$. The space $\sigma_\psi^{(1)}(\tau)$ is the space of automorphic functions on $\widetilde{\text{Sp}}_4(\mathbb{A}_F) = \widetilde{\text{SL}}_2(\mathbb{A}_F)$ spanned by

$$\tilde{g} \mapsto \int_{N_3(\mathbb{F}) \backslash N_3(\mathbb{A}_F)} \text{Res}_{s=1} E(\xi_{\tau,s}, uj'(g))\theta_{\omega_{\psi^{-1}}}^\phi(j'(u))\psi_2^{-1}(u)du.$$
Here $\varphi \in S(\mathbb{A}_F)$, and $\theta^{\varphi}_{\psi^{-1}}(\cdot)$ is the theta series corresponding to the Weil representation $\omega_{\psi^{-1}}$ of $\tilde{\text{SL}}_2(\mathbb{A}_F) \ltimes H'(\mathbb{A}_F)$, where $H'$ is the Heisenberg group in three variables.

The group $N_3$ is

$$N_3 = \left\{ u = \begin{pmatrix} z & x & y \\ I_2 & z' & z' \\ 1 & 1 & 1 \end{pmatrix} \in \text{Sp}_8 \mid z \in \mathbb{Z}_3 = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\} \right\}.$$  

For $u \in N_3$, as in (4.26), $\psi_2(u) = \psi(z_{12} + z_{23})$, and $j'(u) = (x_{31}, x_{32}; y_{31})$ (the surjection $N_3 \to H'$). Finally, for $g \in \text{SL}_2(\mathbb{A}_F)$, $j'(g) = \begin{pmatrix} I_3 & g \\ g & I_3 \end{pmatrix}$.

There are two standard unipotent radicals of maximal parabolic subgroups of $\text{Sp}_4$:

$$R = \left\{ \begin{pmatrix} 1 & x & y \\ I_2 & z & z' \end{pmatrix} \in \text{Sp}_4 \right\}, \quad S = \left\{ (I_2 x I_2) \in \text{Sp}_4 \right\}.$$

**Proposition 17**

(a) The constant term of elements of $\sigma_{\psi}(\tau)$ along $R$ is a sum of certain integrals of elements of $\sigma_{\psi}^{(1)}(\tau)$.

(b) The constant term of elements of $\sigma_{\psi}(\tau)$ along $S$ is a sum of certain integrals of elements of $\sigma_{\psi}^{(0)}(\tau)$.

We conclude that if $\sigma_{\psi}^{(1)}(\tau) = 0$, then the elements of $\sigma_{\psi}(\tau)$ are cuspidal, in the sense that their constant terms along unipotent radical are all zero. Note, as we explained before that $\sigma_{\psi}^{(0)}(\tau)$ is zero. In general, we may consider $\sigma_{\psi}^{(k)}(\tau)$ for $k \leq 2n$. This is a representation of $\tilde{\text{Sp}}_{2k}(\mathbb{A}_F)$. The constant terms of the elements of $\sigma_{\psi}^{(k)}(\tau)$ along unipotent radicals turn out to be sums of elements of $\sigma_{\psi}^{(j)}(\tau)$, for $j < k$. The tower principle says that there is a first index $k_0$, such that $\sigma_{\psi}^{(k_0)}(\tau) \neq 0$, and then $\sigma_{\psi}^{(k_0)}(\tau)$ is cuspidal. We actually prove that $k_0 = n$.

**Proof of Proposition 17(a).** — Put, for short $e_\tau(h) = \text{Res}_{s=1} E(\xi_{\tau,s}, h)$. We consider

$$c(e_\tau, \phi) = \int_{R_F \setminus R_{s,F}} \int_{N_2(F) \setminus N_2(\mathbb{A}_F)} e_\tau(v j(r)) \theta_{\psi^{-1}}^{\phi}(j(v) r) \psi^{-1}(v) dv dr.$$  

Since $R$ splits in $\tilde{\text{Sp}}_4$, we identify $R$ as a subgroup of $\tilde{\text{Sp}}_4$. Let $\gamma = \begin{pmatrix} I_2 & 1 \\ I_2 & 1 \\ 1 & 1 \end{pmatrix}$. Denote the right $\gamma$-translate of $e_\tau$ by $\gamma \cdot e_\tau$. We have

$$c(e_\tau, \phi) = \int_{R_F \setminus R_{s,F}} \int_{N_2(F) \setminus N_2(\mathbb{A}_F)} \gamma \cdot e_\tau(\gamma v j(r) \gamma^{-1}) \theta_{\psi^{-1}}^{\phi}(j(v) r) \psi^{-1}(v) dv dr.$$  

Consider the group $\gamma N_2 j(R) \gamma^{-1}$. We have

$$\gamma N_2 j(R) \gamma^{-1} = T \cdot L \cdot Z \cdot X.$$
where
\[ T = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 1 & I_2 & 0 \\ 1 & I_2 & 1 \end{pmatrix} \in Sp_8 \right\}, \quad Z = \left\{ \begin{pmatrix} I_2 & 0 \\ 0 & I_2 \\ z & 0 \end{pmatrix} \in Sp_8 \mid z = (\frac{1}{1} \star \frac{1}{1}) \right\} \]
\[ L = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 1 & I_2 & 0 \\ 1 & I_2 & 1 \end{pmatrix} \in Sp_8 \right\}, \quad X = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 1 & I_2 & 0 \\ 1 & I_2 & 1 \end{pmatrix} \in Sp_8 \right\}. \]

The integral (4.27) becomes
\[
\int_{X_F \setminus X_{p,F}} \int_{Z_F \setminus Z_{p,F}} \int_{L_F \setminus L_{p,F}} \int_{T_F \setminus T_{p,F}} \gamma e_F(t \cdot \ell \cdot z \cdot x) \theta_{\psi^{-1}}((0, t_{34}, t_{35}, t_{38}; t_{36})(\ell_{31}, 0, 0, 0; 0)j(x))
\]
\[
\psi^{-1}(z_{23}) \, dt \, d\ell \, dz \, dx.
\]
(4.28)

Here \( j \) is an isomorphism of \( X \) with \( R \). It is the inverse to the conjugation by \( j \) composed with \( j \). The theta series in (4.28) equals
\[
(4.29) \quad \sum_{\eta_1 \in F} \sum_{\eta_2 \in F} \omega_{\psi^{-1}}((\eta_1, 0, 0, 0; 0)(0, t_{34}, t_{35}, t_{38}; t_{36})(\ell_{31}, 0, 0, 0; 0)j(x)) \phi(0, \eta_2).
\]

The inner sum in (4.29), as a function \((t_{34}, t_{35}, t_{38}, t_{36})\), is left \( T_F \) invariant, for fixed \( \eta_1, \ell_{31}, x \). In (4.28), we may interchange the \( T_F \setminus T_{p,F} \) integration and the summation over \( \eta_1 \in F \). Now change variable \( t \mapsto \ell_{n}^{-1} t \ell_{n}^{-1} \), where
\[
\ell_{n} = \begin{pmatrix} 1 & 0 & 1 \\ \eta_1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \in L_F.
\]

In (4.28), \( \gamma e_F(t \cdot \ell \cdot z \cdot x) \) becomes \( \gamma e_F(t \cdot (\ell_{n} \ell) \cdot z \cdot x) \), and in (4.29), the inner sum becomes \( \sum_{\eta_2 \in F} \omega_{\psi^{-1}}((0, t_{34}, t_{35}, t_{38}; t_{36})(\eta_1 + \ell_{31}, 0, 0, 0; 0)j(x)) \phi(0, \eta_2) \). Now collapse \( \int_{L_F \setminus L_{p,F}} \sum_{\eta \in F} \) into \( \int_{L_F \setminus L_{p,F}} \), where
\[
L^1 = \left\{ \begin{pmatrix} 1 & 0 \\ \eta_1 & 1 \\ 1 & \star \end{pmatrix} \in Sp_8 \right\}.
\]

We get
\[
(4.30) \quad \int_{X_F \setminus X_{p,F}} \int_{Z_F \setminus Z_{p,F}} \int_{L_F \setminus L_{p,F}} \int_{T_F \setminus T_{p,F}} \gamma e_F(t \cdot \ell \cdot z \cdot x) \cdot \psi^{-1}(z_{23})
\]
\[
\cdot \sum_{\eta \in F} \omega_{\psi^{-1}}((0, t_{34}, t_{35}, t_{38}; t_{36})(\ell_{31}, 0, 0, 0; 0)j(x)) \phi(0, \eta) dt \, d\ell \, dz \, dx.
\]

Note that
\[
\omega_{\psi^{-1}}(j'(x)) \phi(0, \eta) = \varphi(0, \eta).
\]
We can conjugate $x$ “back to the left” in (4.30) to get

\[ \int \int \int \gamma e_\tau(u \cdot \ell \cdot z) \sum_{\eta \in F} \omega_{\psi^{-1}}((0, u_{34}, u_{35}, u_{36}; (\ell_{31}, 0, 0, 0; 0)) \]

(4.31) \[ \phi(0, \eta)\psi^{-1}(z_{23})dud\ell dz, \]

where $U = T \cdot X$. Now take $\phi = \phi_1 \otimes \phi_2$, $\phi_i \in S(\mathbb{A}_F)$. Denote by $\omega'_{\psi^{-1}}$ the Weil representation of $\widetilde{SL}_2(\mathbb{A}_F) \cdot \mathcal{H}(\mathbb{A}_F)$. Then

\[ \omega_{\psi^{-1}}((0, u_{34}, u_{35}, u_{36}; (\ell_{31}, 0, 0, 0; 0))\phi(0, \eta) = \phi_1(\ell_{31})\omega'_{\psi^{-1}}((u_{34}, u_{35}; u_{36}))[\phi_2(\eta). \]

For such $\phi$, (4.31) equals

\[ \int \int \int \int \gamma e_\tau(u \ell z y) \theta_{\psi^{-1}}(u_{34}, u_{35}; u_{36})\psi^{-1}(z_{23})dud\ell dy. \]

Denote

\[ \phi_1 * (\gamma e_\tau)(h) = \int_{\mathbb{A}_F} \phi_1(y) \cdot (\gamma e_\tau)(h\ell y)dy. \]

Then

\[ \begin{align*}
(4.32) \quad c(e_\tau, \phi_1 \otimes \phi_2) &= \int \int \int \int (\phi_1 * (\gamma e_\tau))(u \ell^1 z) \theta_{\psi^{-1}}(i(u)))\psi^{-1}(z_{23})dud\ell^1 dz
\end{align*} \]

Here $i(u) = (u_{34}, u_{35}; u_{36})$. As we did in the previous section, we can exchange in (4.32) the subgroups $L^1$ and

\[ V = \left\{ \left( \begin{array}{ccc} 1 & 0 & s \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{array} \right) \in \text{Sp}_8 \right\}. \]

Denote $Z' = VZ$ and let $\psi_{Z'}$ denote the character of $Z'_{\mathbb{A}_F}$, which is trivial on $V_{\mathbb{A}_F}$ and takes $z$ in $Z_{\mathbb{A}_F}$ to $\psi(z_{23})$. As in (4.17), we get that

\[ (4.33) \quad c(e_\tau, \phi_1 \otimes \phi_2) = \int L^1_{\mathbb{A}_F} \int_{Z'_{\mathbb{A}_F}} \phi_1 * (\gamma e_\tau)(uz^1\ell^1) \theta_{\psi^{-1}}(i(u))\psi^{-1}(z')dudz'd\ell^1. \]

Consider the function on $F\setminus \mathbb{A}_F$

\[ (4.34) \quad t \mapsto \int \int \phi_1 * (\gamma e_\tau)(uz^1x_{1}t^1) \theta_{\psi^{-1}}(i(u))\psi^{-1}(z')dudz', \]

FROM CLASSICAL GROUPS TO GL_n

381
where
\[
x_t = \begin{pmatrix}
1 & t_1 \\
0 & 1
\end{pmatrix}.
\]

Write the Fourier expansion of (4.34) (evaluated at zero)
\[
(4.35) \quad \sum_{\lambda \in F^*} \int_{N_3(F) \backslash N_3(\mathbb{A}_F)} \phi_1 * (\gamma e_\tau) (u \hat{\lambda} \ell^1) |j'(u)| \psi_{-1}^{-1}(u) du,
\]

where \(\hat{\lambda} = \begin{pmatrix} \lambda & t_6 \\ \lambda^{-1} & -t_1 \end{pmatrix}\). See the paragraph before the statement of Proposition 17 for notation. Note that in (4.35) we did not include the constant coefficient, since it will contain as an inner integration the constant term of \(\phi_1 * (\gamma e_\tau)\) along the radical of the standard parabolic subgroup of \(\text{Sp}_8\), which preserves a line. This constant term is clearly zero. Note that the summand in (4.35), corresponding to \(\hat{\lambda}\), is an element of \(\sigma_{\psi}^{(1)}(\tau)\) evaluated at \(\hat{\lambda}\). We proved
\[
(4.36) \quad c(e_\tau, \phi \otimes \phi_2)
= \sum_{\lambda \in F^*} \int_{L_{\lambda \ell}^1} \int_{N_3(F) \backslash N_3(\mathbb{A}_F)} \phi_1 * (\gamma e_\tau) (u \hat{\lambda} \ell^1) |j'(u)| \psi_{-1}^{-1}(u) du d\ell^1.
\]

This completes the proof of Proposition 17a. \(\square\)

4.4. Vanishing of \(\sigma_{\psi}^{(k)}(\tau)\), for \(k < n\): Case \(H = \text{SO}_8\), \(\tau - \text{on GL}_4(\mathbb{A}_F)\)

Let \(\tau\) be an irreducible, automorphic, cuspidal representation of \(\text{GL}_4(\mathbb{A}_F)\), such that \(L^S(\tau, \lambda^2, s)\) has a pole at \(s = 1\). Let \(H = \text{SO}_8\), and let \(P\) be the Siegel parabolic subgroup of \(H\). Let \(\rho_{\tau, s} = \text{Ind}_{P}^{H} \tau |\text{det} |^{-\frac{s}{4}}\), and consider, as before, the corresponding Eisenstein series \(E(\xi_{\tau, s}, h)\). It has a simple pole at \(s = 1\), as data vary. Recall that the representation \(\sigma_{\psi}(\tau)\) of \(\text{SO}_5(\mathbb{A}_F)\) acts in the space spanned by the functions
\[
(4.37) \quad g \mapsto \int_{N_1(F) \backslash N_1(\mathbb{A}_F)} \text{Res}_{s=1} E(\xi_{\tau, s}, u i(g)) \psi_{1,-1}^{-1}(u) du,
\]

where
\[
(4.38) \quad N_1 = \left\{ u = \begin{pmatrix} 1 & v \\ t_6 & 1 \end{pmatrix} \in \text{SO}_8 \right\}
\]

\(\psi_{1,-1}(u) = \psi(v_3 - v_4)\) (for \(u \in N_1\), as in (4.38)). The isomorphism \(i\) sends \(\text{SO}_5\) onto
\[
\left\{ \begin{pmatrix} 1 & h \\ & 1 \end{pmatrix} \in \text{SO}_8 \mid h \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}.
\]

As explained in Section 1.2 and in the previous section, the constant term on \(\sigma_{\psi}(\tau)\) with respect to the radical (in \(\text{SO}_5\)) \(R = \left\{ \begin{pmatrix} 1 & x \\ t_5 & 1 \end{pmatrix} \in \text{SO}_5 \right\}\) is expressed in terms of \(\sigma_{\psi}^{(1)}(\tau)\), and the constant term on \(\sigma_{\psi}(\tau)\) with respect to the Siegel radical
Theorem 4. We will show

\[
\sigma_\psi^{(1)}(\tau) = 0.
\]

Proof. — The proof is using just the fact that at one unramified place \( \nu \), \( \tau_\nu \) is self-dual, and has a trivial central character. Fix such a place \( \nu \). By Proposition 15, the unramified constituent \( \pi_\nu \) of \( \rho_{r,1} = \text{Ind}_{F_\nu}^{H_{F_\nu}} \tau_\nu | \text{det} |^{-1/2} \) is the unramified constituent of a representation of the form \( \rho_{\mu_1,\mu_2} = \text{Ind}_{Q_{\nu}}^{H_{F_\nu}} \mu_1 \circ \text{det} \otimes \mu_2 \circ \text{det} \). Here \( \mu_1,\mu_2 \) are unramified characters of \( F_{\nu}^* \), such that \( \tau_\nu \) is the unramified constituent of the representation of \( \text{GL}_4(F_{\nu}) \) induced from the standard Borel subgroup and its character defined by

\[
\text{diag}(t_1,\ldots,t_n) \mapsto \mu_1 \left( \frac{t_1}{t_2} \right) \mu_2 \left( \frac{t_2}{t_3} \right).
\]

\( Q \) is the standard parabolic subgroup of \( H \), whose Levi part is isomorphic to \( \text{GL}(2) \times \text{GL}(2) \). If \( \sigma_\psi^{(1)}(\tau) \) is nontrivial, then the Jacquet module with respect to \( (\text{Ind}_{F_\nu}^{H_{F_\nu}}(\psi_\nu))_{2,-1} \), \( J_{\text{Ind}_{Q_{\nu}}^{H_{F_\nu}}(\psi_\nu)}(\rho_{\mu_1,\mu_2}) \) is nontrivial. Thus, the proposition will be proved if we show that

\[
J_{\text{Ind}_{Q_{\nu}}^{H_{F_\nu}}(\psi_\nu)}(\rho_{\mu_1,\mu_2}) = 0.
\]

We use Bruhat theory. Let \( Q_2 \) be the standard parabolic subgroup of \( H \), whose Levi part is isomorphic to \( \text{GL}_2 \times \text{SO}_4 \). We first analyze \( J_{\text{Ind}_{Q_{\nu}}^{H_{F_\nu}}(\psi_\nu)}(\rho_{\mu_1,\mu_2}) \), where \( \eta = \mu_1 \circ \text{det} \) and \( \pi \) is an irreducible representation (later to be specified as \( \text{Ind}_{\mu_2 \circ \text{det}} \)). We apply Bruhat theory to study \( \text{Res}_{Q_2(F_\nu)}(\text{Ind}_{Q_{\nu}}^{H_{F_\nu}}(\psi_\nu)) \). This restriction has a filtration of \( Q_2(F_{\nu}) \)-modules, with subquotients parametrized by \( Q_2 \setminus H/Q_2 \). The quotient \( Q_2 \setminus H \) is isomorphic to the variety \( Y_2 \) of two dimensional isotropic subspaces of the (column) space \( F^8 \) equipped with the quadratic form preserved by \( H \). Let \( \{ e_1,\ldots,e_4,e_{-4},\ldots,e_{-1} \} \) be the standard basis of \( F^8 \). Let \( X^{(2)} = \text{Span}(e_1,e_2) \) be the standard two dimensional isotropic subspace. The isomorphism \( Q_2 \setminus H \cong Y_2 \) is given by \( Q_2 h \mapsto h^{-1} \cdot X^{(2)} \). The orbits of \( Q_2 \) in \( Y_2 \) are parametrized by \( r = \dim(X \cap X^{(2)}) \), and \( s = \dim(X \cap (X^{(2)})^\perp) \), \( X \in Y_2 \). Note that \( 0 \leq r \leq s \leq 2 \). A representative is

\[
X_{r,s} = \text{Span}\{e_1,\ldots,e_r;e_{r+1},\ldots,e_{2+s-r};e_{-(r+1)},\ldots,e_{-(2+s-r)}\}.
\]

Choose (a Weyl element, for example) \( w_{r,s} \in H \), such that \( w_{r,s}^{-1}X^{(2)} = X_{r,s} \). The corresponding subquotients for \( \text{Res}_{Q_2(F_\nu)}(\text{Ind}_{Q_{\nu}}^{H_{F_\nu}}(\psi_\nu)) \) are

\[
\Gamma_{r,s} = \text{Ind}_{w_{r,s}^{-1}Q_2(F_\nu)w_{r,s}Q_2(F_\nu)} \left( (\eta \otimes \pi) | \text{det} \right)^{w_{r,s}} \cdot \delta^{-1/2}.
\]
(The factor $\delta^{-1/2}$ appears in order to make the induction normalized.) Consider, for example, the case $r = 1$, $s = 2$. Here, we have

\begin{equation}
 w_{1,2}^{-1}Q_2(F_v)w_{1,2} \cap Q_2(F_v) = \left\{ \begin{pmatrix} a_{11} & a_{12} & x_{11} & x_{12} & x_{13} & x_{14} & y_{11} & y_{12} \\
 a_{22} & 0 & x_{22} & x_{23} & x_{24} & y_{21} & y_{11} & y_{22} \\
 b_{11} & b_{12} & b_{13} & b_{14} & x_{14} & x_{14} & x_{14} & x_{14} \\
 c_{11} & c_{12} & c_{13} & c_{14} & c_{14} & c_{14} & c_{14} & c_{14} \\
 c_{21} & c_{22} & b_{12} & x_{22} & x_{12} & x_{12} & x_{12} & x_{12} \\
 b_{11} & 0 & x_{11} & x_{11} & x_{11} & x_{11} & x_{11} & x_{11} \\
 a_{22} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} \\
 a_{22} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} & a_{12} \\
 \end{pmatrix} \in H_{F_v} \right\} := L_{12}
\end{equation}

The representation $\xi_{1,2} = (\eta \otimes \pi \cdot \delta_{Q_2}^{1/2})w_{1,2}$ takes elements of the form (4.40) to

\begin{equation}
 |a_{11}b_{11}|^{5/2} \mu_1(a_{11}b_{11})\pi \begin{pmatrix} c_{11} & 0 & x_{23} & c_{12} & x_{22} & a_{22} & y_{21} & x_{23} & 0 & a_{22}^{-1} & 0 \\
 c_{21} & 0 & x_{22} & c_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \end{pmatrix}.
\end{equation}

Let us prove that $J_{N_2(F_v), (\psi_v)_{2, -1}}(\Gamma_{1,2}) = 0$. Fit $\Gamma_{1,2}$ into an exact sequence $0 \to S_2 \to \Gamma_{1,2} \to S_1 \to 0$, where $S_2$ is the subspace of functions in $\Gamma_{1,2}$ supported inside $\Omega$, which consists of all matrices \( \begin{pmatrix} a & * & * \\
 b & a & * \\
 \end{pmatrix} \) in $Q_2(F_v)$, such that $a$ lies in the open Bruhat cell of $GL_2(F_v)$. The support of these functions (in $\Gamma_{1,2}$) is compact modulo $L_{12}(F_v)$. $S_1$ is the space of smooth functions on the complement of $\Omega$ inside $Q_2(F_v)$, where left $L_{1,2}(F_v)$ - translations act by (4.41), and the support is compact modulo $L_{1,2}(F_v)$. Thus, we have to show that $J_{N_2(F_v), (\psi_v)_{2, -1}}(S_i) = 0$; $i = 1, 2$. Let $f \in S_1$. We show that

\begin{equation}
 \int_{N_2(F_v)^-M} (\psi_v)^{-1}_{2, -1}(n)f(x_2(t) \begin{pmatrix} I_2 & t \\
 k & I_2 \end{pmatrix} n)dn = 0,
\end{equation}

for all $k \in SO_4(O_{F_v})$, $t \in F_v$;

$$x_2(t) = \begin{pmatrix} 1 & t \\
 1 & 1 \end{pmatrix}.$$ 

The support of the integrand in $t$ depends on $f$, so we may take $M$ large enough so that, in the support of $f$,

$$\begin{pmatrix} I_2 & t \\
 k & I_2 \end{pmatrix} x_2(t) \begin{pmatrix} I_2 & t^{-1} \\
 k & I_2 \end{pmatrix} \in N_2(F_v)^-M,$$

for all $k \in SO_4(O_{F_v})$. Making a change of variable in $n$, we may assume that $t = 0$ in (4.42). Consider now the subintegration in (4.42) on $x_1(z)$, $|z| \leq q_v^M$, where

$$x_1(z) = \begin{pmatrix} 1 & z \\
 1 & 1 \end{pmatrix}.$$ 

It gives

$$\int_{|z| \leq q_v^M} \psi_v^{-1}(z)f(x_1(z) \begin{pmatrix} I_2 & t \\
 k & I_2 \end{pmatrix} n)dz = \int_{|z| \leq q_v^M} \psi_v^{-1}(z)dz f(\begin{pmatrix} I_2 & t \\
 k & I_2 \end{pmatrix} n) = 0.$$
Here we used that $\xi_{1,2}(x_1(z)) = \text{id}$. This shows that $J_{N_2(F_\nu),\psi_{\nu,1,1}}(\Gamma_{1,2}) = 0$.

Let $f \in S_2$. We have to show that

$$
\int_{N_2(P_{\nu}^M)} (\psi_{\nu})_{2,-1}^{-1}(n)f\left(\mathbf{x}_2(t)\left[\begin{smallmatrix} w & 0 \\ k & w^* \end{smallmatrix}\right]\mathbf{x}_1(b)n\right)dn = 0,
$$

where $w = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$. As before, we may assume that $t = b = 0$. Now consider the subintegration on $y(u) = \left(\begin{smallmatrix} 1 & u & * \\ I & u & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right)$.

The corresponding $du$-integration (with $b = 0$) is

$$
\int_{u \in (P_{\nu}^M)^4} \psi^{-1}(u_2 - u_3)f\left(\mathbf{I}_{k} 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 \right)\left(\begin{smallmatrix} w & 0 \\ k & w^* \end{smallmatrix}\right)n)du
$$

$$
= \int_{u \in (P_{\nu}^M)^4} \psi^{-1}(u_2 - u_3)du \cdot f\left(\mathbf{I}_{k} 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 \right) = 0.
$$

This proves that $J_{N_2(F_{\nu}),\psi_{\nu,2,-1}}(S_2) = 0$. Similar arguments imply that $r$ cannot be 1 or 2. Thus, $r = 0$. Similar arguments imply also that for $r = 0$, $s$ cannot be 0 or 2.

Put $w_{0,1} = w_1$. Then

$$
(4.44) \quad w_1^{-1}Q_2w_1 \cap Q_2 = \left\{ \left(\begin{smallmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & x_{14} & 0 & 0 \\ a_{21} & a_{22} & 0 & x_{23} & x_{24} & x_{14} & 0 & 0 & 0 \\ b_{11} & b_{12} & b_{13} & b_{14} & x_{24} & x_{14} & 0 & 0 & 0 \\ c_{11} & c_{12} & b_{13} & x_{23} & 0 & 0 & 0 & 0 & 0 \\ c_{21} & c_{22} & b_{14} & x_{22} & 0 & 0 & 0 & 0 & 0 \\ b_{11}' & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21}' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21}'' & a_{21}'' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{smallmatrix}\right) \in H \right\} := L_1
$$

The representation $\xi_1 = ((\eta \otimes \pi) \cdot \psi_{\nu}^{1/2})w_1$ takes an element of the form (4.44) to

$$
(4.45) \quad \left| \frac{b_{11}}{a_{11}} \right|^{5/2} \mu_1 \left(\frac{b_{11}}{a_{11}}\right) \pi_\omega \left(\begin{smallmatrix} a_{22} & x_{23} & x_{22} & y \\ c_{22} & c_{21} & x_{22} & x_{23} \\ c_{12} & c_{11} & x_{23} & a_{22} \\ a_{21} & a_{21} & a_{21} & a_{21} \end{smallmatrix}\right).
$$

Here $\omega = \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \\ 1 \end{smallmatrix}\right)$.

As before, we consider appropriate analogs $S_{1}', S_{2}'$ of the spaces $S_1, S_2$, and it remains to show that

$$
(4.46) \quad \int_{N_2(P_{\nu}^M)} (\psi_{\nu})_{2,-1}^{-1}(n)f\left(\mathbf{I}_{k} 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 \right)dn = 0,
$$

for $k \in SO_4(\mathcal{O}_{F_\nu})$, and $\alpha_1 = I_2, \alpha_2 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$; $f$ is in $S_{1}', S_{2}'$ (respectively). In case $i = 2$, we consider the subintegration on $x_1(z), |z| \leq q_{\nu}^M$, and we get $\left(\int_{|z| \leq q_{\nu}^M} \psi_{\nu}^{-1}(z)dz\right)$.
Proposition 19

Here \( \sigma \)

\[
\text{Proposition 19:} \quad \text{We have an isomorphism of } SO_5(F_v)\text{-modules}
\]

\[
J_{N_1(F_v), (\psi), 1, -1} \left( \text{Ind}^{H_{F_v}}_{Q_1(F_v)}(\mu_1 \circ \det \otimes \mu_2 \circ \det) \right) \cong \text{Ind}^{SO_5(F_v)}_{B_v} \mu_1 \otimes \mu_2
\]

Here \( B \) is the standard Borel subgroup of \( SO_5 \).

Proof. — The method is the same as in Section 4.4. Again consider \( \eta = \mu_1 \circ \det \) on \( GL_2(F_v) \) and \( \pi = \text{Ind}^{SO_4(F_v)}_{Q_2(F_v)} \mu_2 \circ \det \). Let \( Q_1 \) be the standard parabolic subgroup of \( H \) which preserves an (isotropic) line. We analyze \( \text{Res}_{Q_1(F_v)} \left( \text{Ind}^{H_{F_v}}_{Q_2(F_v)} \eta \otimes \pi \right) \) using Bruhat theory. So consider \( Q_2 \setminus H/Q_1 \). Identify, as in Sec. 4.3, \( Q_2 \setminus H \cong Y_2 \). The orbits of \( Q_1 \) in \( Y_2 \) are determined by \( f = \dim(X \cap X^{(1)}) \), and \( s = \dim(X \cap (X^{(1)})^\perp) \), \( X \in Y_2 \). Here \( X^{(1)} = F_{e_1} \). Note that \( 0 \leq r \leq 1 \leq s \leq 2 \).

If \( r = 1 \), then \( e_1 \in X \), and since \( X \) is isotropic, we get that \( X \subset (X^{(1)})^\perp \), and so \( s = 2 \). Thus, we may take as a representative \( X = X^{(2)} \). The corresponding subquotient of \( \text{Res}_{Q_1(F_v)} \left( \text{Ind}^{H_{F_v}}_{Q_2(F_v)} \eta \otimes \pi \right) \) is

\[
T_{1,2} = \text{Ind}^{Q_1(F_v)}_{(Q_1 \cap Q_2)(F_v)} ((\eta \otimes \pi) \cdot \delta^{1/2} \delta^{-1/2})
\]
We have

\[(4.49) \quad Q_1 \cap Q_2 = \left\{ \begin{pmatrix} a_1 & * & * & * \\ a_2 & * & * & * \\ a_2^{-1} & * & * & * \\ a_1^{-1} \end{pmatrix} \in H \mid b \in SO_4 \right\},\]

and \((\eta \otimes \pi) \cdot d_{Q_2}^{1/2}\) takes an element of the form (4.49) to

\[|a_1 a_2|^{5/2} \mu_1(a_1 a_2) \pi(b).\]

Clearly, for \(f\) in the space of \(T_{1,2}\), and \(M \gg 0\),

\[(4.50) \quad \int_{N_1(P_{\nu}^{-M})} (\psi_\nu)^{1-1}(n)f\left(\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} n\right)dn = 0,\]

for any \(k \in SO_6(F_\nu)\). Indeed, \(f\left(\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} n\right) = f\left(\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} n\right)\), for any \(n \in N_1(F_\nu)\). This shows that \(J_{N_1(F_\nu)_{\nu}}(T_{1,2}) = 0\). Thus, we may assume that \(r = 0\). If \(s = 2\), we may take the representative \(X = \text{Span}\{e_2, e_3\}\). The corresponding representative in \(Q_2 \setminus H/Q_1\) can be taken to be \(w_2 = \begin{pmatrix} 1 \\ I_3 \\ I_3 \\ 0 \end{pmatrix}\) (so that \(w_2^{-1} X^{(2)} = X\)).

Let \(T_2 = \text{Ind}^c_{Q_2(F_\nu)w_2 \cap Q_1(F_\nu)}((\eta \otimes \pi) d_{Q_2}^{1/2}) w_2^{(-1/2)}\). We have

\[(4.51) \quad w_2^{-1} Q_2 w_2 \cap Q_1 = \left\{ \begin{pmatrix} a & 0 & x & z \\ b & y & v & z' \\ c & y' & x' & b' \\ 0 & 0 & a^{-1} \end{pmatrix} \in H \mid c \in SO_2 \right\}.\]

The representation \(\xi_2 = ((\eta \otimes \pi) d_{Q_2}^{1/2}) w_2\) takes an element of the form (4.51) to

\[(4.52) \quad |\det b|^{5/2} \mu_1(\det b) \pi^\omega\left(\begin{pmatrix} a & x & c \\ b & y & z \end{pmatrix}\right),\]

where \(\omega = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\). Consider, for \(f\) in the space of \(T_2\), \(M \gg 0\), and \(k \in SO_6(O_\nu)\),

\[(4.53) \quad \int_{N_1(P_{\nu}^{-M})} (\psi_\nu)^{-1}(n)f\left(\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} n\right)dn.\]

Consider the subintegration of (4.53) on \(n(v) = \begin{pmatrix} 1 & v & * \\ I_6 & v' & 1 \end{pmatrix},\) where \(v = (0,0,u_3, \ldots, u_6)k,\)

\(|u_i| \leq q_{\nu}^M\). By (4.52), we get

\[(4.54) \quad \int_{|u_i| \leq q_{\nu}^M} \psi_\nu^{-1}(0,0,u_3, \ldots, u_6) k \begin{pmatrix} 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix} \pi^\omega\left(\begin{pmatrix} 1 & u_3 & u_4 \ -u_3 u_4 \\ 1 & 0 & -u_4 \-u_3 \\ 0 \end{pmatrix}\right) f\left(\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} n\right)dn.\]

We must have

\[k \begin{pmatrix} 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} a & * \\ \vdots & * \\ 0 \end{pmatrix},\]
otherwise the $d(u_5, u_6)$-integration results in zero. For such $k$,

$$
k \begin{pmatrix}
0 \\
\vdots \\
1 \\
-1 \\
\vdots \\
0
\end{pmatrix} = \begin{pmatrix}
\ast & a & 0 & 0 & 0 \\
0 & \ast & a & 0 & 0 \\
0 & 0 & \ast & a & 0 \\
0 & 0 & 0 & \ast & a \\
0 & 0 & 0 & 0 & \ast
\end{pmatrix}, \quad |a| = 1.
$$

Thus (4.54) becomes (up to $q_r^{2M}$)

$$
\int_{|u_3| \leq q_r^M} \psi^{-1}_\nu(au_3 - a^{-1}u_4)\pi u \left( \begin{pmatrix}
1 & u_3 & -u_4 & u_3 u_4 \\
1 & 0 & 0 & -u_4 \\
0 & 1 & 0 & -u_3 \\
0 & 0 & 1 & -1
\end{pmatrix} \right) f\left( \begin{pmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) d(u_3, u_4),
$$

which is zero for $M$ large enough, exactly as in the end of Sec. 4.3. (This is a place to apply induction. Recall that $\pi = \text{Ind}_{Q_2(F_r)}^{SO_3(F_r)} \mu_2 \circ \det$.) Note that $k, n, a$ may be taken in compact sets, which depend on $f$ only. Finally, let $r = 0, s = 1$. Here, a corresponding representative is $w_1 = \begin{pmatrix} I_5 & 0 \\ 1 & -1 \end{pmatrix} \left( \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)$. Let

$$
T_1 = \text{Ind}_{w_1^{-1}Q_2(F_r)w_1 \cap Q_1(F_r)}^{\nu \otimes \pi} \mu_1 \otimes \delta^{1/2}_{Q_2} u_1 \delta^{-1/2}.
$$

We have

$$
w_1^{-1}Q_2w_1 \cap Q_1 = \left\{ \begin{pmatrix} a & 0 & 0 & x & 0 \\ b & y & z & x' & 0 \\ c & y' & 0 & x' & 0 \\ b^{-1} & 0 & 0 & a & -1 \
\end{pmatrix} \in H \mid c \in SO_4 \right\}.
$$

The representation $\xi_1 = ((\eta \otimes \pi) \cdot \delta^{1/2}_{Q_2}) w_1$ takes an element of the form (4.56) to

$$
J_{Q_1(F_r), \nu} \left( T_1 \right) = \text{Ind}_{w_1^{-1}Q_2(F_r)w_1 \cap Q_1(F_r)}^{\nu \otimes \pi} \mu_1 \otimes \pi \bigg|_{SO_3(F_r)},
$$

where $Q_1'$ is the standard parabolic subgroup of $SO_5$, which preserves an isotropic line. Finally, it is easy to see that $\pi \bigg|_{SO_3(F_r)} \cong \text{Ind}_{B_r'}^{SO_3(F_r)} \mu_2$, for $\pi = \text{Ind}_{Q_2(F_r)}^{SO_3(F_r)} \mu_2 \circ \det$. Here $B'$ is the standard Borel subgroup of $SO_3$. This completes the proof of Proposition 19. 

\[\Box\]

\textbf{References}


FROM CLASSICAL GROUPS TO GL_n


D. Soudry, School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat-Aviv, 69978 Israel • E-mail : soudry@math.tau.ac.il
ON THE JACQUET-LANGLANDS CORRESPONDENCE
IN THE COHOMOLOGY OF
THE LUBIN-TATE DEFORMATION TOWER

by

Matthias Strauch

Abstract. — Let $F$ be a local non-archimedean field, and let $X$ be a one-dimensional formal $\mathfrak{o}_F$-module over $\mathbb{F}_p$ of height $n$. The formal deformation schemes of $X$ with Drinfeld level structures give rise to a projective system of rigid-analytic spaces $(M_K)_K$, where $K$ runs through the compact-open subgroups of $G = GL_n(F)$. On the inductive limit $H^*_c$ of the spaces $H^*_c(M_K \otimes \mathbb{F}_p^\wedge, \mathbb{Q}_\ell)$ ($\ell \neq p$) there is a smooth action of $G \times B^\times$, $B$ being a central division algebra over $F$ with invariant $1/n$. For a supercuspidal representation $\pi$ of $G$ it follows from the work of Boyer resp. Harris-Taylor that in the Grothendieck group of admissible representations of $B^\times$ one has $\text{Hom}_G(H^*_c, \pi) = n \cdot (-1)^{n-1} \mathcal{JL}(\pi)$, $\mathcal{JL}$ denoting the Jacquet-Langlands correspondence. In this paper we propose an approach that is based on a conjectural Lefschetz trace formula for rigid-analytic spaces, and we calculate the contribution coming from the fixed points.

Résumé (Sur la correspondance de Jacquet-Langlands dans la cohomologie de la tour de déformations de Lubin-Tate)

Soient $F$ un corps local non-archimédiens et $X$ un $\mathfrak{o}_F$-module formel de hauteur $n$ sur $\mathbb{F}_p$. Les schémas de déformations de $X$ munis de structures de niveau de Drinfeld fournissent un système projectif d’espaces analytiques rigides $(M_K)_K$, où $K$ parcourt l’ensemble des sous-groupes compacts ouverts de $G = GL_n(F)$. La limite inductive $H^*_c$ des espaces $H^*_c(M_K \otimes \mathbb{F}_p^\wedge, \mathbb{Q}_\ell)$ ($\ell \neq p$) constitue une représentation virtuelle lisse du groupe $G \times B^\times$, $B$ étant une algèbre à division sur $F$ d’invariant $1/n$. Si $\pi$ est une représentation supercuspidale de $G$, les travaux de Boyer et Harris-Taylor impliquent que dans le groupe de Grothendieck des représentations admissibles de $B^\times$ on a la relation $\text{Hom}_G(H^*_c, \pi) = n \cdot (-1)^{n-1} \mathcal{JL}(\pi)$, $\mathcal{JL}$ désignant la correspondance de Jacquet-Langlands. Dans cet article nous proposons une approche de ce résultat fondé sur une formule des traces à la Lefschetz conjecturale, et nous calculons la contribution venant des points fixes.

2000 Mathematics Subject Classification. — 22E50, 11S37, 14G22, 14L05.

Key words and phrases. — Jacquet-Langlands correspondence, deformations of formal groups, rigid-analytic spaces, Lefschetz trace formula.
1. Introduction

Let $F$ be a non-archimedean local field with ring of integers $\mathfrak{o} = \mathfrak{o}_F$. Let $X$ be a formal $\mathfrak{o}$-module of $F$-height $n$ over the algebraic closure of the residue field of $\mathfrak{o}$. The functor of deformations of $X$ is representable by an algebra of formal power series in $n - 1$ variables over $\widehat{\mathfrak{o}}^{nr}$. Associated to this algebra there is a rigid-analytic space: the open polydisc of dimension $n - 1$. Introducing Drinfeld level structures gives rise to a tower of étale coverings of this space with pro-Galois group $GL_n(\mathfrak{o})$. Moreover, the automorphism group $\text{Aut}_\mathfrak{o}(X)$ of $X$ acts on the deformation space and its coverings, and this action commutes with the action of $GL_n(\mathfrak{o})$. It is convenient to work not only with deformations in the strict sense, i.e. ones equipped with an isomorphism of the special fibre to $X$, but with deformations coming along with a quasi-isogeny of the special fibre to $X$. In this way one obtains an infinite disjoint union of such towers (indexed by the height of the quasi-isogeny), all being non-canonically isomorphic, and on this tower there is then an action of $GL_n(F) \times B^\times$, where $B = \text{End}_\mathfrak{o}(X) \otimes_\mathfrak{o} F$ is a central division algebra over $F$ of dimension $n^2$. The inductive limit $H^i_c$ of the $\ell$-adic étale cohomology groups with compact support of these spaces (after base change to an algebraic closure of $F$) furnish smooth/continuous representations of $GL_n(F) \times B^\times \times W_F$, where $W_F$ is the Weil group of $F$, and $\ell \neq p$.

Carayol’s conjecture predicts that for a supercuspidal representation $\pi$ of $GL_n(F)$ the following relation holds true:

$$\text{Hom}_{GL_n(F)}(H^{n-1}_c, \pi) = J\mathcal{L}(\pi) \otimes \sigma(\pi),$$

where $J\mathcal{L}(\pi)$ is the representation of $B^\times$ that is associated to $\pi$ by the Jacquet-Langlands correspondence, and $\sigma(\pi)$ is, up to twist and dualization, the representation of $W_F$ that is associated to $\pi$ by the local Langlands correspondence for $GL_n$.

Cf. [Ca1], sec. 3.3, for a more precise statement also covering the case of non-cuspidal discrete series representations.

In the equal characteristic case, this conjecture has been proven by P. Boyer [Bo]. In the mixed characteristic case it may be regarded as being true by the work of M. Harris and R. Taylor [HT]. Although they do not state it this way, it seems likely that Carayol’s conjecture follows without difficulty from what has been proven in their book, cf. [Ca2]. Both proofs (equal and mixed characteristic case) use global methods.

In this paper we investigate the alternating sum

$$\text{Hom}_{GL_n(F)}(H^*_c, \pi) := \sum (-1)^i \text{Hom}_{GL_n(F)}(H^i_c, \pi)$$

as a virtual representation of $B^\times$ by a purely local method. We do not obtain any information about the Weil group representation, except its dimension. Moreover, we pay only attention to the part of the correspondence that concerns the supercuspidal representations. Our approach is based on a conjectural Lefschetz trace formula for
rigid analytic spaces, and has its origin in Faltings’ paper [Fa]. Faltings investigated there the corresponding situation of Drinfeld’s symmetric spaces and their coverings. In both cases the problem arises that the spaces under consideration are not proper. Hence we cannot expect to express the alternating sum of traces on the cohomology groups as a sum of fixed point multiplicities. Indeed, simple calculations show that in general there will be an extra term coming from the “boundary”. (In the case considered by Faltings however, the “boundary term” turns out to be zero; this is definitively not true in our case, and this is why the situation considered here seems to be more difficult.) In the case $n = 2$ one can use a trace formula for one-dimensional rigid curves proven by R. Huber, cf. [Hu3]. Huber’s trace formula is applied to certain compactifications (in the category of adic spaces) of quasi-compact subspaces, and Huber’s trace formula gives an expression of the trace in terms of usual fixed point multiplicities and a contribution from the finitely many compactifying points. While trying to extend Huber’s formula to the higher-dimensional case, the author found out that there is another kind of canonical “quasi-compactification” in the category of adic spaces, namely the projective limit of all admissible blow-ups of the corresponding formal schemes representing the deformation functors. The advantage of this compactification is that one has an immediate modular interpretation of the boundary: the boundary has a natural stratification and the geometry and combinatorial structure of the strata can be related to parabolic subgroups of $GL_n(\mathfrak{o}/(\varpi^m))$. Unfortunately, this seems still to be not sufficient to prove that the boundary term (=$\text{actual trace minus the number of fixed points}$) is a sum of parabolically induced virtual characters. We are finally led to consider certain “tubular neighborhoods” of the strata in the boundary. These spaces are insofar interesting as they can be considered as examples of truly non-archimedean spaces over higher-dimensional local fields. But work in this direction has not yet been finished, and hence is not included in this paper. Here we therefore assume that the trace on the alternating sum of the cohomology groups has an appropriate shape, cf. sec. 3.5. This conjecture seems to be geometrically justifiable (cf. sec. 3.10), and it turns out that the correspondence between the representations of $GL_n(F)$ and $B^\times$ is then given by the number of fixed points (at least as long we consider supercuspidal representations of $GL_n(F)$).

Acknowledgements. — It is a pleasure to thank L. Fargues, E. Grosse-Klönne, M. Harris, M. Kisin, R. Huber and P. Schneider for helpful and interesting discussions on this subject. Further, the author wants to express his sincere thanks to the organizers of the “Semestre automorphe” at the Institut Henri Poincaré in the first half of 2000 for the invitation. Participation in this activity was made possible by a scholarship of the European Program Network “Arithmetic Algebraic Geometry”; warmest thanks go to N. Schappacher. Part of this work was done while the author had a post-doctoral position at the Sonderforschungsbereich 478 at the University of Münster; the financial support of the DFG is gratefully acknowledged.
Finally, the author is grateful to V. Berkovich, A. Besser, P. Broussous, H. Carayol, C. Cunningham, I. Fesenko, T. Haines, M. Hien, E. de Shalit and J. Tilouine for giving him the opportunity to talk about work related to this subject in workshops or seminars.

2. Deformation spaces and their cohomology groups

2.1. Let $F$ be a non-archimedean local field with ring of integers $\mathfrak{o}$. Fix a generator $\varpi$ of the maximal ideal of $\mathfrak{o}$, and put $\mathbb{F}_q = \mathfrak{o}/(\varpi)$, $q$ being the cardinality of the residue class field. Moreover, we denote by $F$ the residue field of the maximal unramified extension $\mathfrak{o}^{nr}$ of $\mathfrak{o}$, and we let $v : F^\times \to \mathbb{Z}$ the valuation determined by $v(\varpi) = 1$.

Let $X$ be a one-dimensional formal group over $F$ that is equipped with an action of $\mathfrak{o}$, i.e. we assume given a homomorphism $\mathfrak{o} \to \text{End}_F(X)$ such that the action of $\mathfrak{o}$ on the tangent space is given by the reduction map $\mathfrak{o} \to \mathbb{F}_q \subset F$. Such an object is called a formal $\mathfrak{o}$-module over $F$. Moreover, we assume that $X$ is of $F$-height $n$, meaning that the kernel of multiplication by $\varpi$ is a finite group scheme of rank $q^n$ over $F$.

It is known that for each $n \in \mathbb{Z}_{>0}$ there exists a formal $\mathfrak{o}$-module of $F$-height $n$ over $F$, and that it is unique up to isomorphism [Dr], Prop. 1.6, 1.7.

Let $C$ be the category of complete local noetherian $\hat{\mathfrak{o}}^{nr}$-algebras with residue field $F$. A deformation of $X$ over an object $R$ of $C$ is a pair $(X, \iota)$, consisting of a formal $\mathfrak{o}$-module $X$ over $R$ which is equipped with an isomorphism $\iota : X \to X_F$ of formal $\mathfrak{o}$-modules over $F$, where $X_F$ denotes the reduction of $X$ modulo the maximal ideal $m_R$ of $R$. Sometimes we will omit $\iota$ from the notation.

Following Drinfeld [Dr], sec. 4B, we define a structure of level $m$ on a deformation $X$ over $R$ ($m \geq 0$) as an $\mathfrak{o}$-module homomorphism

$$\phi : (\varpi^{-m}\mathfrak{o}/\mathfrak{o})^n \to m_R,$$

such that $[\varpi]_X(T)$ is divisible by

$$\prod_{a \in (\varpi^{-1}\mathfrak{o}/\mathfrak{o})^n} (T - \phi(a)).$$

Here, $m_R$ is given the structure of an $\mathfrak{o}$-module via $X$, and $[\varpi]_X(T)$ is the power series that gives multiplication by $\varpi$ on $X$ (after having fixed a coordinate $T$).

For each $m \geq 1$ let $K_m = 1 + \varpi^m M_n(\mathfrak{o})$ be the $m$th principal congruence subgroup inside $K_0 = GL_n(\mathfrak{o})$. Define the following set-valued functor $\mathcal{M}^{(0)}_{K_m}$ on the category $C$. For an object $R$ of $C$ put

$$\mathcal{M}^{(0)}_{K_m}(R) = \{(X, \iota, \phi) \mid (X, \iota) \text{ is a def. over } R, \phi \text{ is a level-$m$-structure on } X\}/\simeq,$$

where $(X, \iota, \phi) \simeq (X', \iota', \phi')$ iff there is an isomorphism $(X, \iota) \to (X', \iota')$ of formal $\mathfrak{o}$-modules over $R$, which is compatible with the level structures.
2.2. **Theorem (Drinfeld, [Dr] Prop. 4.3).** — The functor $\mathcal{M}_{K_m}^{(0)}$ is representable by a regular local ring, which is a finite flat algebra over $\widehat{\mathcal{O}}[[u_1, \ldots, u_{n-1}]]$ which itself represents $\mathcal{M}_{K_0}^{(0)}$.

The fact that $\widehat{\mathcal{O}}[[u_1, \ldots, u_{n-1}]]$ represents $\mathcal{M}_{K_0}^{(0)}$ is due to Lubin and Tate (for $F = \mathbb{Q}_p$) [LT]. For this reason $\mathcal{M}_{K_0}^{(0)}$, the deformation space without level structures, is sometimes called the Lubin-Tate moduli space, cf. [HG], [Ch].

2.3. Let $X$ be a formal $\mathcal{O}$-module over $R \in \mathcal{C}$ such that $X_F$ has $F$-height $n$, in which case we say that the formal $\mathcal{O}$-module $X$ has height $n$. As pointed out above, $X_F$ is then isomorphic to $X$. Denote $\text{End}_\mathcal{O}(X)$ by $\mathcal{O}_B$; this $\mathcal{O}$-algebra is the maximal compact subring of $B := \mathcal{O}_B \otimes \mathcal{O} \mathcal{F}$, which is a central division algebra over $\mathcal{F}$ with invariant $1/n$.

Any non-zero element of $\text{Hom}_\mathcal{O}(X, X_F) \otimes \mathcal{O}$ is called an $\mathcal{O}$-quasi-isogeny from $X$ to $X_F$. For such an element $\iota$ we define its $F$-height by

$$ F\text{-height}(\iota) = F\text{-height}(\varpi^r \iota) - nr, $$

where we choose some $r \in \mathbb{Z}$ such that $\varpi^r \iota$ lies in $\text{Hom}_\mathcal{O}(X, X_F)$, and for an element $\iota'$ of this latter set, its $F$-height is $h$ if $\ker(\iota')$ is a group scheme of rank $q^h$ over $\mathcal{F}$.

Define for $h \in \mathbb{Z}$ a set-valued functor $\mathcal{M}_{K_m}^{(h)}$ on $\mathcal{C}$ as follows: for $R \in \mathcal{C}$ the set $\mathcal{M}_{K_0}^{(h)}(R)$ consists of equivalence classes of triples $(X, \iota, \phi)$, where $X$ is a formal $\mathcal{O}$-module of height $n$ over $R$, $\iota$ is an $\mathcal{O}$-quasi-isogeny from $X$ to $X_F$ of $F$-height $h$, and $\phi$ is a level-$m$-structure on $X$. Now put

$$ \mathcal{M}_{K_m} = \coprod_{h \in \mathbb{Z}} \mathcal{M}_{K_m}^{(h)}. $$

By the uniqueness of $X$ (up to isomorphism), we have $\mathcal{M}_{K_m}^{(h)} \simeq \mathcal{M}_{K_m}^{(0)}$, but there is no distinguished isomorphism.

2.4. There is an action of $B^\times$ from the right on the functors $\mathcal{M}_{K_m}$ given by

$$ [X, \iota, \phi], b = [X, \iota \circ b, \phi], $$

where we denote by $[X, \iota, \phi]$ the equivalence class of $(X, \iota, \phi)$, and where $b \in B^\times$. If $[X, \iota, \phi]$ belongs to $\mathcal{M}_{K_0}^{(h)}(R)$, then $[X, \iota, \phi], b$ is an element of $\mathcal{M}_{K_m}^{(h+\varepsilon(N(b)))}(R)$, where $N : B \rightarrow F$ denotes the reduced norm.

Next we will describe the “action” of the group $G = GL_n(\mathcal{O})$ on the tower $(\mathcal{M}_{K_m})_m$.

Let $g \in G$ and suppose first that $g^{-1} \in M_n(\mathcal{O})$. For integers $m \geq m' > 0$ such that

$$ g\mathcal{O}^n \subset \varpi^{-(m-m')}\mathcal{O}^n $$

(this inclusion is meant to be inside $F^n$) we will define a natural transformation

$$ g : \mathcal{M}_{K_m} \rightarrow \mathcal{M}_{K_m}. $$
Let \([X, \iota, \phi] \in \mathcal{M}_{K_m}(R), R \in \mathcal{C}\). The following construction gives an element \([X', \iota', \phi']\) of \(\mathcal{M}_{K_m}(R)\) that is the image under the corresponding map 

\[ g_R : \mathcal{M}_{K_m}(R) \to \mathcal{M}_{K_m'}(R) \]

on \(R\)-valued points and it will be denoted by \([X, \iota, \phi] \cdot g\).

The conditions imposed on \(g\) show that \(g\) contains a subgroup of \(\varpi^{-m}\mathfrak{o}^n/\mathfrak{o}^n\), so we may define a formal \(\sigma\)-module \(X'\) over \(R\) by taking the quotient of \(X\) by the finite subgroup \(\phi(g\mathfrak{o}^n/\mathfrak{o}^n)\) (cf. [Dr], Prop. 4.4):

\[ X' = X/\phi(g\mathfrak{o}^n/\mathfrak{o}^n). \]

Moreover, left multiplication with \(g\) induces an injective homomorphism

\[ \varpi^{-m'}\mathfrak{o}^n/\mathfrak{o}^n \to \varpi^{-m}\mathfrak{o}^n/g\mathfrak{o}^n = (\varpi^{-m}\mathfrak{o}^n/\mathfrak{o}^n)/(g\mathfrak{o}^n/\mathfrak{o}^n) \]

and the composition with the morphism induced by \(\phi\),

\[ (\varpi^{-m}\mathfrak{o}^n/\mathfrak{o}^n)/(g\mathfrak{o}^n/\mathfrak{o}^n) \to X/\phi(g\mathfrak{o}^n/\mathfrak{o}^n) = X', \]

gives by [Dr], Prop. 4.4, a level-\(m'\)-structure

\[ \phi' : \varpi^{-m'}\mathfrak{o}^n/\mathfrak{o}^n \to X'[\varpi^{m'}]. \]

Finally define \(\iota'\) to be the composition of \(\iota\) with the projection

\[ X_{\mathbb{F}} \to (X')\mathbb{F}. \]

One checks easily that this construction is independent of the representative \((X, \iota, \phi)\) and gives indeed a morphism of functors. If \([X, \iota, \phi] \in \mathcal{M}_{K_m}^{(h)}(R)\) then \([X, \iota, \phi] \cdot g\) is an element of \(\mathcal{M}_{K_m}^{(h-v(\det g))}(R)\).

For an arbitrary element \(g \in G\), choose \(r \in \mathbb{Z}\) such that \((\varpi^{-r}g)^{-1} \in M_n(\mathfrak{o})\). Then, for \(m \geq m' \geq 0\) with

\[ \varpi^{-r}g\mathfrak{o}^n \subset \varpi^{-(m-m')}\mathfrak{o}^n \]

and \([X, \iota, \phi] \in \mathcal{M}_{K_m}(R)\), define \([X', \iota', \phi'] = [X, \iota, \phi].(\varpi^{-r}g)\) as above and put

\[ [X, \iota, \phi] \cdot g = [X', \iota' \circ \varpi^{-r}, \phi']. \]

This construction gives natural transformation

\[ g : \mathcal{M}_{K_m} \to \mathcal{M}_{K_m'}, \]

which depends neither on \(\varpi\) nor on the integer \(r\) (among all \(r\)'s with \(\mathfrak{o}^n \subset \varpi^{-r}g\mathfrak{o}^n \subset \varpi^{-(m-m')}\mathfrak{o}^n\)). In particular, one gets for each \(m\) an action of \(GL_n(\mathfrak{o})\) on \(\mathcal{M}_{K_m}\) which commutes with the action of \(B^x\).
2.5. The next step is to introduce the analytic spaces whose ℓ-adic étale cohomology
groups we are going to study in this paper.

There are different possible methods how to construct such spaces, namely as
rigid-analytic spaces, as non-archimedean analytic spaces as defined and studied by
V.G. Berkovich [Be1], or finally as adic spaces in the sense of R. Huber [Hu1]. For
each of these kinds of spaces there has been defined an étale cohomology theory
([dJ-vdP], [Be2], [Hu2]) and there are comparison theorems assuring that the re-
sulting cohomology groups for the spaces considered by us are the same ([Hu2],
sec. 8.3). For the purpose of this paper it is not important with which construction
we actually work. The reader is invited to use the theory he feels most comfortable
with. We will give brief references where the actual constructions have been carried
out.

It follows from Theorem 2.2 that each of the functors $\mathcal{M}(h)_m^K$ is representable by
a regular local $\mathfrak{q}^{nr}$-algebra of Krull dimension $n$, $R(h)_m$ say, which are for varying $h$
(but fixed $m$) non-canonically isomorphic. We give $R(h)_m$ the topology defined by the
maximal ideal, and denote also by $M(h)_m^K$ the formal spectrum $\text{Spf}(R(h)_m)$, and by $M^K_m$
the disjoint union over all $h \in \mathbb{Z}$.

A construction due to P. Berthelot, generalizing Raynaud’s construction for $\wp$-
adic formal schemes, associates a rigid-analytic space to $M^K_m$ ([Ber], ch. 0, or [RZ],
sec. 5.1). In the context of non-archimedean analytic spaces, Berkovich has given a
construction of such spaces associated to formal schemes of this type [Be3]. Finally,
R. Huber defines in [Hu1], sec. 4, an adic space $t(M(h)_m^K)$ associated to $M(h)_m^K$
and to $M^K_m$. The set of points of the underlying topological space consists of all (equiva-
lence classes of) continuous valuations $|\cdot|_v$ on $R(h)_m$ such that $|f|_v \leq 1$ for all $f \in R(h)_m$.
The set of valuations $|\cdot|_v$ with $|\wp|_v = 0$ is a closed subset which we denote by $V(\wp)$.
The open complement inherits the structure of an adic space and we put

$$M^K_m = t(M^K_m) - V(\wp), \quad M^K_m = \coprod_{h \in \mathbb{Z}} M(h)_m^K.$$  

There are obvious canonical maps given by restricting the level structure

$$M^K_m \rightarrow M^K_{m'},$$

for $m \geq m'$, which are étale and galois with Galois group $K_{m'}/K_m$. In particular
the Galois group of $M^K_m$ over $M^K_0$ is $GL_n(\mathfrak{o}/(\wp^m))$. By 2.2 each space $M^K_0$
is isomorphic to an open polydiscs of dimension $n - 1$; in particular:

$$M^K_0(F^\prime) \simeq \{ (z_1, \ldots, z_{n-1}) \in (F^\prime)^{n-1} | \text{for all } i : |z_i| < 1 \}.$$

For an open subgroup $K \subset K_0$ we choose a positive integer $m$ such that $K_m$ is
normal in $K$, and we define

$$M^K = M^K_m/(K/K_m).$$
Note that the action of $K/K_m$ on $M_{K_m}$ respects the components $M_{K_m}^{(h)}$, hence we let $M_{K}^{(h)}$ be the quotient of $M_{K_m}^{(h)}$ by $K/K_m$. Let $K \subset K_0$ be as above, and let $g \in G$ such that $g^{-1}Kg \subset K_0$. Choose $m \geq m' \geq 0$ such there is a morphism
\[ g : M_{K_m} \longrightarrow M_{K_{m'}}, \]
as defined in the preceding section. Assume $K_m$ to be normal in $K$, $K_{m'}$ to be normal in $g^{-1}Kg$, and $g^{-1}Kg \subset K_{m'}$. Then we have an induced morphism
\[ g : M_K = M_{K_{m'}}/(K/K_m) \longrightarrow M_{K_{m'}}/(g^{-1}Kg/K_{m'}) = M_{g^{-1}Kg}, \]
which is in fact an isomorphism and does not depend on the specific choices. This allows us to define $M_K$ for arbitrary compact-open subgroups $K$ of $G$: choose a normal subgroup $K' \subset K$ which lies in $K_0$, hence for each $g \in K$ we have just defined an isomorphism $g : M_{K'} \longrightarrow M'_K$, hence we put
\[ M_K = M_{K'}/(K/K'). \]
Again we can define $M_{K'}^{(h)}$ as the quotient of $M_K^{(h)}$ by $K/K'$. Consequently, for any compact-open subgroup $K$ of $G$ and any $g \in G$ there is an isomorphism
\[ g : M_K \longrightarrow M_{g^{-1}Kg}. \]
Via this construction, the tower $(M_K | K \subset G$ compact-open$)$ is equipped with a natural action of $G \times B^\times$ from the right.

2.6. Finally, we introduce the cohomology groups. We use the étale cohomology theory as developed by Huber [Hu2], respectively Berkovich [Be2]. Because of the comparison theorems in [Hu2], sec. 8.3, we can and will use results of V.G. Berkovich for the étale cohomology of non-archimedean analytic spaces. So far, the cohomology theories and the results concern mostly the cohomology of torsion sheaves, and a general theory of $\ell$-adic cohomology has not been developed yet. Nevertheless, for the spaces considered by us, it is not difficult to show the finiteness of
\[ H^i_c(M_K^{(h)} \otimes \overline{\mathcal{F}_{nr}}, \mathbb{Q}_\ell) := \left( \lim_{\rightarrow} H^i_c(M_K^{(h)} \otimes \overline{\mathcal{F}_{nr}}, \mathbb{Z}/l^r\mathbb{Z}) \right) \otimes_{\mathbb{Z}_l} \mathbb{Q}_\ell \]
as a $\mathbb{Q}_\ell$-vector space (cf. [Be4]). The essential ingredient in showing this is the fact that $M_{K_m}^{(h)}$ is a formal scheme which is the completion of a scheme of finite type over $\mathcal{O}_{\mathcal{S}}$ at a closed point of the special fibre. This in turn follows from the very proof of the representability result 2.2, cf. [Dr], Prop. 4.3. Moreover, these cohomology groups are non-zero only in degree $i$ for $n - 1 \leq i \leq 2(n - 1)$ by [Be3], Th. 6.1, Cor. 6.2.

Next we put
\[ H^i_c(M_K) = H^i_c(M_K \otimes \overline{\mathcal{F}_{nr}}, \mathbb{Q}_\ell) = \bigoplus_{h \in \mathbb{Z}} H^i_c(M_K^{(h)} \otimes \overline{\mathcal{F}_{nr}}, \mathbb{Q}_\ell). \]
On each $\mathbb{Q}_\ell$-vector space $H^i_c(M_K)$ there is an induced action of $K_0 \times B^\times$ and for each $g \in G$ there is an isomorphism

$$H^i_c(M_{g^{-1}Kg}) \longrightarrow H^i_c(M_K).$$

These give rise to a representation of $G \times B^\times$ on

$$H^i_c := \lim_{\longrightarrow} H^i_c(M_K),$$

where the limit is taken over all compact-open subgroups $K$ of $G$.

2.7. Theorem (Berkovich). — The action of $G \times B^\times$ on $H^i_c$ is smooth.

Proof. — Any element of $H^i_c$ lies in a cohomology group $H^i_c(M_K)$ on which the action of $K \subset G$ is trivial. Therefore $G$ acts smoothly. It is a non-trivial result due to Berkovich that the action of $B^\times$ on $H^i_c(M_K)$ is smooth, cf. [Be3], introduction.

Remark. — The inertia group $\text{Gal}(F_{\text{sep}}/F_{\text{nr}})$ acts also on $H^i_c(M_K)$, and this action can be extended to an action of the Weil group $W_F$, cf. [Bo], Prop. 2.3.2, [RZ], sec. 3.48. Then one gets a smooth/continuous action of $G \times B^\times \times W_F$ on $H^i_c$. In this paper however we pay only attention to the representations of $G$ and $B^\times$.

Let $\pi$ be a supercuspidal representation of $G$, and let $JL(\pi)$ be the representation of $B^\times$ associated to $\pi$ by the Jacquet-Langlands correspondence. The following theorem is implied by Boyer’s Theorem, [Bo], Th. 3.2.4, in the equal characteristic case, and it follows from the work of Harris and Taylor [HT] in the mixed characteristic case.

2.8. Theorem. — For each $i$ the representation $\text{Hom}_G(H^i_c, \pi)$ of $B^\times$ is finite-dimensional and smooth, and in the Grothendieck group of admissible representations of $B^\times$ the following equality holds:

$$\sum_i (-1)^i \text{Hom}_G(H^i_c, \pi) = n \cdot (-1)^{n-1} \cdot JL(\pi).$$

As a definition, we put

$$\text{Hom}_G(H^*_c, \pi) := \sum_i (-1)^i \text{Hom}_G(H^i_c, \pi),$$

where we consider the right hand side as an element of the Grothendieck group of admissible representations of $B^\times$.

2.9. In [HT] this result comes only as a by-product of a detailed study of the cohomology groups of certain Shimura varieties attached to unitary groups coming from division algebras, and the precise investigation of the reduction of these varieties at bad primes. Similarly, Boyer’s proof ([Bo]) in the equal characteristic case is based on the study of the bad reduction of Drinfeld modular varieties.

In the next paragraph, we discuss a purely local way towards this theorem, that is based on a Lefschetz type trace formula.
3. The approach via a Lefschetz trace formula

3.1. Let \( \pi \) be a supercuspidal representation of \( G = GL_n(F) \). By the fundamental result of Bushnell-Kutzko [BK] and Corwin [Co], we know that \( \pi \) is induced from a (finite-dimensional) irreducible representation \( \lambda \) of some open subgroup \( K_\pi \subset G \) that contains and is compact modulo the centre of \( G \), cf. [BK], Th. 8.4.1, for a more precise statement. Hence we may write

\[
\pi = c \cdot \text{Ind}_{K_\pi}^G(\lambda) = \text{Ind}_{K_\pi}^G(\lambda),
\]

where the second equality holds by [Bu], Th. 1. Moreover, the character of \( \pi \) is a locally constant function on the set of elliptic regular elements in \( G \) (i.e. those whose characteristic polynomial is separable and irreducible), and for such an element \( g \in G \) we have

\[
\chi_\pi(g) = \sum_{g' \in G/K_\pi, (g')^{-1}gg' \in K_\pi} \chi_\lambda((g')^{-1}gg').
\]

For regular elliptic \( g \) the number of elements \( g' \in G/K_\pi \) such that \( (g')^{-1}gg' \in K_\pi \) is finite. This formula is due to Harish-Chandra, proofs can be found in [He] and [Sa]. For the rest of this section we fix \( \pi, K_\pi, \) and \( \lambda \) with this property.

3.2. For \( \pi \) as above, the representation \( \rho = JL(\pi) \) is characterized by the following identity. Let \( g \in G \) and \( b \in B^\times \) be regular elliptic elements with the same characteristic polynomial. Then the following character relation holds

\[
\chi_\rho(b) = (-1)^{n-1} \cdot \chi_\pi(g),
\]

cf. [DKV], introduction, [Ro]. Th. 5.8., [Ba].

3.3. For \( \pi \) as in (3.1) we will analyze \( \text{Hom}_G(H^*_c, \pi) \) as a representation of \( B^\times \). Note first that by Frobenius reciprocity

\[
\text{Hom}_G(H^*_c, \pi) = \text{Hom}_{K_\pi}(H^*_c, \lambda).
\]

Choose \( c \in \mathbb{Q}_\ell \) such that \( \lambda(\varpi) = c^n \), and define a character \( \zeta \) of \( G \) by \( \zeta(g) = c^{-v(\det(g))} \). Then:

\[
\text{Hom}_{K_\pi}(H^*_c, \lambda) = \text{Hom}_{K_\pi}(H^*_c \otimes \zeta, \lambda \otimes \zeta)
\]

\[
= \text{Hom}_{K_\pi}(\langle \varpi \cdot v | v \in H^*_c, \lambda \otimes \zeta \rangle).
\]

where \( \varpi \cdot v \) denotes the action of \( \varpi \), considered as an element of \( G \), on \( v \), considered as an element of \( H^*_c \).

Next, \( \langle H^*_c \otimes \zeta | v \in H^*_c \rangle \) is isomorphic, as a representation of \( G \times B^\times \), to the natural representations of \( G \times B^\times \) on

\[
\left( \lim_{K^\prime} H^*_c(M_{K}/\varpi^{K}) \right) \otimes \xi,
\]

ASTÉRISQUE 298
where the limit is taken over all compact-open subgroups $K'$ of $G$, and $\xi$ is the character of $B^x$ given by $\xi(b) = e^{-\pi(Nrd(b))}$. The map is defined as follows: an element $\alpha \in H^*_c(M_K^{(h)} \otimes_k)$ is mapped to $e^h \varpi^{-k} \alpha \in H^*_c(M_K^{(h_0)} \otimes_k)$, where $h = h_0 + nk$ with $0 \leq h_0 < n$. It is not difficult to check that this is a $G \times B^x$-equivariant isomorphism.

Hence we get the following identity of representation of $B^x$:

$$\text{Hom}_G(H^*_c, \pi) = \text{Hom}_{K_\pi} \left( \lim_{K'} H^*_c(M_{K'}/\varpi^Z), \chi \otimes \zeta \right) \otimes \xi^\vee.$$ 

Let $K \subset K_0$ be an open subgroup that is normal in $K_\pi$ and such that $\lambda|_K$ is a multiple of the trivial representation of $K$. Then

$$\left( \lim_{K'} H^*_c(M_{K'}/\varpi^Z) \right)^K = \lim_{K'} H^*_c(M_{K'}/\varpi^Z)^K = H^*_c(M_K/\varpi^Z),$$

and therefore

$$\text{Hom}_{K_\pi} \left( \lim_{K'} H^*_c(M_{K'}/\varpi^Z), \chi \otimes \zeta \right) \otimes \xi^\vee = \text{Hom}_\Gamma \left( H^*_c(M_K/\varpi^Z), \chi \otimes \zeta \right) \otimes \xi^\vee,$$

where we have put

$$\Gamma = K_\pi/\varpi^Z K;$$

disfinite group. Note that we finally arrived at an expression for $\text{Hom}_G(H^*_c, \pi)$ that involves only finite-dimensional vector spaces.

**3.4.** The next step is to compute the trace of a regular elliptic element $b \in B^x$ on this space. Using the result of the preceding section, the identity below is elementary:

$$\text{tr}(b \mid \text{Hom}_G(H^*_c, \pi)) = \frac{\xi(b)^{-1}}{\# \Gamma} \sum_{\gamma \in \Gamma} \text{tr}((\gamma, b^{-1}) \mid H^*_c(M_K/\varpi^Z)) \cdot \chi_{\lambda \otimes \zeta}(\gamma^{-1}).$$

Of course, this is the point where the Lefschetz trace formula comes in, because we would like to replace the trace of $(\gamma, b^{-1}) \in \Gamma \times B^x$ on the virtual representation $H^*_c(M_K/\varpi^Z)$ by an expression involving the number of fixed points. Our spaces, like $M_K/\varpi^Z$, are not proper however, so we cannot expect to get an expression involving only the number of fixed points, and indeed, there is in general an additional term.

In the case $n = 2$ one can use a trace formula which has been proved by R. Huber ([Hu3]) to get a manageable description of this “boundary term”. For the general case we have the following

**3.5. Conjecture.** — In the setting and with the notations introduced above, there is a trace formula of the following form:

$$\text{tr}((\gamma, b^{-1}) \mid H^*_c(M_K/\varpi^Z)) = \text{Fix}_K(\gamma, b^{-1}) + \beta_K(\gamma, b^{-1}),$$

where $\text{Fix}_K(\gamma, b^{-1})$ denotes the number (counted with multiplicity) of fixed points of $(\gamma, b^{-1})$ on $(M_K/\varpi^Z)(\overline{F})$ and $\beta_K(\gamma, b^{-1})$ has the property that for $\lambda$ as above

$$\sum_{\gamma \in \Gamma} \beta_K(\gamma, b^{-1}) \cdot \chi_{\lambda \otimes \zeta}(\gamma^{-1}) = 0.$$
3.6. Remark. — At the end of this section will give some heuristic arguments justifying this conjecture. For the moment, let us mention that the last formula has the following representation theoretic meaning. Firstly, for fixed \( b \) the function \( \gamma \mapsto \beta_K(\gamma, b^{-1}) \) is a class function on the finite group \( \Gamma \), and can hence be written as a sum

\[
\beta_K(\cdot, b^{-1}) = \sum_{\tau} \alpha_\tau \chi_\tau,
\]

where \( \tau \) runs over the set of equivalence classes of irreducible representations of \( \Gamma \). Those \( \tau \) with non-zero \( \alpha_\tau \) may be called the representations that occur in the boundary. Therefore the preceding formula signifies that:

No representation that gives rise to a supercuspidal representation occurs in the boundary.

3.7. Let us assume that such a trace formula exists. Then the condition on the boundary term \( \beta_K \) implies that

\[
\text{tr}(b \mid \text{Hom}_G(H^*_c, \pi)) = \frac{\xi(b)^{-1}}{\#\Gamma} \sum_{\gamma \in \Gamma} \text{Fix}_K(\gamma, b^{-1}) \cdot \chi_{\lambda \otimes \zeta}(\gamma^{-1}).
\]

The final step to establish the identity of characters is the

3.8. Theorem (Fixed point theorem). — Let \( g_b \) be in the conjugacy class corresponding to \( b \). Then

\[
\text{Fix}_K(\gamma, b^{-1}) = n \cdot \#\{ g \in G/\varpi^Z K \mid g^{-1} g_b g = \gamma^{-1} \}.
\]

The identity \( g^{-1} g_b g = \gamma^{-1} \) means that for some representative \( \dot{g} \) of \( g \in G/\varpi^Z K \) we have \( \dot{g}^{-1} g_b \dot{g} \in K_\pi \) and the class of \( \dot{g}^{-1} g_b \dot{g} \in \Gamma = K_\pi/\varpi^Z K \) is \( \gamma^{-1} \). By the fact pointed out in 3.1, the number of such \( g \in G/\varpi^Z K \) is always finite.

This formula will be proven in the next section. Putting the expression for the number of fixed points in the identity derived in 3.7 we get

3.9. Theorem. — Suppose the conjecture 3.5 on the Lefschetz trace formula is fulfilled. Let \( \pi \) be a supercuspidal representation of \( G = GL_n(F) \). Then the following holds:

(i) for every \( i \) the representation of \( B^\times \) on \( \text{Hom}_G(H^*_c, \pi) \) is admissible;
(ii) in the Grothendieck group of admissible representation of \( B^\times \) we have:

\[
\text{Hom}_G(H^*_c, \pi) = n \cdot (-1)^{n-1} \mathcal{J}L(\pi),
\]

where \( \text{Hom}_G(H^*_c, \pi) \) is \( \sum_i (-1)^i \text{Hom}_G(H^*_c, \pi) \), as defined after theorem 2.8.

Proof. — The first statement follows from the identity

\[
\text{Hom}_G(H^*_c, \pi) = \text{Hom}_F(H^*_c(M_K/\varpi^Z), \lambda \otimes \zeta) \otimes \xi^\vee,
\]

and the fact that \( H^*_c(M_K/\varpi^Z) \) is a finite-dimensional smooth representation of \( B^\times \), cf. paragraph 2.6 and [Be3], introduction. To prove the second statement, let \( b \in B^\times \)
be regular elliptic. Then the preceding discussion immediately gives
\[
\text{tr}(b | \text{Hom}_G(H^*_\infty, \pi)) = \frac{\xi(b)^{-1}}{\#\Gamma} \sum_{g \in G/\pi^2 K} n \cdot \chi_\lambda \otimes \zeta(g^{-1}gbg)
\]
\[
= n \sum_{g \in G/K_\pi} \chi_\lambda(g^{-1}gbg)
\]
\[
= n \cdot \chi_\pi(g_b) = n \cdot (-1)^{n-1} \chi_{JL}(\pi)(b),
\]
and this proves the assertion.\[\square\]

3.10. On the conjectural shape of the trace formula. — The natural approach to prove a trace formula for \(M := M_{K_m}^{(0)}\), is to consider admissible blow-ups \(M' \to M\) of the corresponding formal scheme \(M := M_{K_m}^{(0)} = \text{Spf}(R_m^{(0)})\), and to use the Grothendieck-Verdier Lefschetz formula on the special fibre of \(M'\) with the complex of \(\ell\)-adic nearby-cycles as coefficients. To do this, we only consider blow-ups to which the group action extends. So we have to study the fixed point locus on the special fibre of such blown-up models. The idea is to investigate the group action on all blow-ups simultaneously to eventually derive assertions about the existence of models where the connected components of the fixed point locus lie in different strata (see below for the definition of this stratification). Thus we are led to consider the projective limit
\[
\lim_{\leftarrow} M'
\]
over all admissible blow-ups. This space can in fact be given the structure of an adic space; it is, in Huber’s notation, \(t(M) - V(m_{R_m^{(0)}})\) (where \(m_{R_m^{(0)}}\) denotes the maximal ideal of \(R_m^{(0)}\) and it contains \(M = t(M) - V(\varpi)\) as an open subspace.

Because it is a projective limit of formal schemes with proper special fibres, we consider it as a kind of compactification of \(M\), and we put
\[
\overline{M} := t(M) - V(m_{R_m^{(0)}}).
\]

The underlying set is the set of all (equivalence classes) of continuous valuations of \(R_m^{(0)}\) which do not factor through the residue field of this local ring.

Now we define a stratification on this space. For \(m = 0\) and \(0 \leq i \leq n - 1\) let \(\partial_i \overline{M}\) be the subspace where the connected part of the universal formal \(\mathfrak{o}\)-module has height \(i\). We have in particular: \(\partial_0 \overline{M} = M\). For \(m > 0\) and any proper \(\mathfrak{o}\)-submodule \(A \subset (\varpi^{-m} \mathfrak{o}/\mathfrak{o})^n\) which is free over \(\mathfrak{o}/\varpi^m\) and a direct summand, define \(\partial_A \overline{M}\) to be the subspace where the the kernel of the universal level structure is equal to \(A\). So we have in particular \(\partial_0 \overline{M} = M\). The stabilizer in \(GL_n(\mathfrak{o}/\varpi^m)\) of the strata are maximal proper parabolic subgroups, the corresponding unipotent radical acts trivially on the stratum. Moreover, for any model \(M'\) of \(M\) the images of the strata induce a stratification of the special fibre: consider first the image of the zero-dimensional stratum, then remove this subscheme from the image of the one-dimensional stratum etc.
The conjecture 3.5 can be proven, if there is a model of $M$ such that

(i) the strata of the special fibre of the model are in bijection with the strata of $\overline{M}$;
(ii) the connected components of the fixed point locus on the special fibre of the model are contained in the relative interior of the strata;
(iii) the connected components of the fixed point locus which lie in the open stratum correspond (via the specialisation map) one-to-one to the fixed points in the interior, and the local term attached to such a component is one;
(iv) the local term attached to a connected component lying in a non-open stratum is invariant under the unipotent radical of the stabilizer of the corresponding stratum of $\overline{M}$.

In proving the first condition one uses the fact that the images $\phi(e_1), \ldots, \phi(e_n)$ of the universal Drinfeld base $\phi$ generate the maximal ideal of $R_m^{(0)}$. Assuming the second condition, the fourth condition should be provable, using a result of Berkovich, cf. [Be2], Th. 4.1. The fourth condition implies that the boundary term is a sum of characters of parabolically induced representations, and these characters are orthogonal to characters which give rise to supercuspidal representations (cf. [Bu]). Using Faltings methods ([Fa]), it seems possible to prove the existence of a model satisfying the third condition. (But the fixed point locus on the special fibre is not the one in the naive sense; one has to compute it using a model of $M \times M$ which is not the product of a model of $M$ with itself). This means that we can separate the interior fixed points from fixed points at the boundary. To prove the second condition means to separate the fixed point locus on the boundary according to the strata. And to do this, one has to understand the group action on the boundary. In the interior, i.e. on $M$, this is done via the period map of Gross and Hopkins. This period map has analogues, not on the boundary itself, but on a “tubular neighborhood” of the boundary. We don’t want to go into the definition of these spaces here, but it is hoped that the study of these maps on the tubular neighborhood of the boundary provides enough information about the group action on the boundary to finally prove the second condition.

4. Fixed points and the period morphism

4.1. To count fixed points we will use the period map from the moduli spaces $M_K$ to a projective space of dimension $n - 1$. This map was first studied by M. Hopkins and B. Gross [HG], and some results in connection with this map have been obtained before by G. Laffaille [La]. Later on, M. Rapoport and Th. Zink introduced these morphisms for moduli spaces for $p$-divisible groups [RZ], thereby giving a unified account of $p$-adic period maps that have been studied before (here one should mention Dwork’s period map from the deformation space of an ordinary elliptic curve to the affine line, cf. [Ka]). The set-up of Gross and Hopkins is insofar closer to our situation as they
work with formal $\mathfrak{g}$-modules (hence treat the mixed and equal characteristic case simultaneously), herein following Drinfeld. On the other hand, Gross and Hopkins only work with one component of the moduli space $\mathcal{M}_{K_0}$, namely the component $\mathcal{M}_{K_0}^{(0)}$ where the quasi-isogeny on the special fibre has height zero. After recalling the main results of [HG] in the next section, we will explain how to define the period map on the whole space $M_{K_0}$.

### 4.2.

Let $\mathcal{X}$ be the universal formal $\mathfrak{g}$-module over the formal scheme $\mathcal{M}_{K_0}^{(0)}$, and denote by $\mathcal{E}$ the universal extension of $\mathcal{X}$ with additive kernel. This is a formal $\mathfrak{g}$-module of dimension $n$ which sits in an exact sequence

$$0 \to \mathcal{V} \to \mathcal{E} \to \mathcal{X} \to 0,$$

where $\mathcal{V} = \mathcal{G}_a \otimes \text{Hom}_R(\text{Ext}(\mathcal{X}, \mathcal{G}_a), R)$ and $\mathcal{M}_{K_0}^{(0)} = \text{Spf}(R)$, so $R = R_0^{(0)}$ with the notation of section 2.5. This exact sequence furnishes an exact sequence

$$0 \to \text{Lie}(\mathcal{V}) \to \text{Lie}(\mathcal{E}) \to \text{Lie}(\mathcal{X}) \to 0,$$

of vector bundles on the formal scheme $\mathcal{M}_{K_0}^{(0)}$, and an analogous sequence

$$0 \to \text{Lie}(\mathcal{V})^{\text{rig}} \to \text{Lie}(\mathcal{E})^{\text{rig}} \to \text{Lie}(\mathcal{X})^{\text{rig}} \to 0,$$

on the generic fibre of this formal scheme, i.e. on the space $M_{K_0}^{(0)}$.

### 4.3. Proposition ([HG], Prop. 22.4, 23.2, 23.4)

(i) There is a basis $c_0, \ldots, c_{n-1}$ of $\text{Lie}(\mathcal{E})^{\text{rig}}$ such that the $\hat{F}^{\text{nr}}$-subspace generated by these global sections is stable by the action of $\mathfrak{o}_B$. More precisely, the canonical map of vector bundles on $\mathcal{M}_{K_0}^{(0)}$

$$\langle c_0, \ldots, c_{n-1} \rangle_{\hat{F}^{\text{nr}}} \otimes \mathcal{O}_{\mathcal{M}_{K_0}^{(0)}} \to \text{Lie}(\mathcal{E})^{\text{rig}}$$

is an $\mathfrak{o}_B$-equivariant isomorphism, where $\mathfrak{o}_B$ acts diagonally on the left hand side. The representation of $\mathfrak{o}_B^\times$ on $\langle c_0, \ldots, c_{n-1} \rangle_{\hat{F}^{\text{nr}}}$ is equivalent to the representation of $\mathfrak{o}_B^\times$ on $\mathcal{V} \otimes \hat{F}^{\text{nr}}$ given by left multiplication (where $F_n$ is the unramified extension of degree $n$ in $\hat{F}^{\text{nr}}$).

(ii) Let $w_i$ be the image of $c_i$ in $\text{Lie}(\mathcal{X})^{\text{rig}}$, $i = 0, \ldots, n-1$, and denote by $W$ the space generated by these global sections over $\hat{F}^{\text{nr}}$. Then, the sections $w_i$ have no common zeroes, and they are linearly independent over $\hat{F}^{\text{nr}}$.

(iii) Denote by $\mathbb{P}(W)$ the projective space of hyperplanes in $W$, and by $\mathbb{P}(W)^{\text{rig}}$ the associated analytic space. Define

$$\pi_{K_0}^{(0)} : M_{K_0}^{(0)} \to \mathbb{P}(W)^{\text{rig}}$$

by sending $x \in M_{K_0}^{(0)}$ to the hyperplane

$$\{ w = \alpha_0 w_0 + \cdots + \alpha_{n-1} w_{n-1} \in W \otimes \hat{F}^{\text{nr}}(x) \mid \alpha_0 w_0(x) + \cdots + \alpha_{n-1} w_{n-1}(x) = 0 \}.$$
This map is a rigid-analytic étale morphism. It is $\varphi^{\times}_B$-equivariant and surjective on $\overline{F}^\times$.

4.4. Choose an element $\varpi_B \in \mathfrak{o}_B$ whose reduced norm is a uniformizer of $F$. The action of $B^\times$ on $M_{K_0}$ furnishes for each $h \in \mathbb{Z}$ an isomorphism

$$\varpi_B^h : \mathcal{M}_{K_0}^{(0)} \rightarrow \mathcal{M}_{K_0}^{(h)}.$$ 

Define $\pi_{K_0}^{(h)} : M_{K_0}^{(h)} \rightarrow \mathbb{P}(W)$ by $\pi_{K_0}^{(h)} = \varpi_B^h \circ \pi_{K_0}^{(0)} \circ \varpi_B^{-h}$. Because of the $\mathfrak{o}_B^\times$-equivariance of $\pi_{K_0}^{(0)}$, this map is does not depend on the choice of $\varpi_B$. Finally we get the \textit{period map} on the whole space $M_{K_0}$ by putting

$$\pi_{K_0} = \prod_{h \in \mathbb{Z}} \pi_{K_0}^{(h)} : M_{K_0} = \prod_{h \in \mathbb{Z}} M_{K_0}^{(h)} \rightarrow \mathbb{P}(W)^{\text{rig}}.$$ 

More generally, for any open subgroup $K \subset K_0$ we let $\pi_K$ be the composition of the projection $M_K \rightarrow M_{K_0}$ with $\pi_{K_0}$, and refer to $\pi_K$ as a period morphism. The proposition above gives immediately the following assertion about the morphisms $\pi_K$.

4.5. \textit{Proposition.} — For any open subgroup $K \subset K_0$

$$\pi_K : M_K \rightarrow \mathbb{P}(W)^{\text{rig}}$$

is an étale morphism of analytic spaces over $\overline{F}^\text{nr}$. Moreover, $\pi_K$ is equivariant with respect to the action of $N_G(K) \times B^\times$, where the normalizer $N_G(K)$ of $K$ in $G$ acts trivially on $\mathbb{P}(W)^{\text{rig}}$ and the action of $B^\times$ on $\mathbb{P}(W)^{\text{rig}}$ is the one that is induced by the action of $B^\times$ on $W$.

4.6. Now we are in a position to count fixed points. Let $b \in B^\times$ be an element which is regular elliptic. Hence $b$ has $n$ distinct simple fixed points on $\mathbb{P}(W)^{\text{rig}} \otimes \overline{F}$. Let $K$ be a compact-open subgroup of $G$ that is contained in $K_0$, and let $\gamma$ be an element of the normalizer of $K$ in $G$. By Proposition 4.5, the action of the pair $(\gamma, b^{-1})$ on $M_K \otimes \overline{F}$ stabilizes the fibre of $\pi_K$ over a fixed point of $b$ on $\mathbb{P}(W)^{\text{rig}} \otimes \overline{F}$. Hence we need a description of the fibres of $\pi_K$ together with the action of $(\gamma, b^{-1})$. The next proposition gives such a description.

4.7. \textit{Proposition}

(i) Let $x \in M_{K_0}(\overline{F})$, and let $[X, \iota]$ be the deformation of $X$ corresponding to $x$. Then, the fibre of $\pi_{K_0}$ through $x$ consists of all deformations which are quasi-isogenous to $X$. More precisely, it consists of those pairs $[X', \iota']$ such that there exists a quasi-isogeny $f : X' \rightarrow X$ with the property that $f_\mathfrak{F} \circ \iota' = \iota$, where $f_\mathfrak{F}$ is the the reduction of $f$.

(ii) The fibre of $\pi_{K_0}$ through $x$ can be identified with the set of lattices in the rational Tate module $V(X) = T(X) \otimes_{\mathfrak{o}} F$, where

$$T(X) = \varprojlim X[\varpi^m](\overline{F}).$$
By fixing an isomorphism \( \phi : F^n \to V(X) \), this set gets identified with \( G/K_0 \). More generally, let \( K \subset K_0 \) be an open subgroup, and let \([X,t,\phi]\) be a point of \( M_K(F^r) \). Then, the fibre of \( \pi_K \) through this point can be identified with the coset \( G/K \).

(iii) Consider an \( \overline{F}^r \)-valued fixed point of \( b \) on \( \mathbb{P}(W)^{rig} \), and choose a base point of the set of \( \overline{F}^r \)-valued points of the fibre of \( \pi_K \) over this point. Using this fixed point, identify this set with \( G/K \), as in part ii). Then there exists \( g_b \in G \) with the same characteristic polynomial as \( b \) such that the action of \( (\gamma, b^{-1}) \) on the (set of \( \overline{F}^r \)-valued points of the) fibre is given, in terms of this identification, by

\[
gK \mapsto g_b \gamma K.
\]

Proof. — The first assertion follows from Prop. 23.28 of [HG]. The relationship between lattices in the rational Tate module and quasi-isogenies in the mixed characteristic case can be found in Lubin’s paper [Lu], Theorem 2.2. The same holds true also in the equal characteristic case, cf. [Yu], sec. 3. The second assertion of ii) follows immediately.

Now we are going to prove part iii). Fix an \( \overline{F}^r \)-valued point of \( M_K \), given by a triple \([X,t,\phi]\). We can consider \( \phi \) as an isomorphism \( \phi^n \to T(X) \) which is determined up to multiplication (from the right) by elements from \( K \). Suppose this point is mapped by \( \pi_K \) onto a fixed point of \( b \). Then it follows from [HG], Prop. 23.28, that \( b \) lifts to an endomorphism \( \tilde{b} : X \to X \) of the formal \( \phi \)-module \( X \) such that \( \tilde{b}_F \circ t = t \circ b \), where \( \tilde{b}_F \) is the quasi-isogeny induced on the special fibre. \( \tilde{b} \) is mapped to \( b \) under the canonical map \( \text{End}_{\phi}(X) \otimes F \hookrightarrow \text{End}_{\phi}(X) \otimes F \). Therefore the characteristic polynomial of \( \tilde{b} \) is the same as that of \( b \). Let \( g_b \in G \) be such that the following diagram is commutative:

\[
\begin{array}{ccc}
F^n & \xrightarrow{\phi} & V(X) \\
g_b \downarrow & & \downarrow V(\tilde{b}) \\
F^n & \xrightarrow{\phi} & V(X)
\end{array}
\]

Let \([X',t',\phi']\) be an element in the fibre of \( \pi_K \). Hence there is a quasi-isogeny \( f : X' \to X \) and an element \( g \in G \) such that the following diagram commutes:

\[
\begin{array}{ccc}
F^n & \xrightarrow{\phi'} & V(X') \\
g \downarrow & & \downarrow V(f) \\
F^n & \xrightarrow{\phi} & V(X)
\end{array}
\]

The class \( gK \in G/K \) corresponds to the point \([X',t',\phi']\). This point is mapped by \( b^{-1} \) to \([X',t' \circ b^{-1},\phi']\). The map \( \tilde{b} \circ f : X' \to X \) is then a quasi-isogeny, if we equip \( X' \) with the map \( t' \circ b^{-1} : \mathbb{X} \to (X')^{\overline{F}} \). Moreover, it is easily checked that the
following diagram commutes:

\[
\begin{array}{ccc}
F^n & \phi' & V(X') \\
\downarrow & & \downarrow \\
F^n & \phi & V(X)
\end{array}
\]

The action of \(b^{-1}\) on the fibre of \(\pi_K\) is thus given by sending \(gK\) to \(g\gamma b g K\). It is straightforward to check that the action of some \(\gamma \in N_G(K)\) on this fibre is given by sending \(gK\) to \(g\gamma K\). This proves the third assertion. \(\square\)

4.8. Proof of the Fixed point theorem 3.8. — Let \(b \in B^\times\) be regular elliptic and consider the fibre of the induced map

\[
(M_K/\varpi^2)(\overline{F}) \longrightarrow \mathbb{P}(W)(\overline{F})
\]

over a fixed point of \(b\). By the preceding proposition, we may identify this set with \(G/\varpi^2 K\) and the action of \((\gamma, b^{-1})\), \(\gamma\) in the normalizer of \(K\) in \(G\), is given by

\[
g \varpi^2 K \mapsto g b \gamma \varpi^2 K,
\]

where \(g_b \in G\) has the same characteristic polynomial as \(b\). Hence the number of fixed points on such a fibre is

\[
\# \{g \in G/\varpi^2 K \mid g^{-1} g_b g = \gamma^{-1}\}.
\]

Because there are \(n\) simple fixed points of \(b\) on \(\mathbb{P}(W)\) and the morphism \(\pi_K\) is étale, all fixed points are simple and the total number of fixed points is

\[
n \cdot \# \{g \in G/\varpi^2 K \mid g^{-1} g_b g = \gamma^{-1}\}.
\]

References


2005


2004


295. Cohomologies $p$-adiques et applications arithmétiques (III), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs

294. SÉMINAIRE BOURBAKI, volume 2002/2003, exposés 909-923

293. M. EMERTON, M. KISIN – The Riemann-Hilbert correspondence for unit $F$-crystals

292. F. LE ROUX – Homéomorphismes de surfaces. Théorèmes de la fleur de Leau-Fatou et de la variété stable

291. L. FARGUES, E. MANTOVAN – Variétés de Shimura, espaces de Rapoport-Zink et correspondances de Langlands locales

2003

290. SÉMINAIRE BOURBAKI, volume 2001/2002, exposés 894-908

289. P. SCOTT, G.A. SWARUP – Regular neighbourhoods and canonical decompositions for groups

288. M. REES – Views of Parameter Space : Topographer and Resident


286. Geometric Methods in Dynamics (I), volume in honor of Jacob Palis, W. DE MELO, M. VIANA, J.-C. Yoccoz, editors

285. PO HU – Duality for Smooth Families in Equivariant Stable Homotopy Theory

284. Autour de l’analyse microlocale, volume en l’honneur de Jean-Michel Bony, G. LEBEAU, éditeur

2002

283. L. ROBBIANO, C. ZUILI – Analytic theory for the quadratic scattering wave front set and application to the Schrödinger equation


280. A. MOKRANE, P. POLO, J. TILOUINE – Cohomology of Siegel varieties

279. Cohomologies $p$-adiques et applications arithmétiques (II), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs

278. Cohomologies $p$-adiques et applications arithmétiques (I), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs

277. B. RÉMY – Groupes de Kac-Moody déployés et presque déployés

276. SÉMINAIRE BOURBAKI, volume 1999/2000, exposés 865-879

2001

275. J.-M. BISMUT, S. GOETTE – Families torsion and Morse functions

274. A. BONNET, G. DAVID – Cracktip is a global Mumford-Shah minimizer


272. M. BOILEAU, J. PORTI – Geometrization of 3-orbifolds of cyclic type (avec la collaboration de M. HEUSENER)

271. M. KASHIWARA, P. SCHAPIRA – Ind-Sheaves

270. M. BONK, J. HEINONEN, P. KOSKELA – Uniformizing Gromov hyperbolic spaces

269. J.-L. WALDSPURGER – Intégrales orbitales nilpotentes et endoscopie pour les groupes classiques non ramifiés
2000
268. J. FRANCHETEAU, G. MÉTIVIER – Existence de chocs faibles pour des systèmes quasi-linéaires hyperboliques multidimensionnels
267. R. CERF – Large Deviations for three Dimensional Supercritical Percolation
266. SÉMINAIRE BOURBAKI, volume 1998/1999, exposés 850-864
265. O. BIQUARD – Métriques d’Einstein asymptotiquement symétriques
263. C. SABBH – Équations différentielles à points singuliers irréguliers et phénomène de Stokes en dimension 2
262. A. VASY – Propagation of singularities in three-body scattering

1999
260. S. D. CUTKOSKY – Local monomialization and factorization of morphisms
259. R. KRIKORIAN – Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts
258. Structure Theory of Set Addition, J.-M. DESHOUILLES, B. LANDREAU and A.A. YUDIN, editors
257. J.-P. LABESSE – Cohomologie, stabilisation et changement de base (avec la collaboration de L. BREEN et L. CLOZEL)
256. F. MOREL – Théorie homotopique des schémas
255. R.E. KOTTWITZ, D. SHELSTAD – Foundations of twisted endoscopy
254. C.J. BUSHNELL, G. HENNIART – Local tame lifting for GL(n) II : wildly ramified supercuspidals
253. B. MAGNERON – Involutions complexes et vecteurs sphériques associés pour les groupes de Lie nilpotents réels

1998
251. Nombre et répartition de points de hauteur bornée, E. PEYRE, éditeur
250. C. BONATTI, R. LANGEVIN – Difféomorphismes de Smale des surfaces (avec la collaboration de E. JEANDENANS)
249. P. AUSCHER, P. TCHAMITCHIAN – Square root problem for divergence operators and related topics
248. P. COLMEZ – Intégration sur les variétés p-adiques
247. G. PISIER – Non-commutative vector valued Lp-spaces and completely p-summing maps

1997
246. V. TARASOV, A. VARCHENKO – Geometry of q-hypergeometric functions, quantum affine algebras and elliptic quantum groups
245. SÉMINAIRE BOURBAKI, volume 1996/1997, exposés 820-834
244. J.-M. BISMUT – Holomorphic families of immersions and higher analytic torsion forms
243. L. LAFFORGUE – Chéroucas de Drinfeld et conjecture de Ramanujan-Petersson
242. N. BURO – Pôles de diffusion engendrés par un coin

1996
240. A. SÀ BARETTO, R. B. MELROSE, M. ZWORSKI – Semi-linear diffraction of conormal waves
239. J.-L. VERDIER – Des catégories dérivées des catégories abéliennes
238. A. BROISE, F. DAL’BO, M. PEIGNÉ – Méthodes des opérateurs de transfert : transformations dilatantes de l’intervalle et dénombrement de géodésiques fermées
236. Hommage à P. A. MEYER et J. NEVEU
235. J.-P. OTAL – Le théorème d’hyperbolisation pour les variétés fibrées de dimension 3
234. A. GENESTIER – Espaces symétriques de Drinfeld