THE LOCAL LANGLANDS CORRESPONDENCE: NOTES OF (HALF) A COURSE AT THE IHP SPRING 2000

Michael Harris

UFR de Mathématiques
Université Paris 7
2 Pl. Jussieu
75251 Paris cedex 05, FRANCE

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,

---

Institut des Mathématiques de Jussieu, U.M.R. 7586 du CNRS. Membre, Institut Universitaire de France.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

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.

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:

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

(0.2) Let χ 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 $\sigma(\pi\otimes\chi\circ\det) = \sigma(\pi)\otimes\chi$. In particular, when n = 1, the bijection is given by local class field theory.

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

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

(0.5) Let $\alpha : K \to K_1$ be an isomorphism of local fields. Then α 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 σ respects the $Gal(K/K_0)$-actions on both sides.

(0.6) 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

$$\sigma(BC(\pi)) = \sigma(\pi)|_{WD_{K'}} \tag{0.6.1}$$

$$\sigma(AI(\pi')) = Ind_{K'/K}\sigma(\pi'), \tag{0.6.2}$$

where $Ind_{K'/K}$ 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

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

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

---

[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.

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 ℓ-adic Fourier transform (the subsets of $\mathcal{A}_0(n,K)$ and $\mathcal{G}_0(n,K)$ 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.[3] 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).[4]

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

[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 IHP 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].

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 Φ_{n-h} with canonical action of $GL(n-h, F_w) \times J_{n-h}$, where J_{n-h} is a specific anisotropic inner form of $GL(n-h)$ over F_w.

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 Φ_{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 Φ_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 $h = 0$, Φ_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 ℓ-adic cohomology of the special fiber by means of Grothendieck's version of the Lefschetz trace formula. In [HT] we also use an ℓ-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 Drinfel'd 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.[5] 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 the time of the course. In particular, §2 and §3 contain a brief review of the deformation theory of one-dimensional formal $\mathcal{O}$-modules, following Lubin-Tate and Drinfel'd.

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 Broué, 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.

Lecture 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 τ_0 of F, and let Σ denote the set of complex embeddings of F with the same restriction to E as τ_0. This Σ is a CM type, and is in bijection with the set of real embeddings of F^+. We consider automorphic representations Π of $GL(n, F)$, or more precisely of $GL(n, \mathbf{A}_F)$. Any such representation can be factored $\Pi = \Pi_\infty \otimes \Pi_f$, where Π_f is an admissible irreducible representation of $GL(n, \mathbf{A}_{F,f})$ and Π_∞ is a Harish-Chandra module for $GL(n, \mathbb{C})^\Sigma$; i.e., an admissible irreducible $(\mathfrak{gl}(n, \mathbb{C})^d, U(n)^d)$-module.

We will only be concerned with Π such that Π_∞ is cohomological. We will also restrict attention to cuspidal Π, though general discrete cohomological Π also play a role in the more detailed results of [HT]. Then Π is generic, by Shalika's theorem. Let (ξ, W_ξ) denote a finite-dimensional irreducible representation of $GL(n)_F$. This is equivalent to giving a pair of finite-dimensional irreducible representations $(\xi_\sigma, \xi_{c\sigma})$ of $GL(n, \mathbb{C})$ for each $\sigma \in \Sigma$. For any representation τ, 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*

$$\xi_c \xrightarrow{\sim} \xi^\vee \tag{1.1.2}$$

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

The above fact is a special case of the construction in [C2,3.5], which covers nearly all generic cohomological $(\mathfrak{gl}(n, \mathbb{C}), U(n))$-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 (Ξ, W_Ξ) a finite-dimensional algebraic representation of G. For any open compact subgroup K of $G(\mathbf{A}_f)$, let*

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

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

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

over $\mathcal{M}_K(G)$ (this is a local system provided the central character of Ξ 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 $G(\mathbf{A}_f)$-equivariant subspace $H^*_{cusp}(\mathcal{L}_\Xi)$ of $\varinjlim_K H^*(\mathcal{M}_K(G), \mathcal{L}_\Xi)$ such that*

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

as $G(\mathbf{A}_f)$-modules, where Π runs through $\mathcal{A}_0(G)$.

(1.1.4) When $G = GL(n)_F$, we assume $\Xi = \otimes_\Sigma \Xi_\sigma$, where each Ξ_σ satisfies condition (1.1.2); recall that Σ indexes embeddings of F^+. Fact (1.1.1) shows that the sum runs over Π such that

$$\Pi_\infty \simeq \Pi_\Xi := \otimes_\Sigma \Pi_{\Xi,\sigma},$$

and this implies that $\Pi_\infty \circ c \xrightarrow{\sim} \Pi_\infty^\vee$. We also make an analogous global restriction:

$$\Pi^c \xrightarrow{\sim} \Pi^\vee. \tag{1.1.5}$$

This is necessary in order to attach compatible families of ℓ-adic representations to Π, 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 ℓ-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 $\mathcal{O}_B \subset B$ such that the involution $*$ restricts to an involution of $\mathcal{O}_{B,p} = \mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p$.

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., $Tr_{F/\mathbb{Q}}(t_B(g \cdot g^*)) > 0$ 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 $(bv, w) = (v, b^* w)$) $V \times V \to \mathbb{Q}$ by

$$(x_1, x_2)_\beta = Tr_{F/\mathbb{Q}}(t_B(x_1 \beta x_2^*)).$$

Then for $b \in B$, $b_{op} \in B^{op}$, we have

$$((b \otimes b_{op})x_1, x_2)_\beta = (x_1, (b^* \otimes b_{op}^{\#_\beta})x_2)_\beta. \tag{1.2.1}$$

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

$$g \cdot g^{\#_\beta} = \lambda.$$

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

We identify

$$B \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\sim} \prod_{\tau \in \Sigma} M(n, F_\tau) \xrightarrow{\sim} M(n, \mathbb{C})^d. \tag{1.2.2}$$

For each $\tau \in \Sigma$, $(\bullet,\bullet)_\beta$ thus defines a $*$-hermitian form $(\bullet,\bullet)_{\beta,\tau}$ on $M(n,\mathbb{C})$. If n is even, we assume $1 + \frac{dn}{2}$ has the same parity as $|S_B|$. Then a calculation in Galois cohomology (cf. [C2,§2]; [HT, Lemma 1.1]) shows that β can be chosen in B^- such that G_β is quasisplit 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^+_{β,σ_0} is isomorphic to $U(1, n-1)$ but $G^+_{\beta,\sigma}$ is a compact unitary group for all real places $\sigma \neq \sigma_0$. We fix such a β and drop it henceforth from the notation.

We write ${}_KSh(G)$ for the locally symmetric space denoted $\mathcal{M}_K(G)$ 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, ${}_KSh(G)$ is in fact projective for all K, and smooth if K is sufficiently small (which we assume). Thus there is no distinction between H^*_{cusp} and H^* in Matsushima's formula. The representation (ξ, W_ξ) defined above gives rise to a representation of G (take the factors in Σ, and regard $G(\mathbb{R})$ as a subgroup of unitary similitudes in $GL(n,\mathbb{C})^\Sigma$). We denote by $\mathcal{L}_\xi$ the corresponding local system on $Sh(G) = \varprojlim_K {}_KSh(G)$.

If p splits in E, we can identify

$$G(\mathbb{Q}_p) \xrightarrow{\sim} \prod_{v|p} B_v^{op,\times} \times \mathbb{Q}_p^\times \tag{1.2.3}$$

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

$$\pi = \pi_\infty \otimes \bigotimes_p{}' \pi_p \in \mathcal{A}_0(G),$$

we can further factor π_p as

$$\pi_p = \otimes_{v|p} \pi_v \otimes \psi_p, \tag{1.2.4}$$

where ψ_p is a character of $\mathbb{Q}_p^\times$.

(1.2.5) Remark. In practice, we will arrange that ψ_p always be an unramified character. We will moreover make a habit of suppressing the effect of ψ_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* (Ξ, W_Ξ) *be as in (1.1.4). Let* $\Pi \subset \mathcal{A}_0(GL(n)_F)$ *be a cuspidal automorphic representation with central character* ψ_Π*, and let* ψ *be a Hecke character of* E*. Suppose*

(a) $\Pi_\infty \simeq \Pi_\xi$.
(b) $\Pi^c \xrightarrow{\sim} \Pi^\vee$;
(c) *For every place* $v \in S(B)$, Π_v *is in the discrete series.*
(d) $\psi_\Pi \mid_{\mathbf{A}_E^\times} = \psi^c/\psi$.
(e) $(\xi \mid_{E_\infty^\times})^{-1} = \psi^c \mid_{E_\infty^\times}$.

Then there exists an automorphic representation π *of* G *whose base change to* $GL(n)_F \times E^\times$ *equals* (Π, ψ)*. Moreover,* π_f *occurs in the cohomology of* $Sh(G)$ *with coefficients in* $\mathcal{L}_\xi$*.*

Conversely, given $\pi \in \mathcal{A}_0(G)$*, cohomological for* ξ*, there exists a pair* (Π, ψ) *satisfying (a)-(e), with* $\psi = \psi_\pi^c \mid_{\mathbf{A}_E^\times}$*, such that* (Π, ψ) *is the base change of* π *at all unramified places and at all places that split in* E*. Moreover, if* π_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* Π *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 ψ satisfying (d) and (e).

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

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

where B_y is of course a product of central simple algebras over the completions of F at places dividing y. Thus given π, we can write $\pi_p = ? \otimes \pi_y$, and define $\Pi_p = \pi_y \otimes \pi_y^\#$, where $\pi_y^\#(g) = \pi_y((g^\#)^{-1})$. This doesn't depend on the choice of y. Moreover, we recover π_y from Π_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 π_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

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

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

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

If π_p is the unramified representation $\pi(\alpha)$ corresponding to α, 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 Π_p (as opposed to the unramified character ψ_p). We have thus defined Π_p for almost all p, and by strong multiplicity one, this suffices to determine Π.

Henceforward, to simplify the exposition and minimize notation, we assume Ξ 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 Ξ.

(1.3) Kottwitz' theorem and its refinements.

We can identify complex cohomology with ℓ-adic cohomology, for example, by choosing an isomorphism $\mathbb{C}\xrightarrow{\sim}\bar{\mathbb{Q}}_\ell$. Thus if Π is as above, we can choose a character ψ and define π such that $\pi_f \subset H^*(Sh(G), \bar{\mathbb{Q}}_\ell)$. It is known that $Sh(G)$ admits a canonical model over F (recalled next week), and thus there is a virtual representation of $\Gamma_F = Gal(\bar{F}/F)$ on the π_f-isotypic subspace

$$R_\ell(\pi, G) = \sum_i (-1)^{n-1+i} Hom_{G(\mathbf{A}_f)}(\pi_f, H^i(Sh(G), \bar{\mathbb{Q}}_\ell)).$$

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

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

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 ℓ, such that π_p is unramified, and for all v dividing p, the local representation $R_\ell(\pi)_v$ is isomorphic to $a(\pi)$ copies of*

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

where the v-component Π_v of Π 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 π_f occurs only in the middle degree $n-1$. Kottwitz' theorem uses the theory of the zeta function of the reduction mod v of $Sh(G)$. In particular, the $\alpha_i(\Pi_v)$ are eigenvalues of Frobenius on $H^*(Sh(G), \mathbb{Q}_\ell)$ (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, Π is cuspidal, hence Π_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 π_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

$$R_\ell(\Pi)_{ss} = R_{\ell,0}(\Pi)^{a(\pi)} \tag{1.3.2}$$

for some n-dimensional semisimple representation $R_{\ell,0}(\Pi)$ of Γ_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

$$r_\ell(\Pi) = (R_{\ell,0}(\Pi)_{ss})^\vee \tag{1.3.3}$$

Applying Kottwitz' theorem, we obtain

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

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

Remark. We are always working with the unitary normalization of the Langlands correspondence. So the L-function of Π has to be shifted by $\frac{n-1}{2}$ in order to obtain the L-function of a Galois representation:

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

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 σ_ℓ is the local Langlands correspondence, so far defined for unramified representations. A compatible family $(r_\ell(\Pi))$ as above is called *weakly associated* to Π. 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 Π (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 Π.

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

(1.3.6) Main Theorem*. [HT] *Let $\Pi \in CU(n, E)$. Then for all places v not dividing ℓ,*

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

Now I have to explain how this theorem implies compatibility with ε factors.

---

*Note added in proof: Taylor and T. Yoshida have just proved the expected strengthening of this theorem, in which semisimplication is replaced by Frobenius-semisimplification.

(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

$$Ind_P^{G_{n+m}} \pi \otimes \pi'.$$

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 $\Gamma = Gal(F_3/F_1)$ is solvable. The goal is to show that certain algebraic Hecke characters χ of F_2 define by induction automorphic representations $I_{F_2/F_1}(\chi)$ of $GL(d,F_1)$, where $d = [F_2 : F_1]$. The meaning is clear. If χ is associated to an ℓ-adic character $r_\ell(\chi)$ of Γ_{F_2}, then $I_{F_2/F_1}(\chi)$ should be weakly associated to $Ind_{F_2/F_1} r_\ell(\chi)$, where Ind_{F_2/F_1} denotes induction from Γ_{F_2} to Γ_{F_1}. Concretely, let v be a place of F_1, unramified in F_2, such that χ_w 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 of characters $\chi_w \circ \gamma$ as γ runs through $Gal(F_{2,w}/F_{1,v})$. 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.

(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, $\chi \cdot |\bullet|^{\frac{d-1}{2}}$, rather than χ, is an algebraic Hecke character.

Suppose we can do this for quite general χ: just how general will become clear in a moment. Suppose moreover that $I_{F_2/F_1}(\chi) \in CU(d,F_1)$, so that we can apply the Main Theorem. Let $F_3' \supset F_2' \supset F_1$ be another triple as above, with $[F_2' : F_1] = d'$, and suppose we have a second character χ'. 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 [JPSS,Sh]

$$(1.4.2)\quad L(s,\Pi(\chi)\otimes\Pi(\chi')) = \prod_v \epsilon_v(s,\Pi(\chi)_v\otimes\Pi(\chi')_v,\psi_v)L(1-s,\Pi(\chi)^\vee\otimes\Pi(\chi')^\vee).$$

The local ε factors are those of the automorphic theory. On the other hand, we have $\Pi(\chi)$ weakly associated to $Ind_{F_2/F_1} r_\ell(\chi)$, and likewise for $\Pi(\chi')$. The tensor product $Ind_{F_2/F_1} r_\ell(\chi) \otimes Ind_{F_2'/F_1} r_\ell(\chi')$ is an ℓ-adic representation corresponding to a complex representation of the global Weil group of F_1, hence there is a functional equation

$$\begin{aligned}(1.4.3)\quad & L(s, Ind_{F_2/F_1} r_\ell(\chi) \otimes Ind_{F_2'/F_1} r_\ell(\chi')) \\ = & \prod_v \epsilon_v(s, Ind_{F_{2,v}/F_{1,v}} r_\ell(\chi_v)\otimes Ind_{F'_{2,v}/F_{1,v}} r_\ell(\chi'_v),\psi_v) L(1-s, Ind_{F_2/F_1} r_\ell(\chi)^\vee \otimes Ind_{F_2'/F_1} r_\ell(\chi')^\vee),\end{aligned}$$

where the local ε 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 ε factors for all v:
(1.4.4)

$$\varepsilon_v(s,\Pi(\chi)_v\otimes\Pi(\chi')_v,\psi_v)=\varepsilon_v(s,Ind_{F_{2,v}/F_{1,v}}r_\ell(\chi_v)\otimes Ind_{F'_{2,v}/F_{1,v}}r_\ell(\chi'_v),\psi_v).$$

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 σ and σ' be two representations of Γ_v factoring through $Gal(F_{3,v}/K)$ and $Gal(F'_{3,v}/K)$, respectively, and let π and π' be the corresponding elements of $\mathcal{A}_0(n,K)$ (resp. $\mathcal{A}_0(n',K)$). Note that for any $\sigma\in\mathcal{G}_0(n,K)$ we can choose a local extension $F_{3,v}/K$ which is solvable and such that σ comes from $Gal(F_{3,v}/K)$ 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 $Gal(F_{3,v}/F_{2,j,v})$, and integers e_j such that

$$\sigma=\sum_j e_j Ind_{F_{2,j,v}/K}\chi_{j,v};$$

likewise for σ'. Applying the above identity of ε factors, and ignoring ψ_v, one obtains

$$\varepsilon_v(s,\sigma\otimes\sigma')=\prod_{j,j'}\varepsilon_v(s,\Pi(\chi_j)_v\otimes\Pi(\chi'_{j'})_v)^{e_j\cdot e_{j'}}.$$

Now we apply the Main Theorem. Say

$$\Pi(\chi_j)_v=\boxplus_i\pi_{i,j};\Pi(\chi'_{j'})_v=\boxplus_{i'}\pi'_{i',j'},$$

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

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

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

$$\begin{aligned}&\prod_{j,j'}\varepsilon_v(s,\Pi(\chi_j)_v\otimes\Pi(\chi'_{j'})_v)^{e_j\cdot e_{j'}}\\&=\prod_{i,j,i',j'}\varepsilon_v(s,\pi_{i,j}\otimes\pi'_{i',j'})^{e_j\cdot e_{j'}}\\&=\prod_{i,j,i',j'}\varepsilon_v(s,\sigma^{-1}(\sigma_{i,j})\otimes\sigma^{-1}(\sigma'_{i',j'}))^{e_j\cdot e_{j'}}\\&=\varepsilon_v(s,\sigma^{-1}(\sigma))\otimes\sigma^{-1}(\sigma'))=\varepsilon_v(s,\pi\otimes\pi')\end{aligned}$$

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

By simple approximation arguments, we see that any local extension K'/K can be realized as an $F_{3,v}/F_{1,v}$ as above. In this way, we find that it suffices to prove that, for any intermediate F_2 in a solvable extension and any character χ_v of $F_{2,v}$, then (up to unramified twists) χ_v can be realized as the local component of a Hecke character χ 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 Π_2 is an automorphic representation of $GL(d, F_2)$ invariant under $Gal(F_2/F_1)$, then Arthur and Clozel prove the existence of Π_1 on $GL(d, F_1)$ whose base change to F_2 is Π_2: this is the descent of Π_2 to F_1.

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

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

The result is invariant under Γ, 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 $F_1 \subset \tilde{E} \subset F_3$ with $F_3/\tilde{E}$ and $\tilde{E}/F_1$ cyclic of prime order; let $C = Gal(F_3/\tilde{E})$. Let $J(\chi, E) = Res_{\Gamma_E} Ind_{F_2/F_1} r_\ell(\chi)$. Then

$$Res_{\Gamma_3} J(\chi, \tilde{E}) = Res_{\Gamma_3}(J(\chi, \tilde{E}) \otimes \beta)$$

for any character β of C. There is a similar ambiguity in descent. Consider the first step: let Π be an automorphic representation of $GL(d, F_3)$ invariant under C, Π_C a descent to $GL(d, \tilde{E})$. For any character β of C, the twist $\Pi_C \otimes \beta$ (β viewed as a Hecke character of $GL(1, \tilde{E})$) is another descent of Π. So the total number of descents is on the order of $|C|$ (some twists may be isomorphic). (Actually, there is more ambiguity: if Π_C is not cuspidal, say $\Pi_C = \boxplus_j \Pi_j$, then each Π_j can be twisted separately by a character β_j of C.) On the other hand, locally everywhere $\Pi_{C,\mathfrak{p}}$ can be identified only up to twist(s) by character(s) $\alpha_{\mathfrak{p},j}$ of the decomposition group $C_{v,\mathfrak{p}} \subset C$ at $\mathfrak{p}$. The general theory thus gives that for each $\mathfrak{p}$ there is a character $\alpha_{\mathfrak{p}}$ of $C_{\mathfrak{p}}$ such that, for almost all $\mathfrak{p}$,

$$\sigma_\ell(\Pi_{C,\mathfrak{p}}) = J(\chi, E)_{\mathfrak{p}} \otimes \alpha_{\mathfrak{p}}.$$

A priori, there is no way to prove that the local characters $\alpha_{\mathfrak{p}}$ fit together to a global character β of C.

On the other hand, if it is known that Π_C has a weakly associated ℓ-adic representation $r_\ell(\Pi_C)$ of $\Gamma_{\tilde{E}}$, then $r_\ell(\Pi_C)$ is a descent to E of $Res_{\Gamma_{F_3}} Ind_{F_2/F_1} r_\ell(\chi)$. If moreover $r_\ell(\Pi_C)$ is irreducible, then the only ambiguity in descents comes from twists by characters of C:

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

for a global character α. So one can replace Π_C by $\Pi_C \otimes \alpha^{-1}$ and continue with the descent. More generally, the analogous argument works if $\Pi_C = \boxplus_j \Pi_j$ 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 $\Pi(\chi, E_i) = \boxplus \Pi_{i,j}$ 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 ∞, 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 χ_v to be general, χ_∞ such that $\Pi_{i,j,\infty}$ is cohomological for all i and j, and χ_w such that every $\Pi_{i,j,w}$ is supercuspidal. Since p splits in E, $v \neq v^c$, so the condition $\chi^{-1} = \chi^c$ imposes no restriction on χ_v. The condition at w is a bit more subtle. By Mackey's theorem, the restriction $Res_{E_i} Ind_{F_2/F_1} r_\ell(\chi)$ breaks up as a sum of constituents that may not necessarily be irreducible:

$$Res_{E_i} Ind_{F_2/F_1} r_\ell(\chi) = \bigoplus_a Ind_{E_{i,a}/E_i} a(\chi)$$

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 χ (conjugated by a). We need to choose χ_w so that each of these Mackey constituents is locally irreducible at w. This is true generically (exercise, cf. [H2, Lemma 4.7]). Then the base change theory implies that $\Pi(\chi, E_i) = \boxplus \Pi_{i,a}$, with $\Pi_{i,a}$ supercuspidal at w.

Supposing we have χ_w, we choose a global Hecke character χ_0, trivial at ∞, 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 χ_∞, we work backwards. For any complex place τ of F_1, let $\tau(k), k = 1, \ldots, d$ be the primes of F_2 dividing τ. Then

$$\Pi(\chi)_\infty = \boxplus_k \chi_{\tau(k)}.$$

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

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

The coefficient system Ξ determines the set $\{\chi_{\tau(k)}^{-1}\}$ for each τ. This is again not a restriction. Let χ_2 be any algebraic Hecke character with $\chi_2^c = \chi_2^{-1}$, with $\chi_{2,\tau(k)}$ as just described for all τ, and such that $\chi_{2,v}$ and $\chi_{2,w}$ are unramified. By allowing sufficient ramification elsewhere, we can easily construct such χ_2. Then we let $\chi = \chi_1 \cdot \chi_2$. This has the right behavior at w and ∞, is conjugate self-dual, and is the desired χ_v up to unramified twist. It suffices to show that $I_{F_2/F_1}\chi$ is automorphic for such χ.

Here is the induction step. Suppose we have $\Pi(\chi, E_i) = \boxplus \Pi_{i,a}$, 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 $Gal(E_i/F_1)$. Finally, we assume that the sum of the corresponding Galois representations is $Res_{E_i} Ind(r_\ell(\chi))$. Let $C_i = Gal(E_i/E_{i-1})$, 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

$$\Pi^{\vee}_{i-1,a',0} = \Pi^{c}_{i-1,a',0}.$$

This is true after base change, so

$$\Pi^{\vee}_{i-1,a',0} = \Pi^{c}_{i-1,a',0} \cdot \eta$$

for some character η of C_i. But E_i/E_{i-1} comes from a cyclic extension of totally real fields E_i^+/E_{i-1}^+, so $\eta = \eta^+ \circ N_{E_{i-1}/E_{i-1}^+}$ for some Hecke character η^+ of E_{i-1}^+, trivial at the archimedean places. This implies that η^+ extends to a finite Hecke character α 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

$$\Pi^{\vee}_{i-1,a'} = \Pi^{c}_{i-1,a'}.$$

Next, if $\{\Pi^t_{i,a} \mid t \in C_i\}$ is a non-trivial orbit, the Langlands sum $\boxplus_t \Pi^t_{i,a}$ 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 χ_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 $Res_{E_{i-1}} Ind(r_\ell(\chi))$. In particular, the set of constituents is invariant under $Gal(E_{i-1}/F_1)$, 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 Π_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 ε 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].

Lecture 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 $h : R_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_{m,\mathbb{C}}) \to G_{\mathbb{R}}$, satisfying a familiar list of axioms [De1]. We will always assume the weight morphism w_h, the restriction of h to $\mathbb{G}_{m,\mathbb{R}}$, 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_∞, 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^{der} of G. Upon extension of scalars to $\mathbb{C}$, an $h \in X$ defines a homomorphism $\mathbb{G}_{m,\mathbb{C}} \times \mathbb{G}_{m,\mathbb{C}} \to G_{\mathbb{C}}$ whose first coordinate is a cocharacter denoted $\mu = \mu_h$.

The G-conjugacy class of μ is independent of h and its field of definition is a number field denoted $E(G, X)$; we will write μ_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 $Sh(G, X)$, whose set of complex points is given by

$$Sh(G,X)(\mathbb{C}) = \varprojlim_{U \subset G(\mathbf{A}_f)} Sh_U(G,X)(\mathbb{C}), \tag{2.1.1}$$

where

$$Sh_U(G,X)(\mathbb{C}) = G(\mathbb{Q})\backslash(X \times G(\mathbf{A}_f)/U) \simeq \mathcal{M}_U(G) \tag{2.1.2}$$

(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, τ_0, σ_0, $p = uu^c$, $w \mid u$, Σ, etc.

Choose an $\mathbb{R}$-algebra homomorphism $h_0 : \mathbb{C} \to B^{op} \otimes_{\mathbb{Q}} \mathbb{R}$ such that $h_0(z)^\# = h_0(\bar{z})$ for all $z \in \mathbb{C}$. 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

$$g \mapsto diag(-1, 1, \ldots, 1)^t g^{-1} diag(-1, 1, \ldots, 1)$$

this means that $h_0(z)$ must be conjugate to $diag(z, \bar{z}, \bar{z}, \ldots, \bar{z})$ (in the τ_0 coordinate) and $\bar{z} \cdot I_n$ (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 $\mathbb{C}$ under τ_0.

Recall the Hodge-theoretic interpretation of $h \in X$. Any irreducible representation of $R_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m)$ is of the form $z \mapsto h_{p,q}(z) = z^{-p}\bar{z}^{-q}$ for $p, q \in \mathbb{Z}$. If $h \in X$ and (ρ, 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: $V^{p,q} = H^q(Y, \Omega^p)$ if V is the complex cohomology of a smooth complex variety Y.)

Let $V = B$ itself. Then

$$B_{\mathbb{C}} = B \otimes_{\mathbb{Q}} (\mathbb{C}) = \oplus_{\tau \in \Sigma} B_\tau \oplus B_{c\circ\tau};$$

moreover, B_v is the v-eigenspace for the action of F for $v \in Hom(F,\mathbb{C})$. 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 $*$, defining (via the trace) a bilinear form that takes rational values on $B(\mathbb{Q})$. Hence, choosing a lattice $\Lambda \in B(\mathbb{Q})$, 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 $Lie(A_0) = B(\mathbb{C})^{-1,0}$ as a sum of v-eigenspaces, we find

$$\dim Lie(A_0)_{\tau_0} = n, \quad \dim Lie(A_0)_\tau = 0, \tau \in \Sigma, \tau \neq \tau_0; \tag{2.1.3}$$

and for all τ,

$$\dim Lie(A_0)_\tau + \dim Lie(A_0)_{c\circ\tau} = n^2 = \dim_F B. \tag{2.1.4}$$

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 $T_f(A)$ denote the direct product of the Tate modules $T_\ell(A)$ over all primes ℓ, $V_f(A) = \mathbb{Q}\otimes T_f(A)$. Let $U \subset G(\mathbf{A}_f)$ be a compact open subgroup. Consider the functor $\mathcal{A}_U(B,*)$ on schemes over F, which to S associates the set of equivalence classes of quadruples (A,λ,i,η), where A is an abelian scheme of dimension dn^2, $\lambda : A \to \hat{A}$ is a polarization, $i : B \hookrightarrow End(A)\otimes\mathbb{Q}$ is an embedding, and $\eta : V\otimes_{\mathbb{Q}}\mathbf{A}_f \xrightarrow{\sim} V_f(A)$ an isomorphism of skew-hermitian (see below) $B\otimes_{\mathbb{Q}}\mathbf{A}_f$-modules, modulo U; here V is the $B\otimes B^{op}$-module B, as in (1.2).

Here is the precise meaning of "modulo U" following Kottwitz [K5, p. 390]). We may assume S connected. The Tate module $T_f(A)$ is a smooth $\mathbf{A}_f$-sheaf on S. Fixing a geometric point $s \in S$, it is thus the $\mathbf{A}_f$-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 $\eta : V\otimes_{\mathbb{Q}}\mathbf{A}_f \xrightarrow{\sim} V_f(A_s)$ 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 $End(A)\otimes\mathbb{Q}$ restricts to the involution $*$ on $i(B)$ and η takes the standard pairing on V to an $\mathbf{A}_f$-multiple of the Weil pairing for λ on $V_f(A)$. Most importantly, i induces an action i_F of the center F of B on the $\mathcal{O}_S$-module $Lie(A)$. For each embedding $\tau : F \to \mathbb{C}$, we let $\mathcal{O}_{S,\tau} = \mathcal{O}_S\otimes_{F,\tau}\mathbb{C}$, and let $Lie(A)_\tau = Lie(A)\otimes_{F,\tau}\mathbb{C}$. We then assume that

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

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

We assume U is sufficiently small; then $\mathcal{A}_U(B,*)$ is represented by a smooth projective scheme over F, also denoted $\mathcal{A}_U(B,*)$. 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 [Hida].)

(2.3) Points over $\mathbb{C}$. Hasse principle and connected components..

Using Riemann matrices, we show that $\mathcal{A}_U(B,*)$ is isomorphic to $|ker^1(\mathbb{Q},G)|$ copies of the canonical model of ${}_USh(G,X)$. I begin by explaining the source of the invariant $ker^1(\mathbb{Q},G) = ker[H^1(\mathbb{Q},G) \rightarrow \prod_v H^1(\mathbb{Q}_v,G)]$. Recall that G is the group of automorphisms of the $R_{F/\mathbb{Q}}B$-module V that preserve the $*$-skew-hermitian pairing $(x_1,x_2)_\beta$ up to a scalar. Here and below, a $*$-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 $V_{\overline{\mathbb{Q}}} \xrightarrow{\sim} V'_{\overline{\mathbb{Q}}}$ as skew-hermitian B-modules, and this gives rise to a class in $H^1(\mathbb{Q},G)$. Now suppose we have a point $x = (A,\lambda,i,\eta) \in \mathcal{A}_U(B,*)(\mathbb{C})$, and let $V' = H_1(A,\mathbb{Q})$. This defines a class $c(x) \in H^1(\mathbb{Q},G)$. Now η defines isomorphisms $V_{\mathbb{Q}_p} \xrightarrow{\sim} V'_{\mathbb{Q}_p}$ 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

$$ker[F^{+,\times}/\mathbb{Q}^\times N_{F/F^+}(F^\times) \;\rightarrow\; \mathbf{A}^\times_{F^+}/\mathbf{A}^\times N_{F/F^+}(\mathbf{A}^\times_F)].$$

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

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

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

Indeed, suppose $x \in S^i_U(B,*)(\mathbb{C})$. 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 ι, the datum η defines a point in $G(\mathbf{A}_f)/U$. On the other hand, the complex structure on $H_1(A,\mathbb{R}) = Lie(A)$ defines a map $h' : R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \rightarrow GL(V)$. Since the complex structure commutes with B, h' takes values in $B^{op,\times}$. Again, the conditions on the B-action on $Lie(A)$ 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 $\tilde{x}(\iota) \in X \times G(\mathbf{A}_f)/U$. The choice of ι is well-defined up to an element of $G(\mathbb{Q})$, thus x gives a well-defined point in $G(\mathbb{Q})\backslash(X \times G(\mathbf{A}_f)/U) = Sh_U(G,X)(\mathbb{C})$. We thus have a map

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

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 $G(\mathbf{A}_f)/U$ gives A in terms of a correct choice of lattice. On the other hand, every point in $X \times G(\mathbf{A}_f)/U$ 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 $\iota^i : H_1(A) \xrightarrow{\sim} V^i$; then the same argument goes through with G replaced by $G^i = Aut_B(V^i, (,)_i)$. Note that nothing changes except the set of ι^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 $G^i(\mathbf{A}_f)/U = G(\mathbf{A}_f)/U$), 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* $\ker^1(\mathbb{Q}, Z_G) \to \ker^1(\mathbb{Q}, G)$ *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 $G^{ad} = Aut(G)^0$, 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 $D = G/G^{der}$. Since G^{der} is an inner form of $SL(n)$, it satisfies the Hasse principle, and it follows from a simple diagram chase that $\ker^1(\mathbb{Q}, G) \to \ker^1(\mathbb{Q}, D)$ is injective. Surjectivity is a bit trickier. Let $T \subset G$ be a maximal torus, elliptic at some finite place, $T_{sc} = T \cap G^{der}$. Then the short exact sequence

$$1 \to T_{sc} \to T \to D \to 1$$

yields a commutative diagram of long exact sequences

(2.3.2)

$$\begin{array}{ccccccccc} \longrightarrow & H^1(\mathbb{Q}, T_{sc}) & \longrightarrow & H^1(\mathbb{Q}, T) & \longrightarrow & H^1(\mathbb{Q}, D) & \longrightarrow & H^2(\mathbb{Q}, T_{sc}) \dots \\ & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \longrightarrow & H^1(\mathbf{A}, T_{sc}) & \longrightarrow & H^1(\mathbf{A}, T) & \longrightarrow & H^1(\mathbf{A}, D) & \longrightarrow & H^2(\mathbf{A}, T_{sc}) \dots \end{array}$$

Now we have the following

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

$$\ker^2(\mathbb{Q}, T) = ker[H^2(\mathbb{Q}, T) \to H^2(\mathbf{A}, T)]$$

vanishes if T is anisotropic locally at one place.

In our situation, T_{sc} is elliptic at some finite place, hence the Lemma applies. It follows that any $y \in \ker^1(\mathbb{Q}, D)$ comes from an $x \in H^1(\mathbb{Q}, T)$ whose image in $H^1(\mathbf{A}, T)$ comes from $H^1(\mathbf{A}, T_{sc})$. Since $H^1(\mathbb{Q}, T_{sc})$ maps onto $H^1(\mathbb{R}, T_{sc})$ (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 $H^1(\mathbb{Q}, G)$. Clearly it maps onto y, and it remains to show $z \in \ker^1(\mathbb{Q}, G)$. By construction, it has trivial component at $\mathbb{R}$, and since $H^1(\mathbb{Q}_v, G^{der}) = 1$ (by Kneser's theorem, since G^{der} is simply connected) it is in fact in $\ker^1(\mathbb{Q}, G)$.

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 $Sh(G^i, X)$ as a canonical model defined over its reflex field F by the general theory of Shimura varieties. In particular, the subvarieties $S^i_U(B,*)$ of $\mathcal{A}_U(B,*)$ 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 $w = w_1, w_2, \dots, w_r$ be the primes of F above u. Write $K = F_w$ Since p splits in E, we can identify

$$G(\mathbb{Q}_p) \xrightarrow{\sim} GL(n,K) \times \prod_{i\,>\,1} B^{op,\times}_{w_i} \times \mathbb{Q}_p^{\times} \tag{2.4.1}$$

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

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 $\mathcal{A}_U(B,*)$ is representable, hence has a model over $Spec(\mathcal{O})$, also denoted $\mathcal{A}_U(B,*)$, that represents a slightly modified version of the functor considered above. First, we always take λ to be a prime-to-p polarization, and the equivalence is up to prime-to-p-isogenies. More importantly, $Lie(A)$ 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

(2.4.2) $Lie(A) \otimes_{\mathcal{O}_{F,p}} \mathcal{O}_{w_i} = 0, i > 1$;
(2.4.3) $Lie(A) \otimes_{\mathcal{O}_{F,p}} \mathcal{O}$ 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, $\mathcal{A}_U(B,*)$ is the union of $|ker^1(F,G)|$ copies of a $\mathcal{O}$-model $S_U(G,X)$ of ${}_KSh(G,X)$.

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

Proof. We follow Carayol [Ca1]. First, $\mathcal{A}_U(B,*)$ 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 $\mathcal{O}$, $S = Spec(R)$, and suppose we have a quadruple (A,λ,i,η) over the generic point $Spec(\mathcal{K})$. 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 $End(A_R) \xrightarrow{\sim} End(A_{\mathcal{K}})$ (functoriality of Néron models). Thus $\mathcal{O}_B$ acts on A_k, hence necessarily on T, hence on the character group $X_*(T)$. This group has $\mathbb{Z}$-rank at most equal to the $\dim A = dn^2$, 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 λ follows similarly by functoriality.

Finally, there is the question of extending η. The components of η away from p extend, because the ℓ-division points are étale over S. So we need only worry about extending η_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 $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p \simeq \prod_i \mathcal{O}_B \otimes_{\mathcal{O}_F} \mathcal{O}_{w_i}$. 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 $A_R[w_i^\infty]$ is a p-divisible group with $\mathcal{O}_{B_i}$ action. The condition that the Rosati involution restricts to the involution $*$ of the second kind on $\mathcal{O}_B$ implies that the polarization identifies

$$A_R[w_i^{c,\infty}] \xrightarrow{\sim} \hat{A}_R[w_i^\infty],$$

where $\hat{}$ denotes Cartier dual.

In general the data "polarization + η_p (mod U_p)" is equivalent to "level structure on $A_R[w_i^\infty]$ (mod U_{w_i}) for all i + trivialization of the Tate module of $\mathbb{G}_m$" (mod a subgroup of $\mathbb{Q}_p^\times$). The fact that η is invariant under the factor $\mathbb{Z}_p^\times \subset U_p$ implies that we only have to consider level structures on $A_R[w_i^\infty]$. The condition on the Lie algebras implies $A_R[w_i^\infty]$ is étale for $i > 1$, so the factor η_{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 $\bar{S} = \bar{S}_U(G, X)$ denote the special fiber of our model, and let $x = (A_x, \lambda_x, i_x, \eta_x) \in \bar{S}(\mathbb{F})$ be a geometric point. Let S be an Artinian local $\mathcal{O}$-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 $x \in \bar{S}(\mathbb{F})$ to an S/I-valued point of $\mathcal{A}_U(B, *)$. Then x' lifts to an S-valued point of $\mathcal{A}_U(B, *)$.*

The deformation property (2.4.5) is in fact a property of the formal completion $\mathcal{A}_U(B, *)\hat{}_x$ of $\mathcal{A}_U(B, *)$ at x, a formal scheme over $Spf(\mathcal{O})$. It therefore suffices to prove that $\mathcal{A}_U(B, *)\hat{}_x$ is formally smooth. In fact, we will prove that $\mathcal{A}_U(B, *)\hat{}_x$ 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 $Art(\mathcal{O}, \mathbb{F})$:

$$S \mapsto (A, \lambda, i, \eta) + j : (A, \lambda, i, \eta)_{\mathbb{F}} \xrightarrow{\sim} (A_x, \lambda_x, i_x, \eta_x)$$

This is represented by the formal completion $\mathcal{A}_U(B, *)\hat{}_x$. Consider the second functor $\mathcal{F}_2$

$$S \mapsto (\mathcal{G}, \lambda, i, \eta^w) + j : (\mathcal{G}, \lambda, i, \eta^w)_{\mathbb{F}} \xrightarrow{\sim} (A_x[p^\infty], \lambda_x[p^\infty], i_x[p^\infty], \eta_x^w)$$

The terms need to be explained. Here $\mathcal{G}$ is a p-divisible group scheme over S, λ an isomorphism $\mathcal{G} \xrightarrow{\sim} \hat{\mathcal{G}}$, i an inclusion $\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_p \to End(\mathcal{G})$, and η^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 $S \in Art(\mathcal{O}, \mathbb{F})$ 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 $\mathcal{A}_U(B,*)$, it suffices to study the functor $\mathcal{F}_2$. Obviously, the structure η^w is étale, as is the deformation of the data $A_x[w_i^\infty]$ for $i > 1$. So it suffices to study the deformation of $A_x[w^\infty]$. Now $\mathcal{O}_{B_w} \xrightarrow{\sim} M(n,\mathcal{O})$, by our original hypothesis. Thus there are n orthogonal idempotents in $\mathcal{O}_{B_w}$ which decompose $A_x[w^\infty]$ into n mutually isomorphic p-divisible groups with $\mathcal{O}$ action; this argument is called "Morita equivalence." Let $\mathcal{G}_x$ denote any of these divisible $\mathcal{O}$-modules, $\iota_x : \mathcal{O} \to End(\mathcal{G}_x)$ the action and let $\mathcal{F}_3$ be the functor

$$S \mapsto (\mathcal{G},\iota) + j : (\mathcal{G},\iota)_{\mathbb{F}} \xrightarrow{\sim} (\mathcal{G}_x,\iota_x). \tag{2.4.7}$$

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 $Lie(A_x) \otimes_{\mathcal{O}_B} \mathcal{O}_{B_w}$ is isomorphic to the sum of n copies of $Lie(\mathcal{G}_x)$ (by Morita equivalence). It follows from the definition of the moduli problem that $Lie(\mathcal{G}_x)$ is a projective (i.e. free) rank 1 $\mathcal{O}_S$-module. On the other hand, the height of the p-divisible group $\mathcal{G}_x$ is $n[K:\mathbb{Q}_p]$ (because $A_x[p]$ is a finite flat p-group scheme of rank $2dimA_x$. (Indeed, the polarization breaks up $A_x[p]$ as $A_1[p] \times A_2[p]$ each of height $dim[A]$, with $A_1[p] = A_x[p] \cap \prod_i A_x[w_i^\infty]$. 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 $n[K:\mathbb{Q}_p]$, as stated.)

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

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

$$0 \longrightarrow \mathcal{G}[\varpi^{m_2}] \longrightarrow \mathcal{G}[\varpi^{m_1}] \xrightarrow{\varpi^{m_2}} \mathcal{G}[\varpi^{m_1-m_2}] \longrightarrow 0$$

is an exact sequence of finite flat group schemes;

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

The height of the p-divisible $\mathcal{O}$-module $\mathcal{G}$ is defined to be the h such that $\mathcal{G}[\varpi]$ 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 $(\mathcal{G}_x,\iota_x)$ as a 1-dimensional height n divisible $\mathcal{O}$-module.

We consider the canonical exact sequence

$$0 \longrightarrow \mathcal{G}_x^0 \longrightarrow \mathcal{G}_x \longrightarrow \mathcal{G}_x^{et} \longrightarrow 0 \tag{2.4.9}$$

and let h denote the height of $\mathcal{G}_x^{et}$, so $\mathcal{G}_x^0$ 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), 23-88), skipping increasingly many details as the argument progresses. We begin with the case $h = 0$, and consider a 1-dimensional formal $\mathcal{O}$-module F over $\mathbb{F}$ of height n. Consider the category $Art(\mathcal{O},\mathbb{F})$ of Artinian local $\mathcal{O}$-algebras R with maximal ideal $\mathfrak{m} = \mathfrak{m}_R$

(containing ϖ) and residue field $\mathbb{F}$, and consider the functor of deformations of F on $Art(\mathcal{O}, \mathbb{F})$; i.e., p-divisible $\mathcal{O}$-modules G over $Spec(R)$ given with isomorphisms $j : G_{\mathbb{F}} \xrightarrow{\sim} F$. Because $Spec(R)$ 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 $G \equiv F \pmod{\mathfrak{m}}$ and $a_G \equiv a_F \pmod{\mathfrak{m}}$ 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 $\mathfrak{m}_R^{n-1}$ and the set of deformations of F to R.*

Note that $\mathfrak{m}_R^{n-1}$ 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 $Spf(\mathcal{O}[[t_1, \ldots, t_{n-1}]])$; i.e., $\mathcal{F}_3$ is prorepresented by a power series ring. In particular, taking F to be the formal group $\mathcal{G}_x$ above, we see that $\mathcal{F}_i$ is formally smooth for $i = 1, 2, 3$, which implies that $\mathcal{A}_U(B, *)$ is smooth at any point x where $\mathcal{G}_x^{et} = 0$.

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 Δ and δ_a have the form

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

where Δ_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 $\mathfrak{m}_R$, for $i = 1, \ldots, n-1$, and that they determine the remainder of the deformation. Explicitly, if (Δ, δ_a) is a cocycle, congruent to $\pmod{deg\ n+1}$, 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 $0 \pmod{deg\ n+1}$), whereas

$$(\Delta, \delta_a) \equiv (c\frac{p}{\varpi}[(X+Y)^n - X^n - Y^n], c\frac{a^n - a}{\varpi}X^n) \pmod{deg\ n+1}$$

if n is a power of q. This is exactly as in Lubin-Tate except for the presence of the δ_a.

By writing down the power series, we obtain a universal deformation over $Spf(R_{n,\mathcal{O}})$. Taking successive subgroups of ϖ^m-division points, we obtain a p-divisible $\mathcal{O}$-module $\tilde{\Sigma}_{K,n}$ over $Spf(R_{n,\mathcal{O}})$. It is not hard to see, because it is a direct limit of finite flat group schemes over $Spf(R_{n,\mathcal{O}})$, that in fact $\tilde{\Sigma}_{K,n}$ is actually a p-divisible $\mathcal{O}$-module over $Spec(R_{n,\mathcal{O}})$, 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 $\mathcal{O}_{D_{\frac{1}{n}}}$, the maximal order in the central division algebra $D_{\frac{1}{n}}$ over K with invariant $\frac{1}{n}$. This can be proved by explicit power series calculations, using the techniques of the Lubin-Tate theory; see Drinfel'd's paper for such a proof. One construction is by taking the reduction mod ϖ 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 $Spf(R_{n-h,\mathcal{O}})$. Let $R \in Art(\mathcal{O}, \mathbb{F})$. Over $\mathbb{F}$, we have an isomorphism

$$F \xrightarrow{\sim} F^0 \times (K/\mathcal{O})^h$$

where $F^0 \xrightarrow{\sim} \Sigma_{K,n-h}$.

(2.4.11) 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 $Spf(R_{n-h,\mathcal{O}})$.*

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):

$$0 \to G^0 \to G \to G^{et} \to 0.$$

Here G^0 is a deformation of $\Sigma_{K,n-h}$ and G^{et} 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 $(K/\mathcal{O})^h$. Formally, the short exact sequence

$$0 \to \mathcal{O}^h \to K^h = \varinjlim_{p^n} \mathcal{O}^h \to (K/\mathcal{O})^h \to 0 \tag{2.4.12}$$

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

$$\varprojlim_{p^n} Hom_{\mathcal{O}}(\mathcal{O}^h, G^0) \to Hom_{\mathcal{O}}(\mathcal{O}^h, G^0) \xrightarrow{\delta_R} Ext^1((K/\mathcal{O})^h, G^0) \to Ext^1(\varinjlim \mathcal{O}^h, G^0).$$

Now $Hom_{\mathcal{O}}(\mathcal{O}^h, G^0)$ is represented by $Hom_{\mathcal{O}}(\mathcal{O}^h, \mathfrak{m}_R)$. Since multiplication by p is contracting on $\mathfrak{m}_R$ and $\mathfrak{m}_R$ is nilpotent, the inverse limit is zero.

To conclude, it suffices to show that the map

$$Ext^1(K/\mathcal{O}, G^0) \to Ext^1(\varinjlim \mathcal{O}, G^0)$$

is zero, in other words, that any extension $\mathcal{G}$ 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

$$Hom_{\mathcal{O}}(\mathcal{G}, K/\mathcal{O} \times G^0) \otimes K \to Hom_{\mathcal{O}}(\mathcal{G}_{\mathbb{F}}, K/\mathcal{O} \times \Sigma_{K,n-h}) \otimes K$$

(restriction to the closed point) is an isomorphism.

More generally, we have

(2.4.13) 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):*

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

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 $Spec(\mathcal{O})$. Suppose $U \supset U'$ are two open compact subgroups of $G(\mathbf{A}_f)$ with $U_w = U'_w = GL(n,\mathcal{O})$ 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

$$G(\mathbf{A}_f^w) = G(\mathbf{A}_f^p) \otimes \prod_{i\,>\,1} B_{w_i}^{op,\times}, \tag{2.5.1}$$

so that

$$G(\mathbf{A}_f) = G_w \times \mathbb{Q}_p^\times \times G(\mathbf{A}_f^w) = GL(n,K) \times \mathbb{Q}_p^\times \times G(\mathbf{A}_f^w). \tag{2.5.2}$$

Thus any admissible irreducible representation π of $G(\mathbf{A}_f)$ can be factored

$$\pi = \pi_w \otimes \psi \otimes \pi^w, \tag{2.5.3}$$

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

Now suppose $g \in G(\mathbf{A}_f^w)$. 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

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

There is then a map (Hecke correspondence)

$$T(g) = p_{2,*}p_1^* : H^\bullet(\mathcal{A}_U(B,*),\mathbb{Q}_\ell) \to H^\bullet(\mathcal{A}_U(B,*),\mathbb{Q}_\ell)$$

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.

Lecture 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 $\wp_K$ 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 $\hat{K}^{nr}$ in [HT], and let σ denote the Frobenius (relative to $\mathbb{Q}_p$) acting on $\mathcal{K}$. We let $\mathcal{K}_K = \mathcal{K} \cdot K$ 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, ϕ) where N is a finite-dimensional $\mathcal{K}$-vector space and ϕ is a σ-linear bijection $N \to N$, in the sense that

$$\phi(av) = \sigma(a)\phi(v), a \in \mathcal{K},\ v \in N.$$

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 $\frac{r}{s}$ where $s = \dim N_{\frac{r}{s}}$ and $\phi^s(M) = p^r M$ for some W-lattice $M \subset N_{\frac{r}{s}}$. If $N \xrightarrow{\sim} \oplus N_{\frac{r_i}{s_i}}$ then the $\frac{r_i}{s_i}$ are the *slopes* of N.

To any p-divisible group $\mathcal{G}$ over $\mathbb{F}$ one can associate its (contravariant) Dieudonné module $\mathbb{D}(\mathcal{G})$, and its isocrystal $N(\mathcal{G}) = \mathbb{D}(\mathcal{G}) \otimes_W \mathcal{K}$. $\mathbb{D}(G)$ 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

$$\mathcal{G} \mapsto \mathbb{D}(\mathcal{G})$$

is an anti-equivalence of categories with the category of $W[F, V]$-modules as above, and $N(\mathcal{G}) \xrightarrow{\sim} N(\mathcal{G}')$ as isocrystals if and only if $\mathcal{G}$ and $\mathcal{G}'$ are isogenous.

There is a similar classification of divisible $\mathcal{O}$-modules. Let σ_q denote the lift of the Frobenius $Frob_q \in Gal(\mathbb{F}/k)$ to $Gal(\mathcal{K}_K/K)$, 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 $\mathcal{K}$-vector space N with σ_q-linear bijective morphism Φ). If $\mathcal{G}$ is a divisible $\mathcal{O}$-module, then $\mathbb{D}(\mathcal{G})$ 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 $\frac{r}{s}$; here $N_{r,s}$ has $F^s(M) = \varpi^r M$ for an appropriate lattice.

We have the relations

$$height(\mathcal{G}) = rank_{\mathcal{O}}\mathbb{D}(\mathcal{G}); \dim \mathcal{G} = \dim_{\mathbb{F}} V\mathbb{D}(\mathcal{G})/\varpi\mathbb{D}(\mathcal{G}). \tag{3.1.1}$$

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

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

$$b = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ \varpi & 0 & 0 & \dots & 0 \end{pmatrix} \tag{3.1.2}$$

Then $(b\sigma_q)^g = \varpi \cdot \sigma_q^g$ which implies that the slope is $\frac{1}{g}$. This already shows uniqueness of $\Sigma_{K,g}$ 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:

(3.1.3) Proposition [Dr,Prop. 1.7]**.** *(i) The algebra* $End(N,\phi)^{op} = End(\Sigma_{K,g}) \otimes \mathbb{Q}_p$ *is isomorphic to the central division algebra* $D_g = D_{K,g}$ *over* K *with invariant* $\frac{1}{g}$.

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

Let $N : D_{K,g}^\times \to K^\times$ be the reduced norm, and let $\Pi \in \mathcal{O}_{K,g} := \mathcal{O}_{D_{K,g}} = End(\Sigma_{K,g})$ 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 $Frob_q : \Sigma_{K,g} \to \Sigma_{K,g}^{(q)}$ is a map of $\mathcal{O}$-modules. Thus $\ker Frob_q = \ker \Pi$, which means that

$$\Sigma_{K,g}^{(q)} \xrightarrow{\sim} \Sigma_{K,g}/(\ker \Pi) \xrightarrow{\sim} \Sigma_{K,g}. \tag{3.1.4}$$

We have already defined quasi-isogenies: If $A \in Art(\mathcal{O},\mathbb{F})$ 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 $Hom_A(H_1,H_2) \otimes \mathbb{Q}$ 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 $D_g^\times \times GL(h,K)$, where D_g is the central division algebra over K with invariant $\frac{1}{g}$. 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 Π^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 $QDef(\Sigma_{K,g,h})$ from $Art(\mathcal{O},\mathbb{F})$ to $\{Sets\}$:

$$A \mapsto (H/A, j : \Sigma_{K,g,h} \to H_{\mathbb{F}}) \tag{3.1.5}$$

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

$$\breve{M}_g \xrightarrow{\sim} \breve{M}_g(0) \times \mathbb{Z} \tag{3.1.6}$$

where $\breve{M}_g(0) = Spf(R_{g,K})$ 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 $\mathbb{Z} \times GL(h,K)/GL(h,\mathcal{O})$ (quasi-isogenies of $(K/\mathcal{O})^h$ are indexed by lattices in K^h). We let $(\tilde{\Sigma}_{K,g,h},\tilde{j})$ denote the universal pair over $\breve{M}_{g,h}$.

We need something slightly more general: Let $\widehat{\quad}$ denote Cartier dual, and consider $\Sigma^+_{K,g,h} = \Sigma_{K,g,h} \times \widehat{\Sigma_{K,g,h}}$. This p-divisible group has a canonical polarization $\psi : \Sigma^+_{K,g,h} \times \Sigma^+_{K,g,h} \mapsto \mu_{p^\infty}$, where μ_{p^∞} denotes the p-divisible group of $\mathbb{G}_m$. The functor $QDef(\Sigma^+_{K,g,h})$ classifies pairs $(H^+/A, j : \Sigma^+_{K,g,h} \to H^+_{\bar{k}})$ 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 $\breve{M}^+_{g,h}$, which can be split canonically as $\breve{M}_{g,h} \times \mathbb{Q}_p^\times/\mathbb{Z}_p^\times$, with the second factor for the polarization. The universal pair over $\breve{M}^+_{g,h}$ is denoted $(\tilde{\Sigma}^+_{K,g,h},\tilde{j}^+)$. By analogy with (3.1.6), there is a non-canonical isomorphism

$$\breve{M}^+_{g,h} \xrightarrow{\sim} \breve{M}^+_{g,h}(0) \times \mathbb{Z} \times GL(h,K)/GL(h,\mathcal{O}) \times \mathbb{Q}_p^\times/\mathbb{Z}_p^\times, \tag{3.1.7}$$

where $\breve{M}^+_{g,h}(0)$ 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 $\mathcal{A}_U(B,*)$ at a point x of the special fiber is isomorphic to $\breve{M}_{g,h}$ 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 $\mathbb{F}_q$. This can be done first on the special fiber; the Galois action lifts uniquely.
(2) Next, we add (Drinfel'd) 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 $\bar{S}_U$ the special fiber of $\mathcal{A}_U(B,*)$. Let $\bar{S}_U^{(h)}$, or just $\bar{S}^{(h)}$, be the set of points $x \in \bar{S}_U(\mathbb{F})$ such that $\mathcal{G}_x^{et}$ 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 $x \in \bar{S}_U^{(h)}$, and consider the set $\bar{S}(x) = \bar{S}_U(x)$ of points in the isogeny class of x. It is obviously contained in $\bar{S}^{(h)}$. Thus $\bar{S}(x)$ 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 ϕ breaks up into the w-component and the prime-to-w component. The latter is a lattice in $V_f^w(A)$, isomorphic via η^w to $V(\mathbf{A}_f^w)$. Since ϕ respects the other structures, this gives a well-defined point in $G(\mathbf{A}_f^w)/U^w$. The w-component is the same as an isogeny of $\mathcal{O}$-modules $\mathcal{G}_x \mapsto \mathcal{G}_{x'}$. Taking $\mathcal{G}_x$ as our model for $\Sigma_{K,g,h}$, with $g = n - h$, we thus obtain a point of $\breve{M}_{g,h}(\mathbb{F})$. Thus x' corresponds to a pair $(m, g^w) \in \breve{M}_{g,h}(\mathbb{F}) \times G(\mathbf{A}_f^w)/U^w$. 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, λ, i), acting diagonally on the two data. In this way, we obtain a bijection (uniformization of an isogeny class):

$$\Theta : I_x(\mathbb{Q})\backslash \breve{M}_{g,h}^{+}(\mathbb{F}) \times G(\mathbf{A}_f^w)/U^w \to \bar{S}(x). \tag{3.2.1}$$

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:

$$\Theta : I_x(\mathbb{Q})\backslash \breve{M}_{g,h}^{+} \times G(\mathbf{A}_f^w)/U^w \to \mathcal{A}_U(B,*)_{\widehat{\bar{S}(x)}}. \tag{3.2.2}$$

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 $R \in Art(\mathcal{O}, \mathbb{F})$, and (m, g^w) a point in $\breve{M}_{g,h}(R) \times G(\mathbf{A}_f^w)/U^w$. Thus m corresponds to a pair $(H/R, j : \Sigma_{K,g,h} \to H_{\mathbb{F}})$. Recall that $\Sigma_{K,g,h}$ is identified with $\mathcal{G}_x$ for the fixed basepoint. Lift (A, λ, 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, λ_1, i_1) comes from a certain CM type. Let $\mathcal{G}_{x,1}$ be the corresponding lifting of $\mathcal{G}_x$.

By rigidity of quasi-isogenies (2.4.13), the map j lifts to a quasi-isogeny $j_1 : \mathcal{G}_{x,1} \to H$. The kernel of this quasi-isogeny defines a (virtual) subgroup scheme $S_m \subset A_1[w^\infty]$, whereas g^w defines a lattice $T_{g^w} \subset V_f^w(A_1)$. Suppose for simplicity that S_m is a genuine subgroup scheme and $T_{g^w} \supset T_f^w(A_1)$. 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 ϖ^m structures in characteristic 0.

We want Θ to be rational over $\mathbb{F}_q$ in a certain sense. Rapoport and Zink construct a "Weil descent datum" on $\breve{M}_{g,h}$, as follows. Let σ_q denote the (arithmetic)

Frobenius automorphism in $Gal(\mathbb{F}/\mathbb{F}_q)$, and let $Frob_q : \tilde{\Sigma}_{K,g,h} \to (\sigma_q)^*(\tilde{\Sigma}^+_{K,g,h})$ denote the Frobenius morphism of the (polarized) p-divisible $\mathcal{O}$-module as above. (We will need the polarization in what follows.) Let $R \in Art(\mathcal{O}, \mathbb{F})$, 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 $\breve{M}_{g,h}$. Define $H^\alpha = H$, and let j^α be the morphism

$$j \circ \phi^*(Frob_q^{-1}) : (\sigma_q)^*(\tilde{\Sigma}^+_{K,g,h})_{\mathbb{F}} \to (\tilde{\Sigma}_{K,g,h})_{\mathbb{F}} \to \Sigma_{K,g,h}.$$

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

$$\breve{M}_{g,h}(A) \to \breve{M}_{g,h}(A_{[\sigma_q]}); (H, j) \mapsto (H^\alpha, j^\alpha). \tag{3.2.3}$$

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 σ_q on roots of unity).

If $\breve{M}_{g,h}$ had a $\mathbb{F}_q$-rational structure, then α would correspond to the action of σ_q on points (say over $\mathbb{F}$). The fact that Θ commutes with the action of σ_q comes down, after verification, to the fact that Frobenius on (A, λ) corresponds to Frobenius on $\mathcal{G}_x \times \hat{\mathcal{G}}_x$ together with the polarization.

To any such deformation problem, Rapoport and Zink associate a pair of groups (G, J) over $\mathbb{Q}_p$, with J an inner form of a Levi factor of a rational parabolic subgroup of G. For $\breve{M}^+_{n-h,h}$, the group G is $GL(n, K) \times \mathbb{Q}_p^\times$, and $J = J_{n-h,h,+} = D^\times_{1/n-h} \times GL(h, K) \times \mathbb{Q}_p^\times$. In any case, J is the group of self-quasi-isogenies of the relevant divisible $\mathcal{O}$-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 $m \geq 0$, $\varpi \in \mathcal{O}$ a uniformizing parameter, and consider Drinfel'd ϖ^m-level structures.

(3.3.1) Definition. *Let $R \in Art(\mathcal{O}, \mathbb{F})$, and $(H, j) \in \breve{M}_{g,h}(R)$. A map of groups*

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

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

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

(3.3.3) *The induced map $(\varpi^{-m}\mathcal{O}/\mathcal{O})^n/M \to H[\varpi^m](R)/H^0[\varpi^m](R)$ 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 $Art(\mathcal{O},\mathbb{F})$ 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 $\breve{M}_{g,h}$ by a formal scheme $\breve{M}_{g,h;m}$. We can do the same with $+$; however, we always consider polarizations only up to $\mathbb{Z}_p^\times$-multiples. One of the main results of [Dr] is that the formal scheme $\breve{M}_{g,h;m}$ is flat over $\breve{M}_{g,h}$ and is regular; however, it has bad singularities in characteristic p. Its rigid generic fiber is precisely ${}_{U(m)}\breve{M}_{g,h}^{rig}$, with $U(m) \subset GL(g+h,\mathcal{O})$ the principal congruence subgroup of level ϖ^m. Note that the free rank g summand M in the previous paragraph is a discrete invariant of the triple (H,j,p); thus

$$\breve{M}_{g,h;m} = \coprod_M \breve{M}_{g,h;M}$$

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

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

The action of G, previously defined on ${}_\infty\breve{M}_{g,h}^+$, also extends to the family of integral models $\breve{M}_{g,h;m}$. Here is the construction. Let $(H,j,p) \in \breve{M}_{g,h;m}(A)$ 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 $\gamma^{-1} \in M_g(\mathcal{O})$ and that $\gamma \cdot \mathcal{O}^{g+h} \subset \varpi^{m'-m}\mathcal{O}^{g+h} \subset \varpi^{-m}\mathcal{O}^{g+h}$. Suppose in addition that $\gamma \in P_{M_0}$, and let (γ_g,γ_h) denote its projection on $GL(g) \times GL(h)$. Then γ 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^γ a Drinfel'd level m'-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" $p(\gamma \cdot \mathcal{O}^{g+h}) \subset H[\varpi^m](A)$. (This notion will be defined more generally in the global setting.) Now let $\Sigma(\gamma_g) = \Sigma_{K,g,h}/\ker(Frob_K^{-v_K(\det(\gamma_g))})$. Then j^{-1} identifies $H^\gamma_{\bar{k}}$ with

$$\Sigma^\gamma_{K,g,h} = \Sigma_{K,g,h}/\ker(Frob_K^{-v_K(\det(\gamma_g))}) \times (K^h/\gamma_h \cdot \mathcal{O}^h)$$

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 $\ker(Frob_q^d)$ for some d. We obtain j^γ by composing

$$\Sigma_{K,g} \times (K/\mathcal{O})^h \overset{(Frob_K^{-v_K(\det(\gamma_g))},\gamma_h)}{\longrightarrow} \Sigma^\gamma_{K,g,h} \times (K^h/\gamma_h\cdot\mathcal{O}^h) \overset{j}{\longrightarrow} H^\gamma_{\bar{k}}.$$

Finally, p^γ is just $p \circ \gamma$ where γ is viewed as an embedding

$$\varpi^{-m'}\mathcal{O}^{g+h}/\mathcal{O}^{g+h} \subset \varpi^{-m}\mathcal{O}^{g+h}/\gamma\cdot\mathcal{O}^{g+h}.$$

Letting m and m' vary, we obtain an action of γ satisfying the above properties on the tower $\{\breve{M}_{g,h;p^{-m}M_0/M_0}\}$ 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 $(x^{-1}, x^{-1}) \in P_{M_0} \times J$ 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 $\{\breve{M}_{g,h;p^{-m}M_0/M_0}\}$, factoring through $(P_{M_0} \times J)/K^\times \times W_K$, where $K^\times$ 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 $(G \times J)/K^\times \times W_K$ on the tower $\{\breve{M}_{g,h;m}\}$, which we denote $\breve{M}_{g,h;\infty}$ and view as the projective limit of the $\breve{M}_{g,h;m}$. Note that covering the isomorphism (3.1.7) (and ignoring the +) we have an isomorphism of (ind-profinite) schemes over k:

$$\breve{M}_{g,h;\infty,red} \xrightarrow{\sim} \mathbb{Z} \times GL(h, K) \tag{3.3.4}$$

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 $\pi : \breve{M}^+_{g,h;M} \to \breve{M}^+_{g;m}$. Here if $(H, j, p) \in \breve{M}^+_{g,h;M}(A)$ 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 $H^0_{\bar{k}}$, and p^0 the restriction of p to M. We can factor $\pi = \pi_3 \circ \pi_2 \circ \pi_1$. Here, letting $\breve{M}_{g,h;m,0}$ denote the moduli space over $\breve{M}_{g,h}$ of Drinfel'd level ϖ^m-structures on the connected subgroup of the variable height $g+h$-divisible $\mathcal{O}$-module, we have

$$\pi_1 = \pi_{1,M} : \breve{M}^+_{g,h;M} \to \breve{M}_{g,h;m,0} \times \breve{M}^+_{0,h}$$

takes (H, j, p) to (H, j, p^0, p^{et}), with p^{et} the induced level structure

$$(\varpi^{-m}\mathcal{O}/\mathcal{O})^{g+h}/M \to H/H^0[\varpi^{-m}].$$

Moreover, $\pi_2(H, j, p^0, p^{et}) = (H, j, p^0)$ (forget p^{et}), and π_3 is the base change to $\breve{M}^+_{g;h;0}$ of an analogous map $\pi'_3 : \breve{M}^+_{g,h} \to \breve{M}^+_{g,0}$. (i.e., forget the étale part altogether).

(3.3.5) Proposition. *The map π_2 (resp. π'_3, resp. π_1) is étale, (resp. smooth, resp. radicial over the special fiber).*

Proof. The statement concerning π_2 is easy. The smoothness of π'_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 π_1 is left as an exercise.

(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 $\mathcal{O}$-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 $p : (\varpi^{-m}\mathcal{O}/\mathcal{O})^n \to H[\varpi^m](S)$. Every point $p(x)$ is then an S-subscheme of $H[\varpi^m](S)$. The set $\{p(x), x \in (\varpi^{-m}\mathcal{O}/\mathcal{O})^n\}$ is a "full set of sections" of $H[\varpi^m]$ if, for any affine S-scheme $Spec(R)$ and every function $\phi \in B = H^0(H[\varpi^m]_R, \mathcal{O}_{H[\varpi^m]_R})$, there is the equality of characteristic polynomials in $R[T]$:

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

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

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

is represented by $\mathcal{S} = H[\varpi^m]^{q^{mn}}$. 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 $\mathcal{S} = Spec(R)$, $H[\varpi^m] = Spec(B)$ 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 $\mathcal{S}$. 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

$$Norm_{B[T_1,\ldots,T_M]/R'}(\Phi) = \prod_i f(P_i)$$

i.e.

$$Norm(\sum T_i b_i) = \prod_i(\sum T_i b_i(P_i)).$$

Both sides are homogeneous forms of degree M in $T_1, \dots, T_M$, 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 $P_1, \dots, P_M$ 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 $B[T_1, \dots, T_M]/R'$, 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 $M \geq 1$, $a_1, \dots, a_M$ 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 ϕ and the a_i as independent variables in a big field K, and

$$K[X]/\prod (X - a_i) \xrightarrow{\sim} \prod K[X]/(X - a_i)$$

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:

$$f_{\varpi^m}(T) = g(T) \prod_x (T - T(p(x))),$$

for some power series g. Over $\mathbb{F}$, we may assume $f_{\varpi^m}(T) = T^{q^{mg}}$ (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

$$B = R[[X]]/(f_{\varpi^m}) \xrightarrow{\sim} R[X]/(F)$$

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 Θ with Lemma 3.4.7, we obtain morphisms of all levels. Let $x \in \bar{S}_U^{(h)}$, and let $\bar{S}(x, m)$ denote the inverse image of the isogeny class $\bar{S}(x)$ in $\mathcal{A}_{U(m)}(B,*)$. Then because the Drinfel'd basis depends only on the p-divisible group, we can lift Θ to

$$\Theta_m : I_x(\mathbb{Q})\backslash \breve{M}^+_{n-h,h;m} \times G(\mathbf{A}_f^w)/U^w \xrightarrow{\sim} \mathcal{A}_{U(m)}(B,*)\widehat{\ }_{\bar{S}(x_m)}.$$

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

$$\Theta_\infty : I_x(\mathbb{Q})\backslash \breve{M}^+_{n-h,h;\infty} \times G(\mathbf{A}_f^w) \xrightarrow{\sim} \varprojlim_{U^w,m} \mathcal{A}_{U(m)}(B,*)\widehat{\ }_{\bar{S}(x_m)}. \tag{3.4.9}$$

This commutes with the Weil group action, as before. Note that the action of $I_x(\mathbb{Q})$ on $\breve{M}_{n-h,h;\infty}$ is given by associating to a self-quasi-isogeny of A_x a self-quasi-isogeny of $\mathcal{G}_x$. In other words, it factors through a homomorphism $I_x(\mathbb{Q}) \rightarrow J = J_{n-h,h} = D^\times_{\frac{1}{n-h}} \times GL(h,K)$. Write $G^{(h)}(\mathbf{A}_f) = G(\mathbf{A}_f^w) \times J_{n-h,h}$ (an abuse of notation, because $G^{(h)}(\mathbf{A}_f)$ is not the group of $\mathbf{A}_f$-points of something called $G^{(h)}$). Then (3.4.9) can be rewritten

$$[\breve{M}^+_{n-h,h;\infty} \times (I_x(\mathbb{Q})\backslash G^{(h)}(\mathbf{A}_f))]/J_{n-h,h} \xrightarrow{\sim} \varprojlim_{U^w,m} \mathcal{A}_{U(m)}(B,*)\widehat{\ }_{\bar{S}(x_m)}. \tag{3.4.10}$$

where the $J_{n-h,h}$-action on the left hand side is diagonal (on the left on $\breve{M}_{n-h,h;\infty}$ 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 $G(\mathbf{A}_f^w)$ 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) Θ_∞ 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 $\mathcal{G}$ 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 $m \geq m'$. Note that if (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 $m-m' >> 0$. Under these assumptions we will define a morphism

$$(g_0,g) : \mathcal{A}_{U(m)}(B,*) \rightarrow \mathcal{A}_{U(m)}(B,*).$$

It will send (A,λ,i,η^w,p) over T to $(A/(C \oplus C^\perp), p^{val_p(g_0)}\lambda, i, \eta^w, p\circ g)$, where

(3.5.1) $C_1 \subset \mathcal{G}[\varpi^m]$ is the unique closed subgroup scheme for which the set of $p(x)$ with $x \in g\cdot(\mathcal{O}^n)/\mathcal{O}^n$ is a full set of sections;
(3.5.2) $C = (\mathcal{O}^n_{F,w} \otimes_{\mathcal{O}_{F,w}} C_1) \subset A[\varpi^{-val_p(g_0)}]$;

(3.5.3) $C^{\perp}$ is the annihilator of $C \subset A[\varpi^{-val_p(g_0)}]$ inside $A[(u^c)^{-val_p(g_0)}]$ under the λ-Weil pairing;

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

$$\begin{array}{ccc} A & \xrightarrow{p^{-val_p(g_0)}\lambda} & A^{\vee} \\ \downarrow & & \uparrow \\ A/(C \oplus C^{\perp}) & \xrightarrow{p^{val_p(g_0)}\lambda} & (A/(C \oplus C^{\perp}))^{\vee}; \end{array}$$

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

$$\begin{array}{ccc} \varpi^{-m'}(\mathcal{O}^n)/\mathcal{O}^n & \xrightarrow{p\circ g} & \mathcal{G}[\varpi^m]/C_1(T) \\ \downarrow & & \downarrow \\ \varpi^{-m'}g(\mathcal{O}^n)/g(\mathcal{O}^n) & \longrightarrow & (\mathcal{G}[\varpi^{\infty}]/C_1)[\varpi^{m'}](T) \\ \downarrow & & \downarrow \\ \varpi^{-m'}(\mathcal{O}^n)/g(\mathcal{O}^n) & \longrightarrow & (\mathcal{G}[\varpi^m]/C_1)(T) \\ \uparrow & & \uparrow \\ \varpi^{-m'}(\mathcal{O}^n)/\mathcal{O}^n & \xrightarrow{p} & \mathcal{G}[\varpi^m](T); \end{array}$$

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 $p \in G(\mathbf{A}_f^w)$ – over the generic fiber, which is Zariski dense in the integral model, hence over the whole scheme. Indeed, the diagonal element $p \in Z_G(\mathbb{Q})$ acts trivially on the Shimura variety, but it is the product of $p \in G(\mathbf{A}_f^w)$ 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 $G(\mathbf{A}_f)$ on the tower of moduli schemes $\mathcal{A}_{U(m)}(B, *)$ over $\mathcal{O}$, in such a way that the uniformization map (3.4.9) is $W_K \times G(\mathbf{A}_f)$-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].

Lecture 4. Stratification and vanishing cycles

The present lecture continues the study of the stratification of the special fiber $\bar{S}_U^{(h)}$ of our Shimura variety by isomorphism type of isocrystal, which in the present simple situation corresponds 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 $U_w = GL(n, \mathcal{O})$, and U^w sufficiently small, so that $\mathcal{A}_U(B,*)$ has a smooth model over $\mathcal{O}$. We return to the stratum $\bar{S}_U^{(h)}$ defined last time; this is the subset where $\mathcal{G}_x^{et}$ is of height h. We prove that each $\bar{S}_U^{(h)}$ is smooth of dimension h.

In fact, we can replace $\mathcal{A}_U(B,*)_{\mathbb{F}}$ by any smooth locally noetherian scheme S over $\mathbb{F}$, and consider a one-dimensional divisible $\mathcal{O}$-module H/S of height n. We know that, when $S = \mathcal{A}_U(B,*)_{\mathbb{F}}$, then, for every $s \in S(\mathbb{F})$, the formal completion $S_s^{\hat{}}$ is isomorphic to the universal formal deformation space (over $\mathbb{F}$) 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 $S^{[h]}(\mathbb{F}) \subset S(\mathbb{F})$ be the subset where the height of H^{et} is $\leq h$, $S^{(h)}(\mathbb{F}) = S^{[h]}(\mathbb{F}) - S^{[h-1]}(\mathbb{F})$.

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

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

$$0 \to H^0 \to H \to H^{et} \to 0$$

where H^0 is a one-dimensional formal $\mathcal{O}$-module of height $n-h$ and H^{et} 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^{et} étale, and since both H^0 and H^{et} 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 $S = Spec(R)$ where R is a noetherian ring and $Lie(H)$ 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 $Fr_S : S \to S$ denote the absolute Frobenius morphism. The superscript $^{(p)}$, for schemes over S, denotes pullback with respect to Fr_S. Let

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

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

$$\underline{V}_* : \mathcal{H}^{(p)} \;\to\; \mathcal{H}.$$

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

$$F_{\mathcal{H}} : \mathcal{H} \to Fr_S^*(\mathcal{H})$$

we obtain

$$V_* = \underline{V}_* \circ F_{\mathcal{H}} : \mathcal{H} \;\to\; \mathcal{H}, \tag{4.1.2}$$

a $Frob$-linear version V_* of $\underline{V}_*$:

$$V_*(ay) = a^p V_*(y),\;\; a \in R,\;\; y \in \Gamma(*, \mathcal{H}).$$

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

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

Proof. 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

$$\mathcal{H}_s \xrightarrow{\sim} Hom(H_s[p], \mathbb{G}_a). \tag{4.1.4}$$

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

$$H_s[p](k(s)) \times \mathcal{H}_s \;\to\; \mathbb{G}_a(k(s)) = k(s),$$

which restricts to a pairing

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

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

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

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 $\mu_p^h \hookrightarrow H_s[p]^\vee$. Then there is an embedding

$$Lie(\mu_p)^h \hookrightarrow \mathcal{H}_s,$$

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

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

which yields the desired bound.

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

$$V_*(e_i) = \sum_j v_{ij} e_j.$$

Then $\mathcal{H}^{V_*=1}$ is identified with the subscheme of $\mathbb{A}_R^m$ 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

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

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 $S = Spec(R)$ in a finite separable extension of $Frac(R)$ where the étale covering is trivialized, so $\mathcal{H}_{k(T)}$ 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 ≤ 1.

Step 4. It remains to prove smoothness. This is more subtle, and requires Drinfel'd'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

$$Spf(R_{K,n-h,h}) \overset{\phi}{\underset{\sim}{\longleftarrow}} S_s^{\hat{}} \overset{f}{\longleftarrow} S^{(h)}{}_s^{\hat{}} \overset{cl}{\longrightarrow} Spf(R_{K,n-h}).$$

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

Let P be a minimal prime of $\mathcal{O}_{\widehat{S,s}}$ 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 $R_{K,n-h} \to \mathcal{O}_{\widehat{S,s}}/P$.

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 $(K/\mathcal{O})^h$). We will show that the parameters $t_1, \ldots, t_{n-h-1}$ of Drinfel'd map to zero in $\mathcal{O}_{\widehat{S,s}}/P$. Assuming this, we conclude as follows. It follows that the canonical classifying map $cl_h : S^{(h)}\hat{}_s \xrightarrow{cl} Spf(R_{K,n-h,h})$, corresponding to the deformation of H_s over $S^{(h)}\hat{}_s$, corresponds to a homomorphism

$$R_{K,n-h,h}/(t_1, \ldots, t_{n-h-1}) = \mathcal{O}[[t_1, \ldots, t_{n-h-1}, u_1, \ldots, u_h]]/(t_1, \ldots, t_{n-h-1}) \to \mathcal{O}_{\widehat{S^{(h)},s}}$$

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 $S^{(h)}\hat{}_s$ 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_ϖ (= multiplication by ϖ 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 $T = Spec(R)$, R a Henselian discrete valuation ring, with generic point η 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_η and S_t, and geometric generic fiber $S_{\bar{\eta}}$. Let $\mathcal{F}$ be a constructible sheaf on S_η in $\mathbb{Q}_\ell$-vector spaces, with $\ell \neq p$. The point of vanishing cycles is to calculate $H^\bullet(S_{\bar{\eta}}, \mathcal{F})$ as the hypercohomology of a complex $R\Psi(\mathcal{F})$ on S_t. There is an action of $Gal(k(\bar{\eta})/k(\eta))$ on $H^\bullet(S_{\bar{\eta}}, \mathcal{F})$, hence one wants $R\Psi(\mathcal{F})$ to be endowed with an action of $Gal(k(\bar{\eta})/k(\eta))$ (= inertia). The recipe is formal. One considers the canonical morphisms $\tilde{j} : S_{\bar{\eta}} \to S$ and $i : S_t \to S$; then

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

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

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

Then the Leray spectral sequence becomes

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

More generally, one starts with $k(s)$ perfect (e.g. finite) and takes base change over T by the Witt vectors $W(k(\bar{s}))$; then the spectral sequence becomes equivariant for $Gal(k(\bar{\eta})/k(\eta))$ covering the action of $Gal(k(\bar{s})/k(t))$. 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 [SGA 7]; however, Illusie's article [Il] 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, $T = Spec(\mathcal{O})$, $S = \mathcal{A}_{U(m)}(B,*)$, and we restrict attention for simplicity to $\mathcal{F} = \mathbb{Q}_\ell$. Write $\bar{S}_{U(m)} = \bar{\mathcal{A}}_{U(m)}(B,*)$, the special fiber of S, 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

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

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

$$\varinjlim_{U^w,m} H^p(\bar{S}_{U(m)}, R^q\Psi) \Rightarrow \varinjlim_{U^w,m} H^{p+q}(\mathcal{A}_{U(m)}(B,*)_{\bar{K}}, \mathbb{Q}_\ell). \tag{4.2.2}$$

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

Write $[\pi] = \sum n_{\Pi,\sigma}(\pi)\Pi \otimes \sigma$, and define

$$[H(\mathcal{A}(B,*))] = \sum_j (-1)^j [\varinjlim_{U^w,m} H^j(\mathcal{A}_{U(m)}(B,*)_{\bar{K}}, \mathbb{Q}_\ell)] \tag{4.2.3}$$

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

$$[H(\mathcal{A}(B,*))] = \sum_{p,q} (-1)^{p+q} [\varinjlim_{U^w,m} H^p(\bar{S}_{U(m)}, R^q\Psi)]. \tag{4.2.5}$$

Here we are making use of the fact [F] that the action of $G(\mathbf{A}_f)$ extends canonically to an action on $R\Psi$ by cohomological correspondences, covering the action on $\varprojlim_{U,m} \bar{S}_{U(m)}$.

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

$$\ldots \to H^p_c(\bar{S}^{(h)}_{U(m)}, \Phi) \to H^p(\bar{S}^{[h]}_{U(m)}, \Phi) \to H^p(\bar{S}^{[h-1]}_{U(m)}, i^*_{h-1}\Phi) \to \ldots \tag{4.2.6}$$

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

$$[H(\mathcal{A}(B,*))] = \sum_{p,q,h} (-1)^{p+q} [\varinjlim_{U^w,m} H_c^p(\bar{S}^{(h)}_{U(m)}, i_h^* R^q\Psi))]. \tag{4.2.7}$$

We drop the U and m and just write the right-hand side

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

The stability of each $\bar{S}^{(h)}$ under $G(\mathbf{A}_f)$ follows from the fact that $G(\mathbf{A}_f)$ 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 $G(\mathbf{A}_f)$-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

$$[\varinjlim_m H_c^p(\bar{S}^{(h)}_{U(m)}, i_h^* R^q\Psi))]^{\Gamma_m} = H_c^p(\bar{S}^{(h)}_{U(m)}, i_h^* R^q\Psi))$$

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 $H_c^p(\bar{S}^{(h)}_{U(m')}, i_h^* R^q\Psi))^{\Gamma_m}$ for all $m' > m$. More generally, if $f : Z' \to Z$ is a quotient by a finite group Γ, and if L' is a constructible sheaf on Z' with compatible Γ-action, we have $H_c^p(Z,(f_*L')^\Gamma) \xrightarrow{\sim} H_c^p(Z',L')^\Gamma$. So it suffices to prove the

(4.2.9) Continuity lemma.

$$R^q\Psi(m) \xrightarrow{\sim} f_{m',m,*} R^q\Psi(m')^{\Gamma_m}.$$

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 $\mathcal{X}$ is a "special" formal scheme over $Spf(\mathcal{O})$, Berkovich has constructed a vanishing cycles functor $R\Psi^{form}$ 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 $\breve{M}_{n-h,h;m}$ of Drinfel'd level structures are special in Berkovich's sense; we have seen that their connected components are isomorphic to the formal completion of $\mathcal{A}_{U(m)}(B,*)$ along points in $\bar{S}^{(h)}$. 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 $Spf(R_{n-h;m})$). More generally, the special fiber is a union of connected components of the form $Spf(R_{n-h,h;m})$ 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

$$\Psi^i_{K,\ell,n-h,h,m} = H^0(\breve{M}_{n-h,h;m,red}, R^i\Psi^{form}\mathbb{Q}_\ell)$$

for the formal scheme $\breve{M}_{n-h,h;m}$; here $\breve{M}_{n-h,h;m,red}$ 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, $\breve{M}_{n-h,h;m,red}$ is just a countable disjoint union of profinite schemes.) We let

(4.3.1)

$$\Psi^i_{K,\ell,n-h,h} = \Psi^i_{n-h,h} = \varinjlim_m \Psi^i_{K,\ell,n-h,h,m} = \varinjlim_m H^0(\breve{M}_{n-h,h;m,red}, R^i\Psi^{form}\mathbb{Q}_\ell).$$

Then each $\Psi^i_{n-h,h}$ has an action of $G\times J\times W_K$. When $h=0$, we have $G = GL(n,K)$, $J = D^\times_{\frac{1}{n}}$, and then $\Psi^i_n = \Psi^i_{n,0}$ is called the **fundamental local representation** of $G\times J\times W_K$. More precisely, the virtual representation

$$[\Psi_n] = \sum_i (-1)^i \Psi^i_n$$

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

Let $h=0$, and identify $\breve{M}_{n-h,0;\infty,red}$ with $\mathbb{Z}$ as in (3.3.4); let $x_0 \in \breve{M}_{n-h,0;\infty,red}$ correspond to $0\in\mathbb{Z}$ (quasi-isogenies of height 0). The stalk $\Psi^i_{n-h,0,x_0}$ of $R^i\Psi^{form}\mathbb{Q}_\ell$ at x_0 inherits a representation of the isotropy subgroup at x_0

(4.3.2) $$A_{K,n-h} \subset GL(n-h,K)\times J_{n-h}\times W_K.$$

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

$$\delta : GL(g,K)\times J_g\times W_K \ \rightarrow\ \mathbb{Z}$$

defined by

(4.3.3) $$\delta(\gamma,j,\sigma) = w_K(det(\gamma)) - w_K(N(j)) - w(\sigma)$$

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

(4.3.4) $$\Psi^i_{n-h,0} = c-Ind_{A_{g,K}}^{GL(g,K)\times J_g\times W_K}\Psi^i_{n-h,0,x_0},$$

where $c-Ind$ denotes induction with compact support.

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

$$J_{n-h,h,+} = D^\times_{\frac{1}{n-h}}\times GL(h,K)\times\mathbb{Q}_p^\times.$$

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

$$E_2^{p,q} = H^p(Z_s, R^q\Psi(\mathbb{Q}_\ell)) \Rightarrow H^{p+q}(Z_{\bar{\eta}}, \mathbb{Q}_\ell)), \tag{4.3.5}$$

where now $Z_{\bar{\eta}}$ 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 $\mathcal{X}$ is a single point $x \in Z_s$.

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

(4.3.6) Theorem Berkovich, [B, II, Theorem 3.1]**.** *There is a canonical isomorphism*

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

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

$$\bar{S}(x) = \varprojlim_{U,m} \bar{S}_{U,m}(x) \subset \varprojlim_{U,m} \mathcal{A}_{U(m)}(B,*).$$

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 $G(\mathbf{A}_f) \times W_K$-equivariant sheaves on $\bar{S}(x)$:

$$[\Psi^i_{n-h,h,+} \times (I_x(\mathbb{Q})\backslash G^{(h)}(\mathbf{A}_f))]/J_{n-h,h,+} \xrightarrow{\sim} R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)} . \tag{4.3.7}$$

(4.3.8) Remarks. (i) When $h = 0$, the set $\bar{S}(x)$ maps bijectively to $\varprojlim_U \bar{S}_{U,0}(x)$. This is because the group $\mathcal{G}_x$ is connected and because, over a reduced base, $\mathcal{G}_x[\varpi^m]$ has a unique Drinfel'd basis, namely the trivial one (exercise). Hence we may view $R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)}$ as a sheaf on the $h = 0$ stratum in $\varprojlim_U \mathcal{A}_{U(0)}(B,*)$, though the vanishing cycles themselves require $m \to \infty$.

(ii) For $h > 0$, this is no longer true. On the one hand, the set $\breve{M}_{n-h,h;M}$ maps to a product, as we will see, of $\breve{M}_{n-h,0}$ and $\breve{M}_{0,h}$. 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 Ψ^i_n 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:*

$$\Psi^i_{K,\ell,n,0,m} = (\Psi^i_n)^{\Gamma_m}$$

where $\Gamma_m \subset G$ is the principal congruence subgroup.

Proof. The admissibility is a consequence of

(1) Each $\Psi^i_{K,\ell,n,0,m}$ 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* $\bar{S}^{(0)}$ 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.

(4.3.9.1) 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 $\breve{M}_n$ and hence has no finite-dimensional invariant subspaces. The correct statement would be that, for any character ξ of the center Z_G of G, the maximal quotient of Ψ^i_n on which Z_G acts as ξ 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

$$\Psi^i_{c,g} = c - Ind_{A_{g,K}}^{GL(g,K)\times J_g \times W_K}(\Psi^i_{g,0,x_0})^\vee. \tag{4.3.10}$$

Thus

$$\Psi^i_{c,g} = \varinjlim_m H^0(\breve{M}_{g,0;m,red}, (R^i\Psi^{form}\mathbb{Q}_\ell)^\vee)$$

is just the cohomology of the dual of $R^i\Psi^{form}\mathbb{Q}_\ell$. 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]:

(4.3.11) Theorem [HT]. *Let g be a non-negative integer. Let $\pi \in \mathcal{A}_0(g,K)$, and let $JL(\pi)$ denote the corresponding representation of $J = D^\times_{\frac{1}{g}}$ under the Jacquet-Langlands correspondence (A.1.13).*

(i) *We have*

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

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 $\pi' \neq \pi \in \mathcal{A}(g,K)$ be a discrete series representation. Then for all i, $Hom_J(\Psi^i_{c,g}, JL(\pi'))$ contains no G-subquotients isomorphic to π.*

(iii) *Finally, $\sigma_\ell(\pi)$, defined by*

$$\sigma_\ell(\pi) = r_\ell(\pi \otimes |\bullet|^{\frac{g-1}{2}}),$$

satisfies all the conditions of the local Langlands correspondence (except possibly compatibility with ε factors).

This was conjectured by Carayol [Ca3], following earlier conjectures of Deligne. The proof of Theorem 4.3.11 is given in §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 $M \subset \varpi^{-m}\mathcal{O}^n/\mathcal{O}^n$ is a direct summand isomorphic to $\varpi^{-m}\mathcal{O}^{n-h}/\mathcal{O}^{n-h}$, and p is a Drinfel'd level structure of type M on $\mathcal{G}_x$ if $p\mid_M$ is a Drinfel'd structure on $\mathcal{G}_x^0$ and $p \pmod{M}$ induces a Drinfel'd level structure on the étale quotient. We have the decomposition

$$\breve{M}_{g,h;m} = \coprod_M \breve{M}_{g,h;M}$$

We write

$$\Psi^i_{K,\ell,n-h,h,m} = \oplus_M \Psi^i_{K,\ell,n-h,h,M}$$

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 $\mathcal{O}_m = \mathcal{O}/\varpi^m\mathcal{O}$. There is an isomorphism

$$Ind_{P_{h,0}(\mathcal{O}_m)}^{GL(n,\mathcal{O}_m)} \Psi^i_{K,\ell,n-h,h,M_0(m)} \xrightarrow{\sim} \Psi^i_{K,\ell,n-h,h,m} \tag{4.3.12}$$

that sends a function $f : GL(n,\mathcal{O}_m) \to \Psi^i_{K,\ell,n-h,h,M_0(m)}$ satisfying $f(pg) = pf(g)$ for $p \in P_{h,0}(\mathcal{O}_m)$ to

$$[GL(n,\mathcal{O}_m) : P_{h,0}(\mathcal{O}_m)]^{-1} \sum_{g \in P_{h,0}(\mathcal{O}_m)\backslash GL(n,\mathcal{O}_m)} g^{-1}f(g).$$

It is easy to check that this is an isomorphism of $GL(n,\mathcal{O}_m)$-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 $V_m = \Psi^i_{K,\ell,n-h,h,M_0(m)}$, $H_m = \Psi^i_{K,\ell,n-h,h,m}$, $I_m = Ind_{P_{h,0}(\mathcal{O}_m)}^{GL(n,\mathcal{O}_m)}$. The denominator (which doesn't work integrally!!) makes the following diagram commute:

$$\begin{array}{ccc} I_mV_m & \longrightarrow & H_m \\ \downarrow & & \downarrow \\ I_{m'}V_{m'} & \longrightarrow & H_{m'} \end{array}$$

for $m' > m$, where the right-hand side is just pullback and the left-hand side identifies I_mV_m with functions on $\mathcal{O}_{m'}$ that pullback from functions on $\mathcal{O}_m$ and take values in the image of V_m in $V_{m'}$ under the natural pullback. Thus in the limit this defines an isomorphism

$$\varinjlim_m I_mV_m \xrightarrow{\sim} \varinjlim H_m = \Psi^i_{n-h,h}. \tag{4.3.13}$$

Let

$$\Psi^i_{K,\ell,n-h,h,M_0} = \varinjlim_m \Psi^i_{K,\ell,n-h,h,M_0(m)}$$

(4.3.14) Proposition. *There is a canonical isomorphism of $G \times J \times W_K$-modules*

$$Ind_{P_{h,0}(K)}^{GL(n,K)} \Psi^i_{K,\ell,n-h,h,M_0} \xrightarrow{\sim} \Psi^i_{n-h,h}.$$

Proof. Given the above constructions, it remains to identify the left-hand side via

$$\varinjlim_m I_m V_m \xrightarrow{\sim} Ind_{P_{h,0}(K)}^{GL(n,K)} \varinjlim_m V_m.$$

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 $Ind_{P_{h,0}(\mathcal{O})}^{GL(n,\mathcal{O})} \varinjlim_m V_m$, and the fact that locally constant functions come from the left-hand side.

For convenience we have ignored the polarization datum (the $^+$); now we put it back. We consider an individual $\Psi^i_{K,\ell,n-h,h,M_0(m),+}$ as $P_{h,0}(\mathcal{O}_m)$-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 $\bar{S}^{(h)}$) of the global vanishing cycles on the Shimura variety. On the other hand it follows as before, from the continuity lemma, that

$$\Psi^i_{K,\ell,n-h,h,M_0(m),+} = (\Psi^i_{K,\ell,n-h,h,M_0})^{K_P(m)}.$$

Thus $\Psi^i_{K,\ell,n-h,h,M_0}$ is an *admissible* $P_{h,0}(K)$-module. But by a standard lemma this implies that

(4.3.15) Lemma. *The unipotent radical of $P_{h,0}(K)$ acts trivially on $\Psi^i_{K,\ell,n-h,h,M_0}$.*

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 $P = P_{h,0}(K)$, $N = R_uP$, 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

(4.3.15.1) $$\dots z^{-n}Uz^n \subset \dots \subset z^{-1}Uz \subset U \subset zUz^{-1} \subset \dots z^nUz^{-n} \dots$$

and such that

(4.3.15.2) $$\bigcup_{n \geq 0} z^n U_N z^{-n} = N.$$

For $n \in \mathbb{Z}$, let V_n denote the subspace of V fixed by $z^{-n}Uz^n$. Thus

(4.3.15.3) $$V_n \subset V_{n+1}$$

for all n, and $v \in V_0$. On the other hand, the action of z on V defines an isomorphism $V_n \xrightarrow{\sim} V_{n-1}$. 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.

(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 $\Psi^i_{K,\ell,n-h,h,M_0}$, 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 $\bar{S}^{(h)}_{U(m)}$ 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 $(\bar{S}^{(h)}_{U(m)})_{red}$; 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 $\varprojlim_m \mathcal{A}_{U(m)}(B,*)$, as in §3.5, it is not necessary to define a separate adelic group action on the inverse limit of the strata $\bar{S}^{(h)}_{U(m)}$. 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

$$\pi_1 = \pi_{1,M_0(m)} : \breve{M}^+_{n-h,h;M_0(m)} \to \breve{M}^+_{n-h;m,0} \times \breve{M}_{0,h;m}.$$

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-subschemes. We write $R^i\Psi_{K,\ell,n-h,h,M_0(m),+}$ (resp. $R^i\Psi_{K,\ell,n-h,0;m}$) for Berkovich's vanishing cycles sheaf $R^i\Psi^{form}\mathbb{Q}_\ell$ over $\breve{M}^+_{n-h,h;M_0(m),red}$ (resp. over $\breve{M}^+_{n-h;m}$). 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 $R^i\Psi_{K,\ell,n-h,h,M_0(m),+}$ is the pullback via π_1 of the formal vanishing cycles of $\breve{M}^+_{n-h;m} \times \breve{M}_{0,h;m}$. But $\breve{M}_{0,h;m}$ is étale (even discrete) and it follows from Proposition 3.3.5 that $\breve{M}^+_{n-h;m,0}$ is smooth over $\breve{M}^+_{n-h}$. But smooth morphisms preserve vanishing cycles. Let $(\mathbb{Q}_\ell)_{0,h;m}$ denote the constant sheaf $\mathbb{Q}_\ell$ over the discrete scheme $\breve{M}_{0,h;m}$. 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

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

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

Here the action of the unipotent radical is trivial on the right-hand side, and $\boxtimes$ is the external tensor product over the product $\breve{M}^+_{n-h;m} \times \breve{M}_{0,h;m}$.

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

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

of $P \times J \times W_K$-equivarant 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 $GL(h,K) \to GL(h,K) \times GL(h,K) \subset L \times J$.

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

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

Here $c - Ind$ denotes compact induction, $\Psi^i_{n-h,0,x_0}$ 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 $C_c^\infty(GL(h,K))$.

(4.3.18) Corollary. *The $G \times J \times W_K$ module $\Psi^i_{n-h,h,+}$ 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 $+$).*

(4.3.19) Remark 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) $\bar{S}^{(h)}$:

$$[Ind^{GL(n,K)}_{P_{h,0}(K)}(R^i\Psi_{n-h,0} \boxtimes (\mathbb{Q}_\ell)_{0,h})_+) \times (I_x(\mathbb{Q})\backslash G^{(h)}(\mathbf{A}_f))]/J_{n-h,h,+} \xrightarrow{\sim} R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)} . \tag{4.3.20}$$

(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 $\bar{S}^{(h)}_{U(m)}$ is the disjoint union:

$$\bar{S}^{(h)}_{U(m)} = \coprod_M \bar{S}_{U,M}$$

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

$$\bar{S}^{(h)}_{U(m)} = \coprod_{\gamma\in GL(n,\mathcal{O}_m)/P_{h,0}(\mathcal{O}_m)} \gamma(\bar{S}_{U,M_0(m)}). \tag{4.4.1}$$

(4.4.2) Lemma. *For each fixed U^w, the $P_{h,0}(K)$-module $H = \varinjlim_m H_c^p(\bar{S}_{U,M_0(m)}, R^q\Psi)$ is admissible.*

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

$$V_m = H^{\Gamma_m^h}.$$

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

$$R^q\Psi(m) \xrightarrow{\sim} f_{m',m,*}\Psi(m')^{\Gamma_m^h}.$$

This is a stalkwise calculation, hence we are reduced by uniformization to the corresponding continuity lemma for $\Psi^q_{n-h,h,+}$. 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*

$$H_c^p(\bar{S}^{(h)}, R^q\Psi) \xrightarrow{\sim} Ind_{P_{h,0}(K)}^{GL(n,K)} \varinjlim_{U^w,m} H_c^p(\bar{S}_{U,M_0(m)}, R^q\Psi).$$

(Note: this is non*-normalized induction.) In particular, $H_c^p(\bar{S}^{(h)}, R^q\Psi)$ is an admissible $G(\mathbf{A}_f)$-module.*

Proof. We first observe that the action of $P_{h,0}(K)$ on $\varprojlim_m \bar{S}^{(h)}_{U(m)}$ stabilizes $\varprojlim_m \bar{S}_{U,M_0(m)}$. Indeed, recall that the action of $GL(n,K)$ on $\breve{M}_{n-h,h}$ was defined by inducing from that of $P_{h,0}(K)$ (which was denoted $P_{\tilde{M}}$) on $\{\breve{M}_{g,h;p^{-m}\tilde{M}/\tilde{M}}\}$. The same argument works globally. On the other hand, for each level m, the stabilizer in $GL(n,\mathcal{O})$ of $\bar{S}_{U,M_0(m)}$ 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 $\varprojlim_m \bar{S}_{U,M_0(m)}$ 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 $\varprojlim_m \bar{S}_{U,M_0(m)}$. From this it follows formally that $H_c^p(\bar{S}^{(h)}, R^q\Psi)$ is the representation compactly induced from the action of $P_{h,0}(K)$ on $\varinjlim_m H_c^p(\bar{S}_{U,M_0(m)}, R^q\Psi)$. 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

$$H_c^p(\bar{S}^{(h)}_{M_0}, R^q\Psi) = \varinjlim_{U^w,m} H_c^p(\bar{S}_{U,M_0(m)}, R^q\Psi)$$

We thus obtain the following formula for the cohomology as $G(\mathbf{A}_f)\times W_K$-module:

(4.4.4) First Basic Identity. *The following identity holds in $Groth(G(\mathbf{A}_f) \times W_K)$:*

$$[H(\mathcal{A}(B,*))] = \sum_{p,q,h} (-1)^{p+q}[Ind_{P_h(K)}^{GL(n,K)} H_c^p(\bar{S}^{(h)}_{M_0}, R^q\Psi)].$$

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

(4.4.5)

$$[(R^i\Psi_{n-h,0} \boxtimes (\mathbb{Q}_\ell)_{0,h;m})_+) \times (I_x(\mathbb{Q})\backslash G^{(h)}(\mathbf{A}_f))]/J_{n-h,h,+} \xrightarrow{\sim} R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)_{M_0}} .$$

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

$$I_x(\mathbb{Q})\backslash(\mathbb{Z} \times GL(h,K) \times (\mathbb{Q}_p^\times/\mathbb{Z}_p^\times) \times G(\mathbf{A}_f^w)) \tag{4.4.6}$$

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 $R^i\Psi\mathbb{Q}_\ell\mid_{\bar{S}(x)_{M_0}}$ at a point, say x, in $\bar{S}(x)_{M_0}$. 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

$$R^i\Psi\mathbb{Q}_\ell\mid_x \xrightarrow{\sim} \Psi^i_{n-h,0,x_0} \tag{4.4.7}$$

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

$$J^0_{n-h} = \ker w \circ N : D^\times_{\frac{1}{n-h}} \to \mathbb{Z}.$$

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 $\tilde{x} \in \breve{M}_{n-h,n,+,red} \times I_x(\mathbb{Q})\backslash G^{(h)}(\mathbf{A}_f)$ above x for the uniformization (3.4.10). It follows easily that

(4.4.8) Lemma. *$R^i\Psi\mathbb{Q}_\ell$ is the sheaf on $\bar{S}(x)_{M_0}$ associated to the representation of the isotropy group J^0_{n-h} on $\Psi^i_{n-h,0,x_0}$.*

Lecture 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 (5.2.13.1)) 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 $[H(\mathcal{A}(B,*))]_0$ denote the formal sum of all $G(\mathbf{A}_f) \times W_K$ modules in $[H(\mathcal{A}(B,*))]$ 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:

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

where the subscript 0 on the right also means supercuspidal, in this case under the action of $GL(n,K)^0 = kerw \circ \det \subset A_{K,n}$. 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 $\bar{S}^{(0)}$ is of dimension zero, we just find

$$\sum_i(-1)^i[H^i(\mathcal{A}(B,*))]_0 = \sum_q(-1)^q[H^0(\bar{S}^{(0)},(R^q\Psi)_0)]. \tag{5.1.2}$$

Indeed, there is a stronger assertion. The spectral sequence for vanishing cycles, applied to the supercuspidal part, has the form

$$(E_2^{p,q})_0 = \varinjlim_{U^w,m} H^p(\bar{S}_{U(m)},R^q\Psi)_0.$$

But the same dévissage shows that

$$H^p(\bar{S}_{U(m)},R^q\Psi)_0 = H^p((\bar{S}^{(0)},(R^q\Psi)_0) = 0 \text{ unless } p=0. \tag{5.1.3}$$

Thus the spectral sequence degenerates at E_2 and we have

(5.1.4) $$H^i(\mathcal{A}(B,*),\mathbb{Q}_\ell)_0 \xrightarrow{\sim} H^0((\bar{S}^{(0)},(R^i\Psi)_0), i=1,\ldots,2n-2.$$

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) $$|ker^1(\mathbb{Q},G)|\cdot\bigoplus_\pi H^i(\mathfrak{g},Z_G(\mathbb{R})\cdot K_\infty;\pi_\infty)\otimes\pi_f.$$

Here π runs through automorphic representations of G that are supercuspidal at w. Recall the base change map from the first lecture. From π, one can find a pair (Π,ψ), with $\Pi\subset\mathcal{A}_0(GL(n)_F)$, which is a base change at all unramified places (for π) and all places that split in E. In particular, Π_w is supercuspidal, hence Π is cuspidal.

Clozel's purity lemma then implies that π_∞ 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 $0<i\le n-1$ such that

$$H^a(\mathfrak{g},Z_G(\mathbb{R})\cdot K_\infty;\pi_\infty)\neq 0 \Leftrightarrow a\in\{n-1-i,n-1-i+2,\ldots,n-1+i\}.$$

Thus $H^a(\mathcal{A}(B,*),\mathbb{Q}_\ell)$ contains π_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 $H^a(\mathcal{A}(B,*),\mathbb{Q}_\ell)$ at unramified places v have complex absolute values $q_v^{\frac{a}{2}}$; thus at unramified places v that split in E, say, the Satake parameters of Π_v have several distinct complex absolute values of the form $q_v^{\frac{a}{2}}$. But Π is cuspidal, hence every Π_v is generic by Shalika's theorem. Moreover, Π_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 $\le q_v^{\frac{1}{2}}$). This completes the argument.

Thus we have

(5.1.6) $$H^{n-1}(\mathcal{A}(B,*),\mathbb{Q}_\ell)_0 \xrightarrow{\sim} H^0((\bar{S}^{(0)},(R^{n-1}\Psi)_0);$$

(5.1.7) $$(R^i\Psi)_0=0,\ i\neq n-1.$$

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

$$\Theta : [I_x(\mathbb{Q})\backslash \breve{M}_{n,0,+}(\mathbb{F}) \times G^{(0)}(\mathbf{A}_f)]/J_{n,+} \xrightarrow{\sim} \bar{S}(x). \tag{5.2.1}$$

Here $J_{n,+} = J_n \times \mathbb{Q}_p^\times$ and $\breve{M}_{n,0,+}(\mathbb{F})$ 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, λ) be a polarized abelian variety over $\mathbb{F}_q$, with $F \subset End^0(A)$, such that the Rosati involution induces complex conjugation on F. Let $(N, \mathbf{F})$ 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_i^c!!] 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 $Frob_A$ over $\mathbb{F}_q$ belongs to F.*

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

Proof. For each i there is a $W(\mathbb{F})$-latice $M_i \subset N_i$, stable under $\mathbf{F}$ and $\mathbf{V}$, such that $\mathbf{F}^{s_i} M_i = p^{r_i} M_i$. We may assume that M_i is fixed by $\mathcal{O}_{F_{w_i}}$ (this is obvious in our case, since we are starting from an $\mathcal{O}$-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 $O_K \subset End(A)$. Without loss of generality we may assume all $s_i = s$ and $q = p^s$. So then $\mathbf{F}_q M_i = p^{r_i} M_i$. Let ord_i be the valuation on K_i with $ord_i(p) = 1$.

Consider the following problem in algebraic number theory: Find an element $u \in K$ that is a unit away from p and such that

$$ord_i(u) = r_i;\ uu^c = q.$$

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 $\mathbf{F}_q$ 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

$$ord_i u' = s + r_i - r_{i^c} = 2r_i; u'(u')^c = q^2.$$

So up to replacing q by q^2, we have solved the equation. Now $\varepsilon = u^{-1}Frob_A$ is an automorphism of A (because it fixes $\oplus_i M_i$, by the first equation) that fixes the polarization (by the second equation). Hence by Serre's lemma we conclude that some power of ε equals 1.

(5.2.3) Corollary. *Let (A, λ) and (A', λ') be two abelian varieties over $\mathbb{F}$ satisfying the same assumptions. Then*

$$Hom_K^0(A, A') \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \xrightarrow{\sim} Hom_K(V_\ell(A), V_\ell(A')).$$

Proof. This follows from the proposition and Tate's theorem

$$Hom^0(A, A') \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \xrightarrow{\sim} Hom_{\mathbb{Q}_\ell[Frob]}(V_\ell(A), V_\ell(A')).$$

Now return to $(A_x, \lambda_x, i_x) \in \bar{S}^{(0)}$. Recall the data $(B, *, V)$ of our original moduli problem. Let $C = End^0_B(V) = B^{op}$, $C_x = End^0_B(A_x)$. 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 λ_x. By the Corollary, we have that

$$C_x \otimes \mathbb{Q}_\ell \xrightarrow{\sim} End_B(V_\ell(A_x)).$$

Since $\ell \neq p$, there is a level structure, i.e. a B-invariant symplectic similitude $V_\ell(A_x) \xrightarrow{\sim} V \otimes \mathbb{Q}_\ell$, well defined $\pmod{U_\ell}$. Thus

$$(C_x \otimes \mathbb{Q}_\ell, \#_x) \xrightarrow{\sim} (C \otimes \mathbb{Q}_\ell, \#)$$

as $F \otimes \mathbb{Q}_\ell$-algebras with involution. Therefore there is an isomorphism

$$(C_x \otimes \overline{\mathbb{Q}}, \#_x) \xrightarrow{\sim} (C \otimes \overline{\mathbb{Q}}, \#)$$

which induces a $\overline{\mathbb{Q}}$ isomorphism between G and

$$I_x = \{\gamma \in C_x^\times \mid \gamma \cdot \gamma^{\#_x} \in \mathbb{Q}^\times\}. \tag{5.2.4}$$

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 $ab = n$.)

Note that $I_{x,p} \xrightarrow{\sim} \prod_i I_{x,w_i} \times \mathbb{Q}_p^\times$, 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 $i > 1$, $I_{x,w} \subset J_n$. It follows (by dimension considerations) that these inclusions are isomorphisms. The group $G^{(0)}(\mathbf{A}_f)$ is then just $I_x(\mathbf{A}_f)$, and (5.2.1) becomes

$$\Theta : [I_x(\mathbb{Q}) \backslash \breve{M}_{n,0,+}(\mathbb{F}) \times I_x(\mathbf{A}_f)]/J_{n,+} \xrightarrow{\sim} \bar{S}(x). \tag{5.2.5}$$

How many basic isogeny classes are there? By the above corollary, we see that, if $x, x' \in \bar{S}^{(0)}$, 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) Lemma. *Up to isomorphism, the group I_x is independent of the point $x \in \bar{S}^{(0)}$.*

Proof. The proof of Lemma (2.3.1) applies to the group I_x. (See also Lemma (6.6.8), below.)

Two polarizations λ and λ' from A to $A^\vee$ are equivalent (as B-morphisms inducing $*$) if and only if there exists $d \in C_x = C_{x'} = B^{op}$ and $a \in \mathbb{Q}^\times$ such that $\lambda' = ad^\vee \lambda d$ ($d^\vee$ 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 λ' and λ

implies that $\delta = \delta^*$ ($* =$ Rosati involution); since λ' is a polarization, δ must be totally positive. Then $\lambda \circ \delta = ad^\vee \lambda d$ if and only if

$$\delta = a(\lambda^{-1} d^\vee \lambda) d = a d^* d$$

has a solution (a, d). The set of solutions of this equation is a *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 $H^1(\mathbb{Q}, I_x)$. But there are solutions locally for all primes $\ell \neq p$, by the existence of the level structure; at ∞ because δ 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 Φ_b of basic isogeny classes is mapped by this construction to a subset of $ker^1(\mathbb{Q}, I_x)$. We will see in Lecture 6, using Honda-Tate theory, that this map is *surjective.* (We still haven't shown that $\bar{S}^{(0)}$ is non-empty!) Assume this for now.

(5.2.7) Fact. *The cardinality of* $\ker^1(\mathbb{Q}, G)$ *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

$$[\Psi^i_{n,+} \times (I_x(\mathbb{Q})\backslash I_x(\mathbf{A}_f))]/J_{n,+} \xrightarrow{\sim} R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)} . \tag{5.2.8}$$

It follows formally that

$$R^i\Psi\mathbb{Q}_\ell \mid_{\bar{S}(x)} \xrightarrow{\sim} Hom_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I_x/I_x(\mathbb{R}), \mathbb{Q}_\ell)). \tag{5.2.9}$$

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 $+$ to the definition (4.3.10)).

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

(5.2.10) Proposition. *Let $I = I_x$ for any $x \in \Phi_b$. Let (ρ, ψ) be a representation of $J_{n,+}$ (with $\rho \in \hat{J}_n$, ψ an unramified character of $\mathbb{Q}_p^\times$). Assume $JL(\rho)$ is a supercuspidal representation of G_n. Consider representations π^w of $G(\mathbf{A}^w_f) \xrightarrow{\sim} I(\mathbf{A}^w_f)$. Then there is an isomorphism*

$$H^{n-1}(S(B,*), \mathbb{Q}_\ell)[\pi^w \otimes JL(\rho) \otimes \psi] \xrightarrow{\sim} \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)[\pi^w \otimes \rho \otimes \psi]^n.$$

Moreover, for $i \neq n-1$,

$$H^i(S(B,*), \mathbb{Q}_\ell)[\pi^w \otimes JL(\rho) \otimes \psi] = 0.$$

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

Now it follows from the remarks preceding the proposition that

$$\begin{aligned} H^0(\bar{S}^{(0)}, R^i\Psi\mathbb{Q}_\ell) &\xrightarrow{\sim} \sum_{\Phi_b} Hom_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I_x/I_x(\mathbb{R}), \mathbb{Q}_\ell)) \\ &\xrightarrow{\sim} Hom_{J_{n,+}}(\Psi^i_{c,n,+}, \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell))^{|\ker^1(\mathbb{Q},G)|}. \end{aligned} \tag{5.2.11}$$

Here I is any I_x for $x \in \Phi_b$.

Recalling that $H^{n-1}(\mathcal{A}(B,*),\mathbb{Q}_\ell) = H^{n-1}(S(B,*),\mathbb{Q}_\ell)^{|\ker^1(\mathbb{Q},G)|}$, it follows from the First Basic Identity (4.4.4) and (5.1.6) that, up to semi-simplification,

$$H^{n-1}(S(B,*),\mathbb{Q}_\ell)_0 \xrightarrow{\sim} Hom_{J_{n,+}}((\Psi^{n-1}_{c,n,+})_0, \mathcal{A}(I/I(\mathbb{R}),\mathbb{Q}_\ell)). \tag{5.2.12}$$

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 $(\Psi^{n-1}_{c,n,+})_0$ 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

$$\bigoplus_{\pi \in \mathcal{A}G_0} H^{n-1}(\mathfrak{g}, Z_G(\mathbb{R})K_\infty; \pi_\infty) \otimes \pi_f.$$

Fix ρ as above and π_f with component $JL(\rho)$ (supercuspidal) at w. For given ρ, this is always possible (see (5.2.15), below). Let $R_\ell(\pi_f)$ denote the semisimplified representation of $Gal(\overline{\mathbb{Q}}/F)$ on $Hom_{G(\mathbf{A}_f)}(\pi_f, H^{n-1}(S(B,*),\mathbb{Q}_\ell)_0)$. 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 $\psi \circ N_{K/\mathbb{Q}_p}$ as in §(1.3) to remove the contribution of ψ. Combining the above identities, we find

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

$$\begin{aligned} JL(\rho)\otimes[r_\ell(\pi_f)\mid_{Gal(\bar{K}/K)}]^\vee &\xrightarrow{\sim} [Hom_{J_{n,+}}((\Psi^{n-1}_{c,n,+})_0, \rho\otimes\psi)\otimes r_\ell(\psi\circ N^{-1}_{K/\mathbb{Q}_p})] \\ &\xrightarrow{\sim} \sum_i (-1)^{n-1+i}[Hom_{J_{n,+}}((\Psi^i_{c,n,+}),\rho\otimes\psi)\otimes r_\ell(\psi\circ N^{-1}_{K/\mathbb{Q}_p})] \\ &\xrightarrow{\sim} \sum_i (-1)^{n-1+i}[Hom_{J_n}((\Psi^i_{c,n}),\rho)]. \end{aligned}$$

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

$$\begin{aligned} &[Hom_{J_{n,+}}((\Psi^{n-1}_{c,n,+})_0,\rho\otimes\psi)\otimes r_\ell(\psi\circ N^{-1}_{K/\mathbb{Q}_p})] \\ &\qquad \xrightarrow{\sim} \sum_i (-1)^{n-1+i}[Hom_{J_{n,+}}((\Psi^i_{c,n,+})_0,\rho\otimes\psi)\otimes r_\ell(\psi\circ N^{-1}_{K/\mathbb{Q}_p})], \end{aligned}$$

follows from the vanishing of $(\Psi^i_{c,n,+})_0$ 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 ψ, the virtual G_n-module

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

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

$$[M^h] = \sum_{p,q} (-1)^{p+q} [Ind_{P_h(K)}^{GL(n,K)} H_c^p(\bar{S}_{M_0}^{(h)}, R^q\Psi)].$$

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

In other words, only "automorphic" π^w can contribute to the virtual module $[M^h]$. However, neither this lemma nor the Second Basic Identity determines the individual spaces $[Ind_{P_h(K)}^{GL(n,K)} H_c^p(\bar{S}_{M_0}^{(h)}, R^q\Psi)]$. Note that the uniqueness of π_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 π_w is determined uniquely.

We admit Lemma 5.2.13.1 for the moment. For our given π^w, we thus have $\pi_w = JL(\rho)$. Now the First Basic Identity yields

(5.2.13.2)
$$[H(\mathcal{A}(B,*))][\pi^w] = \sum_h [M^h][\pi^w]$$

in $Groth(G_n)$. Since $\pi_w = JL(\rho)$ is supercuspidal, all the Jacquet modules $(\pi_w)_{R_u P_h}$ vanish for $h > 0$, thus (5.2.13.2) simplifies to yield

(5.2.13.3)
$$[H(\mathcal{A}(B,*))][\pi^w] = [M^0][\pi^w].$$

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

(5.2.13.4)
$$[M^0][\pi^w] = \sum_i (-1)^{n-1+i} [Hom_{J_{n,+}}((\Psi^i_{c,n,+}), \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)[\pi^w])^{|\ker^1(\mathbb{Q},G)|}]$$
$$= \sum_i (-1)^{n-1+i} [Hom_{J_{n,+}}((\Psi^i_{c,n,+}), \rho \otimes \psi) \otimes \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)[\pi^w \otimes \rho \otimes \psi])^{|\ker^1(\mathbb{Q},G)|}]$$
$$\xrightarrow{\sim} [\Psi_n(\rho)] \otimes \mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)[\pi^w \otimes \rho \otimes \psi]^{|\ker^1(\mathbb{Q},G)|},$$

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 $[r_\ell(\pi_f) \mid_{Gal(\bar{K}/K)}]$ is purely local at w; i.e., it depends only on $\pi_w = JL(\rho)$. It also calculates the supercuspidal part of Ψ_n^{n-1} (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 $r_\ell(\pi) \otimes |det|^{\frac{n-1}{2}}$ of W_K defined in this way. Not every π can be realized as a local component of a cohomological automorphic representation of G. Our hypotheses imply that the central character of π is of finite order. Conversely, assume the central character of π to be of finite order. Then

(5.2.14) 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 $GL(n, K)$ as the local component of the right kind of G;

(5.2.15) Given such G and $\pi \in \mathcal{A}_0(n, K)$, it follows from a theorem of Clozel [C1] that one can always find a cohomological representation Π of G with local component π at w, unramified outside some fixed (non-empty) set.

(5.2.16) To extend the correspondence to general π, one notes that any π is of the form $\pi_0 \otimes \psi \circ \det$, where π_0 has central character of finite order and ψ is some character of $K^\times$. So one defines $\sigma_\ell(\pi) = \sigma_\ell(\pi_0) \otimes \psi$ viewing ψ 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

$$\sigma_\ell(\pi \otimes \psi \circ \det) = \sigma_\ell(\pi) \otimes \psi \tag{5.2.17}$$

when ψ is a character of finite order. This follows by applying Kottwitz' theorem to the representation $r_\ell(\pi_f)$ of $Gal(\overline{\mathbb{Q}}/F)$. Indeed, Kottwitz shows that $r_\ell(\pi_f \otimes \chi) = r_\ell(\pi_f) \otimes \chi$ whenever χ 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 ψ, not necessarily of finite order. We have thus verified that the correspondence σ_ℓ 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:

(5.2.18) Proposition. *Let $\pi' \in \mathcal{A}(g, K)$ be a discrete series representation which is not supercuspidal. Then for all i, $Hom_J(\Psi^i_{c,g}, JL(\pi'))$ contains no G-subquotients isomorphic to a supercuspidal representation π.*

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 π^w, as in the statement of Lemma 5.2.13.1, such that $\pi^w \otimes \pi'$ occurs with non-zero multiplicity in $\mathcal{A}(I/I(\mathbb{R}), \mathbb{Q}_\ell)$. Suppose π does occur as a G-subquotient of $Hom_J(\Psi^i_{c,g}, JL(\pi'))$. It then follows from (5.2.12) that $\pi^w \otimes \pi$ occurs with non-zero multiplicity in $H^{n-1}(S(B,*), \mathbb{Q}_\ell)_0$. 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 Π, resp. Π', of $GL(n, F)$. By strong multiplicity one $\Pi = \Pi'$, hence $\pi = \pi'$, contradiction.

(5.2.19) 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 $\sigma_\ell : \mathcal{A}_0(n, K) \to \mathcal{G}(n, K)$ is compatible with character twists. One shows similarly it is compatible with contragredients. Moreover, because the construction is purely local, σ_ℓ commutes with automorphisms of K. These are three properties required of a local Langlands correspondence (cf. (0.6)).

We also need to know that σ_ℓ commutes with cyclic base change and local automorphic induction. Having established these properties, it follows by an argument due to Henniart [BHK][6] that σ_ℓ is a bijection $\mathcal{A}_0(n, K) \to \mathcal{G}_0(n, K)$ for all K, and that it preserves conductors. In other words, it satisfies all the requirements of the local Langlands correspondence except preservation of ε 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) $\pi \subset \mathcal{A}(G)$, with fixed π_w, let Π 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 $\mathcal{A}(G')$. 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 ∞. These are easy to verify [H1,§4]. Applying Kottwitz' theorem and Chebotarev density, we see that

$$\sigma_\ell(BC_{K'/K}(\pi_w)) = \sigma_\ell(\pi_w) \,|_{W_{K'}} \tag{5.3.1}$$

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 ∞. This can be remedied by twisting Π by an appropriate Hecke character χ of F'. If the infinity types of χ are chosen appropriately, and if $\chi \circ c = \chi^{-1}$, then $AI_{F'/F}(\Pi \otimes \chi) \in CU(n[F' : F], F)$. This has the inevitable effect of replacing the initial π_w by an unramified twist, which is not a problem. Again, the details can be found in [H1] (proof of Lemma 5).

[6] 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 π 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 π_{K_n} of π under successive cyclic base change is a principal series representation. The analogous property for $\mathcal{G}_0(n, K)$ is obvious.

(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 henceforward 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 $G(\mathbf{A}_f)$-modules. One could also work with a fixed central character.

If $\gamma \in G(\mathbb{Q})$ and $\phi \in C_c^\infty(G(\mathbf{A}_f))$, we define the orbital integral

$$O_\gamma(\phi) = \int_{G(\mathbf{A}_f)/Z_G(\gamma)(\mathbf{A}_f)} \phi(g\gamma g^{-1}) d\dot{g}. \tag{5.4.1}$$

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 $\phi \in C_c^\infty(G(\mathbf{A}_f))$. Then*

$$Tr(\phi \mid [H(\mathcal{A}(B,*))]) = n\kappa_B \sum_\gamma e(\gamma)[F(\gamma):F]^{-1} vol(Z_G(\gamma)(\mathbb{R})_0^1)^{-1} O_\gamma(\phi).$$

Note that we have written the formula for $\mathcal{A}(B,*)$, 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 $G(\mathbf{A}_f)$ is arbitrary (it appears on both sides).

(5.4.2.2) $e(\gamma) = (-1)^{n/[F(\gamma):F]-1}$ is the Kottwitz sign; γ is regular if and only if $[F(\gamma):F] = n$, and then $e(\gamma) = 1$.

(5.4.2.3) $\kappa_B = 1$ if $4 \mid [B:\mathbb{Q}]$ and equals 2 otherwise.

(5.4.2.4) γ runs over a set of representatives of $G(\mathbf{A})$-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 γ.
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) $Z_G(\gamma)(\mathbb{R})_0$ denotes the compact mod center inner form of $Z_G(\gamma)(\mathbb{R})$ and

$$Z_G(\gamma)(\mathbb{R})_0^1 = \ker|\nu| : Z_G(\gamma)(\mathbb{R})_0 \to \mathbb{R}_{>0}^\times.$$

(5.4.2.7) Let $dz(\gamma)_f$ be the measure used to define the orbital integral, $(dz(\gamma)_\infty)^1_0$ the measure used to define $vol(Z_G(\gamma)(\mathbb{R})^1_0)$, and $(dz(\gamma)_\infty)_0$ the measure on $Z_G(\gamma)(\mathbb{R})_0$ defined by $(dz(\gamma)_\infty)^1_0$ and dt/t on $\mathbb{R}^\times_{>0}$. Let $dz(\gamma)_\infty$ be the measure on $Z_G(\gamma)(\mathbb{R})$ 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

$$\sum_\gamma v(Z_G(\gamma)(\mathbb{Q})\backslash Z_G(\gamma)(\mathbf{A}))O_\gamma(\phi).$$

Here γ runs over $G(\mathbb{Q})$-conjugacy classes. For the volume term one can take Tamagawa measure. To get cohomology, one takes ϕ_∞ to be a sum of discrete series pseudocoefficients (over the set of discrete series with cohomology in the trivial representation); this restricts attention to γ elliptic at ∞, and the orbital integral of ϕ_∞ 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 $G(\mathbf{A})$-conjugacy classes in $G(\mathbb{Q})$ by counting the number of the former in the latter.[7] This number turns out to be related to $\kappa_B/ker^1(Z_G(\gamma))$, 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.

One gets a completely analogous formula when G is replaced by I:

(5.4.3)

$$|ker^1(\mathbb{Q},G)|Tr(\phi^I \mid \mathcal{A}(I/I(\mathbb{R}))) = \kappa_B \sum_\gamma [F(\gamma):F]^{-1} vol(Z_G(\gamma)(\mathbb{R})^1_0)^{-1}O_\gamma(\phi^I).$$

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 ∞). There is no restriction on γ (all elements are elliptic in $I(\mathbb{R})$).

(5.4.4) Lemma. *The set of $I(\mathbf{A})$-conjugacy classes in $I(\mathbb{Q})$ is in bijection with the set of $G(\mathbf{A})$-conjugacy classes in $G(\mathbb{Q})$ elliptic at ∞ and at w. This bijection preserves orbital integrals away from w, and takes $\gamma^I_w \in J_n$ to the conjugacy class in G_n with the same characteristic polynomial.*

Proof. Let $\gamma \in I(\mathbb{Q})$. It is elliptic at ∞ and at w, hence transfers to a conjugacy class in $G(\mathbf{A})$, and the question is whether it has a representative in $G(\mathbb{Q})$. This is a consequence of a general principle: if G and G' are inner forms and $T \subset G$ is a torus that transfers locally to G' everywhere, then T transfers globally to G

[7] Full stabilization goes one step further, replacing $G(\mathbf{A})$-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).

provided a certain cohomological invariant, defined by Langlands and Kottwitz (cf. [K1,§9]), vanishes (in H^2). But this invariant vanishes if T is elliptic (cf. Lemma (2.3.3)).

To make effective use of pseudocoefficients, we fix a compact open subgroup $U^w \subset G(\mathbf{A}_f^w)$ 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 $\Gamma_U = U^w \cap Z_G(\mathbb{Q}) = U^w \cap Z_I(\mathbb{Q})$ is a cocompact subgroup of $Z_G(K) = Z_{G_w}$. Hence we can take ϕ_w to be a pseudocoefficient of a chosen supercuspidal π_w, relative to the set $\mathcal{A}_{d,fin}$ (A.1.3) of representations with central character trivial on Γ_U . This has the effect on the trace side of isolating representations π_f with component π_w at w. As we have seen in (5.1.6), these occur only in H^{n-1}, hence for such ϕ, we have

$$(5.4.5) \quad Tr(\phi \mid [H^{n-1}(\mathcal{A}(B,*))]) = \\ (-1)^{n-1} n \kappa_B \sum_{\gamma} (-1)^{q(\gamma)} [F(\gamma):F]^{-1} vol(Z_G(\gamma)(\mathbb{R})_0^1)^{-1} O_\gamma(\phi).$$

We take $\phi^I = \phi^w \otimes \phi_w^I$, where ϕ_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) Fact. $O_\gamma(\phi_w^I) = (-1)^{n-1} e(\gamma) O_{\gamma^G}(\phi_w)$.

When γ 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 $\mathcal{A}(I/I(\mathbb{R})\cdot U^w)_{JL(\pi_w)}$. 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 $\bar{S}(x)_{M_0}$ (cf. Lemma 4.4.8). We replace n by g and work with the version $\Psi_{c,g}^{g-1}$. We write $G = G_g = GL(g,K)$, $J = J_g$ as before. Recall from (4.3.10) that

$$\Psi_{c,g}^{g-1} = c - Ind_{A_{g,K}}^{G\times J\times W_K} \Psi_{c,g,x_0}^{g-1}.$$

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 $\delta: G\times J\times W_K \to \mathbb{Z}$ with the map $\mathbb{Z} \to \mathbb{Z}/g\mathbb{Z}$, 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

$$(5.5.1) \qquad T'_0 = A'_{g,K} \cap (G\times J);\ T_0 = A_{g,K} \cap (G\times J)$$

Recall from §3.3 that the subgroup

$$Z_0 = \{(x,x) \in K^\times \times K^\times \simeq Z \times Z_J\} \subset T_0$$

acts trivially on the moduli space, hence on the stalks Ψ^i_{c,g,x_0} of the vanishing cycle sheaves for any i. The only representations τ of T_0 such that $Hom_{T_0}(\Psi^i_{c,g,x_0}, \tau) \neq 0$ are thus those on which Z_0 acts trivially. Let $G_{g,0}$, (resp. $J_{g,0}$), denote the kernel of $\delta_G : G_g \to \mathbb{Z}$ (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 τ is an irreducible representation of T_0, trivial on Z_0, and ψ runs through the set of characters of $T_0/T_{00} \cdot Z_0 \simeq \mathbb{Z}/n\mathbb{Z}$. The set of inertial equivalence classes of T_0 is denoted $\mathcal{I}_g$. If τ is a representation of T_0/Z_0, we let $[\tau]$ denote its inertial equivalence class. Then we have a discrete decomposition for each i

(5.5.2)
$$\Psi^i_{c,g,x_0} = \bigoplus_{[\tau]\in\mathcal{I}_g} \Psi^i_{c,g,x_0}[\tau]$$

where $\Psi^i_{c,g,x_0}[\tau]$ is the sum of the τ_j-isotypic components for $\tau_j \in [\tau]$.

We also define inertial equivalence for representations of J and G. Let $\rho \in \mathcal{A}(J)$, with central character ψ_ρ. The *inertial equivalence class* of ρ, denoted $[\rho]$, is the set of representations $\rho \otimes \psi \circ \det$, where ψ runs over unramified characters. The *strong inertial equivalence class* of ρ is the set of $\rho \otimes \psi \circ \det$ where ψ runs over unramified characters of finite order dividing g; this is the set of representations of J_g inertially equivalent to ρ and with central character ψ_ρ. The same terminology is used for discrete series representations of G_g. The cardinality of the strong inertial equivalence class of ρ is an important invariant of ρ: it equals $\frac{g}{c(\rho)}$, where $c(\rho)$ is the number of distinct unramified characters ψ 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 ρ. 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. $[\mathcal{A}]_d(g, K)$).

We write $\rho \sim_i \rho'$ if ρ and ρ' 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 ρ 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 ρ will be called *nearly equivalent*, and we say they belong to the *near equivalence class* $N(\rho)$ of ρ.

(5.5.3) Lemma. *Let τ 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*

$$\tau\ |_{T_{00}} = \oplus(\alpha_i)^\vee \otimes \beta_i$$

where each α_i (resp. β_i) is an irreducible representation of $G_{g,0}$ (resp. $J_{g,0}$). Then

(i) The various α_i (resp. β_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 β_i belong to the near equivalence class of $JL(\pi)$. Then the α_i belong to the near equivalence class of π.

Proof. Part (i) is obvious, and part (ii) follows from (4.3.4) and Theorem 4.3.11.

Let $[\tau]$ and π 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 β_i-isotypic components of Ψ^{g-1}_{c,g,x_0}, for $\beta_i \in N(JL(\pi))$, or as the sum of its α_i-isotypic components, for $\alpha_i \in N(\pi^\vee)$. This justifies writing

$$\Psi^{g-1}_{c,g,x_0}[\tau] = \Psi^{g-1}_{c,g,x_0}[JL(\pi)] = \Psi^{g-1}_{c,g,x_0}[\pi] \tag{5.5.4}$$

(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 $\Psi^{i}_{c,g,x_0}[\tau]$ vanishes for $i \neq g-1$. We thus put

$$[\Psi]_{c,g,x_0} = \sum_i (-1)^k [\Psi^i_{c,g,x_0}] = \bigoplus_{[\rho]\in[\mathcal{A}](J_g)} [\Psi]_{c,g,x_0}[\rho], \tag{5.5.5}$$

where ρ runs alternatively over $[\mathcal{A}](J_g)$, as indicated, or over $[\mathcal{A}]_d(g,K)$. 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 ξ of $K^\times$, with restriction ξ_0 to $\mathcal{O}^\times$. The maximal compact subgroup $\mathcal{O}^\times \times \mathcal{O}^\times \subset Z \times Z_J$ is contained in T_{00}. Let $\Psi^{g-1}_{c,g,x_0}(\xi_0) \subset \Psi^{g-1}_{c,g,x_0}$ denote the subspace on which $(u,u') \in \mathcal{O}^\times \times \mathcal{O}^\times$ 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 $\Psi^{g-1}_{c,g,x_0}(\xi_0)$ extends uniquely to a representation, denoted $\Psi^{g-1}_{c,g,x_0,\xi}$, of $A'_{g,K}$, such that $(x,x') \in Z \times Z_J$ acts as $\xi(x)^{-1}\xi(x')$. Let $\Psi^{g-1}_{c,g,\xi}$ denote the maximal quotient of $\Psi^{g-1}_{c,g}$, on which $(x,x') \in Z \times Z_J$ acts as $\xi(x)^{-1}\xi(x')$. Then there is a canonical isomorphism

$$\Psi^{g-1}_{c,g,\xi} \xrightarrow{\sim} c - Ind_{A'_{g,K}}^{GL(g,K)\times J_g \times W_K} \Psi^{g-1}_{c,g,x_0,\xi}. \tag{5.5.6}$$

Combined with (5.5.2) and (5.5.4), we thus obtain a canonical decomposition

$$\Psi^{g-1}_{c,g,\xi,scusp} \xrightarrow{\sim} \bigoplus_\rho \Psi^{g-1}_{c,g,\xi}[\rho], \tag{5.5.7}$$

$$\Psi^{g-1}_{c,g,\xi}[\rho] \xrightarrow{\sim} c - Ind_{A'_{g,K}}^{G\times J \times W_K} \Psi^{g-1}_{c,g,x_0,\xi}[\rho]. \tag{5.5.8}$$

where the subscript "scusp" designates the G-supercuspidal part, and ρ runs over $\mathcal{A}_0(g,K)$. By Proposition 5.2.18, the sum can also be taken over $\rho \in \mathcal{A}(J)$ with $JL(\rho)$ supercuspidal. The component $\Psi^{g-1}_{c,g,\xi}[\rho]$ is non-trivial if and only if the central character ξ_ρ of ρ equals ξ. There is a similar decomposition for the alternating sum $[\Psi_{c,g,\xi}]$.

(5.5.9) Lemma. *Write $G = GL(g,K)$, $J = J_g$. Let $T'_0 = A'_{g,K} \cap (G \times J)$. Let $\tau_{\xi,0}$ (resp. τ_ξ) denote the restriction to T'_0 of the representation of $A'_{g,K}$ on $\Psi^{g-1}_{c,g,x_0,\xi}$ (resp. of the representation of $G \times J$ on $\Psi^{g-1}_{c,g,\xi}$). Then*

(i) The representations τ_ξ and $\tau_{\xi,0}$ are admissible.

(ii) For any $a \in G \times J$, the representation $\tau^a_{\xi,0}$ of T'_0, defined by $\tau^a_{\xi,0}(x) = \tau_{\xi,0}(axa^{-1})$, has the same character as $\tau_{\xi,0}$.

(iii) In the Grothendieck group of T_0', we have

$$\tau_\xi = g \cdot \tau_{\xi,0}$$

Moreover, the character of $\Psi^{g-1}_{c,g,\xi}$, 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

$$trace(\tau_\xi)(\phi) = g \cdot trace_{Z,\xi^{-1}}(\tau_{\xi,0})(\phi_\xi). \tag{5.5.9.1}$$

Here ϕ is a compactly supported function on T_0, transforming with respect to $\xi_0 \otimes \xi_0^{-1}$ under $\mathcal{O}^\times \times \mathcal{O}^\times \subset Z \times Z_J$, ϕ_ξ the extension of ϕ to a function on T_0' transforming under $\xi \otimes \xi_0^{-1}$ under $Z \times \mathcal{O}^\times \subset Z \times Z_J$. 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 ϕ_ξ 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., φ is an extension of Frobenius to the algebraic closure of K. There is a homomorphism $h_\varphi : G \times J \to A_{g,K}$ given by

$$h_\varphi(\gamma, j) = (\gamma, j, \varphi^d) \text{ where } d = w_K(\det(\gamma)) - w_K(N(j)).$$

The restriction of h_φ 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 $G \times J$.

(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

$$\Psi^{g-1}_{c,g,x_0,\xi}[\rho] \mid_{T_{00}} = \frac{g}{c(\rho)} \bigoplus_{i,j} \alpha_i \otimes \beta_j$$

where α_i (resp. β_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 $\frac{g}{c(\rho)}$, 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 $\Psi^{g-1}_{c,g,x_0,\xi}[\rho]$ 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:

$$[\Psi_{c,g}] \xrightarrow{\sim} \bigoplus_{\rho} [\Psi_{c,g}][\rho], \tag{5.5.11}$$

where

$$[\Psi_{c,g}][\rho] = c - Ind_{A_{g,K}}^{G \times J \times W_K} [\Psi_{c,g,x_0}][\rho],$$

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 ℓ-adic) representations of I_K to be the set of irreducible components of the restriction to I_K of a finite-dimensional continuous irreducible ℓ-adic representation of W_K. The decomposition (5.5.2) represents the $G_{g,0}$-supercuspidal part $(\Psi^i_{c,g,x_0})_0 \subset (\Psi^i_{c,g,x_0})$ 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 $\pi \in \mathcal{A}_0(g,K)$ (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 $r_\ell(\pi)^\vee$. Note that σ 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 π, 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 π supercuspidal, $\Psi^{g-1}_{c,g,x_0}[\pi]$ is the sum of the isotypic subspaces of Ψ^{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 π 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

$$\Psi^{g-1}_{c,g,x_0}[\pi] = \Psi^{g-1}_{c,g,x_0}[JL(\pi)] = \sum_i (-1)^{g-1+i} \Psi^i_{c,g,x_0}[JL(\pi)]$$

for supercuspidal π.

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 $r_\ell(\pi)^\vee$, 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 $\mathcal{H}(\Xi_0 // U \times I(\pi))$.

Lecture 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 $\bar{S}^{(h)}_{M_0}$. 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 also let J_{n-h} be the twisted inner form $D^\times_{\frac{1}{n-h}}$ 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} : $Groth(G_n) \to Groth(L_h)$ denote the standard (normalized) Jacquet functor

$$r_{G_n,L_h}\pi = \pi_{N_h} \otimes \delta_h^{-\frac{1}{2}}.$$

The Jacquet functor for the opposite parabolic is denoted $r^{op}_{G_n,L_h}$. We define a renormalized Jacquet functor

$$re - r^{op}_{G_n,L_h} = r^{op}_{G_n,L_h} \otimes \delta^{1/2}_{P_h};$$

since $\delta^{1/2}_{P_h} = \delta^{-1/2}_{P_h^{op}}$, this means it has been normalized twice.

In what follows, we fix h and let $\rho \in \mathcal{A}(J_{n-h})$. First assume $JL(\rho)$ is supercuspidal. We define a map

$$red^{(h)}_\rho : Groth(G_n) \to Groth(G_h)$$

as the composition of

$$re - r^{op}_{G_n,L_h} : Groth(G_n) \to Groth(G_{n-h} \times G_h)$$

and the map c_ρ that sends $[\alpha \otimes \beta]$, with $\alpha \in \mathcal{A}(n-h,K)$ and $\beta \in \mathcal{A}(h,K)$, to 0 if $\alpha \neq JL(\rho)$, and to $[\beta]$ otherwise. In other words, $red^{(h)}_\rho$ is the renormalized Jacquet functor followed by projection on the $JL(\rho)$-component in the first variable.

For general $\rho \in \mathcal{A}(J_{n-h})$, we replace c_ρ in the preceding definition by the map c'_ρ that sends $[\alpha \otimes \beta]$ to

$$Tr(\alpha)(\phi_{JL(\rho),\omega}) \cdot [\beta].$$

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 ω.

In what follows, we fix a level subgroup U^w away from w, and write $H_c^p(\bar{S}_{M_0}^{(h)}, R^q\Psi)$ for $H_c^p(\bar{S}_{M_0}^{(h)}, R^q\Psi)^{U^w}$, $[H_c(\bar{S}_{M_0}^{(h)}, R\Psi)] = \sum_{p,q}(-1)^{p+q}[H_c^p(\bar{S}_{M_0}^{(h)}, R^q\Psi)]$. 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 $\mathcal{A}(J_{n-h})_{fin} \subset \mathcal{A}(J_{n-h})$ such that*

$$(6.1.2.1) \qquad n \cdot [H_c(\bar{S}_{M_0}^{(h)}, R\Psi)] = \bigoplus_{\rho \in \mathcal{A}(J_{n-h})_{fin}} red_\rho^{(h)}[H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho)]$$

in $Groth(G(\mathbf{A}_f^w)) \times L_h \times W_K)$.

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 $[H(\mathcal{A}(B,*))]$ is ignored.*

(ii) The set $\mathcal{A}(J_{n-h})_{fin}$ can be chosen so that, for any $\rho \in \mathcal{A}(J_{n-h})_{fin}$, the intersection

$$\mathcal{A}(J_{n-h})_{fin}[\rho] = [\rho] \cap \mathcal{A}(J_{n-h})_{fin}$$

of $\mathcal{A}(J_{n-h})_{fin}$ with the inertial equivalence class $[\rho]$ of ρ, defined as in §5.5, is a finite set.

Write $\mathcal{A}(J_{n-h})_{fin} = \mathcal{A}_{n-h}^0 \coprod \mathcal{A}'_{n-h}$ where $\mathcal{A}_{n-h}^0$ is the subset of ρ such that $JL(\rho)$ is supercuspidal, and $\mathcal{A}'_{n-h}$ are the others. Write

$$(6.1.2.2) \qquad \begin{aligned} \mathcal{R}^{h,0} &= \bigoplus_{\rho \in \mathcal{A}_{n-h}^0} red_\rho^{(h)}[H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho)]; \\ \mathcal{R}^{h,\prime} &= \bigoplus_{\rho \in \mathcal{A}'_{n-h}} red_\rho^{(h)}[H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho)] \end{aligned}$$

(iii) **(Second Basic Identity, second version):** *For any $\rho \in \mathcal{A}(J_{n-h})_{fin}$, with $JL(\rho)$ supercuspidal, we have the following identity in $Groth(G(\mathbf{A}_f^w) \times L_h \times W_K)$:*

$$(6.1.2.3) \qquad \mathcal{R}^{h,0} = \bigoplus_{\rho \in \mathcal{A}_{n-h}^0} red_\rho^{(h)}[H(\mathcal{A}(B,*))] \otimes [JL(\rho) \otimes r_\ell(\rho)^{\vee,+}].$$

Here $r_\ell(\rho)^{\vee,+}$ is $r_\ell(\rho)^\vee$ twisted by the contribution $\psi \circ N_{K/\mathbb{Q}_p}$ of $\mathbb{Q}_p^\times$, 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 $H_c^p(\bar{S}_{M_0}^{(h)}, R\Psi)$ 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 $Groth(G(\mathbf{A}_f^w)) \times L_h)$, i.e. ignoring the Galois action. In §6.3 it will be shown that (6.1.2.1) in $Groth(G(\mathbf{A}_f^w)) \times L_h)$, 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)

$$n[H(\mathcal{A}(B,*))] = \sum_h \Big[\bigoplus_{\rho \in \mathcal{A}^0_{n-h}} Ind^{G_n}_{P_h(K)}\big(red^{(h)}_\rho H(\mathcal{A}(B,*)) \otimes JL(\rho) \otimes r_\ell(\rho)^{\vee,+}\big)\Big]$$

$$\bigoplus \sum_h \bigoplus_{\rho \in \mathcal{A}'_{n-h}} Ind^{G_n}_{P_h(K)}\big(red^{(h)}_\rho [H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho)]\big).$$

Write

$$[H(\mathcal{A}(B,*))] = \sum_{\pi_f} [\pi_f] \otimes [R_\ell(\pi_f)]$$

where $[R_\ell(\pi_f)] \in Groth(Gal(\bar{F}/F))$. 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 $G(\mathbf{A}_f^w)$ 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 $G(\mathbf{A}_f)$. Then in $Groth(GL(n,K) \times W_K)$ we have*

$$n[\pi_w] \otimes [R_\ell(\pi_f) \mid_{W_K}] =$$

$$(\dim R_\ell(\pi_f)) \sum_{h, \rho \in \mathcal{A}^0_{n-h}} n - Ind^{GL(n,K)}_{P_h(K)} n - red^{(h)}_\rho [\pi_w] \otimes JL(\rho) \otimes r_\ell(\rho)^{\vee,+}(-\frac{h}{2})$$

$$\bigoplus \sum_{h, \rho \in \mathcal{A}'_{n-h}} Ind^{G_n}_{P_h(K)}\big[red^{(h)}_\rho Hom_{G(\mathbf{A}_f^w)}(\pi^w, [H(\mathcal{A}(B,*))]) \otimes [\Psi_{n-h}(\rho)]\big].$$

Here we define $n - red^{(h)}_\rho$ by replacing $re - r^{op}$ by the normalized Jaquet functor r^{op}, which is just a twist by $\delta_{P_h}^{-\frac{1}{2}}$. Similarly, $n - Ind$ is normalized induction. The twist of $re - r^{op}$ by $\delta_{P_h}^{-\frac{1}{2}}$ cancels the opposite twist in n-Ind but introduces a new twist in the second step of the definition of red_ρ, which accounts for the twist by the unramified character $|\bullet|^{-\frac{h}{2}}$, 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

$$red^{(h)}_\rho [H(\mathcal{A}(B,*))] = \sum_{\pi_f} \pi_f^w \otimes red^{(h)}_\rho [\pi_w] \otimes [R_\ell(\pi_f)],$$

where here the term $[R_\ell(\pi_f)]$ is just a vector space without structure: all the Galois action is on the $r_\ell(\rho)^{\vee,+}(-\frac{h}{2})$! 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 ρ-contribution to $n\cdot[H_c(\bar{S}^{(h)}_{M_0}, R\Psi)]$ for any ρ, 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 $|\ker^1(\mathbb{Q}, G)|$ copies (for the different Shimura varieties in $\mathcal{A}(B,*)$ 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 $(\dim R_\ell(\pi_f)) = n|\ker^1(\mathbb{Q}, G)|$ and the formula in Theorem 6.1.5 becomes
(6.2.1)

$$[\pi_w] \otimes [R_0(\pi_f) \mid_{W_K}] = \sum_{h,\rho} n - Ind_{P_h(K)}^{G_n} n - red_\rho^{(h)}[\pi_w] \otimes JL(\rho) \otimes r_\ell(\rho)^{\vee,+}(-\frac{h}{2}).$$

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 $n - Ind_{P_h(K)}^{G_n} n - red_\rho^{(h)}[\pi_w]$, as h varies.

We know π_w is generic and unitary. Thus there is a parabolic subgroup $P = P_\nu$, with $\nu = (n_1, \dots, n_r)$, and an r-tuple of discrete series representations $\tau_1, \dots, \tau_r$ such that

(split case) $$\pi_w = n - Ind_P^{G_n} \tau_1 \otimes \cdot \otimes \tau_r.$$

As explained in (1.4.5), we restrict our attention to the case where each τ_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 $n - Ind$ and $n - red$ designate normalized induction and restriction, respectively, whereas $re - r^{op}$ denotes RE-normalized restriction!!!

We first recall the following theorem due to Bernstein and Zelevinski:

(6.2.2) Geometric Lemma (Bernstein-Zelevinski).

$$[r^{op}_{G_n,L_h}(n - Ind_P^{G_n}(\tau_1 \otimes \cdots \otimes \tau_r))] = \sum_{\nu=\nu_I \coprod \nu_{II}, \sum_{i\in\nu_{II}} n_i = h} [n - Ind_{P_I}^{G_{n-h}} \otimes_{i\in\nu_I} \tau_i \otimes n - Ind_{P_{II}}^{G_h} \otimes_{j\in\nu_{II}} \tau_j].$$

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

$$Ind_{P_h(K)}^{G_n}\left[red_\rho^{(h)}\pi_w \otimes [\Psi_{n-h}(\rho)]\right].$$

where $JL(\rho)$ is not supercuspidal. By (A.1.5.iii), the definition of $red_\rho^{(h)}$, 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

$$[n-Ind_{P_h(K)}^{G_n} n-red_\rho^{(h)} n-Ind_P^{G_n}\tau_1\otimes\cdots\otimes\tau_r].$$

First, apply the (normalized) Jacquet functor relative to N^{op} to $n-Ind_P^{G_n}\tau_1\otimes\cdots\otimes\tau_r$. 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 π_w is irreducible (and unitary) implies, by the Bernstein-Zelevinski classification of the discrete series [BZ,Z], that the $JL(\rho)^\vee$-isotypic component of the term corresponding to $\nu=\nu_I\coprod\nu_{II}$ is trivial unless

(i) $JL(\rho)$ is supercuspidal, and
(ii) ν_I is a single element i.

In other words, projection on $JL(\rho)$ picks out those $n_i=n-h$ and those $\tau_i=JL(\rho)^\vee$. Thus, letting $\nu^i=(n_1,\ldots,\widehat{n_i},\ldots,n_r)$. we have

$$(6.2.3)\quad [n-red_\rho^{(h)} n-Ind_P^{G_n}\tau_1\otimes\cdot\otimes\tau_r]=\sum_{n_i=n-h,\tau_i=JL(\rho)}[n-Ind_{P_{\nu^i}}^{G_h}\otimes_{j\in\nu^i}\tau_j]$$

Now comparing this with our original formula, and using transitivity of induction (first from $G_{n-h}\times P_{\nu^i}$ to L_h, then from P_h to G_n) we have
(6.2.4)

$$[\pi_w]\otimes[R_0(\pi_f)\mid_{W_K}]=\sum_{h,\rho}\sum_{n_i=n-h,\tau_i=JL(\rho)}[n-Ind_P^{G_n}JL(\rho)\otimes\bigotimes_{j\in\nu^i}\tau_j]\otimes r_\ell(\rho)^{\vee,\prime}(-\frac{h}{2}).$$

But each term on the right hand side of (6.2.4) is of the form $[\pi_w]\otimes r_\ell(\rho)^{\vee,\prime}(-\frac{h}{2})$ where $[\pi_w]$ is fixed, ρ runs through the τ_i and each τ_i occurs once (for $n-h=n_i$). Thus we can cancel the $[\pi_w]$ from both sides and obtain

$$(6.2.5)\qquad [R_0(\pi_f)\mid_{W_K}]=\sum_i r_\ell(\tau_i)^{\vee,\prime}(\frac{n_i-n}{2}).$$

If we define

$$(6.2.6)\qquad r_\ell(\pi_w)=\oplus_i\, r_\ell(\tau_i)^\vee(\frac{n_i-n}{2}),$$

then we conclude

$$(6.2.7)\qquad [R_0(\pi_f)\mid_{W_K}]=[r_\ell(\pi_w)]\otimes(\psi\circ N_{K/\mathbb{Q}_p}^{-1})$$

where for a change I put back the contribution of $\mathbb{Q}_p^\times$. 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 $[H_c(\bar{S}_{M_0}^{(h)},R\Psi)]$, 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

$$[\breve{M}^+_{n-h,h} \times (I_x(\mathbb{Q})\backslash G(\mathbf{A})^{(h)})]/J_{n-h,h}.$$

For general Shimura varieties, the term $\breve{M}^+_{n-h,h}$ 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 $|\ker^1(\mathbb{Q}, I_x)|$. 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 ℓ-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 $I_x(\mathbf{A})$-conjugacy, resp. $G(\mathbf{A})$-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 $G(\mathbf{A})$-conjugacy.

Kottwitz' Ann Arbor article is predicated upon taking the calculation one step further, classifying the contributions up to stable conjugacy (which is $G(\overline{\mathbb{Q}})$-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 $\pi^\vee$. Then*

(i) For all i $R^i\Psi[\pi]$ is a pro-constructible sheaf on $\bar{S}^{(h)}_{M_0}$, indeed is isomorphic to $\pi \otimes R^i\Psi_\pi$ where $R^i\Psi_\pi$ 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 $\rho = JL(\pi) \in \mathcal{A}(J)$.

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 $\mathcal{H}(G^1//U)$, 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 to a decomposition

$$R^i\Psi = \oplus_{[\rho]\in[\mathcal{A}](J_{n-h})} R^i\Psi[\rho]$$

of lisse sheaves on $\bar{S}^{(h)}_{M_0}$. 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 $z \in \bar{S}^{(h)}_{M_0}$,

$$R^i\Psi[\pi]_z \subset R^i\Psi_z[JL(\pi)] \tag{6.3.2}$$

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 ρ, 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, $[H_c(\bar{S}^{(h)}_{M_0}, R\Psi)][\pi]$ is isotypic for $r_\ell(\rho)^{\vee,+}$.*

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 $\mathcal{A}(B,*)$ 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

$$[H_c(\bar{S}^{(h)}_{M_0}, R\Psi)][\pi] = [H_c(\bar{S}^{(h)}_{M_0}, R\Psi[\pi])] = [H_c(\bar{S}^{(h)}_{M_0}, R^{n-h-1}\Psi[\pi])].$$

It then follows from (6.3.1)(iii) and (5.5.4) (which depends on Theorem 4.3.11 (i)) that $[H_c(\bar{S}^{(h)}_{M_0}, R^{n-h-1}\Psi)][\pi]$ is at least isotypic for the inertial equivalence class of $r_\ell(\rho)^{\vee,+}$.

Since $R^{n-h-1}\Psi[\pi]$ is constructible for any open , $\bar{S}^{(h)}_{M_0}$ 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 dévissage – we are working in the Grothendieck group – we may replace $\bar{S}^{(h)}_{M_0}$ 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 inertial equivalence class of π, 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 Ξ_0-equivariantly isomorphic to $(\Psi^{n-h-1}_{c,n-h,x_0})_0[\rho]$. Let $I(\rho) \subset I_K$ denote a subgroup of finite index acting trivially on $(\Psi^{n-h-1}_{c,n-h,x_0})_0[\rho]$. 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 $(\Psi^{n-h-1}_{c,n-h,x_0})_0[\rho]$. 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 $[H_c(\bar{S}^{(h)}_{M_0}, R^{n-h-1}\Psi)][\pi]$, is again $\mathcal{H}(\Xi_0//U \times I(\rho))$-isotypic for subquotients of $(\Psi^{n-h-1}_{c,n-h,x_0})_0[\rho]$.

Since this action is semisimple (cf. (5.5.12)), we can replace the word "subquotients" by "quotients." Then by Frobenius reciprocity, applied to $c - Ind_{\Xi_0}^{G\times W_K}(\bullet)$, the action of the Hecke algebra $\mathcal{H}(G//U \times W_K/I(\pi))$ on $[H_c(\bar{S}^{(h)}_{M_0}, R^{n-h-1}\Psi)[\pi]]$ 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 (post 4.3.11)) $$n \cdot [H_c(\bar{S}^{(h)}_{M_0}, R\Psi)]_\rho = (n-h) \cdot red^{(h)}_\rho H(\mathcal{A}(B,*))$$

in $Groth(G(\mathbf{A}^w_f) \times G_h)$, for all $\rho \in \mathcal{A}(J_{n-h})_{fin}$. Here the $n-h$ on the right-hand side comes from forgetting the $n-h$-dimensional representation $r_\ell(\rho)^{\vee,+}$ of W_K. Corresponding to the version (6.1.2.1), we just have

(6.3.4 (pre 4.3.11))

$$n \cdot [H_c(\bar{S}^{(h)}_{M_0}, R\Psi)] = \bigoplus_{\rho \in \mathcal{A}(J_{n-h})_{fin}} red^{(h)}_\rho [H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho)] \in Groth(G(\mathbf{A}^w_f)) \times G_h);$$

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 $\mathbb{F}$. 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 $\mathbb{F}$ is determined by its Frobenius endomorphism $\pi_A : A \to A$, where A is defined over some $\mathbb{F}_q$ and π_A is the q-th power of $Frob : A \to A^{(p)}$. Since A is also defined over any extension of $\mathbb{F}_q$, π_A is only well-defined up to powers; i.e., in the group $\overline{\mathbb{Q}}^\times/\mu_\infty$, where μ_∞ denotes roots of 1. But we also know that π_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 $q^{\frac{1}{2}}$. So π_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 π_A is obtained (by reducing abelian varieties with CM). This justifies the following definition:

(6.4.1) Definition. *Let M be a CM field, and let $\mathbb{Q}[\mathfrak{P}_M]$ be the $\mathbb{Q}$-vector space with basis the places of M above p. For any fractional ideal $I \subset M$, we let $[I] = \sum_{v|p} v(I) \cdot v \in \mathbb{Q}[\mathfrak{P}_M]$. A p-adic type for M is an $\eta \in \mathbb{Q}[\mathfrak{P}_M]$ such that $\eta + c_*(\eta) = [p]$. Two pairs (M, η) and (M', η') are isomorphic if there is an isomorphism of fields $M \xrightarrow{\sim} M'$ taking η to η'.*

A finite extension of CM fields $i : M \to N$ induces maps in both directions

$$i_* : \mathbb{Q}[\mathfrak{P}_M] \to \mathbb{Q}[\mathfrak{P}_N];\ i^*\mathbb{Q}[\mathfrak{P}_N] \to \mathbb{Q}[\mathfrak{P}_M].$$

via $i_*(v) = \sum_{v'|v} e_{v'/v} v'$; $i^*(v') = f_{v'/v} v$ if $v = v' \mid_M$. Let $\sim$ denote the equivalence relation on pairs (M, η) generated by $(M, \eta) \sim (N, i_*\eta)$. A p-adic type is an equivalence class of (M, η).

(6.4.2) Exercise. *Every p-adic type has a unique minimal representative, up to isomorphism.*

If $q = p^r$ and π is a q-number, let $\mathfrak{b}(\pi)$ be the p-adic type equivalent to $(\mathbb{Q}(\pi), \frac{1}{r}[\pi])$. Because π is determined (mod roots of unity) by $[\pi]$, it is easy to see that any sufficiently divisible power of π 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 $\mathbb{F}_q$ is sufficiently big). The preceding discussion shows that

(6.4.3) Theorem **Honda-Tate, [Ta].** *There is an equivalence between p-adic types and isogeny classes of simple abelian varieties over $\mathbb{F}$.*

Moreover, we can determine the invariants of $A_\mathfrak{b}$ as follows. Let $\mathfrak{b}$ be a p-adic type with minimal representative (M, η). Then:

(6.4.3.1) $End^0(A_\mathfrak{b})$ is the division algebra with center M and invariants $\frac{1}{2}$ at real primes, 0 at finite primes away from p, and $\eta_v f_{v/p}$ for v dividing p;

(6.4.3.2) $\dim A_\mathfrak{b} = \frac{1}{2}[M : \mathbb{Q}][End^0(A_\mathfrak{b}) : M]^{\frac{1}{2}}$

(6.4.3.3) For any $v \mid p$, $A_\mathfrak{b}[v^\infty]$ is a p-divisible group of height $[M_v : \mathbb{Q}_p][End^0(A_\mathfrak{b}) : M]^{\frac{1}{2}}$ and its Dieudonné module is isoclinic with slope $\eta_v/e_{v/p}$.

The fact that the minimal $M = \mathbb{Q}[\pi_A]$ is the center of $End^0(A_{\mathfrak{b}})$ – note that these are endomorphisms over $\mathbb{F}$ – follows from Tate's theorem that π_A generates the center of $End^0(A)$ [Ta] and our choice of π_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, η) 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 $i : B \hookrightarrow End^0(A)$. 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 $\mathfrak{b}$ be a p-adic type over F with minimal representative (M, η), and let $(A_{\mathfrak{b}}, i_{\mathfrak{b}})$ denote the corresponding simple object in the category of abelian varieties up to isogeny with B-action. Then:

(6.4.4.1) $End^0(A_{\mathfrak{b}})$ is the division algebra with center M and invariants $\frac{1}{2} - inv_v(B \otimes_F M)$ if v is real, $-inv_v(B \otimes_F M)$ at finite primes away from p, and $\eta_v f_{v/p} - inv_v(B \otimes_F M)$ for v dividing p;

(6.4.4.2) $\dim A_{\mathfrak{b}} = \frac{1}{2} n \cdot [M : \mathbb{Q}][End^0_B(A_{\mathfrak{b}}) : M]^{\frac{1}{2}}$

(6.4.4.3) For any $v \mid p$, $A_{\mathfrak{b}}[v^\infty]$ is a p-divisible group of height $[M_v : \mathbb{Q}_p][B : F]^{\frac{1}{2}}[End^0(A_{\mathfrak{b}}) : M]^{\frac{1}{2}}$ and its Dieudonné module is isoclinic with slope $\eta_v / e_{v/p}$.

Henceforward we fix an $h \in 0, \ldots, n-1$. The goal is to classify isogeny classes $[x] \subset \bar{S}^{(h)}(\mathbb{F})$. 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 $(A, i) = (A_{\mathfrak{b}}, i_{\mathfrak{b}})$ with minimal representative (M, η) as above. Let (A', i') be a simple factor, $C' = End^0_B(A')$; 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^c (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

$$inv_{v'}(C') = -inv_{v'}(B_M) = -[M_{v'} : F_v] inv_v B.$$

Since C' is a division algebra, $[C' : M]^{\frac{1}{2}}$ is at least the denominator of $-[M_{v'} : F_v] inv_v B$, and since B_v is a division algebra, the denominator of $inv_v B$ equals n. So

$$[C' : M] \geq (n/[M_{v'} : F_v])^2 \geq n^2/[M : F]^2;$$

$$\dim A' = \frac{1}{2} n \cdot [M : \mathbb{Q}][C' : M]^{\frac{1}{2}} \geq [F^+ : \mathbb{Q}] n^2 = \dim A,$$

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 $C = End^0_B(A)$, then $n = [M : F][C : M]^{\frac{1}{2}}$.*

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 η 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 η_{v^c}. 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

$$[M_{\tilde{w}}:\mathbb{Q}_p][B:F]^{\frac{1}{2}}[C:M]^{\frac{1}{2}} = n[K:\mathbb{Q}_p][M_{\tilde{w}}:K][C:M]^{\frac{1}{2}};$$

i.e.

(6.4.6) $$n-h = [M_{\tilde{w}}:K][C:M]^{\frac{1}{2}}.$$

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, $A[v^\infty]$ is again étale. Next

(6.4.8) Lemma. *M embeds over F in B (or in B^{op}).*

Proof. We consider the invariants of C at places v of M. For finite v not dividing w or w^c, $A[v^\infty]$ is either étale or multiplicative, hence we have $inv_v(C) = -inv_v(B_M)$. 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 $\mathbb{R}$.

Finally, we obtain the following result:

(6.4.9) Lemma. *There is a bijection betwen 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,η) translates into minimality of $\tilde{w}$, since η 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,η) over F with

(6.4.10) $$\eta_{\tilde{w}} = e_{\tilde{w}/w}/((n-h)f_{w/p});\ \eta_v = 0 \text{ if } v \mid u, v \neq \tilde{w}.$$

This determines η uniquely, and one checks that the corresponding (A,i) is of type h.

(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 polarisation $\lambda_0 : A \to A^\vee$ whose Rosati involution stabilizes $B \otimes M$ and induces $* \otimes c$, and a finitely generated $B \otimes M$ module W_0 with $* \times c$-hermitian alternating pairing $<,>_0: W_0 \otimes W_0 \to \mathbb{Q}$, such that there are*

1. An isomorphism of $B \otimes M \otimes_F \mathbf{A}_f^w$-modules

$$W_0 \otimes \mathbf{A}_f^w \xrightarrow{\sim} V^w(A)$$

taking $<,>_0$ to an $\mathbf{A}_f^w$-multiple of the Weil pairing induced by λ_0, and

2. An isomorphism $W_{0,\mathbb{R}} \xrightarrow{\sim} V_{\mathbb{R}}$ of $B_{\mathbb{R}}$-modules taking $<,>_0$ to an $\mathbb{R}^\times$ multiple of the standard pairing $(,)$ on $V_{\mathbb{R}}$.

Recall that this means in particular that the signatures of $<,>_0$ are $(1, n-1)$ at τ_0 and so on.

The existence of such an embedding (a #-embedding of M in B^{op}) 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 $\mathcal{A}(B, *)$. 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 $B^{op} = End_B(W_0)$ induced by the pairing $<,>_0$, $G_0 = GU(W_0, \#_0)$. 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 $\mathcal{A}(B, *)$ 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

$$V \otimes \mathbf{A}_f^w \xrightarrow{\sim} V^w(A)$$

as $B \otimes_F \mathbf{A}_f^w$-modules, compatible with the polarizations as before. In particular, the $*$-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 $S^i(B, *)$, 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 $D = End_{B \otimes M}(V)$, so $D = Cent_{B^{op}}(M)$, and let $G_{[x]} \subset D$ be the unitary similitude group of $(D, \#)$. Thus $G_{[x]} \subset G$. Let $*_{[x]}$ be the Rosati involution on $C = Aut(A_x, i_x)$, and let $I_{[x]} = Aut((A_x, i_x, \lambda_x)) = GU(C, *_{[x]})$. Then $I_{[x]}$ and $G_{[x]}$ are inner forms of each other; indeed they are locally isomorphic everywhere except p and ∞, because $W_{\mathbf{A}_f^p}$ and $V^p(A_x)$ are isomorphic as $* \otimes c$-hermitian $B_M(\mathbf{A}_f^p)$-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 $* \otimes c$-hermitian $\mathbb{Q}$-alternating form such that

(6.5.2) $\qquad (V', (,)')$ is equivalent to $(V, (,))$

as $*$-hermitian B-modules? On the one hand, the equivalence classes of pairs $(V', (,)')$ without condition corresponds to $\phi_x \in H^1(\mathbb{Q}, G_{[x]})$; the condition (6.5.2) means ϕ_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 $H^1(\mathbb{Q}, G_{[x]})(0)$. 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 $G(\overline{\mathbb{Q}})$-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 $G(\mathbf{A})$-conjugacy.

(6.6.1) Lemma. *The bijection above induces a bijection between*

(1) $G(\mathbf{A}_f^p)$*-conjugacy classes of* $\#$*-embeddings* $j : M \to B^{op}$*;*
(2) *The kernel* $H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}_f^p))(0)$ *of the map*

$$H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}_f^p)) \to H^1(\mathbb{Q}, G(\bar{\mathbf{A}}_f^p)).$$

Proof. It is clear from the preceding that elements of (1) correspond one-to-one to the images $x \in H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}_f^p))$ of elements $y \in H^1(\mathbb{Q}, G_{[x]})(0)$. So we must show that the restriction to kernels of the localization map

$$H^1(\mathbb{Q}, G_{[x]})(0) \to H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}_f^p))(0)$$

is surjective.

Now it follows from [K2, Prop. 2.6] that there is a commutative diagram with exact rows:
(6.6.2)

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \ker^1(\mathbb{Q}, G_{[x]}) & \longrightarrow & H^1(\mathbb{Q}, G_{[x]}) & \xrightarrow{f} & H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}})) & \longrightarrow & A(G_{[x]}) \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \ker^1(\mathbb{Q}, G) & \longrightarrow & H^1(\mathbb{Q}, G) & \xrightarrow{g} & H^1(\mathbb{Q}, G(\bar{\mathbf{A}})) & \longrightarrow & A(G) \end{array}$$

The group $A(G)$ is what Labesse, in [L], calls $H^1_{ab}(\mathbf{A}/\mathbb{Q}, G)$, 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 $\ker^1(\mathbb{Q}, Z_G) \to \ker^1(\mathbb{Q}, G)$ 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 $\pi_0(Z(\hat{G})^{Gal(\overline{\mathbb{Q}}/\mathbb{Q})})$.)

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 λ 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 $\mu \in \mathbb{Q}^\times_{>0}$ such that $\lambda' = \mu\delta^\vee\lambda\delta$.*

(2) *Equivalence classes of non-zero $\#_{[x]}$-fixed totally positive elements $\gamma \in C$ (thus $\gamma = \delta^{\#_{[x]}}\delta$ over $C_\mathbb{R}$), where equivalence is given by the equation $\gamma' = \mu\delta^{\#_{[x]}}\gamma\delta$. (We have already seen this equivalence in the supersingular case, via $\gamma \to \lambda_0 \circ \gamma$.)*

(3) *Same as (2), but where γ 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 γ 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 (γ, μ) to $(\mu(\gamma)^{\#_{[x]},-1}, \mu)$.

Thus there is a bijection between

(1) Equivalence classes of polarizations λ 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) $\ker[H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{R}, I_{[x]})] \cap H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}}^p_f))(0)$ where $H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}}^p_f))(0) = H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}^p_f))(0)$ via the isomorphism $I_{[x],\mathbf{A}^p_f} \xrightarrow{\sim} G_{[x],\bar{\mathbf{A}}^p_f}$.

Say λ and λ' are *nearly equivalent* if they are equivalent over $\mathbf{A}^p_f$.

(6.6.4) Lemma. *There are bijections between the following sets*

(1) *Near equivalence classes of polarizations λ 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) $H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}}^p_f))(0)$

(3) *$G(\mathbf{A}^p_f)$-conjugacy classes of $\#$-embeddings $j : M \to B^{op}$*

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

$$H^1(\mathbb{Q}, I_{[x]}) \to H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}})) = H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}}^p_f)) \oplus H^1(\mathbb{R}, I_{[x]}) \to A(I_{[x]});$$

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 $(M, \tilde{w}, [j])$ where $(M, \tilde{w})$ is a Hodge type of type h and $[j]$ is a $G(\mathbf{A}^p_f)$-conjugacy class of $\#$-embeddings $j : M \to B^{op}$.*

The above lemmas show that there is a surjective map from the set of isogeny classes $[x] = [(A_x, i_x, \lambda_x)]$ 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* $\ker^1(\mathbb{Q}, I_{[x]})$ *isogeny classes.*

Proof. The fiber is the set of equivalence 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 $\mathbf{A}_f^p$ lies in

$$H^1(\mathbb{Q}, I_{[x]}(\mathbf{A}_f^p))(0) = \ker[H^1(\mathbb{Q}, G_{[x]}(\bar{\mathbf{A}}_f^p)) \to H^1(\mathbb{Q}, G(\bar{\mathbf{A}}_f^p))].$$

Two equivalence classes are nearly equivalent if they map to the same element of $H^1(\mathbb{Q}, I_{[x]}(\bar{\mathbf{A}}_f^p))(0)$. But since they are already map to zero $inH^1(\mathbb{R}, I_{[x]})]$, 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 $\ker^1(\mathbb{Q}, I_{[x]})$. On the other hand, $\ker^1(\mathbb{Q}, I_{[x]})$ is a finite group (it can be identified with the image of $\ker^1(\mathbb{Q}, Z_{I_{[x]}})$, as before, for instance) that acts faithfully on $H^1(\mathbb{Q}, I_{[x]})$, so the cardinality is as indicated.

Recall the fixed complex embedding τ_0 of F. If $z = (M, \tilde{w}, [j]) \in PHT^{(h)}$ and $j \in [j]$, there is a unique distinguished $\tilde{\tau}_0$ of τ_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_σ, b_σ) for every real embedding σ of M^+. For only one such $\tilde{\sigma}_0$ is this signature indefinite, it restricts to our chosen σ_0 on F^+, and we let $\tilde{\tau}_0 = \tilde{\tau}_0(j)$ be its extension to M lifting τ_0. (Think of breaking up V_{τ_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 $\tilde{\tau}_0(j) = \tilde{\tau}_0(j')$, 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 $G(\mathbf{A}_f^p)$-conjugacy class means that $V_j(\mathbf{A}_f^p) \xrightarrow{\sim} V_{j'}(\mathbf{A}_f^p)$ (with their hermitian forms); and at p there is no possible difference. In the end, we have isomorphisms $G_{j,\mathbf{A}} \xrightarrow{\sim} G_{j',\mathbf{A}}$ where G_j is the unitary similitude group of V_j, and this isomorphism is canonical up to $G_{j',\mathbf{A}}$-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* $j \to \tilde{\tau}_0(j)$ *is surjective.*

(2) If $\tilde{\tau}_0(j) = \tilde{\tau}_0(j')$ *then the isomorphism* $G_{j,\mathbf{A}} \xrightarrow{\sim} G_{j',\mathbf{A}}$ *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 $\ker[H^1(\mathbb{R}, G_j) \to H^1(\mathbb{R}, G)]$ which correspond precisely to the extensions of τ_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 $\mathbf{A}$ 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* $I_{[x']}(\mathbf{A}_f^p)\xrightarrow{\sim}G_j(\mathbf{A}_f^p)\xrightarrow{\sim}I_{[x]}(\mathbf{A}_f^p)$.

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 $U^{w,h} \subset G(\mathbf{A}_f^w) \times G_h \times \mathbb{Q}_p^\times$ (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 $\bar{S}^{(h)}$:

(6.6.9)

$$\bar{S}^{(h)}_{M_0,U^{w,h}}(\mathbb{F}) = \coprod_{z \in PHT^{(h)}} ([\breve{M}_{n-h,+} \times (I_z(\mathbb{Q})\backslash G^{(h)}/U^{w,h})]/J_{n-h,+})^{|\ker^1(\mathbb{Q},I_z)|}.$$

This decomposition is compatible with the action of Frobenius (on the pro-discrete set $\breve{M}_{n-h,+}$, and it factors through a finite Galois group), and of the Hecke algebra of $G(\mathbf{A}_f^w) \times L_{n-h,h} \times \mathbb{Q}_p^\times$. Here $L_{n-h,h} = G_{n-h} \times G_h$ acts as follows: G_{n-h} acts on $\breve{M}_{n-h,+}$ 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.

Lecture 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 $G(\mathbf{A}_f^w) \times L_{n-h,h,+}$ on the cohomology $[H_c(\bar{S}^{(h)}_{M_0}, R\Psi)]$. Recall the description from the last time:

$$\bar{S}^{(h)}_{M_0}(\mathbb{F}) = \coprod_{z \in PHT^{(h)}} ([\breve{M}_{n-h,+} \times (I_z(\mathbb{Q})\backslash G^{(h)})]/J_{n-h,+})^{|\ker^1(\mathbb{Q},I_z)|}.$$

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.

(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^{op}. We return to this later. An element is *very h-regular* if the difference in p-adic valuations is $>> N$ for some large integer N determined by the problem.

(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]

(7.1.3) Lemma. *Let $\gamma \in I_z(\mathbb{Q})$. The number of $I_z(\mathbb{Q})$-conjugacy classes in the $I_z(\mathbf{A})$-conjugacy class of γ equals $|\ker^1(\mathbb{Q}, Z_{I_z}(\gamma))|/|\ker^1(\mathbb{Q}, I_z)|$.*

Proof. If γ and $\gamma' \in I_z(\mathbb{Q})$ are conjugate over $\mathbf{A}$, their centralizers in I_z are inner forms of one another that become isomorphic over $\mathbf{A}$. In this way one sees that the number is the cardinality of

$$\ker[\ker^1(\mathbb{Q}, Z_{I_z}(\gamma)) \to \ker^1(\mathbb{Q}, I_z)].$$

The Lemma follows from the surjectivity of this map, which follows from the fact (already used in §6.6) that

$$\ker^1(\mathbb{Q}, Z_{I_z}) \to \ker^1(\mathbb{Q}, I_z)$$

is an isomorphism, and likewise for $Z_{I_z}(\gamma)$.

Recall that the group I_z depends only on z (up to isomorphism and $G^{(h)}(\mathbf{A})$-conjugacy), whereas the inner forms $G_{[x]} \subset G$ depend also on the choice of an extension $\tilde{\tau}_0$ of τ_0 to M (up to isomorphism and $G(\mathbf{A})$-conjugacy). We let $G_{z,\tilde{\tau}_0}$ denote this $\mathbb{Q}$-group.

(7.1.4) Now we discuss transfer from I_z to $G_{z,\tilde{\tau}_0}$ to G. Note that we can discuss h-regular elements in $G_{z,\tilde{\tau}_0}(\mathbb{Q})$, since it comes with an embedding in $L_{h,n-h}$ at w. Consider the following three sets:

(7.1.4.1) The set $\mathcal{I}^{(h)}$ of pairs $(z, [a])$ where $z \in PHT^{(h)}$ and $[a]$ is an h-regular $I_z(\mathbf{A})$-conjugacy class in $I_z(\mathbb{Q})$.

(7.1.4.2) The set $\mathcal{G}^{(h)}$ of triples $(z, \tilde{\tau}_0, [\gamma])$ where $z = (M, \tilde{w}, [j]) \in PHT^{(h)}$, $\tilde{\tau}_0$ is as above, and $[\gamma]$ is an h-regular $G_{z,\tilde{\tau}_0}(\mathbf{A})$-conjugacy class in $G_{z,\tilde{\tau}_0}(\mathbb{Q})$ 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 $(\gamma, \tilde{w})$ where $\gamma \in G(\mathbb{Q})$ is an h-regular $\mathbb{R}$-elliptic element and where $\tilde{w}$ is a place of $F(\gamma)$ above w such that

$$(n-h)[F(\gamma) : F] = n[F(\gamma)_{\tilde{w}} : F_w]. \tag{7.1.4.4}$$

The pairs $(\gamma, \tilde{w})$ and $(\gamma', \tilde{w}')$ are equivalent if γ and γ' are conjugate by an element of $G(\mathbf{A})$ inducing an isomorphism $F(\gamma)_w \xrightarrow{\sim} F(\gamma')_w$ identifying $\tilde{w}$ with $\tilde{w}'$.

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 (7.4.1.3) can embed in $G^{(h)}(\mathbf{A})$, as follows: The embedding of $G(\mathbb{Q})$ in $G(\mathbf{A}_f^p)$ 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 $F(\gamma)_{\tilde{w}}$ embeds in D_{n-h}, and this follows from the equality of degrees (7.1.4.4).

Recall that the group $G_{z,\tilde{\tau}_0}$ comes with an embedding in G.

(7.1.5) Lemma. *The map $\mathcal{G}^{(h)} \to FP^{(h)}$, sending $(z, \tilde{\tau}_0, [\gamma])$ with $z = (M, \tilde{w}, [j])$ to $(\gamma, \tilde{w}')$, where $\tilde{w}'$ is the unique place of $M(\gamma) = F(\gamma)$ above the place $\tilde{w}$ of M, is a bijection.*

Proof. First note that γ being h-regular, $M(\gamma) = F(\gamma)$. The G_{n-h}-ellipticity implies that $F(\gamma) \otimes_M M_{\tilde{w}}$ is a field, hence that $\tilde{w}'$ exists. Let $(\gamma, \tilde{w}) \in FP^{(h)}$, and let $M \subset F(\gamma)$ be the minimal subfield containing F for which $\tilde{w}$ is inert from M to $F(\gamma)$. (The existence of such an M is left as an exercise). Then $(M, \tilde{w})$ 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 γ already in G, this module has an $* \otimes c$-hermitian $\mathbb{R}$-alternating pairing. The invariants (a_τ, b_τ) of this pairing, for $\tau : M^+ \to \mathbb{R}$, pick out a unique complex place $\tilde{\tau}_0$ except in the excluded case $n = 2$. This defines an element $((M, \tilde{w}, [j]), \tilde{\tau}_0, [\gamma]) \in \mathcal{G}^{(h)}$ above $(\gamma, \tilde{w})$, and it is clearly unique.

The other comparison is deeper. There is a map $\phi : \mathcal{G}^{(h)} \to \mathcal{I}^{(h)}$ sending $(z, \tilde{\tau}_0, [\gamma])$ to $(z, [a])$ where $a \in I_z(\mathbb{Q})$ is conjugate to γ in $I_z(\mathbf{A}_f^w) \xrightarrow{\sim} G_{z,\tilde{\tau}_0}(\mathbf{A}_f^w)$. 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 $I_z(\mathbf{A})$-conjugacy.

(7.1.6) Lemma. *The map ϕ is well-defined; i.e., the $\mathbf{A}_f^w$-conjugacy class of γ has a representative in $I_z(\mathbb{Q})$. Moreover, the map ϕ 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 $a_{\mathbf{A}} \subset I_z(\mathbf{A})$ to $(z, \tilde{\tau}_0, [\gamma])$. Away from ∞ and w there is nothing to say. Since γ 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 γ in L_{n-h} transfers to a well-defined conjugacy class in $D_{n-h}^\times \times G_h$. But this follows from the hypothesis that the image of γ in G_{n-h} is elliptic.

We view $a_{\mathbf{A}}$ as an $I_z(\mathbf{A})$ conjugacy class that contains a representative in $I_z(\overline{\mathbb{Q}}) = G_{z,\tilde{\tau}_0}(\overline{\mathbb{Q}})$, namely γ. 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 $obs([\gamma], [a_{\mathbf{A}}])$ where the first term is an $H(\overline{\mathbb{Q}})$-conjugacy class and the second an $H(\mathbf{A})$-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 $\mathfrak{K}(I^0/\mathbb{Q})$, 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 $G_{z,\tilde{\tau}_0}$) of γ. (By a theorem of Kottwitz, γ always transfers to the quasi-split inner form.) But this is precisely the group that vanishes for every possible I_z^0, 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 $\mathcal{G}^{(h)}$ to the set of $(z, \tilde{\tau}_0, [\gamma])$ 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 $(z, [a]) \in \mathcal{I}^{(h)}$. Now we have to determine the cardinality of its inverse image under ϕ. 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 $G_{z,\tilde{\tau}_0}(\mathbf{A})$-conjugacy classes, not $G_{z,\tilde{\tau}_0}(\mathbb{Q})$-conjugacy classes! But the $I_z(\mathbf{A})$-conjugacy class of $[a]$ determines the $G_{z,\tilde{\tau}_0}(\mathbf{A}_f)$-conjugacy class uniquely. Indeed, the groups only differ at w, but there the transfer from $D_{n-h}^\times$ to G_{n-h} is injective. So the only ambiguity is at ∞, 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

$$\ker[H^1(\mathbb{R}, Z_{G_{z,\tilde{\tau}_0}}(a)) \to H^1(\mathbb{R}, G_{z,\tilde{\tau}_0}(a))],$$

(kernel as map of pointed sets), and this set also parametrizes equivalence classes of $* \otimes c$-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 $U_h^w = U^w \times U_h$ for some compact open subgroup $U_h \subset G_h = L_{0,h}$. We introduce a class of acceptable functions

$$\phi \in C_c^\infty(G(\mathbf{A}_f^w) \times L_{n-h,h}//U_h^w)$$

where the symbol // designates bi-invariance. These functions act as correspondences on $\bar{S}_{U_h^w}^{(h)}$ and on the complex $R\Psi$, hence define operators on $[H_c(\bar{S}_{M_0}^{(h)}, R\Psi)]^{(U_h^w)}$. We assume ϕ factors as $\phi^w \otimes \phi_w$, with $\phi_w = \phi_{w,n-h} \otimes \phi_{w,h}$. Say ϕ is h-regular (resp. very h-regular) if ϕ_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 ϕ on $[H_c(\bar{S}_{M_0}^{(h)}, R\Psi)]^{(U_h^w)}$ for all ϕ. 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 $\phi \in C_c^\infty(G(\mathbf{A}_f^w) \times L_{n-h,h}//U_h^w)$, and suppose ϕ is very h-regular (depending on U^w and U_h). Then*

$$\begin{aligned} &Tr(\phi \mid [H_c(\bar{S}_{M_0}^{(h)}, R\Psi)]^{(U_h^w)}) \\ &= \sum_{z \in PHT^{(h)}} |\ker^1(\mathbb{Q}, I_z)| Tr(\phi \mid H^0([\breve{M}_{n-h,+} \times (I_z(\mathbb{Q}) \backslash G^{(h)}/U_h^w)]/J_{n-h,+}, [R\Psi]_z)). \end{aligned}$$

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 $(G(\mathbf{A}_f^w) \times L_{n-h,h}//U_h^w) \times W_K$-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

$$H^0([\breve{M}_{n-h,+} \times (I_z(\mathbb{Q}) \backslash G^{(h)}/U_h^w)]/J_{n-h,+}, [R\Psi]_z)$$

are smooth, but not generally admissible, representations of $G^{(h)}$. However, under the hypothesis that ϕ is very h-regular, the set of fixed points of ϕ on $[\breve{M}_{n-h,+} \times (I_z(\mathbb{Q}) \backslash G^{(h)}/U_h^w)]/J_{n-h,+}$ 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 ϕ 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 ϕ, twisted by a suficiently 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 $[\rho] \in [\mathcal{A}](J_{n-h})$, and the corresponding decomposition (5.5.11) for the cohomology. There is also a version $[\rho, +]$ incorporating the action of the extra factor $\mathbb{Q}_p^\times/\mathbb{Z}_p^\times$, 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 $\mathcal{S}(z)$, and hence an expression for the cohomology space on the right-hand side of (7.2.1):

$$H^0(\mathcal{S}(z), [R\Psi]_z) = \bigoplus_{[\rho]\in[\mathcal{A}](J_{n-h})} H^0(\mathcal{S}(z), [\Psi][\rho]). \tag{7.2.2}$$

We rewrite Fujiwara's trace formula accordingly:

(7.2.3) Corollary. *Under the hypotheses of Theorem 7.2.1,*

$$Tr(\phi \mid [H_c(\bar{S}^{(h)}_{M_0}, R\Psi)]^{(U_h^w)}) = \sum_{z\in PHT^{(h)}} \sum_{[\rho]\in[\mathcal{A}](J_{n-h})} |\ker^1(\mathbb{Q}, I_z)| t_{z,[\rho]}(\phi)$$

where

$$t_{z,[\rho]}(\phi) = Tr(\phi\ |H^0([\breve{M}_{n-h,+} \times (I_z(\mathbb{Q})\backslash G^{(h)}/U_h^w)]/J_{n-h,+}, [R\Psi]_z[\rho])).$$

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

$$(7.3.1)\quad n[H_c(\bar{S}^{(h)}_{M_0}, R\Psi)] =$$
$$\bigoplus_{[\rho]\in[\mathcal{A}](J_{n-h})} \bigoplus_{\rho'\in\mathcal{A}(J_{n-h})_{fin}[\rho]} red^{(h)}_{\rho'}[H(\mathcal{A}(B,*))]\otimes[\Psi_{n-h}(\rho')] \in Groth(G(\mathbf{A}_f^w))\times G_h);$$

Formula (7.3.1) is understood as an equality in $Groth(G(\mathbf{A}_f^w) \times L_{n-h,h})$. 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 ϕ^w and $\phi_{w,h}$ are chosen arbitrarily, whereas $\phi_{w,n-h}$ has to be chosen so that the resulting ϕ 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 $\mathcal{A}(J_{n-h})_{fin}[\rho]$ 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 $\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]$, relative to the set $JL(\mathcal{A}(J_{n-h})_{fin}) \subset \mathcal{A}_d(n-h, K)$ (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 ϕ be very h-regular.

Fix $\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]$ and let $\pi' = JL(\rho')$. To fix ideas, and to simplify the formulas the first time around, we assume

(7.3.2) Hypothesis. *π' is supercuspidal and $\phi_{w,n-h}$ is a pseudocoefficient for π', denoted $\phi_{\pi';\omega}$ in the notation of (A.1).*

Here ω 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 π' among its unramified twists occurring in $JL(\mathcal{A}(J_{n-h})_{fin})$. Set $\phi_h^w = \phi^w \otimes \phi_h$. A test function of the form $\phi = \phi_h^w \otimes \phi_{\pi';\omega}$ as above – in particular, satisfying the h-regularity condition – will be called *acceptable for ρ'*. We verify (7.3.1) by proving equality of traces for all test functions acceptable for ρ', for all $\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]$. 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 ρ is an arbitrary representation of J_{n-h}.

For $a \in I_z(\mathbb{Q})$, define the orbital integral

$$O_{[a]}^h(\phi_h^w) = \int_{Z(a)\backslash G(\mathbf{A}_f^w)\times G_h} \phi_h^w(gag^{-1})d\dot{g}. \tag{7.3.3}$$

Here $Z(a)$ is the centralizer of a in $G^{(h)}(\mathbf{A}_f)$. 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[8] to see that this implies

$$Z_G(a) = Z_{I_z}(a) \tag{7.3.4}$$

[HT,Lemma V.2.2], so $Z(a)$ is the adelization of the $\mathbb{Q}$-group $Z_{I_z}(a)$ (though $G^{(h)}(\mathbf{A}_f)$ 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 $J_{n-h} \xrightarrow{\sim} G_{n-h}$ over $\overline{K}$. (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 vol. Here is an expression for the contribution of $z \in PHT^{(h)}$ to the trace formula:

[8] 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.

(7.3.5) Theorem. *Fix $\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]$, and let $\phi = \phi_h^w \otimes \phi_{\pi';\omega}$ be a test function acceptable for ρ'. Then*

$$\begin{aligned} &Tr(\phi \mid H^0([\breve{M}_{n-h,+} \times S^h_{U_h^w}(z)]/J_{n-h,+}, [R\Psi]_z[\rho])) \\ &\quad = (n-h)\sum_{[a]} e(\gamma(a))O^h_{[a]}(\phi^w_h)\cdot O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi';\omega})v(Z_{I_z}(a)(\mathbb{Q})\backslash Z_{I_z}(a)(\mathbf{A}_f)). \end{aligned}$$

Here $[a]$ runs through h-regular $I_z(\mathbb{Q})$-conjugacy classes in $I_z(\mathbb{Q})$, and the volume $v(Z_{I_z}(a)(\mathbb{Q})\backslash Z_{I_z}(a)(\mathbf{A}_f))$ 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 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.

The first subtlety involves rewriting the volume factor, using Kottwitz' results on Tamagawa numbers [K3]. The formula is

$$v(Z_{I_z}(a)(\mathbb{Q})\backslash Z_{I_z}(a)(\mathbf{A}_f)) = \kappa_B|\ker^1(\mathbb{Q}, Z_{I_z}(a))|^{-1}v(Z_{I_z}(a)(\mathbb{R})^1)^{-1}$$

where κ_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)

$$= (n-h)\kappa_B\sum_{[a]} e(\gamma(a))O^h_{[a]}(\phi^w_h)\cdot O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi',\omega})|\ker^1(\mathbb{Q}, Z_{I_z}(a))|^{-1}v(Z_{I_z}(a)(\mathbb{R})^1)^{-1}$$

Next, to rewrite Theorem 7.3.5 as a sum over $I_z(\mathbf{A})$-conjugacy classes, we note that if $a, a' \in I_z(\mathbb{Q})$ are $I_z(\mathbf{A})$-conjugate, then their centralizers are inner forms of each other that become isomorphic over $\mathbf{A}$, and their Tamagawa measures agree under this isomorphism. Thus

$$\text{(7.3.7)} \qquad O^h_{[a]}(\phi^w_h)vol(Z_{I_z}(a)(\mathbb{R})^1)^{-1} = O^h_{[a']}(\phi^w_h)vol(Z_{I_z}(a')(\mathbb{R})^1)^{-1}.$$

So it suffices to count the number of $I_z(\mathbb{Q})$-conjugacy classes in an $I_z(\mathbf{A})$-conjugacy class, and this is

$$\text{(7.3.8)} \qquad |\ker[\ker^1(\mathbb{Q}, Z_{I_z}(a)) \to \ker^1(\mathbb{Q}, I_z)]| = |\ker^1(\mathbb{Q}, Z_{I_z}(a))|/|\ker^1(\mathbb{Q}, I_z)|$$

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, $[a]/\mathbf{A}$ for $I_z(\mathbf{A})$-conjugacy classes, and

write $v(a) = vol(Z_{I_z}(a)(\mathbb{R})^1)$. Then
(7.3.9)

$$
\begin{aligned}
&Tr(\phi \mid [H_c(\bar{S}^{(h)}_{M_0}, [R\Psi]]^{(U^w_h)}) = \\
&= \sum_{z \in PHT^{(h)}} \sum_{[\alpha]\in[\mathcal{A}](J_{n-h})} |\ker^1(\mathbb{Q}, I_z)| t_{z,[\alpha]}(\phi) \\
&= \sum_{z \in PHT^{(h)}} |\ker^1(\mathbb{Q}, I_z)| t_{z,[\rho]}(\phi) \\
&= (n-h)\kappa_B \sum_{z,[a]} |\ker^1(\mathbb{Q}, I_z)|/|\ker^1(\mathbb{Q}, Z_{I_z}(a))| e(\gamma(a)) O^h_{[a]}(\phi^w_h) \cdot O^{G_{n-h}}_{[\gamma(a)]}(\phi_{n-h}) \\
&= (n-h)\kappa_B \sum_{(z,[a])\in\mathcal{I}^{(h)}} v(a)^{-1} e(\gamma(a)) O^h_{[a]}(\phi^w_h) \cdot O^{G_{n-h}}_{[\gamma(a)]}(\phi_{n-h}).
\end{aligned}
$$

The first equality is (7.2.3), and the third is (7.3.5). The second follows from our choice of $\phi_{w,n-h}$ to be a pseudocoefficient for π', 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 §7.6. The final line summarizes the discussion following (7.3.5). Note that the passage from $[a]/\mathbb{Q}$ to $[a]/\mathbf{A}$ 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 $\mathcal{G}^{(h)}$, and then with $FP^{(h)}$, to rewrite this as

$$
(7.3.10) \qquad (n-h)\kappa_B \sum_{(\gamma,\tilde{w})\in FP^{(h)}} [F(\gamma):F]^{-1} e(\gamma) O^h_{[a]}(\phi^w) \cdot O^{G_{n-h}}_{[\gamma]}(\phi_{n-h}) v(a)^{-1}
$$

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 $a \in I_z(\mathbb{Q})$ which transfers to γ. To remove all trace of a, we consider these terms in turn. First, $v(a) = v(\gamma) = vol(Z_G(\gamma)(\mathbb{R})^1_0)$ where $Z_G(\gamma)$ is of course the centralizer of γ in G, $Z_G(\gamma)(\mathbb{R})_0$ is the compact mod center inner form of $Z_G(\gamma)(\mathbb{R})$,

$$
Z_G(\gamma)(\mathbb{R})^1_0 = \ker|\nu| : Z_G(\gamma)_0 \to \mathbb{R}^\times_{>0}.
$$

Moreover, $Z_G(\gamma)$ is given Tamagawa measure as before. Next, we can obviously replace the orbital integral over $[a]$ in $G(\mathbf{A}^w_f) \times G_h$ 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

$$
\begin{aligned}
(7.3.11) \quad & Tr(\phi| \, [H_c(\bar{S}^{(h)}_{M_0}, R\Psi)]) = \\
&= (n-h)\kappa_B \sum_{(\gamma,\tilde{w})\in FP^{(h)}} [F(\gamma):F]^{-1} v(\gamma)^{-1} e(\gamma) O^{G(\mathbf{A}^w_f)}_{[\gamma]}(\phi^w) \cdot O^{L_{n-h,h}}_{[\gamma]}(\phi_w)
\end{aligned}
$$

We have removed the superscript $^{(U^w_h)}$ because it is built into our choice of functions ϕ. By definition of $FP^{(h)}$, the γ'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.

$$Tr(\Phi \mid [H(\mathcal{A}(B,*))]) = n\kappa_B \sum_{\gamma} e(\gamma)[F(\gamma):F]^{-1} v(\gamma)^{-1} O_{[\gamma]}(\Phi) \tag{7.4.1}$$

Here $\Phi = \Phi_w \otimes \phi^w \in C_c^\infty(G_n \times G(\mathbf{A}_f^w))$ and we have written $v(\gamma)$ for $vol(Z_G(\gamma)(\mathbb{R})_0^1)$, as in (7.3.11). To compare this with our final version (7.3.11) of the trace formula for the stratum $\bar{S}_{M_0}^{(h)}$, we need a way to compare orbital integrals on G_n with orbital integrals on $L_{n-h,h}$. 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 ϕ_{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 $O_\gamma^{G_n}(\Phi_w) = 0$;*

(7.4.2.2) *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 γ (i.e., $s \in G_n$ takes γ to $L_{n-h,h}$) such that the G_{n-h}-factor of $s(\gamma)$ is elliptic;

(7.4.2.3) *Let π be an irreducible admissible representation of G_n, with*

$$[r_{G_n,L_h}^{op}(\pi)] = \sum m_{\alpha,\beta}[\alpha \otimes \beta]$$

for $\alpha \in \mathcal{A}(n-h,K)$ and $\beta \in \mathcal{A}(h,K)$. Then

$$Tr(\pi)(\Phi_w) = \sum m_{\alpha,\beta} Tr(\alpha)(\phi_{n-h}) Tr(\beta)(\phi_h).$$

The ellipticity hypothesis in the above proposition is superfluous, but is satisfied in our present situation. The existence of Φ_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 Φ_w, for $\phi_{n-h} = \phi_{\pi';\omega}$ a test function acceptable for ρ', we find

$$(7.4.3) \quad (n-h)Tr(\Phi \mid [H(\mathcal{A}(B,*))]) =$$
$$n(n-h)\kappa_B \sum_{\gamma} e(\gamma)[F(\gamma):F]^{-1} v(\gamma)^{-1} \sum_{s} O_{[\gamma]}^{G(\mathbf{A}_f^w)}(\phi^w) O_{s(\gamma)}^{L_{n-h,h}}(\phi_{n-h} \otimes \phi_h)$$

Now note that there is a one-to-one correspondence between $\tilde{w}$ as in (7.3.11) and s as in (7.4.2.2): each s defines the subfield $F_w(\gamma) \subset M_{n-h}(K)$ – a subfield because of the ellipticity condition – hence a completion $F(\gamma)_{\tilde{w}}$ above w that satisfies the degree condition.

Formula (7.4.3) does not require π' to be supercuspidal. If we now return to Hypothesis 7.3.2 – in particular, π' is supercuspidal – we can combine (7.4.3) with (7.3.11), and obtain

$$(n-h)Tr(\Phi \mid [H(\mathcal{A}(B,*))]) = n \cdot Tr(\phi| \; [H_c(\bar{S}^{(h)}_{M_0}, R\Psi)^U[\rho]]). \tag{7.4.4}$$

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 $[H(\mathcal{A}(B,*))]$ coming from $[\rho]$. Indeed, if

$$[H(\mathcal{A}(B,*))] = \sum_{\tau} a(\tau_f)\tau_f^w \otimes \tau_w \in Groth(G(\mathbf{A}_f))$$

where τ runs through a set of cohomological automorphic representations, then the trace relation (7.4.2.3) implies that

$$Tr(\Phi \mid [H(\mathcal{A}(B,*))]) = \sum_{\tau}^{\pi'} Tr(\phi^w \mid \tau_f^w) Tr(\Phi_w \mid \tau_w).$$

Here the symbol $\sum_{\tau}^{\pi'}$ indicates that the sum is taken over those τ such that $[r^{op}_{G_n,L_h}(\tau_w)] = \sum m_{\alpha,\beta}[\alpha \otimes \beta]$ such that, for some α that occurs,

$$Tr(\alpha)(\phi_{\pi';\omega}) \neq 0; \tag{7.4.5}$$

in other words, such that α is inertially equivalent to π'. But since we are working at finite level, the set of all α's that arise in this way is finite. Hence, by expanding $\mathcal{A}(J_{n-h})_{fin}$ if necessary, we can arrange that (7.4.5) only holds for $\alpha = \pi'$, and in that case, as we know, $Tr(\pi')(\phi_{\pi';\omega}) = 1$. It then easily follows that

(7.4.6)

$$Tr(\Phi \mid [H(\mathcal{A}(B,*))]) = \sum_{\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]} Tr(\phi^w \otimes \phi_w \mid JL(\rho') \otimes red^{(h)}_{\rho'} H(\mathcal{A}(B,*))$$

Indeed, only $JL(\rho') = \pi'$, for our chosen π', gives a non-zero contribution to the right-hand side of (7.4.6). By varying ρ', we see that the identity (7.4.6) is valid for every $\rho' \in \mathcal{A}^0_{n-h}$. 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 ϕ_{n-h} be arbitrary subject to the h-regularity condition. As noted, Proposition 7.4.2 holds without the ellipticity hypothesis, and we let Φ_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

$$\begin{aligned}
&Tr(\phi^w \otimes \phi_w \mid red^{(h)}_{\rho'}[H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho')]) \\
&\quad = Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]) \cdot Tr(\phi^w \otimes \phi_h \mid red^{(h)}_{\rho'}[H(\mathcal{A}(B,*))]) \\
&\quad = Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]) \cdot Tr(\phi^w \otimes \phi_{\pi',\omega} \otimes \phi_h \mid r^{op}_{G_n,L_h}[H(\mathcal{A}(B,*))]) \\
&\quad = Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]) \cdot Tr(\Phi_w(\pi',\omega) \otimes \phi^w \mid [H(\mathcal{A}(B,*))])
\end{aligned}$$

Returning to (7.4.3), we thus have

(7.4.3 (pre 4.3.11))

$$\sum_{\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]} Tr(\phi^w \otimes \phi_w \mid red^{(h)}_{\rho'}[H(\mathcal{A}(B,*))] \otimes [\Psi_{n-h}(\rho')] =$$

$$n\kappa_B \sum_{\gamma} e(\gamma)[F(\gamma):F]^{-1} v(\gamma)^{-1} \sum_s O^{G(\mathbf{A}^w_f)}_{[\gamma]}(\phi^w) O^{L_{n-h,h}}_{s(\gamma)}(\phi_{\pi',\omega} \otimes \phi_h) \cdot Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')])$$

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 ϕ with the term $(n-h)O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi';\omega})$ replaced by

$$\sum_{\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]} O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi';\omega}) \cdot Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]). \tag{7.4.7}$$

For a general function ϕ_{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, and J, and a discrete group I that embeds (discretely) in $Y \times J$. There is also a discrete abelian group Δ, 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 $\overline{\mathbb{F}}_p$, 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 δ_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_α of any connected component M_α of any M_i is an open compact subgroup of J, and that the action of J_α on M_α factors through a finite quotient. We also assume there is an action of G on M that factors through δ_G, compatible with δ_M. It follows that the stabilizer in G of any M_α 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 ℓ-adic complex $\Psi^\bullet$ on M. This means that, for any open compact subgroup $\mathcal{U} \subset G(0)$, the sheaf $\mathcal{H}^j(\Psi^\bullet)^{\mathcal{U}}$ is a constructible J-equivariant sheaf on M. Then the action of J_α on $\mathbb{H}^\bullet(M_\alpha, \Psi^\bullet)$ factors through a finite quotient.

To simplify notation, and because this is the only case we need, we assume $M = M_i$, a principal homogeneous space for Δ, with fixed component M_α denoted M_0; we write $J(0)$ for J_α. In our applications, J is the compact mod center group $D^\times_{\frac{1}{n-h}}$, $J(0)$ its unique maximal compact subgroup, M is zero-dimensional, and $J/J(0) \xrightarrow{\sim} \mathbb{Z}$ acts transitively on M. However, the arguments presented below can be applied simplicially to the Čech complex of the more general M discussed above. Similarly, we replace the complex $\Psi^\bullet$ by one of the cohomology sheaves $\mathcal{H}^j(\Psi^\bullet)$, which we denote simply Ψ, or by the alternating sum $[\Psi] = \sum_j (-1)^j [\mathcal{H}^j(\Psi^\bullet)]$, a virtual representation of $G \times J$. Additional properties of Ψ, satisified in the applications, will be specified below.

For any open compact subgroup $U \subset Y$, let $\mathcal{S}_U = [M \times (I\backslash Y \times J/U)]/J$, with the profinite topology. The group G acts on $\mathcal{S}_U$ via the action on M. For simplicity, we write δ instead of δ_G, δ_J. Let $y \in Y$, f the characteristic function of the double coset UyU. Let $\phi \in C_c^\infty(G)$.

First, assume ϕ supported on $G(d) = \delta^{-1}(d)$ for some fixed $d \in \Delta$. The pair (f, ϕ) defines a *Hecke correspondence* on $\mathcal{S}_U \times \mathcal{S}_U$: it is the set of pairs of classes of points

$$([\delta, x, j], [d \cdot \delta, xy, j]) \in \mathcal{S}_U \times \mathcal{S}_U,\ \delta \in M,\ x \in Y,\ j \in J \tag{7.5.1}$$

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 $s \in \mathcal{S}_{U(y)}$. We may as well take $j = 1$. A fixed point of the correspondence is a class $[\delta, x, 1]$ such that $[d \cdot \delta, xy, 1] = [\delta, x, 1]$; 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

$$a = j^{-1} = x(yu^{-1})x^{-1}, \delta(a)d = 1. \tag{7.5.2}$$

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 β 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 $[\delta(j)\delta, \beta xv, \beta j]$ 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 $Fix(f \otimes \phi, a)$ 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 $Fix(f \otimes \phi, a)$ is empty. If $\delta(a)d = 1$, one checks easily that

$$|Fix(f \otimes \phi, a)| \ = \ |M/(Z_I(a) \cap U(y)) \times Z_I(a)\backslash X_1(g, a)]/U(y)|, \tag{7.5.3}$$

where $X_1(g, a) = \{x \in Y \mid x^{-1}ax \in yU\}$ and $U(y) = U \cap yUy^{-1}$.

Remark. Suppose we are in the setting of §§6, (7.3); i.e., $Y = G(\mathbf{A}_f^w) \times G_h$, $U = U_h^w$, $G = G_{n-h}$, $J = J_{n-h}$, $\Delta = \mathbb{Z} = J/J(0)$ acts simply transitively on the set $M = \breve{M}_{n-h,+}$. The set on the right-hand side of (7.5.3) is then the same as

$$Z_I(a)\backslash\{x \in Y \times J \mid x^{-1}ax \in yU \times J(d)\}/(U(y) \times J(0)), \tag{7.5.4}$$

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).

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

$$vol(U(y)J(0))^{-1}\cdot vol(\{x\in Z_I(a)\backslash(Y\times J)\mid x^{-1}ax\in yU\times J(0)\}). \tag{7.5.5}$$

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

$$vol(U(y)J(0))^{-1}\cdot vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^{Y\times J}_{[a]}(\chi_{y,U}\cdot\chi_d) \tag{7.5.6}$$

where χ_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 ϕ is supported on $G(d)$ and f is the characteristic function of UyU, the number of fixed points $s\in\mathcal{S}_U$ with $[a(s)]=[a]$ equals*

$$[vol(U)vol(J(0))]^{-1}\cdot vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^{Y\times J}_{[a]}(f\cdot\chi_d).$$

More generally, let $\omega=[a,b]\subset\mathbb{Z}$ be an interval as in (A.1.10), and assume ϕ 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

$$[vol(U)vol(J(0))]^{-1}\cdot vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^{Y}_{[a]}(f)O^{J}_{[a]}(\chi_\omega).$$

Here χ_ω is the characteristic function of $J(\omega)=\delta_J^{-1}(\omega)$.

The formula in the final paragraph follows by linearity.

For fixed a and ω, the orbital integral $O^J_{[a]}(\chi_\omega)$ is given as follows:

$$O^J_{[a]}(\chi_\omega)=vol(J/Z_J(a)),\delta(a)\in\omega;\;\; O^J_{[a]}(\chi_\omega)=0,\delta(a)\notin\omega. \tag{7.5.8}$$

The measure on $Z_J=K^\times$ is normalized by (A.1.8), and one sees readily that

(7.5.9)

$$vol(J(0))^{-1}vol(J/Z_J(a))=[J:Z_J\cdot J(0)]\cdot vol(Z_J(a)/Z_J)^{-1}=g\cdot vol(Z_J(a)/Z_J)^{-1}.$$

Thus the cardinality in (7.5.7) can be rewritten:

(7.5.10)

$$|Fix(f\otimes\phi,a)|=g\cdot vol(U)^{-1}vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^Y_{[a]}(f)\cdot vol(Z_J(a)/Z_J)^{-1}.$$

Now the $G\times J$-equivariant constructible ℓ-adic sheaf Ψ on M descends to a constructible ℓ-adic sheaf, still denoted Φ, on $\mathcal{S}_U$. The function $f\otimes\phi$ acts as a *Hecke operator* on Ψ over $\mathcal{S}_U$. 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 $vol(U)$. Let

$$[H(\mathcal{S}_U,[\Psi])] = \sum_{i,j}(-1)^{i+j}H^i(\mathcal{S}_U,\mathcal{H}^j(\Psi^\bullet)).$$

In the application to (7.3.5), and indeed under our assumption that Δ 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 Ψ:

$$(7.5.11) \quad Tr(f \otimes \phi \mid [H(\mathcal{S}_U,[\Psi])]) = vol(U) \sum_{[a]\in I(\mathbb{Q})} \sum_{s \in Fix(f\otimes\phi,a)} Loc_s(f \otimes \phi,[\Psi]).$$

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 $Loc_s(f \otimes \phi,[\Psi])$ 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 $s \in Fix(f \otimes \phi, a)$, and indeed is independent of f (since Y acts trivially on $[\Psi]$). For fixed a, the local term is given by

$$(7.5.12) \qquad Loc_s(f \otimes \phi,[\Psi]) = trace(\phi \otimes a|[\Psi]_0)$$

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 ϕ 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)

$$(7.5.13) \quad Tr(f \otimes \phi \mid [H(\mathcal{S}_U,[\Psi])]) =$$
$$g \cdot \sum_{[a]\in I(\mathbb{Q})} vol(Z_I(a)\backslash Z_{Y\times J}(a)) \cdot O^Y_{[a]}(f) \cdot vol(Z_J(a)/Z_J)^{-1} \cdot trace(\phi \otimes a \mid [\Psi]_0).$$

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 $\phi = \phi_{\pi';\omega}$, $f = \phi^w_h$, as in (A.1), and taking the alternating sum $[\Psi][\rho]$ for Ψ. 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 π' 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 $(-1)^{n-h-1}\Psi^{n-h-1}[\rho]$. Thus it follows from (A.1.12) and (5.5.9) that

(7.6.1 (post 4.3.11))

$$(n-h)\cdot trace(\phi\otimes a \mid \Psi^{n-h-1}[\rho]_0) = (-1)^{n-h-1} trace_{Z_g,\xi}(\phi_\xi\otimes a \mid \Psi^{n-h-1}[\rho]_\xi)$$

for any appropriate central character ξ. Here the extension of the compactly supported function $\phi\otimes a$ to the function $\phi_\xi\otimes a$, compactly supported modulo Z_g, is as in (5.5.9.1). Indeed, if $\xi\neq\psi_{\pi'}$, then

$$trace(\phi\otimes a \mid \Psi[\rho]_\xi) = 0$$

because ϕ is a pseudocoefficient relative to $\mathcal{A}(n-h,K)_{fin}$. On the other hand, if $\xi=\psi_{\pi'}$, then the formula above holds (cf. (5.5.9.1)) and the right hand side can be simplified:

(7.6.2 (post 4.3.11)) $$trace_{Z_g,\xi}((\phi_{\pi';\omega})_\xi\otimes a|\Psi^{n-h-1}[\rho]_\xi) = \frac{n-h}{m}\chi_{JL(\pi',\vee)}(a)$$

Here the coefficient $\frac{n-h}{m}$ 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 $\phi_{\pi';\omega}$ by $(\phi_{\pi';\omega})_\xi$ 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

(7.6.3 (post 4.3.11)) $$\Psi^{n-h-1}[\rho]_\xi = (n-h)\bigoplus_{\rho'\in[\rho]_\xi}\rho'^{,\vee}\otimes JL(\rho'),$$

as representation of $J_{n-h}\times G_{n-h}$; this is just Theorem 4.3.11 with the action of W_K forgotten.

Comparing (7.6.1) and (7.6.2), the specialization of (7.5.13) to the situation of Theorem 7.3.5 becomes

(7.6.4 (post 4.3.11))

$$\begin{aligned}
&Tr(f\otimes\phi \mid H^0([\breve{M}_{n-h,+}\times S^h_{U^w_h}(z)]/J_{n-h,+},[R\Psi]_z[\rho]) \;=\; Tr(f\otimes\phi \mid [H(\mathcal{S}_U,[\Psi])]) =\\
&(-1)^{n-h-1}(n-h)\sum_{[a]\in I(\mathbb{Q})} vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^Y_{[a]}(f)\cdot vol(Z_J(a)/Z_J)^{-1}\cdot\frac{1}{m}\chi_{JL(\pi',\vee)}(a)\\
&\qquad = (n-h)\sum_{[a]\in I(\mathbb{Q})} e(\gamma(a))vol(Z_I(a)\backslash Z_{Y\times J}(a))\cdot O^Y_{[a]}(f)\cdot O_{[\gamma]}(\phi)
\end{aligned}$$

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):

$$(n-h)\sum_{[a]\in I_z(\mathbb{Q})} e(\gamma(a))O^h_{[a]}(\phi^w_h)\cdot O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi';\omega})v(Z_{I_z}(a)(\mathbb{Q})\backslash Z_{I_z}(a)(\mathbf{A}_f)).$$

By the choice of ϕ_w 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 $trace(\phi \otimes a \mid [\Psi]_0)$.

In (7.5.13) we take $[R\Psi]_z[\rho]$ for $[\Psi]_0$. It suffices to show that, for general $\phi = \phi_{n-h}$, assumed to have zero trace on any $\pi \in \mathcal{A}(n-h, K)_{fin}$ with $\psi_\pi \neq \xi$, we have

(7.6.5)
$$(n-h)trace(\phi \otimes a \mid [\Psi]_0) = \sum_{[\rho]} \sum_{\rho' \in \mathcal{A}(J_{n-h})_{fin}[\rho]} O^{G_{n-h}}_{[\gamma(a)]}(\phi_{\pi';\omega}) \cdot Tr(\phi_{n-h} \mid [\Psi_{n-h}(\rho')]),$$

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 (pre 4.3.11))
$$[\Psi][\rho]_\xi = \bigoplus_{\rho' \in [\rho]_\xi} \rho'^{,\vee} \otimes [\Psi_{n-h}(\rho')].$$

Using Proposition (A.1.12 bis), we now find that

(7.6.2 (pre 4.3.11))
$$trace(\phi_{n-h} \otimes a | [\Psi]^U [\rho]_\xi) = \frac{1}{m} \sum_{\rho' \in [\rho]_\xi} \chi_{JL(\pi'^{,\vee})}(a) Tr((\phi_{n-h} \mid [\Psi_{n-h}(\rho')]).$$

We conclude as above. This completes the proof of the Second Basic Identity.

Lecture 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 $\mathcal{K}$ the fraction field of the Witt vectors of the algebraic closure of $\mathbb{F}_p$, and let σ denote the Frobenius acting on $\mathcal{K}$. If G is a reductive group over $\mathbb{Q}_p$, let $B(G)$ denote the set of σ-conjugacy classes in $G(\mathcal{K})$, i.e., equivalence classes for the relation

$$b \sim h \cdot b \cdot \sigma(h)^{-1}, h \in G(\mathcal{K})$$

For any $\mathbb{Q}_p$-rational representation (τ, V) of G, an element $b \in B(G)$ defines a structure of isocrystal on $N_\tau = V \otimes_{\mathbb{Q}_p} \mathcal{K}$ by defining

$$\phi = \tau(b) \otimes \sigma : V \otimes_{\mathbb{Q}_p} \mathcal{K} \to V \otimes_{\mathbb{Q}_p} \mathcal{K}. \tag{8.1.1}$$

If $G = GL(V)$, then any isocrystal with underlying vector space $V \otimes_{\mathbb{Q}_p} \mathcal{K}$ arises this way; b is the matrix of ϕ with respect to some basis of $V \otimes_{\mathbb{Q}_p} \mathcal{K}$, and the σ-linearity of ϕ implies that changing the basis replaces b by a σ-conjugate matrix. For general G and τ, the isocrystal N_τ has "additional structure" in the sense that invariants of G in tensor powers of V give rise to ϕ-fixed vectors ("crystalline Tate classes") in the corresponding tensor powers of N_τ. When G is the similitude group of a non-degenerate bilinear form on V, then (N_τ, ϕ) has a polarization of the corresponding type in the category of isocrystals; when V is a C-module for some $\mathbb{Q}_p$-algebra C, and $G \subset GL_C(V)$, then one obtains a map $C \to End(N_\tau, \phi)$. Combining these two kinds of structure, one obtains the sort of isocrystals arising from the Dieudonné modules of abelian varieetics 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 fibers 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 $h : R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \to G_{\mathbb{R}}$ 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 $K \in G(\mathbf{A}_f)$, where $\mathbf{A}_f$ defines the ring of finite adeles,

$${}_KSh(G, X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbf{A}_f)/K.$$

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 $G(\mathbf{A}_f)$ on

$$Sh(G,X)(\mathbb{C}) = \varprojlim_K {}_K Sh(G,X)(\mathbb{C})$$

is rational over E. In particular, for any irreducible admissible representation π of $G(\mathbf{A}_f)$, the π-isotypic component of the $Hom_{G(\mathbf{A}_f)}(\pi, H^i(Sh(G,X),\mathbb{Q}_\ell))$ (étale cohomology) is naturally a representation space $H^i[\pi]$ for $Gal(\bar{E}/E)$, easily seen to be finite-dimensional.

For any point $h \in X$, we let $\mu_h : \mathbb{G}_{m,\mathbb{C}} \to G_{\mathbb{C}}$ denote the (complex) cocharacter associated to h: identifying the complexification of $R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m)$ with $\mathbb{C}^\times \times \mathbb{C}^\times$, we have

$$\mu_h(z) = h_{\mathbb{C}}(z,1).$$

The conjugacy class of μ_h depends only on X, and its field of definition is precisely $E(G,X)$. We may regard μ_h, or simply μ, as a character of a maximal torus of the complex dual group $\hat{G}$ of G, hence as an extreme weight, necessarily minuscule, of a certain irreducible representation of $\hat{G}$. Let r_μ denote the representation of the L-group ${}^LG$ (relative to $E(G,X)$) constructed by Langlands [La]; its restriction to $\hat{G}$ is just the minuscule representations with extreme weight μ. In [La] Langlands expressed the expectation that most of the middle-dimensional ℓ-adic cohomology of $Sh(G,X)$ would break up as a sum in $Groth(G(\mathbf{A}_f) \times Gal(\bar{E}/E))$:

$$H^{\dim X}(Sh(G,X),\bar{\mathbb{Q}}_\ell) = \oplus \pi_f \otimes r_\ell(\pi_f) \oplus \text{endoscopic contributions} \tag{8.1.2}$$

where the sum on the right is taken over those admissible irreducible representations of $G(\mathbf{A}_f)$ occurring in stable cohomological L-packets (the meaning of "most" above) and $r_\ell(\pi_f)$ is a $\bar{\mathbb{Q}}_\ell$-valued representation of $Gal(\bar{E}/E))$ of dimension $\dim r_\mu$. Moreover, at a place v where π_f is unramified, the local component π_v of π_f is classified, via the Satake isomorphism, by a semi-simple conjugacy class $s(\pi_v) \in {}^LG(\bar{\mathbb{Q}}_\ell)$, and up to conjugacy, geometric Frobenius is given by the formula

$$r_\ell(\pi_f)(Frob_v) = r_\mu(s(\pi_v)). \tag{8.1.3}$$

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 π_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 $Sh(G,X)$. 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 $Sh(G,X)$), and if K_p is a hyperspecial maximal compact subgroup of $G(\mathbb{Q}_p)$, then for sufficiently small compact open subgroups $K^p \subset G(\mathbf{A}_f^p)$, one expects ${}_{K_p \cdot K^p}Sh(G,X)$ to have good reduction at v (cf. [K5] for the PEL case). Let ${}_{K_p}\bar{\mathcal{S}}$ denote the special fiber. Guided by our experience with the Shimura varieties treated in [HT], one would then expect ${}_{K_p}\bar{\mathcal{S}}$ to have a stratification in terms of isocrystals. Moreover, assuming the ${}_{K'_p\cdot K^p}Sh(G,X)$ have reasonable integral models for open subgroups $K'_p \subset K_p$, one would expect the stratification to lift to the corresponding special fibers ${}_{K'_p}\bar{\mathcal{S}}$ 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 $\bar{S}^{(h)}$.

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_μ 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 $\Gamma = Gal(\bar{F}/F)$. Let G be a quasi-split reductive group over F, and fix an F-rational Borel subgroup P_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 $\hat{G}$. Let $A \subset T_0$ be a maximal F-split torus, with cocharacter group $X_*(A)$, and let $\Phi_0 \subset Hom(X_*(A),\mathbb{Z})$ denote the set of roots of A in N_0. Define $\mathfrak{A}$, $\bar{C}_{\mathbb{Q}} \subset \mathfrak{A}$, as in [K7, pp. 267-268]:

$$\mathfrak{A}_{\mathbb{Q}} = X_*(A)\otimes_{\mathbb{Z}}\mathbb{Q};\ \mathfrak{A} = X_*(A)\otimes_{\mathbb{Z}}\mathbb{R};$$

$$\bar{C}_{\mathbb{Q}} = \{x \in \mathfrak{A}_{\mathbb{Q}} |\ <\alpha, x> \geq 0 \forall \alpha \in \Phi_0\}.$$

Let

$$\bar{\nu}: B(G) \to \bar{C}_{\mathbb{Q}}$$

be the Newton map, defined as in [RR] and [K6, *loc. cit*]. When $G = GL(n)$, $\bar{\nu}$ 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 $\bar{\nu}$ by embedding G faithfully in an appropriate $GL(n)$ and using Tannakian arguments.

If P is a standard parabolic subgroup, let $A_P \subset A$ be a split component, and define $\mathfrak{A}_P = X_*(A_P)\otimes_{\mathbb{Z}}\mathbb{R}$ and $\mathfrak{A}_{P,\mathbb{Q}}$ as above. Then $\mathfrak{A}_P$ is naturally a subset of $\mathfrak{A}$, and indeed the chamber $\bar{C}_{\mathbb{Q}}$ is a disjoint union over standard parabolics of the corresponding walls $\mathfrak{A}_P^+$ (see [K7, 5.1] for this notation; we omit the subscript $\mathbb{Q}$ for the walls). Let $\bar{\mathfrak{A}}_P^+ \supset \mathfrak{A}_P^+$ 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

$$B(G) \to X^*(Z(G)^\Gamma) = H_0(\Gamma, \pi_1(G))$$

(see [K6,§3] for the first version of the maps, [Mi2,Prop. B.27] for the second) the image of δ is the negative of the class of $-\mu_x$ (see [Mi2, 6.1.4] for an explanation,)

and such that

$$\bar{\nu}(\delta) \leq \mu_{\mathfrak{A}}. \tag{8.2.2}$$

Here $\mu_{\mathfrak{A}} \in \bar{C}_{\mathbb{Q}}$ is what Kottwitz denotes μ_2, and the order $\leq$ is the usual lexicographic order. The $\delta \in B(G,\mu)$ are precisely those such that, up to replacing μ in its conjugacy class, the pair (δ,μ) is *weakly admissible* in the sense of [RZ]. Equivalently, the filtered isocrystal induced by (δ,μ) 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 $\bar{\nu}(\delta) \in \mathfrak{A}^{+,0}_{P_\delta}$. If $P = LU$ is a standard parabolic, let $B(G)_{L,r} = \{\delta \in B(G) \mid P_\delta = P\}$. Then $B(G) = \coprod_L B(G)_{L,r}$, 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 $i_{LG} : B(L) \to B(G)$ [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 $\bar{\nu} : B(L) \to \mathfrak{A}_P$; let $B(L)^+_b$ (resp. $\bar{B}(L)^+_b) \subset B(L)_b$ denote the subset whose image under this Newton map is contained in $\mathfrak{A}^+_P$ (resp. in $\bar{\mathfrak{A}}^+_P$). Then i_{LG} is injective on $\bar{B}(L)^+_b$, 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 μ. Let $\Gamma_E = Gal(\bar{E}/E)$. Consider the Langlands representation r_μ of ${}^LG$, taken relative to E. Let $P = LU \subset G$ be a standard $\mathbb{Q}_p$-rational parabolic. The representation r_μ decomposes, upon restriction to ${}^LL$, as a sum of irreducible components $\mathcal{C}_0(L,\mu)$, each intervening with multiplicity one. Indeed, μ is a minuscule weight, with stabilizer $W_\mu = W_{\mathfrak{q}}$ for a certain parabolic subalgebra $\mathfrak{q} \subset \mathfrak{g}$ defined over $\bar{\mathbb{Q}}$. The irreducible components of r_μ are indexed by the set of Γ_E-orbits in $(W_P\backslash W_G/W_{\mathfrak{q}})$ 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 Γ 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 λ to λ'. 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 Γ_E *orbits in* $W_G(\mu)$ *and the set of* Γ*-orbits in the* $W_G \ltimes \Gamma$*-orbit of* μ *in* $X^*(T^0)$*. Thus* $\mathcal{C}(L,\mu)$ *can be identified with the set of* $W_P \ltimes (W_G(L) \ltimes \Gamma)$ *orbits in the* $W_G \ltimes \Gamma$*-orbit of* μ *in* $X^*(T_0)$*. In particular, we can replace* Γ_E*-orbits by* Γ*-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 μ 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

$$\{\text{minuscule highest weights of } \hat{L}\} \leftrightarrow X^*(Z(\hat{L})). \tag{8.2.4}$$

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 $W_G(L) \ltimes \Gamma$-equivariant (W_P acts trivially on both sides) and induces a bijection, which we denote β_L, between the set $\mathbb{M}(\hat{L})$ of Γ-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 $\mathbb{M}(\hat{L})$.

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* α *of* (G,T_0)*; here* H_α *is the standard vector in* $Lie(T_0)$.

Then the map β_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 β_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 μ. In other words, we need to show that, if λ 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 μ' 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

$$B(G,\mu)_L = B(G,\mu) \cap i_{LG}(B(L)_b) = B(G,\mu) \cap i_{LG}(\bar{B}(L)_b^+).$$

Note that $B(G,\mu)_L$ is *not* generally contained in $B(G)_{L,r}$. Let

$$\alpha_L : B(L)_b \xrightarrow{\sim} X^*(Z(\hat{L})^\Gamma)$$

denote the bijection of [K6, Proposition 5.6]. To any element $\delta \in B(L)_b$ we can associate its Kottwitz invariant $\alpha(\delta) = \alpha_G(i_{LG}(\delta)) \in X^*(Z(\hat{G})^\Gamma)$. 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 $Strat_L : \mathcal{C}(L,\mu) \to B(G,\mu)_L$ uniquely determined by the property that, if $Strat_L(w\mu) = i_{LG}(\delta_L)$, 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 α_L is a bijection on basic classes, this condition certainly determines $Strat_L$ uniquely. It thus remains to be shown that α_L defines a bijection between $\bar{B}(L)_b^+$, which we identify with $B(G,\mu)_L$ via i_{LG}, and $\mathcal{S}(L,\mu)$. It follows from (2) of Lemma 8.2.5 that $\mathcal{S}(L,\mu) \subset \alpha_L(B(G,\mu)_L)$. Moreover, every element of $B(G,\mu)$ satisfies (1) of Lemma 8.2.5. On the other hand, the order on $\bar{B}(L)_b^+$ defined by the Newton map is compatible with that on $X^*(Z(\hat{L})^\Gamma)$, i.e., by pairings with the vectors H_α for simple roots α. Since $\beta_G(\mu) \geq \alpha_L(\delta_L) \geq 0$, for $\delta_L \in \bar{B}(L)_b^+$, with μ minuscule, it follows that $\alpha_L(\delta)$ satisfies (3) as well. This completes the proof.

We now let $\mathcal{C}(\mu) = \coprod_L \mathcal{C}(L,\mu)$, where L runs through the classes of standard Levi subgroups of G.

Corollary 8.2.7. *There is a natural surjective map*

$$Strat : \mathcal{C}(\mu) \to B(G,\mu)$$

given on $\mathcal{C}(L,\mu)$ by $Strat_L$.

Indeed, the map is surjective because

$$B(G) = \cup_L i_{LG}(\bar{B}(L)_b^+) \tag{8.2.8}$$

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 $Rep(b) = Strat^{-1}(b)$; it is a set of pairs (L,λ), with $\lambda \in \mathcal{C}(L,\mu)$, partially ordered by inclusion in the obvious sense. It contains a maximal element $(M = M(b), \lambda_b)$ 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 $Rep(b)$ and the set of $\mathcal{P}(b)$ of standard parabolics $P \subset M = M(b)$ that transfer to the inner form $J(b)$ of M defined by the basic σ-conjugacy class b.*

Proof. We have seen that $Rep(b)$ 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 σ-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

$$j : H^1(\Gamma, M_{ad}) \xrightarrow{\sim} B(M_{ad})_b \tag{8.2.9.1}$$

(cf. [K7, 3.2]) and $j^{-1}(b_{ad})$ is the cohomology class defining the inner form $J(b)$.

One direction is simple. Suppose $b = i_{LG}(b_L)$ 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 σ-conjugate to $b_I \in B(I_0) \cap j_{I_0}(H^1(\Gamma, I_0))$, where for any reductive group H, there is a natural bijection

$$j_H : H^1(\Gamma, H) \to B(H)_b \tag{8.2.9.2}$$

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 $b = i_{I_0 M}(b_I) \in Im[H^1(\Gamma, L) \to B(L)_b \to B(M)]$, hence *a fortiori* belongs to the image under i_{LM} of the image of b_I in $B(L)_b$.

Example (8.2.10) We work out the stratification in the case of a Shimura variety $Sh(G, X)$ 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 $G(\mathbb{R}) \xrightarrow{\sim} GU(k, n-k)$. For an appropriate choice of Shimura datum (G, X), the corresponding representation r_μ of the dual group $\hat{G} \cong GL(n, \mathbb{C}) \times \mathbb{C}^\times$ of G is of the form $\wedge^k St \otimes \nu$, where St is the standard representation of $GL(n)$ and ν 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 ${}_K Sh(G, X)$ 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

$$(k, n) = \sum_{i=1}^{m} (r_i, s_i) \tag{8.2.10.1}$$

where r_i, s_i are non-negative integers satisfying $r_i \leq s_i$ for all i, and the rational numbers $\frac{r_i}{s_i}$ are all distinct. The order in the sum is immaterial. The geometric point $x \in \bar{S}$ belongs to the stratum $\bar{S}(\{(r_i, s_i)\})$ if and only if the p-divisible group $\mathcal{H}_x$ is isogenous to a p-divisible group of the form $\prod_i (\mathcal{H}_{\frac{r_i}{s_i}})^{d_i}$, where $\mathcal{H}_{\frac{r_i}{s_i}}$ is a simple p-divisible group of slope $\frac{r_i}{s_i}$ and d_i is the greatest common divisor of r_i and s_i. The centralizer in $G(\mathcal{K})$ of the corresponding slope morphism is then $M(\{(r_i, s_i)\}) = \prod_{i=i}^m GL(s_i, \mathcal{K}) \times \mathcal{K}^\times$, and the associated twisted form is

$$J(\{(r_i, s_i)\}) = \prod_{i=i}^{m} GL(d_i, D_{\frac{r_i}{s_i}}) \times \mathbb{Q}_p^\times,$$

where $D_{\frac{r_i}{s_i}}$ is the division algebra of dimension $(\frac{s_i}{d_i})^2$ with invariant $\frac{r_i}{s_i}$. 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 (δ_i), where each δ_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 $L = L(\{n_j\}) \equiv \prod GL(n_j, \mathbb{Q}_p) \times \mathbb{Q}_p^\times$ 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 $\hat{L}$ 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 $\hat{L}$-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 $Strat : \mathcal{C}(\mu) \rightarrow B(G, \mu)$ 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 $\frac{k_i}{n_i}$. 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:

$$(1, n) = (1, n - h) + (0, h) \tag{8.2.10.2}$$

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 $Strat^{-1}(1, n - h)$ consists of a single element, whereas $Strat^{-1}(0, h)$ consists of all partitions of h. In other words, $Strat^{-1}(1, n - h)$ 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 ${}_{K_p}\bar{\mathcal{S}}$ at a place v of the reflex field E of the Shimura variety ${}_{K_p \cdot K^p}Sh(G, X)$ with good reduction at v. Let $G_p = G_{\mathbb{Q}_p}$, and let μ be the cocharacter of G associated to the Shimura datum (G, X), viewed as a $\overline{\mathbb{Q}}_p$-cocharacter. Then one expects ${}_{K_p}\bar{\mathcal{S}}$ to decompose as a disjoint union of locally closed reduced subvarieties

$${}_{K_p}\bar{\mathcal{S}} = \coprod_{b \in B(G_p, \mu)} {}_{K_p}\bar{\mathcal{S}}(b). \tag{8.3.1}$$

When (G, X) is a PEL type, ${}_{K_p}\bar{\mathcal{S}}$ 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: ${}_{K_p}\bar{\mathcal{S}}(b)$ 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 ${}_{K_p}\bar{\mathcal{S}}(\geq b)$ denote the closure of ${}_{K_p}\bar{\mathcal{S}}(b)$ in the special fiber. It then follows from the results of [RR] that ${}_{K_p}\bar{\mathcal{S}}(\geq b)$ is a finite union of strata ${}_{K_p}\bar{\mathcal{S}}(b')$, for $b' \in B(G_p, \mu)$ such that $\bar{\nu}(b') \leq \bar{\nu}(b)$ 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 $Sh^{rig}_{K'_p}$, associated to the Shimura variety ${}_{K'_p \cdot K^p}Sh(G,X)$ (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 $Sh^{rig,\geq b}_{K_p} \subset Sh^{rig}_{K_p}$ denote the (open) tube over the closed subvariety ${}_{K_p}\bar{\mathcal{S}}(\geq b)$ of the special fiber: $Sh^{rig,\geq b}_{K_p}$ is the set of points of $Sh^{rig}_{K_p}$ whose specialization lies in ${}_{K_p}\bar{\mathcal{S}}(\geq b)$. For any open subgroup $K'_p \subset K_p$, we define $Sh^{rig,\geq b}_{K'_p}$ to be the fiber product of $Sh^{rig,\geq b}_{K_p}$ with $Sh^{rig}_{K'_p}$ over $Sh^{rig}_{K_p}$; note that this can be defined without reference to an integral model of ${}_{K'_p \cdot K^p}Sh(G,X)$. We let $Sh^{rig,b}_{K_p}$ denote the complement of $Sh^{rig,\geq b'}_{K_p}$, for $\bar{\nu}(b') < \bar{\nu}(b)$, in $Sh^{rig,\geq b}_{K_p}$.

Let π_f be a representation of $G(\mathbf{A}_f)$ contributing to non-endoscopic cohomology in (8.1.2). We will soon assume π_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 π_p is isomorphic to the representation induced from a discrete series representation τ_p of L. Then the Langlands parameter attached to π_p is (conjecturally) given by a homomorphism $\sigma(\pi_p) : WD_{E_v} \to {}^L L(\bar{\mathbb{Q}}_\ell)$. Compatibility of local and global correspondences, generalizing Theorem 1.3.6, amounts to the hypothesis that the restriction to WD_{E_v} of $r_\ell(\pi_f)$ to WD_{E_v} is equivalent to $r_\mu \circ \sigma(\pi_p)$. In particular, by the discussion preceding Remark 8.2.3,

$$(8.3.2) \qquad r_\ell(\pi_f)|_{WD_{E_v}} = \oplus_{\lambda \in \mathcal{C}(L,\mu)} r_\ell(\pi_f)_\lambda,$$

where we have grouped together irreducible summands that are associate.

Let $[H^\bullet(Sh(G,X), \bar{\mathbb{Q}}_\ell)]$ denote the direct limit, over $K'_p \cdot K^p$, of the alternating sum of the ℓ-adic cohomology groups of ${}_{K'_p \cdot K^p}Sh(G,X)$. We define $[H^\bullet_c(Sh^{rig,b}, \bar{\mathbb{Q}}_\ell)]$ analogously, using this time the ℓ-adic cohomology of the indicated rigid space. Roughly speaking, the stratification gives rise to an identity in the Grothendieck group of $G(\mathbf{A}_f) \times WD_{E_v}$, analogous to the First Basic Identity (4.4.4):

$$(8.3.3) \qquad [H^\bullet(Sh(G,X), \bar{\mathbb{Q}}_\ell)] = \sum_{b \in B(G,\mu)} [H^\bullet_c(Sh^{rig,b}, \bar{\mathbb{Q}}_\ell)].$$

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 $\varinjlim_{K'_p} H^\bullet_c(Sh^{rig,b}_{K_p}, \bar{\mathbb{Q}}_\ell)$. In [HT] the partition $(1,n) = (1, n-h) + (0,h)$ of (8.2.10.2) corresponds to the stratum here denoted $\bar{S}^{(h)}$. The Second Basic Identity, and more precisely the calculations (6.2.3)-(6.2.7), show that the π_p that contribute to the cohomology of $\bar{S}^{(h)}$ are precisely those for which π_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 π_p is supercuspidal, there is only one λ in the sum (8.3.2)[9], namely $(L = G, w\mu = \mu)$, and $Strat(\lambda)$ is the basic stratum. The heuristic expectation is then the

[9]This does *not* mean the representation $r_\ell(\pi_f)|_{WD_{E_v}}$ is necessarily irreducible, or even indecomposable.

(8.3.4) Conjecture. *Let $b_0 \in B(G,\mu)$ denote the basic class. If π_p is supercuspidal, then $r_\ell(\pi_f)$ comes exclusively from the contribution of $[H_c^\bullet(Sh^{rig,b_0}, \bar{\mathbb{Q}}_\ell)]$ to the right-hand side of (8.3.3). In other words, for $b \neq b_0$,*

$$\sum_i (-1)^i Hom_{G(\mathbf{A}_f)}(\pi_f, H_c^\bullet(Sh^{rig,b_0}, \bar{\mathbb{Q}}_\ell)) = 0$$

in $Groth(G(\mathbf{A}_f))$.

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:

(8.3.5) Theorem (Fargues). *Suppose G is the unitary similitude group of a division algebra B of degree n^2 over a CM field of the form $F = F^+E$, as in (1.1). Suppose B is locally everywhere either split or a division algebra. Let p be a prime unramified in F. Suppose either*

(1) *p splits in E; or*
(2) *p is inert in E and $n = 3$*

Then Conjecture (8.3.4) is true.

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 ϕ against the sum $\sum_i (-1)^i H_c^\bullet(Sh^{rig,b}, \bar{\mathbb{Q}}_\ell)$ 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 ϕ on $[H^\bullet(Sh(G,X), \bar{\mathbb{Q}}_\ell)]$ equals the trace of ϕ on $[H_c^\bullet(Sh^{rig,b_0}, \bar{\mathbb{Q}}_\ell)]]$.

The trace of ϕ on $[H^\bullet(Sh(G,X), \bar{\mathbb{Q}}_\ell)]$ is given by the cohomological trace formula (5.4.2). To compute the trace of ϕ on $[H_c^\bullet(Sh^{rig,b_0}, \bar{\mathbb{Q}}_\ell)]]$, 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_w} 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 $r_\ell(\pi_f)|_{WD_{F_w}}$ depends only on π_p, which allows him to "twist by Frobenius" for fixed π_p at the level of the cohomological trace formula.

For general PEL types of type A and C, Rapoport and Zink have shown in [RZ] that the basic stratum Sh^{rig,b_0} admits a rigid-analytic uniformization by a tower of moduli spaces $\breve{M}(G_p,\mu)_{K'_p}$; a special case of this uniformization is (3.4.10). Using this uniformization, Fargues determines the trace of $[H_c^\bullet(Sh^{rig,b_0}, \bar{\mathbb{Q}}_\ell)]]$ 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 $[H^{\bullet}_c(\breve{M}(G_p,\mu),\bar{\mathbb{Q}}_\ell)]$, as defined by Berkovich or Huber. These cohomology groups are smooth modules for $G_p \times J_{b_0} \times WD_{F_w}$, 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 case (1) of Theorem 8.3.5, J_{b_0} is the multiplicative group of a division algebra with invariant $\frac{r}{n}$ 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* π *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*

$$\sum_i (-1)^i Hom_{J_{b_0}}(H^i_c(\breve{M}(G_p,\mu),\bar{\mathbb{Q}}_\ell),\pi) = [JL(\pi)] \otimes r_\mu \circ \tilde{\sigma}_\ell(JL(\pi)).$$

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.

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 reprsentations of $G_g = GL(g,K)$, $\mathcal{A}_d(g,K)$ (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 $Tr(\pi)$ is a distribution, defined on $C_c^\infty(G_g)$ as the trace of the finite rank operator

$$\pi(\phi) = \int_{G_g} \phi(g)\pi(g)dg. \tag{A.1.1}$$

Note that the trace depends linearly on the choice of Haar measure. It is known thanks to Harish-Chandra that $Tr(\pi)$ is represented by a locally L^1-function χ_π, defined on the regular semi-simple elements G_g^{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 $\sum_{\pi\in A} a_\pi Tr(\pi) = 0$, where $A \subset \mathcal{A}(g,K)$ is a finite subset and $a_\pi \in \mathbb{C}$, is trivial.*

Let $\mathcal{A}_{d,fin} \subset \mathcal{A}_d(g,K)$ 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 ψ of $K^\times$ such that $\pi\otimes\psi\circ\det \in \mathcal{A}_{d,fin}$ is finite.*

In other words, for any $\pi \in \mathcal{A}_d(g,K)$, the intersection $\mathcal{A}_{d,fin}(\pi)$ of $\mathcal{A}_{d,fin}$ with the inertial equivalence class of π is finite. Let If $\pi \in \mathcal{A}_{d,fin}$, a *pseudocoefficient* for π, relative to $\mathcal{A}_{d,fin}$, is a function $\phi_\pi \in C_c^\infty(G_g)$ such that

$$\begin{aligned} \text{(A.1.4)} \quad & Tr(\pi)(\phi_\pi) = 1; \\ & Tr(\pi')(\phi_\pi) = 0 \text{ if either } \pi' \in \mathcal{A}_{d,fin}, \pi' \neq \pi \text{ or } \pi' \in \mathcal{A}_t(g,K), \pi' \notin \mathcal{A}_{d,fin} \end{aligned}$$

(A.1.5) Theorem. [DKV, HT I.3]

(i) For any set $\mathcal{A}_{d,fin}$ satisfying (A.1.3) and any $\pi \in \mathcal{A}_{d,fin}$, a pseudocoefficient ϕ_π for π (relative to $\mathcal{A}_{d,fin}$) exists.

(ii) If π is supercuspidal, then $Tr(\pi')(\phi_\pi) = 0$ for any $\pi' \neq \pi$ (not necessarily tempered).

(iii) For general $\pi \in \mathcal{A}_{d,fin}$, let $\pi' \in \mathcal{A}(g,K)$ be a non-tempered representation such that $Tr(\pi')(\phi_\pi) \neq 0$. Then π' belongs to the **block** *of π; i.e., there is an unramified character ψ of $K^\times$, a proper standard parabolic subgroup $P \subset G_g$, and an irreducible admissible representation τ of P such that $\pi'\otimes\psi\circ\det$ and π are Jordan-Hölder constituents of $n-Ind_P^{G_g}\tau$.*

The pseudocoefficients, and the block of π, can also be defined cohomologically, as Euler-Poincaré functions; cf. [SS], Proposition III.4.1 and Corollary III.4.8.

Without the restriction to $\mathcal{A}_{d,fin}$, the theorem is false, because a given π 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 π, in which case the construction is relatively simple. Let Z_g denote the center of G_g. Let $\phi_{v',v}$ be any matrix coefficient of the contragredient $\pi^\vee$ of π:

$$\phi_{v',v}(g) =< \pi^\vee(g)v', v >$$

for some $v' \in \pi^\vee$, $v \in \pi$ such that $< v', v > \neq 0$. Let ψ_π denote the central character of π; then

$$\phi_{v',v}(zg) = \psi_\pi^{-1}(z)\phi_{v',v}(g), z \in Z_g. \tag{A.1.6}$$

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 π is supercuspidal, the matrix coefficient $\phi_{v',v}$ belongs to $C_c^\infty(G_g, \psi_\pi^{-1})$.

If π' 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 π' by the formula

$$\pi'(f) = \int_{G_g/Z_g} f(g)\pi'(g)d\dot{g}. \tag{A.1.7}$$

Here $d\dot{g}$ is an invariant measure on G_g/Z_g. We write $Tr_{Z_g,\psi}(\pi) = Tr_{Z_g,\psi_\pi}(\pi)$ 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

$$\int_{\mathcal{O}^\times} dz = 1, \tag{A.1.8}$$

and define $d\dot{g}$ to be the quotient measure dg/dz. Then (cf. [DKV], A.3.g)

$$Tr_{Z_g,\psi_\pi}(\pi)(\phi_{v',v}) = d(\pi)^{-1}\phi_{v',v}(1) \tag{A.1.9}$$

where $d(\pi)$ is the formal degree (which depends on the choice of Haar measure on G_g). Thus by choosing v' and v appropriately, we may assume $Tr_{Z_g,\psi_\pi}(\pi)(\phi_{v',v}) = 1$; we then write $\phi_Z = \phi_{v',v}$.

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 ω denote the interval $[a, b]$. For any locally constant function f on G_g, we define the ω-truncation $t_\omega(f) \in C_c^\infty(G_g)$ by

$$t_\omega(f)(g) = f(g), \delta(g) \in \omega;\ t_\omega(f)(g) = 0 \text{ otherwise }. \tag{A.1.10}$$

Then it is easy to see that, for any interval $[a, b]$ as above with m sufficiently large, relative to the set $\mathcal{A}_{d,fin}$, the function

$$\phi = \phi_{\pi;\omega} = \frac{1}{m}t_\omega(\phi_Z) \tag{A.1.11}$$

is a pseudocoefficient for π relative to $\mathcal{A}_{d,fin}$. (In any case, $\phi_{\pi;\omega}$ has trace zero on any tempered representation not inertially equivalent to π, and for large enough m, $\phi_{\pi;\omega}$ separate elements of $\mathcal{A}_{d,fin}(\pi)$.) In particular, we can assume all ϕ_π 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 π 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 $f \in C_c^\infty(G_g)$, the orbital integral $O_\gamma(f) = O_\gamma^{G_g}(f)$ 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 ϕ_π), and on the Haar measure on the centralizer $Z(\gamma) \subset G_g$. Let $d\dot{z}_\gamma$ denote the quotient measure on the quotient $Z(\gamma)/Z_g$ (recall (A.1.9)). Then

(A.1.12) Proposition. *The orbital integrals $O_\gamma(\phi)$ of the pseudocoefficient $\phi = \phi_{\pi;\omega}$ vanish on all non-elliptic semisimple regular $\gamma \in G_g$. For $\gamma \in G_g$ regular elliptic,*

$$O_\gamma(\phi) = vol(Z(\gamma)/Z_g)^{-1}\frac{1}{m}t_\omega(\chi_{\pi^\vee})(\gamma),$$

where

$$vol(Z(\gamma)/Z_g) = \int_{Z(\gamma)\backslash G_g} 1 d\dot{z}_\gamma.$$

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 $\frac{1}{m}$ as well as the truncation.

The *Jacquet-Langlands correspondence* is a bijection

$$\mathcal{A}_d(G_g) \overset{JL}{\leftrightarrow} \mathcal{A}(J_g). \qquad \text{(A.1.13)}$$

The notation JL designates the bijection in either direction. It is characterized by the following character identity

$$\chi(JL(\pi))(a) = (-1)^{g-1}\chi(\pi)(\gamma). \qquad \text{(A.1.14)}$$

if γ is an elliptic regular element and $a \in J_g$ transfers to γ. Thus the expression in Proposition (A.1.12) can be rewritten

$$O_\gamma(\phi) = (-1)^{g-1}\frac{1}{m}vol(Z(\gamma)/Z_g)^{-1}t_\omega(\chi_{JL(\pi^\vee)})(a)$$

for γ 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 γ, the following identity holds*

$$O_\gamma(\phi) = (-1)^{g-1}e(\gamma)vol(Z(\gamma)/Z_g)^{-1}\frac{1}{m}t_\omega(\chi_{JL(\pi^\vee)})(a).$$

Here $e(\gamma)$ is the Kottwitz sign (cf. [L, 1.7.1]).

We let $\mathcal{A}(J_g)_{fin}$ denote the image under JL of $\mathcal{A}_{d,fin}$. Then the analogue of Theorem A.1.5 holds for $\mathcal{A}(J_g)_{fin}$; indeed, the pseudocoefficients can be constructed starting from matrix coefficients just as for $GL(g,K)$.

If π is an admissible representation of $G(\mathbf{A})$, or of $G(\mathbf{A}_f)$, then $Tr(\pi)$, defined just as in the local case, exists as a distribution on $C_c^\infty(G(\mathbf{A}))$. 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

$$Tr(\pi)(\phi) = \prod_v Tr(\pi_v)(\phi_v). \tag{A.1.14}$$

A.2. L and ε 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 $\mathcal{A}_0(GL(n,F)\backslash GL(n,\mathbf{A}_F))$ of global cusp forms.

Let $\mathcal{A}_{gen}(n,K)$ 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 $\pi \in \mathcal{A}(n,K)$, $\pi' \in \mathcal{A}(m,K)$. Let $\Pi \in \mathcal{A}_0(n,F)$, $\Pi' \in \mathcal{A}_0(m,F)$. We fix an global additive character $\psi : ad_F/F \to \mathbb{C}^\times$, and a local additive character $\psi_K : K \to \mathbb{C}^\times$; if $K = F_v$ we assume ψ_K to be the restriction of ψ 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 σ_0, which is a product of Γ-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 Σ_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 Σ_0, and let ψ_v denote the restriction of the additive character ψ. Then there is a functional equation

$$L(s,\Sigma_0) = \varepsilon(s,\Sigma_0)L(1-s,\hat{\Sigma}_0);\ \varepsilon(s,\Sigma_0) = \prod_v \varepsilon(s,\Sigma_{0,v},\psi_v) \tag{A.2.1}$$

Note that the product of the local ε-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

$$\varepsilon(s,\sigma_0 \oplus \sigma_1,\psi_K) = \varepsilon(s,\sigma_0,\psi_K) \cdot \varepsilon(s,\sigma_1,\psi_K), \tag{A.2.2}$$

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 σ' is a virtual representation of dimension zero of $WD_{K'}$, then

$$\varepsilon(s,\sigma',\psi_K \circ Tr_{K'/K}) = \varepsilon(s,Ind_{K'/K}(\sigma'),\psi_K). \tag{A.2.3}$$

These properties are used in (1.4).

Now suppose $\sigma \in \mathcal{G}(n,K)$, $\sigma' \in \mathcal{G}(m,K)$, 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 π, π' as above. This can be done, and with the notation introduced above, there is a global functional equation

$$L(s, \Pi \otimes \Pi') = \prod_v \epsilon_v(s, \Pi_v \otimes \Pi'_v, \psi_v) L(1-s, \Pi^\vee \otimes \Pi'^{,\vee}) \tag{A.2.4}$$

already encountered in (1.4.2). Moreover, the local epsilon factors of pairs satisfy the following analogue of (A.2.2):

$$\varepsilon(s, \pi \boxplus \pi', \psi_K) = \varepsilon(s, \pi, \psi_K) \cdot \varepsilon(s, \pi', \psi_K), \tag{A.2.4}$$

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 Π_v and Π'_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 π_1 and π_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)
$$\prod_{w \in S} \gamma_w(s, \Pi(\chi)_w \otimes \Pi(\chi')_w, \psi_w) = \prod_{w \in S} \gamma_w(s, Ind_{F_{2,w}/F_{1,w}} r_\ell(\chi_w) \otimes Ind_{F'_{2,w}/F_{1,w}} r_\ell(\chi'_w), \psi_w).$$

Here S is the finite set of primes where the data are ramified (including all places w at which either of the local ε-factors is non-trivial and all places where one doesn't know *a priori* that $L_w(s, \Pi(\chi) \otimes \Pi(\chi')) = L_w(s, Ind_{F_2/F_1} r_\ell(\chi) \otimes Ind_{F'_2/F_1} r_\ell(\chi')))$, and

$$\gamma_w(?) = \frac{\varepsilon_w(s, ?, \psi_w) L_w(1 - s, \hat{?})}{L(s, ?)}.$$

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 ε 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 ε 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 ε factors at v.

Referemces

[A] Arthur, J.: The L^2 Lefschetz numbers of Hecke operators, *Invent. Math.*, **97**, 257-290 (1991).

[AC] Arthur, J. and L. Clozel: *Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula*, *Annals of Math. Studies*, **120**, Princeton University Press (1989).

[B] Berkovich, V.G.: Vanishing cycles for formal schemes, *Invent. Math.*, **115**, 549-571 (1994); II, *Invent. Math.*, **125**, 367-390 (1996).

[BZ] Bernstein I. N. and A.V.Zelevinsky: Induced representations of reductive p-adic groups I, *Ann. Sci. ENS (4)*, **10** (1977), 441-472.

[Bu] Bourbaki, N.: *Groupes et algèbres de Lie*, Chap. 7 et 8, Paris: Hermann (1975).

[Bo] Boyer, P.: Mauvaise réduction de variétés de Drinfel'd et correspondance de Langlands locale, *Invent. Math.*, **138**, 573-629 (1999).

[BHK] Bushnell, C., G. Henniart, and P. Kutzko: Correspondance de Langlands locale pour GL_n et conducteurs de paires, *Ann. Sci. E. N.S.*, **31**, 537-560 (1998).

[Ca1] Carayol, H.: Sur la mauvaise réduction des courbes de Shimura, *Compositio Math.*, **59**, 151-230 (1986).

[Ca2] Carayol, H.: Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert, *Ann. scient. Ec. Norm. Sup*, **19**, 409-468 (1986).

[Ca3] Carayol, H.: Non-abelian Lubin-Tate theory, in L. Clozel and J.S. Milne, eds., *Automorphic Forms, Shimura varieties, and L-functions*, New York: Academic Press, vol II, 15-39 (1990).

[C1] Clozel, L.: On the limit multiplicities of discrete series representations in spaces of automorphic forms, *Invent. Math.*, **83** (1986), 265-284.

[C2] Clozel, L.: Représentations Galoisiennes associées aux représentations automorphes autoduales de GL(n), *Publ. Math. I.H.E.S.*, **73**, 97-145 (1991).

[CL] Clozel, L. and J.-P. Labesse, Changement de base pour les représentations cohomologiques de certains groupes unitaires, Appendix to [L], 119-133.

[De1] Deligne, P.: Travaux de Shimura, Séminaire Bourbaki *Lecture Notes in Math.*, **244**, (1971).

[De2] Deligne, P.: Les constantes des équations fonctionnelles des fonctions L, in P. Deligne, W. Kuyk, eds., *Modular functions of one variable, II*, *Lecture Notes in Math.*, **349**, 501-597 (1973).

[DKV] Deligne, P., D. Kazhdan, and M.-F. Vignéras: Représentations des algèbres centrales simples p-adiques, in J.-N.Bernstein, P.Deligne, D.Kazhdan, M.-F.Vigneras *Représentations des groupes réductifs sur un corps local*, Paris: Hermann (1984).

[Dr] Drinfel'd, V. G.: Elliptic Modules, *Mat. Sbornik*, **94** (1974), 561-592.

[Fa] Fargues, L.: Correspondances de Langlands locales dans la cohomologie des espaces de Rapoport-Zink, thèse de doctorat, Université Paris 7 (2001), in *Variétés de Shimura, espaces de Rapoport-Zink, et correspondances de Langlands locales*, *Astérisque*, **291** (2004).

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

[Ha] Harder, G.: Über die Galoiskohomologie halbeinfacher Matrizengruppen. II, *Math. Zeit.*, **92**, 396-415 (1966).

[H1] Harris, M.: Supercuspidal representations in the cohomology of Drinfel'd upper half spaces; elaboration of Carayol's program, *Invent. Math.*, **129**, 75-119 (1997).

[H2] Harris, M.: The local Langlands conjecture for $GL(n)$ over a p-adic field, $n < p$, *Invent. Math.*, **134**, 177-210 (1998).

[H3] Harris, M.: Local Langlands correspondences and vanishing cycles on Shimura varieties, in *Proceedings of the European Congress of Mathematics*, Barcelona, (2000).

[HT] Harris, M. and R. Taylor: On the geometry and cohomology of some simple Shimura varieties, *Annals of Math. Studies* 151 (2001).

[HT2] Harris, M. and R. Taylor: Deformations of automorphic Galois representations, manuscript (1998).

[He1] Henniart, G.: On the local Langlands conjecture for $GL(n)$: the cyclic case, *Ann. Math.*, **123** (1986), 145-203.

[He2] Henniart, G.: La conjecture de Langlands locale numérique pour $GL(n)$, *Ann. Sci. E.N.S. (4)*, **21** (1988), 497-544.

[He3] Henniart, G.: Une conséquence de la théorie de changement de base pour $GL(n)$, in *Analytic Number Theory*, *Lecture Notes in Math.*, **1434** (1990), 138-142.

[He4] Henniart, G.: Caractérisation de la correspondence de Langlands locale par les facteurs ε de paires, *Invent. Math.*, **113** (1993), 339-350.

[He5] Henniart, G.: Une preuve simple des conjectures de Langlands pour GL_n sur un corps p-adique *Invent. Math.*, **139** 439-455 (2000).

[HH] Henniart, G. and R. Herb: Automorphic induction for $GL(n)$ (over local non-archimedean fields), *Duke Math. J.*, **78** (1995) 131-192.

[Hida] Hida, H.: *p-adic Automorphic Forms on Shimura Varieties*, Springer Monographs in Mathematics, New York (2004).

[HG] Hopkins, M. J. and B. Gross, Equivariant vector bundles on the Lubin-Tate moduli space, *Contemporary Math.*, **158** (1994), 23-88.

[Il] Illusie, L.: Autour du théorème de monodromie locale, in: *Périodes p-adiques*, *Astérisque*, **223**, 9-57 (1994).

[JPSS] Jacquet, H., I. I. Piatetski-Shapiro, and J. Shalika: Rankin-Selberg convolutions, *Am. J. Math.*, **105**, 367-483 (1983).

[Ka] Katz, N.: Serre-Tate local moduli, in *"Surfaces Algébriques"*, *Lecture Notes in Math.*, **868**, 141-143 (1981).

[KM] Katz, N. and B. Mazur: *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies **108** (1985).

[K1] Kottwitz, R.: Stable trace formula: cuspidal tempered terms, *Duke Math. J.* 51 (1984), 611-650.

[K2] Kottwitz, R.: Stable trace formula: elliptic singular terms, *Math. Ann.* 275 (1986), 365-399.

[K3] Kottwitz, R.: Tamagawa Numbers, *Annals of Math.*, **127**, (1988), 629-646.

[K4] Kottwitz, R.: On the λ-adic representations associated to some simple Shimura varieties, *Invent. Math.*, **108** (1992), 653-665.

[K5] Kottwitz, R.: Points on some Shimura varieties over finite fields, *Jour. of the AMS* **5** (1992), 373-444.

[K6] Kottwitz, R.: Isocrystals with additional structure, *Compositio Math.*, **56**, 201-220 (1985).

[K7] Kottwitz, R.: Isocrystals with additional structure II, *Compositio Math.*, **109**, 255-339 (1997).

[L] Labesse, J.-P.: Cohomologie, stabilisation, et changement de base, *Astérisque*, **257** (1999).

[La] Langlands, R.P.: Automorphic representations, Shimura varieties, and motives. Ein Märchen, in *Automorphic forms, representations, and L-functions*, *Proc. Symp. Pure Math. AMS*, **33**, part 2, 205-246 (1979).

[LRS] Laumon, G., M. Rapoport, and U. Stuhler, $\mathcal{D}$-elliptic sheaves and the Langlands correspondence, *Invent. Math.*, **113**, 217-338 (1993).

[LT] Lubin, J. and J. Tate: Formal moduli for one-parameter formal Lie groups, (*Bull. SMF*, **94**, 49-59 (1966).

[M] Mantovan, E.: On certain unitary group Shimura varieties, Harvard Ph. D. thesis (2002), in *Variétés de Shimura, espaces de Rapoport-Zink, et correspondances de Langlands locales*, *Astérisque*, **291** (2004).

[Me] Messing, W.: The crystals associated to Barsotti-Tate groups: with applications to abelian schemes, *Lecture Notes in Math.*, **264**, (1972).

[Mi1] Milne, J. S.: Canonical models of mixed Shimura varieties and automorphic vector bundles, in L. Clozel and J.S. Milne, eds., *Automorphic Forms, Shimura varieties, and L-functions*, New York: Academic Press, vol I, 283-414 (1990).

[Mi2] Milne, J. S.: The points on a Shimura variety modulo a prime of good reduction, in R. P. Langlands and D. Ramakrishnan, eds., *The Zeta Functions of Picard Modular Surfaces*, Montréal: Les publications CRM, 151-253 (1992).

[Mu] Mumford, D.: *Abelian Varieties*, Oxford University Press (1974).

[Oo] Oort, F.: Subvarieties of moduli space, *Invent. Math.*, **24** (1974), 95-119.

[RR] Rapoport, M. and M. Richartz: *On the classification and specialization of F-isocrystals with additional structure*, *Compositio Math.*, **103** 153-181 (1996).

[RZ] Rapoport, M. and T. Zink: *Period Spaces for p-divisible Groups*, Princeton: Annals of Mathematics Studies **141** (1996).

[Ro] Rogawski, J: Representations of $GL(n)$ and division algebras over a p-adic field, *Duke Math. J.*, **50**, 161-169 (1983).

[SGA 7] Grothendieck, A., P. Deligne, and N. Katz: *Groupes de monodromie en géométrie algébrique*, *Lecture Notes in Math.*, **288, 240** (1972, 1973).

[Sh] Shahidi, F.: Local coefficients and normalization of intertwining operators for $GL(n)$, *Comp. Math.*, **48**, 271-295 (1983).

[SS] Schneider, P., and U. Stuhler: Representation theory and sheaves on the Bruhat-Tits building, *Publ. Math. IHES*, **85**, 97-191 (1997).

[Ta] Tate, J.: Classes d'isogénie des variétés abéliennes sur un corps fini, Séminaire Bourbaki 352, *Lecture Notes in Math.*, **179**, (1971).

[V] Vignéras, M.-F.: Correspondance de Langlands semi-simple pour $GL(n, F)$ modulo $\ell \neq p$, Prépublication, Institut de Mathématiques de Jussieu (1999).

[Z] Zelevinsky, A. N.: Induced representations of reductive p-adic groups II: on irreducible representations of $GL(n)$, *Ann. Sci. E.N.S.*, (4) **13**, 165-210 (1980).

[Zi] Zink, T.: Über die schlechte Reduktion einiger Shimuramannigfaltigkeiten, *Comp. Math.*, **45**, 15-107 (1981).