THE FUNDAMENTAL LEMMA OF JACQUET AND RALLIS

ZHIWEI YUN

Appendix by Julia Gordon

Abstract

We prove both the group version and the Lie algebra version of the fundamental lemma appearing in a relative trace formula of Jacquet and Rallis in the function field case when the characteristic is greater than the rank of the relevant groups. In the appendix by Gordon, our results are transferred to the p-adic field case, for sufficiently large p.

Contents

1. Introduction

1.1. The conjecture of Jacquet and Rallis and its variant

In [9], Jacquet and Rallis proposed an approach to the Gross-Prosad conjecture for unitary groups using the relative trace formula. In establishing the relative trace formula, they needed a form of the fundamental lemma comparing the orbital integrals of the standard test functions on the symmetric space $\mathrm{GL}_n(E)/\mathrm{GL}_n(F)$ and on the

DUKE MATHEMATICAL JOURNAL
Vol. 156, No. 2, © 2011 DOI 10.1215/00127094-2010-210
Received 25 May 2009. Revision received 13 April 2010.
2010 *Mathematics Subject Classification.* Primary 14H60, 22E35; Secondary 14F20.

unitary group $\mathrm{U}_n(F)$ (here E/F is an unramified extension of a local field F with odd residue characteristic). They explicitly stated (up to sign) a Lie algebra version of this fundamental lemma as a conjecture and verified it for $n \leq 3$ by explicit computation. Following this idea, Zhang [13] stated the group version of this fundamental lemma as a conjecture and verified it for $n \leq 3$.

Let σ be the Galois involution of E fixing F. Let $\eta_{E/F} : F^\times \to \{\pm 1\}$ be the quadratic character associated to the extension E/F. Let $n \geq 2$ be an integer. Let $\mathrm{S}_n(F)$ be the subset of $\mathrm{GL}_n(E)$ consisting of A such that $A\sigma(A) = 1$. Let $\mathrm{U}_n(F)$ be the unitary group associated with the Hermitian space E^n with trivial discriminant. We also need the Lie algebra counterparts of the above spaces. Let $\mathfrak{s}_n(F)$ be the set of n-by-n matrices with entries in E^- (purely imaginary elements in the quadratic extension E/F), and let $\mathfrak{u}_n(F)$ be the set of n-by-n skew-Hermitian matrices with entries in E (the Hermitian form on E^n has trivial discriminant). For a subset of $K \subset \mathfrak{gl}_n(E)$ or $\mathrm{GL}_n(E)$, let $\mathbf{1}_K$ denote the characteristic function of K. The two versions of the Jacquet and Rallis conjecture are the following identities of orbital integrals.

CONJECTURE 1.1.1 (Jacquet and Rallis; see [9], [13])

(1) *For* strongly regular semisimple elements *(see Definition 2.2.1) $A \in \mathfrak{s}_n(F)$ and $A' \in \mathfrak{u}_n(F)$ that* match each other *(see Definition 2.5.1), we have*

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) = (-1)^{v(A)}\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}). \tag{1.1.1}$$

(2) *For strongly regular semisimple elements $A \in \mathrm{S}_n(F)$ and $A' \in \mathrm{U}_n(F)$ that match each other (same notion as above), we have*

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathrm{S}_n(\mathcal{O}_F)}) = (-1)^{v(A)}\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathrm{U}_n(\mathcal{O}_F)}). \tag{1.1.2}$$

Here, for elements $A, A' \in \mathfrak{gl}_n(E)$ or $\mathrm{GL}_n(E)$ and for smooth compactly supported functions f, f' on $\mathfrak{gl}_n(E)$ or $\mathrm{GL}_n(E)$, we have

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(f) = \int_{\mathrm{GL}_{n-1}(F)} f(g^{-1}Ag)\eta_{E/F}\big(\det(g)\big)\, dg,$$

$$\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(f') = \int_{\mathrm{U}_{n-1}(F)} f'(g^{-1}A'g)\, dg.$$

For the definition of $v(A)$, see Definition 2.2.2. The Haar measures dg on $\mathrm{GL}_{n-1}(F)$ and $\mathrm{U}_{n-1}(F)$ are normalized so that $\mathrm{GL}_{n-1}(\mathcal{O}_F)$ and $\mathrm{U}_{n-1}(\mathcal{O}_F)$ have volume 1.

(3) *In the above two situations, if A does not match any A', then the left-hand side of (1.1.1) and (1.1.2) are zero.*

Let $\mathrm{H}_n(F)$ be the set of Hermitian matrices in $\mathrm{GL}_n(E)$, and let $\mathfrak{h}_n(F)$ be the set of n-by-n Hermitian matrices in $\mathfrak{gl}_n(E)$ with respect to the chosen Hermitian form with trivial discriminant. We have the following variant of Conjecture 1.1.1.

CONJECTURE 1.1.2

(1) *For strongly regular semisimple elements $A \in \mathfrak{gl}_n(F)$ and $A' \in \mathfrak{h}_n(F)$ that* match each other *(in the similar sense as Definition 2.5.1), we have*

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{gl}_n(\mathcal{O}_F)}) = (-1)^{v(A)} \mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{h}_n(\mathcal{O}_F)}). \tag{1.1.3}$$

(2) *For strongly regular semisimple elements $A \in \mathrm{GL}_n(F)$ and $A' \in \mathrm{H}_n(F)$ that match each other, we have*

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathrm{GL}_n(\mathcal{O}_F)}) = (-1)^{v(A)} \mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathrm{H}_n(\mathcal{O}_F)}). \tag{1.1.4}$$

(3) *In the above two situations, if A does not match any A', then the left-hand side of (1.1.3) or (1.1.4) is zero.*

1.2. Main results

The purpose of the article is to prove the above conjectures in the case F is a local function field (i.e., of the form $k((\varpi))$ for some finite field k) and in the case $\mathrm{char}(F) > n$ (see Corollary 2.7.2). In the appendix by Gordon, it is shown that the transfer principle of Cluckers and Loeser, which relies on model-theoretic methods, applies to our situation, and therefore our results on local function fields imply the validity of the above conjectures for any local field of sufficiently large residue characteristic.

We first do some reductions. In fact, as observed by Xinyi Yuan, (1.1.3) is simply equivalent to (1.1.1) because multiplication by a purely imaginary element in $E^-\backslash 0$ interchanges the situation. In Proposition 2.6.1, we show that the group version (1.1.2) follows from the Lie algebra version (1.1.1) for any F (of any characteristic); the same argument shows that (1.1.4) follows from (1.1.3). Therefore, the orbital integral identity (1.1.1) for Lie algebras implies all the other identities for any F. Moreover, the vanishing result in Conjectures 1.1.1(3) and 1.1.2(3) follows from a cancellation argument (see Lemma 2.5.3).

To prove (1.1.1) in the case $\mathrm{char}(F) > n$, we follow the strategy of the proof of the Langlands-Shelstad fundamental lemma in the Lie algebra and function field case (recently finished by Ngô [12]), building on the work of many mathematicians over the past thirty years. The geometry involved in the Langlands-Shelstad fundamental lemma consists of a local part—the affine Springer fibers (see [5], [6])—and a global part—the Hitchin fibration (see Ngô [11], [12]). Roughly speaking, the motives of the affine Springer fibers, after taking Frobenius traces, give the orbital integrals. The motives of the Hitchin fibers can be written as a product of the motives of affine

Springer fibers. The advantage of passing from local to global is that we can control the "bad" (noncomputable) orbital integrals by "nice" (computable) orbital integrals using global topological machinery such as perverse sheaves.

In the following two sections, we reformulate Conjecture 1.1.1(1) using local and global moduli spaces, and we indicate the main ideas of the proof.

1.3. The local reformulation

As the first step toward a local reformulation, we translate the problem of computing orbital integrals into that of counting lattices. In [9], the authors introduced $(2n-1)$-invariants associated to an element $A \in \mathfrak{gl}_n(E)$ with respect to the conjugation action of $\mathrm{GL}_{n-1}(E)$. Here the embedding $\mathrm{GL}_{n-1}(E) \hookrightarrow \mathrm{GL}_n(E)$ is given by a splitting of the vector space $E^n = E^{n-1} \oplus E$. The splitting gives a distinguished vector e_0 (which spans the one-dimensional direct summand) and a distinguished covector e_0^* (the projection to Ee_0). In this article, we use a different (but equivalent) set of invariants $a = (a_1, \dots, a_n)$ and $b = (b_0, \dots, b_{n-1})$ for A, where the a_i's are the coefficients of the characteristic polynomial of A, and $b_i = e_0^* A^i e_0$ (so that $b_0 = 1$). We say that $A \in \mathfrak{s}_n(F)$ and $A' \in \mathfrak{u}_n(F)$ *match each other* if they have the same collection of invariants viewed as elements in $\mathfrak{gl}_n(E)$.

We fix a collection of invariants (a, b) which is integral and strongly regular semisimple (see Definition 2.2.1 and Lemma 2.2.4). Then we can associate a finite flat $\mathcal{O}_F$-algebra R_a (see Section 2.2). The invariants b give an R_a-linear embedding $\gamma_{a,b} : R_a \hookrightarrow R_a^\vee = \mathrm{Hom}_{\mathcal{O}_F}(R_a, \mathcal{O}_F)$. We can rewrite the left-hand side of (1.1.1) as

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) = \pm \sum_i (-1)^i \# M_{i,a,b}^{\mathrm{loc}},$$

where each $M_{i,a,b}^{\mathrm{loc}}$ is the set of R_a-lattices Λ such that $R_a \subset \Lambda \subset R_a^\vee$ and such that $\mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : \Lambda) = i$ (see Notation 1.6.1).

The E-vector space $R_a(E) = R_a \otimes_{\mathcal{O}_F} E$ carries a natural Hermitian form given also by b. Similarly, we can rewrite the right-hand side of (1.1.1) as

$$\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}) = \# N_{a,b}^{\mathrm{loc}},$$

where $N_{a,b}^{\mathrm{loc}}$ is the set of $R_a(\mathcal{O}_E)$-lattices Λ' which satisfy $R_a(\mathcal{O}_E) \subset \Lambda' \subset R_a^\vee(\mathcal{O}_E)$ and are self-dual under the Hermitian form.

When F is a function field with residue field k, there are obvious moduli spaces of lattices $\mathcal{M}_{i,a,b}^{\mathrm{loc}}$ and $\mathcal{N}_{a,b}^{\mathrm{loc}}$ defined over k such that $M_{i,a,b}^{\mathrm{loc}}$ and $N_{a,b}^{\mathrm{loc}}$ are the set of k-points of $\mathcal{M}_{i,a,b}^{\mathrm{loc}}$ and $\mathcal{N}_{a,b}^{\mathrm{loc}}$. By the Lefschetz trace formula for schemes over k, Conjecture 1.1.1 is a consequence of the following theorem, which is the main local result of the article.

MAIN THEOREM (Local part)
Suppose that $\mathrm{char}(F) = \mathrm{char}(k) > n$ *and* $\eta_{E/F}(\Delta_{a,b}) = 1$. *Then there is an isomorphism of graded* Frob_k*-modules:*

$$\bigoplus_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} H^*\big(\mathcal{M}^{\mathrm{loc}}_{i,a,b} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell(\eta_{k'/k})^{\otimes i}\big) \cong H^*(\mathcal{N}^{\mathrm{loc}}_{a,b} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell).$$

See Section 2.7 for more details on the notation. This theorem is deduced from the global main theorem, of which we shall give an overview in the next section.

We point out that several easy cases of Conjecture 1.1.1 can be verified without any restriction on F (see Section 2.5).

1.4. The global approach

The global geometry related to the Jacquet and Rallis fundamental lemma is derived from modified versions of Hitchin fibrations for the groups GL_n and U_n. The geometry of Hitchin fibrations is studied in great detail by Ngô [12] and Laumon-Ngô [10]; the latter treats the unitary group case, which is especially important for the purpose of this article.

We fix a smooth projective and geometrically connected curve X over k; to study the unitary group, we also fix an étale double cover $X' \to X$. We fix two effect divisors D and D_0 on X of large enough degree.

We introduce the stack $\mathcal{M}$, classifying quadruples $(\mathcal{E}, \phi, \lambda, \mu)$ where $\mathcal{E}$ is a vector bundle of rank n, ϕ is a Higgs field on $\mathcal{E}$, $\lambda : \mathcal{O}_X(-D_0) \to \mathcal{E}$ is the global counterpart of the distinguished vector e_0, and $\mu : \mathcal{E} \to \mathcal{O}_X(D_0)$ is the global counterpart of the distinguished covector e_0^*. The stack $\mathcal{M}$ is the disjoint union of $\mathcal{M}_i$ ($i \in \mathbb{Z}$), according to the degree of $\mathcal{E}$.

We also introduce the stack $\mathcal{N}$, classifying quadruples $(\mathcal{E}', h, \phi', \mu')$, where $\mathcal{E}'$ is a vector bundle of rank n on X', h is a Hermitian structure on $\mathcal{E}'$, ϕ' is a skew-Hermitian Higgs field on $\mathcal{E}'$, and $\mu' : \mathcal{E} \to \mathcal{O}_{X'}(D_0)$ is the distinguished covector (the distinguished vector is determined by μ' using the Hermitian structure).

Both $\mathcal{M}_i$ and $\mathcal{N}$ fiber over the *Hitchin base* $\mathcal{A} \times \mathcal{B}$, classifying global invariants (a, b). In Section 3, we prove the following geometric properties of these moduli spaces.

- Propositions 3.2.6 and 3.3.2: Over the locus $\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$, $\mathcal{M}_i^{\mathrm{int}}$ and $\mathcal{N}^{\mathrm{int}}$ are smooth schemes over k and the "Hitchin maps" $f_i^{\mathrm{int}} : \mathcal{M}_i^{\mathrm{int}} \to \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$ and $g^{\mathrm{int}} : \mathcal{N}^{\mathrm{int}} \to \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$ are proper.
- Proposition 3.5.2: Fix a Serre invariant $\delta \geq 1$ (see Section 3.5). For $\deg(D)$ and $\deg(D_0)$ large enough, the restrictions of f_i^{int} and g^{int} to $\mathcal{A}^{\leq\delta} \times \mathcal{B}^\times$ are *small* (see Notation 1.6.6).

- Propositions 3.4.1 and 3.4.2: The fibers $\mathcal{M}_{i,a,b}$ and $\mathcal{N}_{a,b}$ can be written as products of local moduli spaces $\mathcal{M}^x_{i_x,a_x,b_x}$ and $\mathcal{N}^x_{a_x,b_x}$, defined in a similar way as $\mathcal{M}^{\mathrm{loc}}_{i,a,b}$ and $\mathcal{N}^{\mathrm{loc}}_{a,b}$.

The global part of the main theorem is as follows.

MAIN THEOREM (Global part)
Fix $\delta \geq 1$. For $\deg(D)$, $\deg(D_0)$ *large enough there is a natural isomorphism in $D^b_c(\mathcal{A}^{\leq\delta} \times \mathcal{B}^\times)$:*

$$\bigoplus_{i=-d}^{d} f^{\mathrm{int}}_{i,*} L_{d-i}|_{\mathcal{A}^{\leq\delta}\times\mathcal{B}^\times} \cong g^{\mathrm{int}}_{*}\overline{\mathbb{Q}}_\ell|_{\mathcal{A}^{\leq\delta}\times\mathcal{B}^\times}. \tag{1.4.1}$$

Here $d = n(n-1)\deg(D)/2 + n\deg(D_0)$, and L_{d-i} is a local system of rank one and order two on $\mathcal{M}^{\mathrm{int}}_i$ (see Section 4.1), which is a geometric analogue of the factor $\eta_{E/F}(\det(g))$ appearing in the orbital integral $\mathbf{O}^{\mathrm{GL}_{n-1},\eta}_A(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)})$.

The smallness of f^{int}_i and g^{int} over $\mathcal{A}^{\leq\delta} \times \mathcal{B}^\times$ and the smoothness of $\mathcal{M}^{\mathrm{int}}_i$ and $\mathcal{N}^{\mathrm{int}}$ are the key geometric properties that enable one to prove the above theorem. In fact, these two properties imply that both sides are middle extensions of some local system on some dense open subset of $\mathcal{A}^{\leq\delta} \times \mathcal{B}^\times$, and to verify (1.4.1), we only need to restrict to any open dense subset where both sides are explicitly computable.

Finally, we use the global part of the main theorem to prove the local part. We identify F with the local field associated to a k-point x_0 on the curve X. For local invariants (a^0, b^0) around x_0, we may approximate them by global invariants (a, b) such that the global moduli spaces $\mathcal{M}_{i,a,b}$ and $\mathcal{N}_{a,b}$, when expressed as a product of local moduli spaces, are very simple away from x_0. Taking the Frobenius traces of the two sides of (1.4.1) and using the product formulas, we get a formula of the form

$$\mathrm{Tr}(\mathrm{Frob}_k, M_{x_0}) \cdot \prod_{x\neq x_0} \mathrm{Tr}(\mathrm{Frob}_k, M_x) = \mathrm{Tr}(\mathrm{Frob}_k, N_{x_0}) \cdot \prod_{x\neq x_0} \mathrm{Tr}(\mathrm{Frob}_k, N_x),$$

where M_x and N_x are the cohomology groups of the relevant local moduli spaces. It is easy to show that $\mathrm{Tr}(\mathrm{Frob}_k, M_x) = \mathrm{Tr}(\mathrm{Frob}_k, N_x)$ for all $x \neq x_0$, but in order to conclude that $\mathrm{Tr}(\mathrm{Frob}_k, M_{x_0}) = \mathrm{Tr}(\mathrm{Frob}_k, N_{x_0})$, which is what we need, we have to make sure that the terms $\mathrm{Tr}(\mathrm{Frob}_k, N_x)$ are nonzero for all $x \neq x_0$. This is a technical difficulty of the article, and is responsible for the length of Section 5.2.

1.5. Plan of the article

In Section 2, we first fix notations and introduce the invariants (a, b). We then reformulate the problem into one of lattice counting in Sections 2.3 and 2.4. We verify a few simple cases of Conjecture 1.1.1 in Section 2.5, including Conjecture 1.1.1(3). We deduce the group version Conjecture 1.1.1(2) from the Lie algebra version Conjecture 1.1.1(1) in Section 2.6. Except for Section 2.7, we work with no restrictions on F.

In Section 3, we introduce the global moduli spaces $\mathcal{M}_i$ and $\mathcal{N}$, and we study their geometric properties.

In Section 4, we formulate and prove the global part of the main theorem. To this end, we need to study perverse sheaves on the symmetric powers of curves, especially the "binomial expansion" formula (Lemma 4.2.3), which demystifies the decomposition (1.4.1).

In Section 5, we deduce the local part of the main theorem from the global part.

1.6. Notations

1.6.1

Let $\mathcal{O}$ be a commutative ring. For a scheme X over $\mathcal{O}$ or an $\mathcal{O}$-module M, and for an $\mathcal{O}$-algebra R, we let $X(R)$ be the R-points of X and $M(R) := M \otimes_{\mathcal{O}} R$.

For a discrete valuation ring (DVR)$\mathcal{O}$ and two full-rank $\mathcal{O}$-lattices Λ_1, Λ_2 in some $\mathrm{Frac}(\mathcal{O})$-vector space V, we define the relative length $\mathrm{leng}_{\mathcal{O}}(\Lambda_1 : \Lambda_2)$ to be

$$\mathrm{leng}_{\mathcal{O}}(\Lambda_1 : \Lambda_2) := \mathrm{leng}_{\mathcal{O}}(\Lambda_1/\Lambda_1 \cap \Lambda_2) - \mathrm{leng}_{\mathcal{O}}(\Lambda_2/\Lambda_1 \cap \Lambda_2),$$

where $\mathrm{leng}_{\mathcal{O}}(-)$ denotes the usual length of a torsion $\mathcal{O}$-module.

1.6.2

Coherent sheaves are denoted by the calligraphic letters $\mathcal{E}, \mathcal{F}, \mathcal{L}, \ldots$; constructible $\overline{\mathbb{Q}}_\ell$-complexes are denoted by the capital letters $L, K, \ldots$.

1.6.3

From Section 3, we will work with a fixed smooth base curve X over a field k. For a morphism $p : Y \to X$ and for a closed point $x \in X$, we denote by $\mathcal{O}_{Y,x}$ the completed semilocal ring of $\mathcal{O}_Y$ along $p^{-1}(x)$.

If Y is a Gorenstein curve, let $\omega_{Y/X}$ be the relative dualizing sheaf

$$\omega_{Y/X} = \omega_{Y/k} \otimes p^*\omega_{X/k}^{-1}.$$

For a coherent sheaf $\mathcal{F}$ on Y, let $\mathcal{F}^\vee = \underline{\mathrm{Hom}}_Y(\mathcal{F}, \omega_{Y/X})$ be the (underived) Grothendieck-Serre dual. When we work over an extra parameter scheme S so that $p : Y \to X \times S$, then $^\vee$ means $\underline{\mathrm{Hom}}_Y(-, \omega_{Y/X\times S})$.

1.6.4

For an étale double cover (a finite étale map of degree 2) $\pi : X' \to X$ of a scheme X, we decompose the sheaf $\pi_*\overline{\mathbb{Q}}_\ell$ into ± 1-eigenspaces under the natural action of

$\mathbb{Z}/2 \cong \mathrm{Aut}(X'/X)$ to get

$$\pi_*\overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell \oplus L.$$

The rank one local system L satisfies $L^{\otimes 2} \cong \overline{\mathbb{Q}}_\ell$. We call L the local system *associated to* the étale double cover π.

1.6.5

We use the terminology *middle extension* in a nonstrict way. If K is a $\overline{\mathbb{Q}}_\ell$-complex on a scheme X, we say that K is a *middle extension on* X if, for some (and hence any) open dense subset $j : U \hookrightarrow X$ over which K is a local system placed at degree n, we have

$$K \cong j_{!*}(j^*K[n + \dim X])[-n - \dim X].$$

1.6.6

Recall from [7, Section 6.2] that a proper surjective morphism $f : Y \to X$ between *irreducible* schemes over an algebraically closed field Ω is called *small* if, for any $r \geq 1$, we have

$$\mathrm{codim}_X \left\{ x \in X \,\middle|\, \dim f^{-1}(x) \geq r \right\} \geq 2r + 1. \tag{1.6.1}$$

We will use this terminology in a loose way: we will not require Y to be irreducible (but we will require all other conditions). This is just for notational convenience. The main property that we will use about small morphisms is the following.

If Y/Ω is smooth and equidimensional, then $f_*\overline{\mathbb{Q}}_\ell$ is a middle extension on X.

1.6.7

For a finite field k, we denote by Frob_k the *geometric* Frobenius element in $\mathrm{Gal}(\overline{k}/k)$.

2. Local formulation

2.1. The setting

Let F be a non-Archimedean local field with valuation ring $\mathcal{O}_F$, uniformizing parameter ϖ, and residue field $k = \mathbb{F}_q$ such that q is odd. Let E be either the unramified quadratic extension of F (in which case we call E/F *nonsplit*) or $E = F \times F$ (in which case we call E/F *split*). Let $\mathcal{O}_E$ be the valuation ring of E, and let k' be its residue algebra. Let σ be the generator of $\mathrm{Gal}(E/F)$. Let $\eta_{E/F} : F^\times/\mathrm{Nm}E^\times \to \{\pm 1\}$ be the quadratic character associated to the extension E/F: this is trivial if and only if E/F is split. We decompose E and $\mathcal{O}_E$ according to eigenspaces of σ to get

$$E = F \oplus E^- \qquad \text{and} \qquad \mathcal{O}_E = \mathcal{O}_F \oplus \mathcal{O}_E^-,$$

where σ acts on E^- and $\mathcal{O}_E^-$ by -1.

Let $n \geq 2$ be an integer. We fix a free $\mathcal{O}_F$-module W of rank $(n-1)$. Let $V = W \oplus \mathcal{O}_F \cdot e_0$ be a free $\mathcal{O}_F$-module of rank n with a distinguished element e_0, and let $e_0^* : V \to \mathcal{O}_F$ be the projection along W such that $e_0^*(e_0) = 1$. Let $W^\vee = \mathrm{Hom}_{\mathcal{O}_F}(W, \mathcal{O}_F)$, and let $V^\vee = \mathrm{Hom}_{\mathcal{O}_F}(V, \mathcal{O}_F)$.

Let $\mathrm{GL}_{n-1} = \mathrm{GL}(W)$, and let $\mathrm{GL}_n = \mathrm{GL}(V)$ be the general linear groups over $\mathcal{O}_F$. We have the natural embedding $\mathrm{GL}_{n-1} \hookrightarrow \mathrm{GL}_n$ as block diagonal matrices:

$$A \mapsto \begin{pmatrix} A & \\ & 1 \end{pmatrix}.$$

Let $\mathfrak{gl}_n$ be the Lie algebra (over $\mathcal{O}_F$) of GL_n consisting of $\mathcal{O}_F$-linear operators on V. Let

$$\mathfrak{s}_n(\mathcal{O}_F) := \left\{ A \in \mathfrak{gl}_n(\mathcal{O}_E) \,\middle|\, A + \sigma(A) = 0 \right\}.$$

Then $\mathfrak{s}_n(F)$ is the subset of $\mathfrak{gl}_n(E)$ consisting of matrices with entries in E^-. The group GL_n acts on $\mathfrak{s}_n$ by conjugation.

Let $(\cdot,\cdot)$ be a Hermitian form on W with trivial discriminant, and extend this to a Hermitian form on V by requiring that $(W, e_0) = 0$ and that $(e_0, e_0) = 1$. These Hermitian forms define unitary groups $\mathrm{U}_{n-1} = \mathrm{U}(W, (\cdot,\cdot))$ and $\mathrm{U}_n = \mathrm{U}(V, (\cdot,\cdot))$ over $\mathcal{O}_F$. We also have the natural embedding $\mathrm{U}_{n-1} \hookrightarrow \mathrm{U}_n$ as block diagonal matrices:

$$A \mapsto \begin{pmatrix} A & \\ & 1 \end{pmatrix}.$$

Let $\mathfrak{u}_n$ be the Lie algebra (over $\mathcal{O}_F$) of U_n, that is,

$$\mathfrak{u}_n(\mathcal{O}_F) = \left\{ A \in \mathfrak{gl}_n(\mathcal{O}_E) \,\middle|\, A + A^\# = 0 \right\},$$

where $A^\#$ is the adjoint of A under the Hermitian form $(\cdot,\cdot)$. Then U_n acts on $\mathfrak{u}_n$ by conjugation.

2.2. The invariants

For any $A \in \mathfrak{gl}_n(E)$ and for $i = 1, \ldots, n$, let

$$a_i(A) := \mathrm{Tr}\left(\bigwedge^i A \right)$$

be the coefficients of the characteristic polynomial of A, that is,

$$\det(t \,\mathrm{id}_V - A) = t^n + \sum_{i=1}^n (-1)^i a_i(A) t^{n-i}.$$

For $i = 0, \dots, n-1$, let

$$b_i(A) = e_0^*(A^i e_0),$$

so that $b_0 = 1$. The $2n$-tuple $(a(A), b(A))$ of elements in E are called the *invariants* of $A \in \mathfrak{gl}_n(E)$. It is easy to check that these are invariants of $\mathfrak{gl}_n(E)$ under the conjugation action by $\mathrm{GL}_{n-1}(E)$.

Definition 2.2.1
An element $A \in \mathfrak{gl}_n(E)$ is said to be *strongly regular semisimple with respect to the* $\mathrm{GL}_{n-1}(E)$*-action* if
(1) A is regular semisimple as an element of $\mathfrak{gl}_n(E)$;
(2) the vectors $\{e_0, Ae_0, \dots, A^{n-1}e_0\}$ form an E-basis of $V(E)$;
(3) the vectors $\{e_0^*, e_0^*A, \dots, e_0^*A^{n-1}\}$ form an E-basis of $V^\vee(E)$.

Definition 2.2.2
For $A \in \mathfrak{gl}_n(E)$, define $v(A) \in \mathbb{Z}$ to be the F-valuation of the n-by-n matrix formed by the row vectors $\{e_0^*, e_0^*A, \dots, e_0^*A^{n-1}\}$ under an $\mathcal{O}_F$-basis of $V^\vee$.

For a collection of invariants $(a, b) \in E^{2n}$ (we allow general $b_0 \in E$, not just 1), we introduce a finite E-algebra:

$$R_a(E) := E[t]\big/\big(t^n - a_1 t^{n-1} + \cdots + (-1)^n a_n\big).$$

Let $R_a^\vee(E) := \mathrm{Hom}_E(R_a(E), E)$ be its linear dual, which is naturally an $R_a(E)$-module. The data of b gives the following element $b' \in R_a^\vee(E)$:

$$\begin{aligned} b' : R_a(E) &\to E \\ t^i &\mapsto b_i, \quad i = 0, 1, \dots, n-1, \end{aligned} \tag{2.2.1}$$

which induces an $R_a(E)$-linear homomorphism

$$\gamma'_{a,b} : R_a(E) \to R_a^\vee(E).$$

In other words, $\gamma'_{a,b}$ is given by the pairing

$$\begin{aligned} R_a(E) \otimes_E R_a(E) &\to E, \\ (x, y) &\mapsto b'(xy). \end{aligned} \tag{2.2.2}$$

Definition 2.2.3
The Δ-invariant $\Delta_{a,b}$ of the collection $(a, b) \in E^{2n}$ is the determinant of the map $\gamma'_{a,b}$ under the E-basis $\{1, t, \dots, t^{n-1}\}$ of $R_a(E)$ and the corresponding dual basis of $R_a^\vee(E)$.

It is easy to see that, for $A \in \mathfrak{gl}_n(E)$ with invariants (a, b), $\Delta_{a,b}$ is the determinant of the matrix $(e_0^* A^{i+j} e_0)_{0 \le i, j \le n-1}$.

LEMMA 2.2.4
Let $A \in \mathfrak{gl}_n(E)$ with invariants $(a, b) \in E^{2n}$. Then A is strongly regular semisimple if and only if
(1) *$R_a(E)$ is an étale algebra over E;*
(2) *$\Delta_{a,b} \neq 0$; that is, $\gamma'_{a,b} : R_a(E) \to R_a^\vee(E)$ is an isomorphism.*

Proof
Condition (1) is equivalent to the first condition in Definition 2.2.1. Condition (2) is equivalent to the condition that the matrix $(e_0^* A^{i+j} e_0)_{0 \le i, j \le n-1}$ is nondegenerate. Since this matrix is the product of the two matrices $(e_0^*, e_0^* A, \ldots, e_0^* A^{n-1})$ and $(e_0, Ae_0, \ldots, A^{n-1} e_0)$, the nondegeneracy of $(e_0^* A^{i+j} e_0)_{0 \le i, j \le n-1}$ is equivalent to the nondegeneracy of $(e_0^*, e_0^* A, \ldots, e_0^* A^{n-1})$ and $(e_0, Ae_0, \ldots, A^{n-1} e_0)$, which are the last two conditions of Definition 2.2.1. $\square$

Remark 2.2.5
From the above lemma, we see that the strong regular semisimplicity of $A \in \mathfrak{gl}_n(E)$ is in fact a property of its invariants (a, b). Therefore, we call a collection of invariants (a, b) *strongly regular semisimple* if it satisfies the conditions in Lemma 2.2.4.

If the invariants (a, b) are elements in $\mathcal{O}_E$, we get a canonical $\mathcal{O}_E$-form $R_a(\mathcal{O}_E)$ of $R_a(E)$ by setting

$$R_a(\mathcal{O}_E) := \mathcal{O}_E[t] / \big(t^n - a_1 t^{n-1} + \cdots + (-1)^n a_n\big). \tag{2.2.3}$$

Let $R_a^\vee(\mathcal{O}_E) = \mathrm{Hom}_{\mathcal{O}_E}(R_a(\mathcal{O}_E), \mathcal{O}_E)$ be its dual. Then $\gamma'_{a,b}$ restricts to an $R_a(\mathcal{O}_E)$-linear map:

$$\gamma'_{a,b} : R_a(\mathcal{O}_E) \to R_a^\vee(\mathcal{O}_E).$$

For $A \in \mathfrak{s}_n(F)$ or $\mathfrak{u}_n(F)$ with invariants (a, b), it is obvious that $a_i, b_i \in E^{\sigma = (-1)^i}$. Suppose furthermore that $a_i, b_i \in \mathcal{O}_E^{\sigma = (-1)^i}$. We extend the involution σ on $\mathcal{O}_E$ to an involution σ_R on $R_a(\mathcal{O}_E)$ by requiring that $\sigma_R(t) = -t$. The involution σ_R defines an $\mathcal{O}_F$-form of $R_a(\mathcal{O}_E)$:

$$R_a := R_a(\mathcal{O}_E)^{\sigma_R}.$$

Let $R_a^\vee := \mathrm{Hom}_{\mathcal{O}_F}(R_a, \mathcal{O}_F)$. The map $\gamma'_{a,b}$ restricts to an R_a-linear homomorphism

$$\gamma_{a,b} : R_a \to R_a^\vee, \tag{2.2.4}$$

and $\gamma'_{a,b} = \gamma_{a,b} \otimes_{\mathcal{O}_F} \mathcal{O}_E$. Since R_a is an $\mathcal{O}_F$-form of $R_a(\mathcal{O}_E)$ defined before, there is no confusion in using the notations $R_a(\mathcal{O}_E)$ or $R_a(E)$ (see Notation 1.6.1).

2.3. Orbital integrals for $\mathfrak{s}_n(F)$

Let $A \in \mathfrak{s}_n(F)$ be strongly regular semisimple with invariants (a, b). Let

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) := \int_{\mathrm{GL}_{n-1}(F)} \mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}(g^{-1}Ag)\eta_{E/F}\big(\det(g)\big)\, dg,$$

where dg is the Haar measure on $\mathrm{GL}_{n-1}(F)$ such that $\mathrm{vol}(\mathrm{GL}_{n-1}(\mathcal{O}_F), dg) = 1$, and where $\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}$ is the characteristic function of $\mathfrak{s}_n(\mathcal{O}_F) \subset \mathfrak{s}_n(F)$.

Remark 2.3.1

It is easy to see that if the orbital integral $\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) \neq 0$, then $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$.

Now we suppose that (a, b) is strongly regular semisimple and that $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$. We view R_a as a sublattice of $R_a^\vee$ via the map $\gamma_{a,b} : R_a \hookrightarrow R_a^\vee$. Then

$$\mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : R_a) = \mathrm{val}_F(\Delta_{a,b}).$$

For each integer $0 \le i \le \mathrm{val}_F(\Delta_{a,b})$, let

$$M_{i,a,b}^{\mathrm{loc}} := \big\{R_a\text{-lattices } \Lambda \,\big|\, R_a \subset \Lambda \subset R_a^\vee \text{ and } \mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : \Lambda) = i\big\}.$$

PROPOSITION 2.3.2

Let $A \in \mathfrak{s}_n(F)$ be strongly regular semisimple with invariants $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$ (recall that $b_0 = 1$). Then

$$\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) = \eta_{E/F}(\varpi)^{v(A)} \sum_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} \eta_{E/F}(\varpi)^i \# M_{i,a,b}^{\mathrm{loc}}.$$

Proof

By Definition 2.2.1(2), we have a (σ_R, σ)-equivariant E-linear isomorphism:

$$\iota' : R_a(E) \xrightarrow{\sim} V(E), \tag{2.3.1}$$

$$t^i \mapsto A^i e_0. \tag{2.3.2}$$

Therefore, ι' restricts to an F-linear isomorphism:

$$\iota : R_a(F) \xrightarrow{\sim} V(F).$$

We also identify $R_a^\vee(F)$ with $V(F)$ using ι and $\gamma_{a,b}$. We denote this identification also by ι. Define $R_{a,W} := R_a \cap \iota^{-1}(W(F))$ and $R_{a,W}^\vee := R_a^\vee \cap \iota^{-1}(W(F))$. We define some auxiliary sets

$$
\begin{aligned}
X_{i,A} &:= \left\{ g \in \mathrm{GL}_{n-1}(F)/\mathrm{GL}_{n-1}(\mathcal{O}_F) \,\middle|\, g^{-1}Ag \in \mathfrak{s}_n(\mathcal{O}_F),\ \mathrm{val}_F(\det(g)) = i \right\},\\
X'_{i,A} &:= \left\{ \mathcal{O}_F\text{-lattices } L \subset W(F) \,\middle|\, A(L) \subset \mathcal{O}_E^- L,\ \mathrm{leng}_{\mathcal{O}_F}(L : W) = i \right\},\\
M^W_{i,a,b} &:= \Big\{ \mathcal{O}_F\text{-lattices } \Lambda_W \subset R_{a,W}(F) \,\Big|\, \Lambda_W \oplus \mathcal{O}_F 1_R \subset R_a(F) \text{ is stable under } R_a,\\
&\qquad\qquad \mathrm{leng}_{\mathcal{O}_F}(R^\vee_{a,W} : \Lambda_W) = i \Big\},
\end{aligned}
$$

where $1_R \in R_a$ is the identity element.

Note that the group $\mathrm{GL}_{n-1}(F)$ acts transitively on the set of $\mathcal{O}_F$-lattices in $W(F)$ by left translation, and the stabilizer of W is equal to $\mathrm{GL}_{n-1}(\mathcal{O}_F)$. Therefore we get a bijection:

$$
\begin{aligned}
X_{i,A} &\xrightarrow{\sim} X'_{i,A},\\
g &\mapsto gW.
\end{aligned}
$$

We identify $R_a \overset{\gamma_{a,b}}{\hookrightarrow} R_a^\vee$ both as $\mathcal{O}_F$-lattices in $V(E)$ via ι. Observe that

$$
\mathrm{leng}_{\mathcal{O}_F}(R^\vee_{a,W} : W) = \mathrm{leng}_{\mathcal{O}_F}(R^\vee_a : V) = v(A).
$$

Therefore, we have a bijection:

$$
\begin{aligned}
M^W_{v(A)-i,a,b} &\xrightarrow{\sim} X'_{i,A},\\
\Lambda_W &\mapsto \iota(\Lambda_W).
\end{aligned}
$$

Finally, we have a bijection:

$$
\begin{aligned}
M^W_{i,a,b} &\xrightarrow{\sim} M^{\mathrm{loc}}_{i,a,b},\\
\Lambda_W &\mapsto \Lambda_W \oplus \mathcal{O}_F 1_R.
\end{aligned}
$$

We check that this is a bijection. On one hand, for $\Lambda_W \in M^W_{i,a,b}$, we have

$$
R_a \subset R_a(\Lambda_W \oplus \mathcal{O}_F 1_R) \subset \Lambda_W \oplus \mathcal{O}_F 1_R.
$$

We also have

$$
b'\big(R_a(\Lambda_W \oplus \mathcal{O}_F 1_R)\big) \subset b'(\Lambda_W \oplus \mathcal{O}_F 1_R) \subset \mathcal{O}_F.
$$

Therefore, $\Lambda_W \oplus \mathcal{O}_F 1_R \subset R^\vee_a$. This verifies $\Lambda_W \oplus \mathcal{O}_F 1_R \in M^{\mathrm{loc}}_{i,a,b}$.

On the other hand, we have to make sure that every $\Lambda \in M^{\mathrm{loc}}_{i,a,b}$ has the form $\Lambda = \Lambda_W \oplus \mathcal{O}_F 1_R$ for some lattice $\Lambda_W \subset R_{a,W}(F)$. But we can factorize the identity map on $\mathcal{O}_F$ as

$$\mathcal{O}_F \xrightarrow{1\mapsto 1_R} R_a \xrightarrow{\gamma_{a,b}} R_a^\vee \xrightarrow{\mathrm{ev}(1_R)} \mathcal{O}_F.$$

Therefore, for $R_a \subset \Lambda \subset R_a^\vee$, $\mathcal{O}_F 1_R$ is always a direct summand of Λ.

Now that we have set up a bijection between $X_{i,A}$ and $M^{\mathrm{loc}}_{v(A)-i,a,b}$, we have

$$\begin{aligned}\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}) &= \sum_i \eta_{E/F}(\varpi)^i \# X_{i,A}\\ &= \eta_{E/F}(\varpi)^{v(A)} \sum_i \eta_{E/F}(\varpi)^i \# M^{\mathrm{loc}}_{i,a,b}.\end{aligned}$$ □

2.4. Orbital integrals for $\mathfrak{u}_n(F)$

Let $A' \in \mathfrak{u}_n(F)$ be strongly regular semisimple with invariants (a, b). Let

$$\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}) := \int_{\mathrm{U}_{n-1}(F)} \mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}(g^{-1}Ag)\, dg,$$

where dg is the Haar measure on $\mathrm{U}_{n-1}(F)$ such that $\mathrm{vol}(\mathrm{U}_{n-1}(\mathcal{O}_F), dg) = 1$, and where $\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}$ is the characteristic function of $\mathfrak{u}_n(\mathcal{O}_F) \subset \mathfrak{u}_n(F)$.

Remark 2.4.1

It is easy to see that if the orbital integral $\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n}(\mathcal{O}_F)) \neq 0$, then $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$.

Now we suppose that (a, b) is strongly regular semisimple and that $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$. We identify $R_a(\mathcal{O}_E)$ as a sublattice of $R_a^\vee(\mathcal{O}_E)$ via $\gamma'_{a,b}$. Recall from (2.3.1) that we have an isomorphism $\iota' : R_a(E) \xrightarrow{\sim} V(E)$. The transport of the Hermitian form $(\cdot, \cdot)$ to $R_a(E)$ via ι' is given by

$$(x, y)_R = b'\big(x\sigma_R(y)\big), \tag{2.4.1}$$

where $b' : R_a(E) \to E$ is defined in (2.2.1).

Remark 2.4.2

Since the Hermitian form $(\cdot, \cdot)$ has trivial discriminant, so does $(\cdot, \cdot)_R$. Therefore, if $A' \in \mathfrak{u}_n(F)$ has invariants (a, b), then $\eta_{E/F}(\Delta_{a,b}) = 1$.

Recall that for an $\mathcal{O}_E$-lattice $\Lambda' \subset R_a(E)$, the *dual lattice under the Hermitian form* $(\cdot, \cdot)_R$ is the $\mathcal{O}_E$-lattice

$$\Lambda'^{\perp} := \big\{x \in R_a(E) \,\big|\, (x, \Lambda')_R \subset \mathcal{O}_E\big\}.$$

Such a lattice is called *self-dual* (under the given Hermitian form) if $\Lambda'^{\perp} = \Lambda'$. By the pairing (2.2.2), it is easy to see that $R_a^{\vee}(\mathcal{O}_E)$ is the dual of $R_a(\mathcal{O}_E)$ under the Hermitian form $(\cdot,\cdot)_R$. We define

$$N_{a,b}^{\mathrm{loc}} := \left\{\text{self-dual } R_a(\mathcal{O}_E)\text{-lattices } \Lambda' \,\middle|\, R_a(\mathcal{O}_E) \subset \Lambda' \subset R_a^{\vee}(\mathcal{O}_E)\right\}.$$

PROPOSITION 2.4.3
Let $A' \in \mathfrak{u}_n(F)$ be strongly regular semisimple with invariants $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$ (recall that $b_0 = 1$). Then

$$\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}) = \# N_{a,b}^{\mathrm{loc}}.$$

Proof
The argument is similar to the proof of Proposition 2.3.2. Let

$$Y_{A'} = \left\{g \in \mathrm{U}_{n-1}(F)/\mathrm{U}_{n-1}(\mathcal{O}_F) \,\middle|\, g^{-1}A'g \in \mathfrak{u}_n(\mathcal{O}_F)\right\}.$$

Then $\mathrm{U}_{n-1}(F)$ acts transitively on the set of self-dual $\mathcal{O}_E$-lattices in $W(E)$ such that the stabilizer of $W(\mathcal{O}_E)$ is $\mathrm{U}_{n-1}(\mathcal{O}_F)$. Therefore, we get a bijection:

$$\begin{aligned} Y_{A'} &\xrightarrow{\sim} N_{a,b}^{\mathrm{loc}}, \\ g &\mapsto \iota'^{-1}\big(gW(\mathcal{O}_E)\big) \oplus \mathcal{O}_E 1_R. \end{aligned}$$

Hence

$$\mathbf{O}_{A'}^{\mathrm{U}_{n-1}}(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}) = \# Y_{A'} = \# N_{a,b}^{\mathrm{loc}}. \qquad \square$$

2.5. The fundamental lemma and simple cases

Definition 2.5.1
Two strongly regular semisimple elements $A \in \mathfrak{s}_n(F)$ and $A' \in \mathfrak{u}_n(F)$ are said to *match each other* if they have the same invariants.

Now we have explained all the notions appearing in Conjecture 1.1.1. By Propositions 2.3.2 and 2.4.3, parts (1) and (3) of Conjecture 1.1.1 are implied by the following.

CONJECTURE 2.5.2
For any strongly regular semisimple collection of invariants (a, b) such that $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$ (in particular, we allow arbitrary $b_0 \in \mathcal{O}_F$), we have

$$\sum_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} \eta_{E/F}(\varpi)^i \# M_{i,a,b}^{\mathrm{loc}} = \# N_{a,b}^{\mathrm{loc}}. \tag{2.5.1}$$

In the rest of this section, we prove some easy cases of Conjecture 2.5.2 by straightforward counting arguments.

LEMMA 2.5.3
Conjecture 2.5.2 is true if $\eta_{E/F}(\Delta_{a,b}) \neq 1$, in which case both sides of (2.5.1) are zero. In particular, Conjecture 1.1.1(3) holds.

Proof
The situation $\eta_{E/F}(\Delta_{a,b}) \neq 1$ happens if and only if E/F is nonsplit and $\mathrm{val}_F(\Delta_{a,b}) = \mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : R_a)$ is odd. In this case, $\mathrm{leng}_{\mathcal{O}_E}(R_a^\vee(\mathcal{O}_E) : R_a(\mathcal{O}_E))$ is odd, and therefore there are no self-dual lattices in $R_a(E)$; that is, $N_{a,b}^{\mathrm{loc}} = \varnothing$.

Now we show that the left-hand side of (2.5.1) is also zero. For an $\mathcal{O}_F$-lattice $\Lambda \in R_a(F)$, the linear dual $\Lambda^\vee = \mathrm{Hom}_{\mathcal{O}_F}(\Lambda, \mathcal{O}_F)$ can be naturally viewed as another $\mathcal{O}_F$-lattice of $R_a(F)$ via the identification $\gamma_{a,b} : R_a(F) \xrightarrow{\sim} R_a^\vee(F)$. It is easy to check that if Λ is stable under multiplication by R_a, then the same is true for $\Lambda^\vee$. This operation sets up a bijection:

$$(-)^\vee : M_{i,a,b}^{\mathrm{loc}} \xrightarrow{\sim} M_{\mathrm{val}_F(\Delta_{a,b})-i,a,b}^{\mathrm{loc}}.$$

Since E/F is nonsplit and $\mathrm{val}_F(\Delta_{a,b})$ is odd, we have

$$\sum_i (-1)^i \# M_{i,a,b}^{\mathrm{loc}} = \sum_{i=0}^{\lfloor \mathrm{val}_F(\Delta_{a,b})/2 \rfloor} (-1)^i (\# M_{i,a,b}^{\mathrm{loc}} - \# M_{\mathrm{val}_F(\Delta_{a,b})-i,a,b}^{\mathrm{loc}}) = 0.$$

Therefore, in this case, both sides of (2.5.1) are zero.

In [9], Jacquet and Rallis showed that every strongly regular semisimple $A' \in \mathfrak{u}_n(F)$ matches some $A \in \mathfrak{s}_n(F)$. Conversely, a strongly regular semisimple $A \in \mathfrak{s}_n(F)$ matches some $A' \in \mathfrak{u}_n(F)$ if and only if $\eta_{E/F}(\Delta_{a,b}) = 1$, where (a, b) are the invariants of A. We have seen from Lemma 2.5.3 that if $A \in \mathfrak{s}_n(F)$ does not match any element in $\mathfrak{u}_n(F)$, then $\mathbf{O}_A^{\mathrm{GL}_{n-1},\eta}(\mathbf{1}_{\mathfrak{s}_n}(\mathcal{O}_F)) = 0$.

Therefore, Conjecture 1.1.1(3) holds. □

LEMMA 2.5.4
Conjecture 2.5.2 is true if E/F is split.

Proof
Fix an isomorphism $E \cong F \oplus F$ such that σ interchanges the two factors. Using this, we can identify $R_a(E)$ with $R_a(F) \oplus R_a(F)$, and the Hermitian form $(\cdot, \cdot)_R$ takes the

form

$$(x \oplus y, x' \oplus y')_R = b'(xy') \oplus b'(x'y) \in E, \quad x, y, x', y' \in R_a(F).$$

Note that b' restricted to $R_a(F)$ takes values in F. Therefore, each $R_a(\mathcal{O}_E)$-lattice $R_a(\mathcal{O}_E) \subset \Lambda' \subset R_a^\vee(\mathcal{O}_E)$ has the form $\Lambda' = \Lambda_1 \oplus \Lambda_2$, with $R_a \subset \Lambda_i \subset R_a^\vee$. The self-duality requirement is equivalent to $\Lambda_2 = \gamma_{a,b}^{-1}(\Lambda_1^\vee)$ (note that $\Lambda_1^\vee \subset R_a^\vee(F)$, and recall the isomorphism $\gamma_{a,b} : R_a(F) \xrightarrow{\sim} R_a^\vee(F)$). In this way, we get a bijection:

$$\coprod_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} M_{i,a,b}^{\mathrm{loc}} \xrightarrow{\sim} N_{a,b}^{\mathrm{loc}},$$

$$\Lambda \mapsto \Lambda \oplus \gamma_{a,b}^{-1}(\Lambda^\vee).$$

Therefore,

$$\sum_i \# M_{i,a,b}^{\mathrm{loc}} = \# N_{a,b}^{\mathrm{loc}},$$

which verifies Conjecture 2.5.2 in the split case because $\eta_{E/F}$ is trivial. □

LEMMA 2.5.5
Conjecture 2.5.2 is true if R_a is a product of DVRs.

Proof
By reducing to the lattice-counting problems, it suffices to deal with the case that R_a is a DVR. Since we already dealt with the split case, we may assume that E/F is nonsplit. Let $k(R_a)$ be the residue field of R_a, and let ϖ_R be a uniformizing parameter of R_a. Then $R_a^\vee = \varpi_R^{-d} R_a$ for some integer $d = \mathrm{leng}_{R_a}(R_a^\vee : R_a)$. We have

$$\mathrm{val}_F(\Delta_{a,b}) = \mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : R_a) = d[k(R_a) : k]. \tag{2.5.2}$$

We already solved the case when $\mathrm{val}_F(\Delta_{a,b})$ is odd in Lemma 2.5.3 and when E/F is split in Lemma 2.5.4. Now, suppose that $\mathrm{val}_F(\Delta_{a,b})$ is even and that E/F is nonsplit. In this case, either d or $[k(R_a) : k]$ has to be even. Let us explicitly count the cardinalities of $M_{i,a,b}^{\mathrm{loc}}$ and $N_{a,b}^{\mathrm{loc}}$.

On one hand, the only R_a-lattices that sit between R_a and $R_a^\vee$ are $\varpi_R^{-j} R_a$ for $0 \le j \le d$ and $\mathrm{leng}_{\mathcal{O}_F}(R_a^\vee : \varpi_R^{-j} R_a) = (d-j)[k(R_a) : k]$. Therefore, we have

$$\sum_i (-1)^i \# M_{i,a,b}^{\mathrm{loc}} = \sum_{j=0}^{d} (-1)^{(d-j)[k(R_a):k]} = \begin{cases} d+1 & [k(R_a) : k] \text{ even}, \\ 1 & [k(R_a) : k] \text{ odd}. \end{cases}$$

On the other hand, if $[k(R_a):k]$ is odd and if d is even, then $R_a(\mathcal{O}_E)$ remains a DVR. Therefore, $\varpi_R^{-d/2}R_a(\mathcal{O}_E)$ is the unique self-dual $R_a(\mathcal{O}_E)$-lattice between $R_a(\mathcal{O}_E)$ and $R_a^\vee(\mathcal{O}_E)=\varpi_R^{-d}R_a(\mathcal{O}_E)$. If $[k(R_a):k]$ is even, then we can identify $R_a(\mathcal{O}_E)\cong R_a\oplus R_a$ such that σ_R acts by interchanging the two factors. In this case, $N^{\mathrm{loc}}_{a,b}$ consists of lattices $\varpi_R^{-j}R_a\oplus\varpi_R^{-d+j}R_a$ for $0\le j\le d$. In any case, we have

$$\sum_i(-1)^i\#M^{\mathrm{loc}}_{i,a,b}=\#N^{\mathrm{loc}}_{a,b}. \qquad \square$$

2.6. From the Lie algebra version to the group version

As mentioned in the Introduction, it is the group version identity (1.1.2) which is directly relevant to Jacquet and Rallis's approach to the Gross-Prasad conjecture for the unitary groups. In this section, we deduce the group version (1.1.2) from the Lie algebra version (1.1.1). The same argument also shows that (1.1.4) follows from (1.1.3).

For an element $A\in\mathrm{GL}_n(E)$, viewed as an element in $\mathfrak{gl}_n(E)$, the invariants $a_i(A)$, $b_i(A)$ and $v(A)$ are defined as in Section 2.2. When $a_i,b_i\in\mathcal{O}_E$, we introduce the $\mathcal{O}_E$-algebra

$$\mathbf{R}_a(\mathcal{O}_E)=\mathcal{O}_E[t,t^{-1}]\big/\big(t^n-a_1t^{n-1}+\cdots+(-1)^na_n\big). \tag{2.6.1}$$

We view $Z'_a=\operatorname{Spec}\mathbf{R}_a(\mathcal{O}_E)$ as a subscheme of $\operatorname{Spec}\mathcal{O}_E\times\mathbb{G}_m$ which is finite flat over $\operatorname{Spec}\mathcal{O}_E$ of degree n. Let θ be the involution on $\operatorname{Spec}\mathcal{O}_E\times\mathbb{G}_m$, which is the product of σ on $\mathcal{O}_E$ and $t\mapsto t^{-1}$ on $\mathbb{G}_m$. The fixed point subscheme under θ is the unitary group $\mathrm{U}_{\mathcal{O}_E/\mathcal{O}_F}(1)$ over $\operatorname{Spec}\mathcal{O}_F$.

Recall that $\mathrm{S}_n(\mathcal{O}_F)=\{A\in\mathrm{GL}_n(\mathcal{O}_E)\,|\,A\sigma(A)=1\}$. For an element A in either $\mathrm{S}_n(\mathcal{O}_F)$ or $\mathrm{U}_n(\mathcal{O}_F)$, the subscheme Z'_a is stable under θ, hence determining a subscheme $Z_a\subset\mathrm{U}_{\mathcal{O}_E/\mathcal{O}_F}(1)$, finite flat of degree n over $\operatorname{Spec}\mathcal{O}_F$. Let $\mathbf{R}_a$ be the coordinate ring of Z_a, which is a finite flat $\mathcal{O}_F$-algebra of rank n satisfying $\mathbf{R}_a\otimes_{\mathcal{O}_F}\mathcal{O}_E=\mathbf{R}_a(\mathcal{O}_E)$. The invariants b_i determine an $\mathbf{R}_a$-linear map $\gamma_{a,b}:\mathbf{R}_a\to\mathbf{R}_a^\vee$, as in (2.2.4).

PROPOSITION 2.6.1

Conjecture 1.1.1(2) follows from Conjecture 1.1.1(1).

Proof

Using the same argument as in Propositions 2.3.2 and 2.4.3, we reduce the orbital integrals in (1.1.2) to a counting of points in the corresponding sets $\mathbf{M}^{\mathrm{loc}}_{i,a,b}$ and $\mathbf{N}^{\mathrm{loc}}_{a,b}$, defined using $\mathbf{R}_a$ instead of R_a. Therefore, it suffices to find $\widetilde{a}_i,\widetilde{b}_i\in\mathcal{O}_E^{\sigma=(-1)^i}$ and an isomorphism of $\mathcal{O}_F$-algebras $\rho:R_{\widetilde{a}}\xrightarrow{\sim}\mathbf{R}_a$ such that the following diagram is commutative:

$$\begin{tikzcd} R_{\widetilde{a}} \arrow[r, "\gamma_{\widetilde{a},\widetilde{b}}"] \arrow[d, "\rho", "\wr"'] & R_{\widetilde{a}}^{\vee} \\ \mathbf{R}_a \arrow[r, "\gamma_{a,b}"] & \mathbf{R}_a^{\vee}. \arrow[u, "\rho^{\vee}"', "\wr"] \end{tikzcd} \tag{2.6.2}$$

Moreover, once we find $\rho : R_{\widetilde{a}} \xrightarrow{\sim} \mathbf{R}_a$, the choice of $\widetilde{b}$ is uniquely determined by the diagram (2.6.2), because the data of $\widetilde{b}$ and $\gamma_{\widetilde{a},\widetilde{b}}$ determine each other, as seen in (2.2.1) and (2.2.2). Therefore, we only need to find $\widetilde{a}$ such that $\mathbf{R}_a$ is isomorphic to $R_{\widetilde{a}}$.

Consider the special fiber of the $\mathcal{O}_F$-scheme, which is a finite subscheme of $\mathrm{U}_{k'/k}(1)$ of degree n. Since $\mathrm{U}_{k'/k}(1)$ is a smooth curve, any subscheme of it can be embedded into $\mathbb{A}^1_k$. In other words, there is a surjection of algebras $k[s] \twoheadrightarrow \mathbf{R}_a \otimes_{\mathcal{O}_F} k$. Lifting the image of the generator s to an element of $\mathbf{R}_a$, we get a surjection of $\mathcal{O}_F$-algebras $\mathcal{O}_F[s] \twoheadrightarrow \mathbf{R}_a$ (surjectivity follows from Nakayama's lemma). In other words, Z_a can be embedded as a subscheme of $\operatorname{Spec} \mathcal{O}_F \times \mathbb{A}^1$. It is well known that any such finite flat $\mathcal{O}_F$-subscheme of $\operatorname{Spec} \mathcal{O}_F \times \mathbb{A}^1$ is defined by one equation of the form $t^n - c_1 t^{n-1} + \cdots + (-1)^n c_n$ for some $c_i \in \mathcal{O}_F$; that is, there is an isomorphism:

$$R_c^F := \mathcal{O}_F[s]\big/\big(t^n - c_1 t^{n-1} + \cdots + (-1)^n c_n\big) \xrightarrow{\sim} \mathbf{R}_a. \tag{2.6.3}$$

Let $\jmath \in \mathcal{O}_E^- \cap \mathcal{O}_E^\times$ be a purely imaginary unit element, and let $\widetilde{a}_i = \jmath^i c_i \in \mathcal{O}_E^{\sigma=(-1)^i}$. Then we have an isomorphism $R_{\widetilde{a}} \xrightarrow{\sim} R_c^F$ by sending $t \mapsto \jmath^{-1} s$. Using (2.6.3), we get the desired isomorphism of $\mathcal{O}_F$-algebras $\rho : R_{\widetilde{a}} \xrightarrow{\sim} \mathbf{R}_a$. This completes the proof. $\square$

2.7. Geometric reformulation

In this section, we assume that $\operatorname{char}(F) = \operatorname{char}(k)$. In this case, we can interpret the sets $M_{i,a,b}^{\mathrm{loc}}$ and $N_{a,b}^{\mathrm{loc}}$ as k-points of certain schemes.

Fix a strongly regular semisimple pair (a, b) such that $a_i, b_i \in \mathcal{O}_E^{\sigma=(-1)^i}$ (we allow any $b_0 \in \mathcal{O}_F$). For $0 \leq i \leq \operatorname{val}_F(\Delta_{a,b})$, consider the following functor

$$S \mapsto \left\{ \begin{matrix} R_a \otimes_k \mathcal{O}_S\text{–} \\ \text{modules } \Lambda_x \end{matrix} \,\middle|\, \begin{matrix} R_a \otimes_k \mathcal{O}_S \subset \Lambda \subset R_a^\vee \otimes_k \mathcal{O}_S, \\ R_a^\vee \otimes_k \mathcal{O}_S/\Lambda \text{ is a vector bundle of rank } i \text{ over } S \end{matrix} \right\}.$$

It is clear that this functor is represented by a projective scheme $\mathcal{M}_{i,a,b}^{\mathrm{loc}}$ over k, and we have

$$M_{i,a,b}^{\mathrm{loc}} = \mathcal{M}_{i,a,b}^{\mathrm{loc}}(k).$$

Similarly, we have a projective scheme $\mathcal{N}_{a,b}^{\mathrm{loc}}$ over k representing the functor

$$S \mapsto \left\{ \begin{matrix} R_a(\mathcal{O}_E)\otimes_k \\ \mathcal{O}_S\text{-modules } \Lambda' \end{matrix} \left| \begin{matrix} R_a(\mathcal{O}_E) \otimes_k \mathcal{O}_S \subset \Lambda' \subset R_a^\vee(\mathcal{O}_E) \otimes_k \mathcal{O}_S, \\ R_a^\vee(\mathcal{O}_E) \otimes_k \mathcal{O}_S/\Lambda' \text{ is a vector bundle over } \mathcal{O}_S, \\ \Lambda' \text{ is self-dual under the Hermitian form } (\cdot,\cdot)_R \end{matrix} \right. \right\}.$$

We also have

$$N_{a,b}^{\mathrm{loc}} = \mathcal{N}_{a,b}^{\mathrm{loc}}(k).$$

Let ℓ be a prime number different from char(k). Let $\overline{\mathbb{Q}}_\ell(\eta_{k'/k})$ be the rank one $\overline{\mathbb{Q}}_\ell$-local system on Spec k associated to the extension k'/k: it is trivial if E/F is split and has order two otherwise.

The local part of the main theorem of the article is the following.

THEOREM 2.7.1
Suppose that $\mathrm{char}(F) = \mathrm{char}(k) > n$ *and that* $\eta_{E/F}(\Delta_{a,b}) = 1$. *Then there is an isomorphism of graded* Frob_k*-modules:*

$$\bigoplus_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} H^*\big(\mathcal{M}_{i,a,b}^{\mathrm{loc}} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell(\eta_{k'/k})^{\otimes i}\big) \cong H^*(\mathcal{N}_{a,b}^{\mathrm{loc}} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell). \tag{2.7.1}$$

Taking Frobenius traces of the isomorphism (2.7.1), we get the following.

COROLLARY 2.7.2
Conjecture 2.5.2, hence Conjecture 1.1.1, is true if $\mathrm{char}(F) = \mathrm{char}(k) > n$.

The proof of Theorem 2.7.1 will be completed in Section 5. At this point, we observe that the k-schemes $\coprod \mathcal{M}_{i,a,b}^{\mathrm{loc}}$ and $\mathcal{N}_{a,b}^{\mathrm{loc}}$ are geometrically isomorphic. More precisely, we have the following.

LEMMA 2.7.3
Let $\Omega = k$ *if* k'/k *is split, or let* $\Omega = k'$ *if* k'/k *is nonsplit, and we have an isomorphism of schemes over* Ω*:*

$$\coprod_{i=0}^{\mathrm{val}_F(\Delta_{a,b})} \mathcal{M}_{i,a,b}^{\mathrm{loc}} \otimes_k \Omega \xrightarrow{\sim} \mathcal{N}_{a,b}^{\mathrm{loc}} \otimes_k \Omega. \tag{2.7.2}$$

Proof
After base change to Ω, we may assume that E/F is split. Then the argument is the same as the proof of Lemma 2.5.4, once we fix an identification $R_a(E) \cong R_a(F) \oplus R_a(F)$. □

3. Global formulation: The moduli spaces

Let $k = \mathbb{F}_q$ be a finite field with $\mathrm{char}(k) > n$. Let X be a smooth, projective, and geometrically connected curve over k of genus g. Let $\pi : X' \to X$ be an étale double cover such that X'/k is also geometrically connected. Let σ denote the nontrivial involution of X' over X. We have a canonical decomposition,

$$\pi_* \mathcal{O}_{X'} = \mathcal{O}_X \oplus \mathcal{L}$$

into ± 1-eigenspaces of σ. Here $\mathcal{L}$ is a line bundle on X such that $\mathcal{L}^{\otimes 2} \cong \mathcal{O}_X$.

Let D and D_0 be effective divisors on X. Assume that $\deg(D) \geq 2g - 1$.

3.1. The moduli spaces associated to $\mathfrak{s}_n$

Consider the functor $\underline{\mathcal{M}} : \mathrm{Sch}/k \to \mathrm{Grpd}$

$$S \mapsto \left\{ (\mathcal{E}, \phi, \lambda, \mu) \,\middle|\, \begin{array}{l} \mathcal{E} \text{ is a vector bundle of rank } n \text{ over } X \times S, \\ \phi : \mathcal{E} \to \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{L}(D), \\ \mathcal{O}_{X \times S}(-D_0) \xrightarrow{\lambda} \mathcal{E} \xrightarrow{\mu} \mathcal{O}_{X \times S}(D_0) \end{array} \right\}.$$

Here, the twisting by (D) or (D_0) means tensoring with the pullback of the line bundles $\mathcal{O}_X(D)$ or $\mathcal{O}_X(D_0)$ to $X \times S$. For each integer i, we define the subfunctor $\underline{\mathcal{M}_i}$ of $\underline{\mathcal{M}}$ by taking only those vector bundles $\mathcal{E}$ such that

$$\chi\big(X \otimes_k k(s), \mathcal{E} \otimes_k k(s)\big) = i - n(g-1)$$

for any geometric point s of S. It is clear that $\underline{\mathcal{M}_i}$ is represented by an algebraic stack $\mathcal{M}_i$ over k locally of finite type; $\underline{\mathcal{M}}$ is represented by $\mathcal{M} = \coprod_i \mathcal{M}_i$.

Let $\mathcal{M}_i^{\mathrm{Hit}}$ be the *Hitchin moduli stack* for GL_n (with the choice of the line bundle $\mathcal{L}(D)$ on X) which classifies only the pairs $(\mathcal{E}, \phi)$ as above. (For more details about this Hitchin stack, we refer the readers to [12, section 4.4]).

Let

$$\mathcal{A} := \bigoplus_{i=1}^{n} H^0\big(X, \mathcal{L}(D)^{\otimes i}\big),$$

$$\mathcal{B} := \bigoplus_{i=0}^{n-1} H^0\big(X, \mathcal{O}_X(2D_0) \otimes \mathcal{L}(D)^{\otimes i}\big)$$

be viewed as affine spaces over k. We have a natural morphism,

$$f_i : \mathcal{M}_i \to \mathcal{A} \times \mathcal{B},$$

which, on the level of S-points, sends $(\mathcal{E}, \phi, \lambda, \mu)$ to $a = (a_1, \dots, a_n) \in \mathcal{A}$ and $b = (b_0, \dots, b_{n-1}) \in \mathcal{B}$, where

$$a_i = \operatorname{Tr}\Big(\bigwedge^i \phi\Big) \in H^0\big(X \times S, \mathcal{L}(D)^{\otimes i}\big)$$

and where $b_i \in H^0(X \times S, \mathcal{O}_{X\times S}(2D_0) \otimes \mathcal{L}(D)^{\otimes i}) = \operatorname{Hom}_{X\times S}(\mathcal{O}_{X\times S}(-D_0), \mathcal{O}_{X\times S}(D_0) \otimes \mathcal{L}(D)^{\otimes i})$ is represented by the following homomorphism:

$$\mathcal{O}_{X\times S}(-D_0) \xrightarrow{\lambda} \mathcal{E} \xrightarrow{\phi^i} \mathcal{E} \otimes \mathcal{L}(D)^{\otimes i} \xrightarrow{\mu} \mathcal{O}_{X\times S}(D_0) \otimes \mathcal{L}(D)^{\otimes i}.$$

3.2. The spectral curves

Following [10, Section 2.5], we define the universal spectral curve $p : Y \to \mathcal{A} \times X$ as follows. For each S-point $a = (a_1, \dots, a_n) \in \mathcal{A}(S)$, define the following scheme, affine over $X' \times S$, as

$$Y'_a := \underline{\operatorname{Spec}}_{X'\times S}\Big(\bigoplus_{i=0}^{n-1} \mathcal{O}_{X'\times S}(-iD)t^i\Big),$$

where the ring structure on the right-hand side is defined by the relation

$$t^n - a_1 t^{n-1} + a_2 t^{n-2} - \cdots + (-1)^n a_n = 0.$$

Let $p'_a : Y'_a \to X' \times S$ be the natural projection. This is a finite flat morphism of degree n. The scheme Y'_a over $X' \times S$ naturally embeds into the total space $\operatorname{Tot}_{X'\times S}(\mathcal{O}(D))$ of the line bundle $\mathcal{O}_{X'\times S}(D)$ over $X' \times S$. The free involution σ on X' extends to a free involution on Y'_a by requiring that $\sigma(t) = -t$. The quotient of Y'_a by σ is the scheme

$$Y_a := \underline{\operatorname{Spec}}_{X\times S}\mathcal{O}^{\sigma}_{Y'_a} = \underline{\operatorname{Spec}}_{X\times S}\Big(\bigoplus_{i=0}^{n-1} \mathcal{L}(-D)^{\otimes i} \boxtimes \mathcal{O}_S t^i\Big).$$

Let $p_a : Y \to X \times S$ be the natural projection. This is a finite flat morphism of degree n. The scheme Y_a naturally embeds into the total space $\operatorname{Tot}_{X\times S}(\mathcal{L}(D))$ of the line bundle $\mathcal{L}(D)$ over $X \times S$. The quotient map $\pi_a : Y'_a \to Y_a$ is an étale double cover.

Let $\mathcal{A}^{\mathrm{int}}$ (resp., $\mathcal{A}^{\heartsuit}$; resp., $\mathcal{A}^{\mathrm{sm}}$) be the open subset of $\mathcal{A}$ consisting of those geometric points a such that Y'_a, and hence Y_a, are integral (resp., reduced; resp., smooth and irreducible). Let $\mathcal{B}^{\times} = \mathcal{B} - \{0\}$. Let $\mathcal{M}_i^{\mathrm{int}}$ be the restriction of $\mathcal{M}_i$ to $\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}$.

LEMMA 3.2.1
The codimension of $\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{int}}$ in $\mathcal{A}^{\heartsuit}$ is at least $\deg(D)$.

Proof
In our situation, $\mathcal{A}$ serves as the Hitchin base for GL_n and U_n at the same time (the unitary Hitchin stack will be recalled in Section 3.3). The locus $\mathcal{A}^{\mathrm{int}}$ is in fact the intersection of two elliptic loci: $\mathcal{A}^{\mathrm{int}} = \mathcal{A}^{\mathrm{ell}}_{\mathrm{GL}_n} \cap \mathcal{A}^{\mathrm{ell}}_{\mathrm{U}_n}$. Here, $\mathcal{A}^{\mathrm{ell}}_{\mathrm{GL}_n}$ is the locus where Y_a is irreducible; $\mathcal{A}^{\mathrm{ell}}_{\mathrm{U}_n}$ is the locus where the set of irreducible components of Y'_a is in bijection with that of Y_a see [10, Section 2.8]. In the U_n case, the elliptic locus $\mathcal{A}^{\mathrm{ell}}_{\mathrm{U}_n}$ is the same as the anisotropic locus considered in [12, section 4.10.5], and by [12, proposition 6.3.6], we have

$$\mathrm{codim}_{\mathcal{A}^{\heartsuit}}(\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{ell}}_{\mathrm{U}_n}) \geq \deg(D).$$

In the GL_n case, the same argument also works to prove that

$$\mathrm{codim}_{\mathcal{A}^{\heartsuit}}(\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{ell}}_{\mathrm{GL}_n}) \geq \deg(D).$$

In fact, we only need to compute the dimension of the Hitchin bases for the Levi subgroups $\mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_r}$.

Therefore,

$$\dim(\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{int}}) = \dim\big((\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{ell}}_{\mathrm{U}_n}) \cup (\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{ell}}_{\mathrm{GL}_n})\big) \leq \dim \mathcal{A}^{\heartsuit} - \deg(D). \quad \square$$

The following lemma is a direct calculation.

LEMMA 3.2.2
For a geometric point $a \in \mathcal{A}^{\mathrm{int}}$, the arithmetic genera of the curves Y'_a and Y_a are

$$\begin{aligned} g'_Y &:= 1 - \chi(Y'_a, \mathcal{O}_{Y'_a}) = n(n-1)\deg(D) + (2g-2)n + 1, \\ g_Y &:= 1 - \chi(Y_a, \mathcal{O}_{Y_a}) = n(n-1)\deg(D)/2 + (g-1)n + 1. \end{aligned}$$

Recall that for a locally projective flat family of geometrically integral curves C over S, we have the compactified Picard stack $\overline{\mathcal{P}\mathrm{ic}}(C/S) = \coprod_i \overline{\mathcal{P}\mathrm{ic}}^i(C/S)$ over S (see [1]) whose fiber over a geometric point $s \in S$ classifies the groupoid of torsion-free coherent sheaves $\mathcal{F}$ of generic rank one on C_s such that $\chi(C_s, \mathcal{F}) = i$. Each $\overline{\mathcal{P}\mathrm{ic}}^i(C/S)$ is an algebraic stack of finite type over S; it is in fact a $\mathbb{G}_m$-gerb over the compactified Picard scheme of $\overline{\mathrm{Pic}}^i(C/S)$. The scheme $\overline{\mathrm{Pic}}^i(C/S)$ is proper over S and contains the usual Picard scheme $\mathrm{Pic}^i(C/S)$ as an open substack.

For each $a \in \mathcal{A}^{\mathrm{int}}(S)$ and $\mathcal{F} \in \overline{\mathcal{P}\mathrm{ic}}(Y_a/S)$, the coherent sheaf $\mathcal{E} = p_{a,*}\mathcal{F}$ is a vector bundle of rank n over $X \times S$, which is naturally equipped with a Higgs field $\phi : \mathcal{E} \to \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{L}(D)$; conversely, every object $(E, \phi) \in \mathcal{M}^{\mathrm{Hit}}(S)$ over $a \in \mathcal{A}^{\mathrm{int}}(S)$

comes in this way. Therefore, we have a natural isomorphism of stacks (see [2]):

$$\overline{\mathcal{P}\mathrm{ic}}(Y/\mathcal{A}^{\mathrm{int}}) \xrightarrow{\sim} \mathcal{M}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}}. \tag{3.2.1}$$

Therefore, we can view $\mathcal{M}_i^{\mathrm{int}}$ as a stack over $\overline{\mathcal{P}\mathrm{ic}}(Y/\mathcal{A}^{\mathrm{int}})$.

LEMMA 3.2.3
The stack $\mathcal{M}_i^{\mathrm{int}}$ represents the following functor:

$$S \mapsto \left\{ (a, \mathcal{F}, \alpha, \beta) \,\middle|\, \begin{array}{l} a \in \mathcal{A}^{\mathrm{int}}(S),\ \mathcal{F} \in \overline{\mathcal{P}\mathrm{ic}}^{\,i-n(g-1)}(Y_a/S), \\ \mathcal{O}_{Y_a}(-D_0) \xrightarrow{\alpha} \mathcal{F} \xrightarrow{\beta} \omega_{Y_a/X\times S}(D_0) \text{ such that } \gamma = \beta \circ \alpha \text{ is} \\ \text{nonzero along each geometric fiber of } Y_a \to S \end{array} \right\}.$$

Proof
For a quadruple $(a, \mathcal{F}, \alpha, \beta)$ as above, we associate $(\mathcal{E} = p_{a,*}\mathcal{F}, \phi) \in \mathcal{M}_i^{\mathrm{Hit}}$ by the isomorphism (3.2.1). By adjunction, we have

$$\begin{aligned} \mathrm{Hom}_X\big(\mathcal{O}_{X\times S}(-D_0), \mathcal{E}\big) &= \mathrm{Hom}_{Y_a}\big(\mathcal{O}_{Y_a}(-D_0), \mathcal{F}\big), \\ \mathrm{Hom}_X\big(\mathcal{E}, \mathcal{O}_{X\times S}(D_0)\big) &= \mathrm{Hom}_{Y_a}\big(\mathcal{F}, \omega_{Y_a/X\times S}(D_0)\big). \end{aligned}$$

Therefore, from (α, β), we can associate a unique pair of homomorphisms:

$$\mathcal{O}_{X\times S}(-D_0) \xrightarrow{\lambda} \mathcal{E} \xrightarrow{\mu} \mathcal{O}_{X\times S}(D_0).$$

Let $b = (b_0, \dots, b_{n-1}) \in \mathcal{B}$ be the second collection of invariants of $(\mathcal{E}, \phi, \lambda, \mu)$. The composition $\gamma = \beta \circ \alpha$ is an element in

$$\mathrm{Hom}_{Y_a}\big(\mathcal{O}_{Y_a}(-D_0), \omega_{Y_a/X\times S}(D_0)\big) = \mathrm{Hom}_{X\times S}\left(\mathcal{O}_{Y_a}, \mathcal{O}_{X\times S}(2D_0)\right),$$

which is given by

$$(b_0, \dots, b_{n-1}) : \mathcal{O}_{Y_a} = \bigoplus_{i=0}^{n-1} \mathcal{L}(-D)^{\otimes i} \to \mathcal{O}_{X\times S}(2D_0).$$

Therefore, for any geometric point s of S, the condition $b(s) \neq 0$ is equivalent to that $\gamma|_{Y_a(s)} \neq 0$. □

Remark 3.2.4
From the proof of the above lemma, we see that for fixed $(a, b) \in \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}$, the homomorphism $\gamma : \mathcal{O}_{Y_a}(-D_0) \to \omega_{Y_a/X}(D_0)$ is independent of the choice of

$(\mathscr{F}, \alpha, \beta) \in \mathscr{M}^{\mathrm{int}}_{i,a,b}$. We denote this γ by $\gamma_{a,b}$. Therefore, we get a morphism,

$$\begin{aligned}\operatorname{coker}(\gamma) : \mathscr{A}^{\mathrm{int}} \times \mathscr{B}^{\times} &\to \operatorname{Quot}^{2d}\big(\omega_{Y/X}(D_0)/Y/\mathscr{A}^{\mathrm{int}}\big)\\ (a,b) &\mapsto \operatorname{coker}(\gamma_{a,b}),\end{aligned}$$

where $\operatorname{Quot}^{2d}(\omega_{Y/X}(D_0)/Y/\mathscr{A}^{\mathrm{int}})$ is the Quot-scheme parameterizing quotients (of length $2d$) of the twisted relative dualizing sheaf $\omega_{Y_a/X}(D_0)$ as a varies in $\mathscr{A}^{\mathrm{int}}$.

Remark 3.2.5
Let

$$d = n\deg(D_0) - n(g-1) + g_Y - 1 = n(n-1)\deg(D)/2 + n\deg(D_0).$$

By the moduli interpretation given in Lemma 3.2.3, $\mathscr{M}^{\mathrm{int}}_i$ is nonempty only if

$$-d-n(g-1) = \chi\big(Y_a, \mathscr{O}_{Y_a}(-D_0)\big) \le i-n(g-1) \le \chi\big(Y_a, \omega_{Y_a/X}(D_0)\big) = d-n(g-1)$$

(here $a \in \mathscr{A}^{\mathrm{int}}$ is any geometric point); that is, $-d \le i \le d$.

PROPOSITION 3.2.6
Suppose that $n(\deg(D_0) - g + 1) \ge g_Y$. Then for $-d \le i \le d$, $\mathscr{M}^{\mathrm{int}}_i$ is a scheme smooth over k and the morphism $f^{\mathrm{int}}_i : \mathscr{M}^{\mathrm{int}}_i \to \mathscr{A}^{\mathrm{int}} \times \mathscr{B}^{\times}$ is proper.

Proof
Using the definition of d, we see that $n(\deg(D_0) - g + 1) \ge g_Y$ is equivalent to $d \ge 2g_Y - 1$. We may assume that $i \ge 0$; the argument for $i \le 0$ is similar.

We have the following Cartesian diagram,

$$\begin{tikzcd} \mathscr{M}^{\mathrm{int}}_i \arrow[r, "r_\beta"] \arrow[d, "r_\alpha"] & \operatorname{Quot}^{d-i}\big(\omega_{Y/X}(D_0)/Y/\mathscr{A}^{\mathrm{int}}\big) \arrow[d, "\mathrm{AJ}_{d-i}"] \\ \operatorname{Quot}^{d+i}\big(\omega_{Y/X}(D_0)/Y/\mathscr{A}^{\mathrm{int}}\big) \arrow[r, "(\mathrm{AJ}_{d+i})^\vee"] & \overline{\mathscr{P}\mathrm{ic}}^{\,i-n(g-1)}(Y/\mathscr{A}^{\mathrm{int}}), \end{tikzcd} \tag{3.2.2}$$

which, on the level of S-points over $a \in \mathscr{A}^{\mathrm{int}}(S)$, is defined as

$$\begin{tikzcd} (\mathscr{F}, \alpha, \beta) \arrow[r, "r_\beta"] \arrow[d, "r_\alpha"] & \operatorname{coker}\big(\mathscr{F} \xrightarrow{\beta} \omega_{Y_a/X\times S}(D_0)\big) \arrow[d, "\mathrm{AJ}_{d-i}"] \\ \operatorname{coker}\big(\omega_{Y_a/X\times S}(D_0) \xrightarrow{\alpha^\vee} \mathscr{F}^\vee\big) \arrow[r, "(\mathrm{AJ}_{d+i})^\vee"] & \mathscr{F}. \end{tikzcd}$$

The only thing that we need to check is that $\operatorname{coker}(\beta)$ and $\operatorname{coker}(\alpha^\vee)$ are finite *flat* $\mathcal{O}_S$-modules of rank $d-i$ and $d+i$, respectively (when S is locally Noetherian). We check this for $\mathcal{Q} := \operatorname{coker}(\beta)$. For any geometric point $s \in S$, the map $\gamma_s : \mathcal{O}_{Y_{a(s)}}(-D_0) \xrightarrow{\alpha_s} \mathcal{F}_s \xrightarrow{\beta_s} \omega_{Y_{a(s)}/X_s}(D_0)$ is nonzero, and hence generically an isomorphism. Therefore, β_s is surjective on the generic point of $Y_{a(s)}$. Since $\mathcal{F}_s$ is torsion-free of rank one, we conclude that β_s is injective. Since $\omega_{Y_a/X\times S}(D_0)$ is flat over $\mathcal{O}_S$, we have

$$\operatorname{Tor}_1^{\mathcal{O}_S}\big(\mathcal{Q}, k(s)\big) = \ker(\beta_s) = 0.$$

This being true for any geometric point $s \in S$, we conclude that $\mathcal{Q}$ is flat over $\mathcal{O}_S$. The rank of $\mathcal{Q}_s$ over $k(s)$ for any geometric point $s \in S$ is

$$\chi\big(Y_{a(s)}, \omega_{Y_{a,s}/X_s}(D_0)\big) - \chi(Y_{a(s)}, \mathcal{F}_s) = d - i.$$

CLAIM 1
The scheme $\operatorname{Quot}^{d-i}(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}})$ *is smooth over* k.

Proof
Let S be the spectrum of an Artinian k-algebra. Fix an S-point $a \in \mathcal{A}^{\mathrm{int}}(S)$ and $(\mathcal{F}, \beta) \in \operatorname{Quot}^{d-i}(\omega_{Y_a/X}(D_0)/Y_a)(S)$. Alternatively, we view this point as $(\mathcal{E}, \phi, \mu)$, where $(\mathcal{E}, \phi)$ is a Higgs bundle and where $\mu : \mathcal{E} \to \mathcal{O}_{X\times S}(D_0)$.

Similar to the deformation theoretic calculation of Biswas and Ramanan [4] for Higgs bundles, the obstruction for the infinitesimal liftings of $(\mathcal{E}, \phi, \mu)$ is $H^1(X \times S, \mathcal{K})$, where $\mathcal{K}$ is the two-step complex (in degree -1 and 0)

$$\operatorname{End}(\mathcal{E}) \xrightarrow{[-,\phi]+\mu^*} \operatorname{End}(\mathcal{E}) \otimes \mathcal{L}(D) \oplus \mathcal{E}^\vee(D_0). \tag{3.2.3}$$

Here the map $\mu^* : \operatorname{End}(\mathcal{E}) \to \mathcal{E}^\vee(D_0)$ sends $\psi \in \operatorname{End}(\mathcal{E})$ to $\mathcal{E} \xrightarrow{\psi} \mathcal{E} \xrightarrow{\mu} \mathcal{O}_{X\times S}(D_0)$. We need to prove that $H^1(X \times S, \mathcal{K}) = 0$. Let $\mathcal{P}$ be the kernel of the transpose of (3.2.3). Then Serre duality implies that

$$H^1(X \times S, \mathcal{K}) = H^0(X \times S, \mathcal{P} \otimes \omega_X)^\vee.$$

Let $\mathcal{P}'$ be the kernel of $[\phi, -] : \operatorname{End}(\mathcal{E}) \otimes \mathcal{L}^\vee(-D) \to \operatorname{End}(\mathcal{E})$ (the transpose of $\operatorname{End}(\mathcal{E}) \xrightarrow{[-,\phi]} \operatorname{End}(\mathcal{E}) \otimes \mathcal{L}(D)$). We have an exact sequence:

$$0 \to \mathcal{P}' \to \mathcal{P} \to \mathcal{E}(-D_0).$$

We claim that $\mathcal{P}'$ is equal to $\mathcal{P}$. In fact, since $\mathcal{P}/\mathcal{P}' \subset \mathcal{E}(-D_0)$ is torsion-free as an $\mathcal{O}_X$-module, it suffices to argue that $\mathcal{P}' = \mathcal{P}$ over $S \times \operatorname{Spec} K$, where $\operatorname{Spec} K$ is a geometric generic point of X. Write $S \times \operatorname{Spec} K$ as $\operatorname{Spec} R$, with R an Artinian K-algebra. Over $\operatorname{Spec} R$, we can identify $\mathcal{E}_R = R^n$, trivialize $\mathcal{L}, \mathcal{O}_{X\times S}(D)$ and

$\mathcal{O}_{X\times S}(D_0)$, and diagonalize $\phi = \mathrm{diag}\{c_1, \ldots, c_n\}$ with $c_i \in R$ (since ϕ is generically regular semisimple). Write $\mu = (\mu_1, \ldots, \mu_n) \in R^{n,\vee}$ under the standard dual basis. Then $\{\mu, \mu\phi, \ldots, \mu\phi^{n-1}\}$ form a basis of $R^{n,\vee}$ (this follows from the fact that $\mu : \mathcal{E} \to \mathcal{O}_{X\times S}(D_0)$ is adjoint to an *injection* $\beta : \mathcal{F} \to \omega_{Y_a/X\times S}(D_0)$). This implies that $\mu_i \in R^\times, \forall i$. The transpose of the map (3.2.3) restricted to Spec R has the form

$$\begin{aligned} \mathfrak{gl}_n(R) \oplus R^n &\to \mathfrak{gl}_n(R), \\ (\psi, \lambda) &\mapsto [\phi, \psi] + {}^t\mu \cdot \lambda. \end{aligned} \tag{3.2.4}$$

Suppose that $[\phi, \psi] + {}^t\mu \cdot \lambda = 0$ for some $\psi \in \mathfrak{gl}_n(R)$, and suppose that $\lambda = (\lambda_1, \ldots, \lambda_n) \in R^n$. Comparing diagonal entries gives $\mu_i\lambda_i = 0$ for all i. Since $\mu_i \in R^\times$, we conclude that $\lambda = 0$. Hence the kernel of (3.2.4) necessarily lies in $\mathfrak{gl}_n(R)$; that is, $\mathcal{P}' = \mathcal{P}$ over Spec R. This implies that $\mathcal{P}' = \mathcal{P}$.

By Serre duality again, $H^0(X \times S, \mathcal{P}' \otimes \omega_X)^\vee = H^1(X \times S, \mathcal{K}')$ calculates the obstruction to lifting the Higgs bundle $(\mathcal{E}, \phi)$, where $\mathcal{K}'$ is the complex $\mathrm{End}(\mathcal{E}) \xrightarrow{[-,\phi]} \mathrm{End}(\mathcal{E}) \otimes \mathcal{L}(D)$ (in degree -1 and 0). Since the Hitchin moduli space $\mathcal{M}_i^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}}$ is smooth by [12, proposition 4.14.1] (here we used the assumption $\deg(D) \geq 2g - 1$), the obstruction $H^0(X \times S, \mathcal{P}' \otimes \omega_X) = 0$. By the above discussions, $H^1(X \times S, \mathcal{K}) = H^0(X \times S, \mathcal{P} \otimes \omega_X)^\vee = H^0(X \times S, \mathcal{P}' \otimes \omega_X)^\vee = 0$, hence there is no obstruction to lifting $(\mathcal{E}, \phi, \mu)$; that is, $\mathrm{Quot}^{d-i}(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}})$ is smooth over k. $\square$

Since $d + i \geq d \geq 2g_Y - 1$, $(\mathrm{AJ}_{d+i})^\vee$ is smooth and schematic by [1, Theorem 8.4(v)]. By Claim 1 above, $\mathrm{Quot}^{d-i}(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}})$ is a smooth scheme over k. By diagram (3.2.2), $\mathcal{M}_i^{\mathrm{int}}$ is a smooth scheme over k.

We have a $\mathbb{G}_m \times \mathbb{G}_m$ action on $\mathcal{M}_i^{\mathrm{int}}$, where (c_1, c_2) acts by changing $(\mathcal{F}, \alpha, \beta)$ to $(\mathcal{F}, c_1\alpha, c_2\beta)$. It is easy to see that (c, c^{-1}) acts trivially so that the action factors through the multiplication map $\mathbb{G}_m \times \mathbb{G}_m \to \mathbb{G}_m$. This $\mathbb{G}_m$-action is free and the quotient $\overline{\mathcal{M}}_i^{\mathrm{int}}$ exists as a scheme. In fact, we can define $\overline{\mathcal{M}}_i^{\mathrm{int}}$ by a similar Cartesian diagram as (3.2.2) (the only difference is that the stack $\overline{\mathcal{P}\mathrm{ic}}^{i-n(g-1)}(Y/\mathcal{A}^{\mathrm{int}})$ is replaced by the scheme $\overline{\mathrm{Pic}}^{i-n(g-1)}(Y/\mathcal{A}^{\mathrm{int}})$):

$$\begin{tikzcd} \overline{\mathcal{M}}_i^{\mathrm{int}} \arrow[r] \arrow[d] & \mathrm{Quot}^{d-i}\big(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}}\big) \arrow[d, "\mathrm{AJ}_{d-i}"] \\ \mathrm{Quot}^{d+i}\big(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}}\big) \arrow[r, "(\mathrm{AJ}_{d+i})^\vee"] & \overline{\mathrm{Pic}}^{i-n(g-1)}(Y/\mathcal{A}^{\mathrm{int}}). \end{tikzcd} \tag{3.2.5}$$

From diagram (3.2.5), we see that $\overline{\mathcal{M}}_i^{\mathrm{int}}$ is proper over $\mathcal{A}^{\mathrm{int}}$ because the Quot-scheme $\mathrm{Quot}^{d+i}(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}})$ is proper over $\mathcal{A}^{\mathrm{int}}$ and the morphism AJ_{d-i} is proper. Moreover, we have a Cartesian diagram,

$$\begin{tikzcd} \mathcal{M}_i^{\mathrm{int}} \arrow[r] \arrow[d, "f_i^{\mathrm{int}}"] & \overline{\mathcal{M}}_i^{\mathrm{int}} \arrow[d, "\overline{f_i^{\mathrm{int}}}"] \\ \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times} \arrow[r] & \mathcal{A}^{\mathrm{int}} \times \mathbb{P}\mathcal{B}^{\times}, \end{tikzcd}$$

where the horizontal maps are $\mathbb{G}_m$ torsors. Since both $\overline{\mathcal{M}}_i^{\mathrm{int}}$ and $\mathcal{A}^{\mathrm{int}} \times \mathbb{P}\mathcal{B}^{\times}$ are proper over $\mathcal{A}^{\mathrm{int}}$, the morphism $\overline{f_i^{\mathrm{int}}}$ is proper. Therefore, f_i^{int} is also proper. This completes the proof. □

3.3. The moduli space associated to $\mathfrak{u}_n$

Consider the functor $\underline{\mathcal{N}} : \mathrm{Sch}/k \to \mathrm{Grpd}$:

$$S \mapsto \left\{ (\mathcal{E}', h, \phi', \mu') \left| \begin{array}{l} \mathcal{E}' \text{ is a vector bundle of rank } n \text{ over } X' \times S, \\ h : \mathcal{E}' \xrightarrow{\sim} \sigma^*(\mathcal{E}')^{\vee} \text{ is a Hermitian form, that is, } \sigma^* h^{\vee} = h, \\ \phi' : \mathcal{E}' \to \mathcal{E}'(D) \text{ such that } \sigma^*\phi'^{\vee} \circ h + h \circ \phi' = 0, \\ \mu' : \mathcal{E}' \to \mathcal{O}_{X' \times S}(D_0) \end{array} \right. \right\}.$$

Here $(-)^{\vee} = \underline{\mathrm{Hom}}_{X'\times S}(-, \mathcal{O}_{X'\times S})$. It is clear that $\underline{\mathcal{N}}$ is represented by an algebraic stack $\mathcal{N}$ locally of finite type.

Recall that we also have the usual Hitchin moduli stack $\mathcal{N}^{\mathrm{Hit}}$ for U_n classifying only the triples $(\mathcal{E}', h, \phi')$ as above. (For details about this Hitchin stack, we refer the readers to [10].)

For $(\mathcal{E}', h, \phi', \mu') \in \mathcal{N}(S)$, since $\sigma^*\phi'^{\vee} = -h \circ \phi' \circ h^{-1}$, we have

$$a_i = \mathrm{Tr}\Big(\bigwedge^i \phi'\Big) \in H^0\big(X' \times S, \mathcal{O}_{X'\times S}(iD)\big)^{\sigma=(-1)^i} = H^0\big(X, \mathcal{L}(D)^{\otimes i}\big).$$

Let $\lambda' = h^{-1} \circ \sigma^*\mu'^{\vee} : \mathcal{O}_{X'\times S}(-D_0) \to \sigma^*\mathcal{E}'^{\vee} \xrightarrow{\sim} \mathcal{E}'$. Consider the homomorphism

$$b_i' : \mathcal{O}_{X'\times S}(-D_0) \xrightarrow{\lambda'} \mathcal{E} \xrightarrow{\phi'^i} \mathcal{E}'(iD) \xrightarrow{\mu'} \mathcal{O}_{X'\times S}(D_0 + iD).$$

We have a canonical isomorphism,

$$\begin{aligned} \iota : \sigma^* b_i'^{\vee} &\cong \sigma^*\lambda'^{\vee} \circ (\sigma^*\phi'^{\vee})^i \circ \sigma^*\mu'^{\vee} \\ &\cong \mu' \circ h^{-1} \circ (-h \circ \phi' \circ h^{-1})^i \circ h \circ \lambda' \\ &= (-1)^i \mu' \circ \phi'^i \circ \lambda' = (-1)^i b_i', \end{aligned}$$

such that $\sigma^*\iota^\vee = \iota$. Therefore, b_i' comes from a homomorphism:

$$b_i : \mathcal{O}_{X\times S}(-D_0) \to \mathcal{O}_{X\times S}(D_0) \otimes \mathcal{L}(D)^{\otimes i}.$$

In other words, we may view b_i as an element in $H^0(X \times S, \mathcal{O}_{X\times S}(2D_0) \otimes \mathcal{L}(D)^{\otimes i})$. The map that sends $(\mathcal{E}', h, \phi', \lambda') \in \mathcal{N}(S)$ to $a = (a_1, \ldots, a_n) \in \mathcal{A}(S)$ and $b = (b_0, \ldots, b_{n-1}) \in \mathcal{B}(S)$ defines a morphism:

$$g : \mathcal{N} \to \mathcal{A} \times \mathcal{B}.$$

Let $\mathcal{N}^{\mathrm{int}}$ be the restriction of $\mathcal{N}$ to $\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$. Similar to Lemma 3.2.3, we can rewrite $\mathcal{N}^{\mathrm{int}}$ in terms of spectral curves.

LEMMA 3.3.1
The stack $\mathcal{N}^{\mathrm{int}}$ represents the following functor

$$S \mapsto \left\{ (a, \mathcal{F}', \beta') \;\middle|\; \begin{array}{l} a \in \mathcal{A}^{\mathrm{int}}(S),\ \mathcal{F}' \in \overline{\mathcal{P}\mathrm{ic}}(Y_a'/S),\ h : \mathcal{F}' \xrightarrow{\sim} \sigma^*(\mathcal{F}'^\vee), \\ \text{such that } \sigma^*h^\vee = h \text{ and } \mathcal{F}' \xrightarrow{\beta'} \omega_{Y_a'/X\times S}(D_0), \\ \text{which is nonzero along each geometric fiber of } Y_a' \to S \end{array} \right\}.$$

For $(\mathcal{F}', h, \beta') \in \mathcal{N}_a^{\mathrm{int}}(S)$, let $\gamma' = \beta' \circ \alpha'$, where $\alpha' = h^{-1} \circ \sigma^*\beta'^\vee : \mathcal{O}_{Y_a'}(-D_0) \to \mathcal{F}'$. Then $\gamma' \in \mathrm{Hom}_{Y_a'}(\mathcal{O}_{Y_a'}(-D_0), \omega_{Y_a'/X\times S}(D_0))$ satisfies $\sigma^*\gamma'^\vee = \gamma'$. Therefore, γ' comes from $\gamma \in \mathrm{Hom}_{Y_a}(\mathcal{O}_{Y_a}(-D_0), \omega_{Y_a/X\times S}(D_0))$ via pullback along the double cover $\pi_a : Y_a' \to Y_a$. It is easy to see that this γ is the same as the $\gamma_{a,b}$ defined in Remark 3.2.4.

PROPOSITION 3.3.2
Suppose that $n(\deg(D_0) - g + 1) \geq g_Y$. Then the stack $\mathcal{N}^{\mathrm{int}}$ is a scheme smooth over k and the morphism $g^{\mathrm{int}} : \mathcal{N}^{\mathrm{int}} \to \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$ is proper.

Proof
By the moduli interpretation given in Lemma 3.3.1, we have a Cartesian diagram,

$$\begin{array}{ccc} \mathcal{N}^{\mathrm{int}} & \xrightarrow{r'_\beta} & \mathrm{Quot}^{2d}\big(\omega_{Y/X}(D_0)/Y'/\mathcal{A}^{\mathrm{int}}\big) \\ \downarrow r & & \downarrow \mathrm{AJ}_{2d} \\ \mathcal{N}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}} & \xrightarrow{u} & \overline{\mathcal{P}\mathrm{ic}}^{-2n(g-1)}(Y'/\mathcal{A}^{\mathrm{int}}), \end{array} \tag{3.3.1}$$

which, on the level of S-points over $a \in \mathcal{A}^{\mathrm{int}}(S)$, is defined as

$$\begin{tikzcd}
(\mathcal{F}', h, \beta') \arrow[r, "r'_\beta"] \arrow[d, "r"] & \mathrm{coker}\big(\mathcal{F}' \xrightarrow{\beta'} \omega_{Y'_a/X\times S}(D_0)\big) \arrow[d, "\mathrm{AJ}_{2d}"] \\
(\mathcal{F}', h) \arrow[r, "u"] & \mathcal{F}'.
\end{tikzcd}$$

Here $\mathcal{F}'$ has Euler characteristic $-2n(g-1)$ along every geometric fiber of $Y'_a \to S$ because $\mathcal{F}' \cong \sigma^*\mathcal{F}'^\vee$.

Since u is schematic and since $\mathrm{Quot}^{2d}(\omega_{Y/X}(D_0)/Y'/\mathcal{A}^{\mathrm{int}})$ is a scheme, we see that $\mathcal{N}^{\mathrm{int}}$ is a scheme. By [10, propositions 2.5.2, 2.8.4], $\mathcal{N}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}}$ is smooth over k and proper over $\mathcal{A}^{\mathrm{int}}$. Moreover, since $d = n\deg(D_0 - g + 1) + g_Y - 1 \geq 2g_Y - 1 = g'_Y$, $2d \geq 2g'_Y$; therefore, AJ_{2d} is smooth by [1, Theorem 8.4(v)]. Therefore, $\mathcal{N}^{\mathrm{int}}$ is a smooth scheme over k.

We have a $\mathbb{G}_m$-action on $\mathcal{N}^{\mathrm{int}}$ by rescaling β'. Unlike the case of $\mathcal{M}_i^{\mathrm{int}}$, this action is not free: the subgroup $\mu_2 \subset \mathbb{G}_m$ acts trivially. The quotient of $\mathcal{N}^{\mathrm{int}}$ by this $\mathbb{G}_m$-action is a Deligne-Mumford stack $\overline{\mathcal{N}}^{\mathrm{int}}$ proper over $\mathcal{N}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}}$, and hence proper over $\mathcal{A}^{\mathrm{int}}$. Let $(\mathbb{P}\mathcal{B}^\times)'$ be the quotient of $\mathcal{B}^\times$ by the square action of the dilation by $\mathbb{G}_m$. This is a separated Deligne-Mumford stack. Therefore, $\overline{\mathcal{N}}^{\mathrm{int}}$ is proper over $\mathcal{A}^{\mathrm{int}} \times (\mathbb{P}\mathcal{B}^\times)'$. We have a Cartesian diagram:

$$\begin{tikzcd}
\mathcal{N}^{\mathrm{int}} \arrow[r] \arrow[d, "g^{\mathrm{int}}"] & \overline{\mathcal{N}}^{\mathrm{int}} \arrow[d, "\overline{g^{\mathrm{int}}}"] \\
\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times \arrow[r] & \mathcal{A}^{\mathrm{int}} \times (\mathbb{P}\mathcal{B}^\times)'.
\end{tikzcd}$$

This implies that $g^{\mathrm{int}} : \mathcal{N}^{\mathrm{int}} \to \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^\times$ is also proper. □

3.4. The product formulas

We want to express the fibers of f_i^{int} in terms of local moduli spaces similar to $\mathcal{M}^{\mathrm{loc}}_{i,a,b}$ defined in Section 2.3. To simplify the notations, we only consider fibers $\mathcal{M}^{\mathrm{int}}_{i,a,b}$ where $(a,b) \in \mathcal{A}^{\mathrm{int}}(k) \times \mathcal{B}^\times(k)$, but the argument is valid for $(a,b) \in \mathcal{A}^{\mathrm{int}}(\Omega) \times \mathcal{B}^\times(\Omega)$ for any field $\Omega \supset k$.

Let $|X|$ be the set of closed points of X. For any $x \in |X|$, let $\mathcal{O}_{X,x}$ be the *completed* local ring of X at x with fraction field F_x and residue field $k(x)$. Then $\mathcal{O}_{X,x}$ is naturally a $k(x)$-algebra. Let $\mathcal{O}_{X',x}$ be the completed semilocal ring of X' along $\pi^{-1}(x)$ (recall Notation 1.6.3), and let E_x be its ring of fractions. Then E_x is an unramified (split or nonsplit) quadratic extension of F_x.

Fix an σ-equivariant trivialization of $\mathcal{O}_{X'}(D)$ along $\operatorname{Spec}\mathcal{O}_{X',x}$, which allows us to identify $\mathcal{L}|_{\operatorname{Spec}\mathcal{O}_{X,x}}$ with $\mathcal{O}_{E_x}^{-}$. Fix a trivialization of $\mathcal{O}_X(D_0)$ over $\operatorname{Spec}\mathcal{O}_{X,x}$. Using these trivializations, we can identify the restriction of (a,b) on $\operatorname{Spec}\mathcal{O}_{X,x}$ with a collection of invariants $a_x = (a_{1,x},\dots,a_{n,x})$ and $b_x = (b_{0,x},\dots,b_{n-1,x})$ such that $a_{i,x}, b_{i,x} \in (\mathcal{O}_{E_x})^{\sigma=(-1)^i}$. Using F_x, E_x, and (a_x,b_x) in place of F, E, and (a,b), respectively, in the discussion of Sections 2.2 and 2.3, we can define the $\mathcal{O}_{F_x}$-algebra R_{a_x}, which is isomorphic to $\mathcal{O}_{Y_a,x}$ (see Notation 1.6.3). The trivializations also identify $R_{a_x}^\vee$ with $\omega_{Y_a/X}(D_0)|_{\operatorname{Spec}\mathcal{O}_{Y_a,x}}$ and $\gamma_{a_x,b_x} : R_{a_x} \to R_{a_x}^\vee$ (defined in (2.2.4)) with $\gamma_{a,b}|_{\operatorname{Spec}\mathcal{O}_{Y,x}}$ (defined in Remark 3.2.4). As in (2.4.1), $R_{a_x}(E_x)$ has a natural Hermitian form under which $R_{a_x}(\mathcal{O}_{E_x})$ and $R_{a_x}^\vee(\mathcal{O}_{E_x})$ are dual to each other.

With these data, we can define the local moduli spaces of lattices $\mathcal{M}^{\mathrm{loc}}_{i_x,a_x,b_x}$ ($i_x \in \mathbb{Z}$) and $\mathcal{N}^{\mathrm{loc}}_{a_x,b_x}$ as in Sections 2.3 and 2.4, which are projective schemes over $k(x)$. To emphasize their dependence on the point x, we denote them by $\mathcal{M}^x_{i_x,a_x,b_x}$ and $\mathcal{N}^x_{a_x,b_x}$.

PROPOSITION 3.4.1

For $(a,b) \in \mathcal{A}^{\mathrm{int}}(k) \times \mathcal{B}^\times(k)$, there is an isomorphism of schemes over k:

$$\mathcal{M}_{i,a,b} \cong \coprod_{\sum[k(x):k]i_x = d-i} \Big(\prod_{x\in|X|} \operatorname{Res}_{k(x)/k} \mathcal{M}^x_{i_x,a_x,b_x} \Big). \tag{3.4.1}$$

Here $|X|$ is the set of closed points of X, and $\operatorname{Res}_{k(x)/k}$ means restriction of scalars. The disjoint union is taken over the set of all assignments $x \in |X| \mapsto i_x \in \mathbb{Z}$ such that $\sum_{x\in|X|}[k(x):k]i_x = d-i$. The product is the fiber product of schemes over k.

Proof

First we remark that the right-hand side in the isomorphism (3.4.1) is in fact finite. Note that for each $x \in |X|$, there are only finitely many i_x such that $\mathcal{M}^x_{i_x,a_x,b_x}$ is nonempty (more precisely, $0 \le i \le \operatorname{val}_x(\Delta_{a_x,b_x})$). For fixed $(a,b) \in \mathcal{A}^{\mathrm{int}}(k) \times \mathcal{B}^\times(k)$, the map $\gamma_{a,b} : \mathcal{O}_{Y_a}(-D_0) \to \omega_{Y_a/X}(D_0)$ is an isomorphism away from a finite subset $Z \subset |X|$, and when γ_x is an isomorphism, $\mathcal{M}^x_{i_x,a_x,b_x} \cong \operatorname{Spec} k(x)$. Therefore, we can rewrite the right-hand side of (3.4.1) as a finite disjoint union of

$$\prod_{x\in Z} \operatorname{Res}_{k(x)/k} \mathcal{M}^x_{i_x,a_x,b_x},$$

which makes sense.

By the Cartesian diagram (3.2.2), the assignment $(\mathcal{F},\alpha,\beta) \mapsto \operatorname{coker}(\beta)$ defines an isomorphism of schemes,

$$\mathcal{M}_{i,a,b} \cong \operatorname{Quot}^{d-i}(\mathcal{Q}/Y_a/k),$$

where $\mathcal{Q} = \operatorname{coker}(\gamma_{a,b})$. Since $\gamma_{a,b}$ is an isomorphism over $X - Z$, $\mathcal{Q}$ is supported over $Z \times S$. Therefore, we get a canonical decomposition,

$$\mathcal{Q} = \bigoplus_{x \in Z} \mathcal{Q}_x,$$

with each $\mathcal{Q}_x$ supported over x. We get a corresponding decomposition of the Quot-scheme

$$\mathrm{Quot}^{d-i}(\mathcal{Q}/Y_a/k) \cong \coprod_{\sum [k(x):k] i_x = d-i} \Big(\prod_{x \in Z} \mathrm{Res}_{k(x)/k} \, \mathrm{Quot}^{i_x}\big(\mathcal{Q}_x/R_{a_x}/k(x)\big)\Big).$$

To prove the isomorphism (3.4.1), it remains to identify $\mathrm{Quot}^{i_x}(\mathcal{Q}_x/R_{a_x}/k(x))$ with $\mathcal{M}^x_{i_x,a_x,b_x}$, but this is obvious from definition. □

Similarly, we have the following.

PROPOSITION 3.4.2
For $(a,b) \in \mathcal{A}^{\mathrm{int}}(k) \times \mathcal{B}^{\times}(k)$, we have the following isomorphism of schemes over k:

$$\mathcal{N}_{a,b} \cong \prod_{x \in |X|} \mathrm{Res}_{k(x)/k} \, \mathcal{N}^x_{a_x,b_x}. \tag{3.4.2}$$

COROLLARY 3.4.3
For any geometric point $(a,b) \in \mathcal{A}^{\mathrm{int}}(\Omega) \times \mathcal{B}^{\times}(\Omega)$, there is an isomorphism of schemes over Ω:

$$\coprod_{i=-d}^{d} \mathcal{M}_{i,a,b} \cong \mathcal{N}_{a,b}.$$

Proof
This follows from the two product formulas and Lemma 2.7.3. □

3.5. Smallness
Recall that in [12, sections 3.6.3, 4.4.3], Ngô defines the local and global Serre invariants for points on the Hitchin base. In our case, for a geometric point $a \in \mathcal{A}(\Omega)$, let $\widetilde{Y}_a \to Y_a$ be the normalization. Then the local and global Serre invariants are

$$\delta(a,x) = \dim_\Omega(\mathcal{O}_{\widetilde{Y}_a,x}/\mathcal{O}_{Y_a,x}),$$
$$\delta(a) = \dim_\Omega H^0(Y_a, \mathcal{O}_{\widetilde{Y}_a}/\mathcal{O}_{Y_a}) = \sum_{x \in X(\Omega)} \delta(a,x).$$

COROLLARY 3.5.1
For any geometric point $(a, b) \in (\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})(\Omega)$, *we have*

$$\dim_{\Omega} \mathcal{N}_{a,b} = \sup_{i} \dim_{\Omega} \mathcal{M}_{i,a,b} \leq \delta(a). \tag{3.5.1}$$

Proof
The first equality follows from Lemma 2.7.3. Now we prove the inequality.

For each $x \in X(\Omega)$, $\mathcal{M}^{x}_{i,a_x,b_x}$ is a subscheme of the affine Springer fiber of GL_n associated to a regular semisimple element with characteristic polynomial a_x. On the other hand, $\delta(a, x)$ is the dimension of that affine Springer fiber (see Bezrukavnikov's dimension formula [3, Main Theorem], and [12, section 3.7.5]). Therefore,

$$\dim_{\Omega} \mathcal{M}^{x}_{i,a_x,b_x} \leq \delta(a, x). \tag{3.5.2}$$

Now the inequality (3.5.1) follows from the product formula (3.4.1) (note that we only formulated the product formula for k-points (a, b) of $\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}$, but it has an obvious version for any geometric point of $\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}$.) □

For each $\delta \geq 0$, let $\mathcal{A}^{\leq\delta}$ (resp., $\mathcal{A}^{\geq\delta}$) be the open (resp., closed) subset of $\mathcal{A}^{\mathrm{int}}$ consisting of those geometric points a such that $\delta(a) \leq \delta$ (resp., $\delta(a) \geq \delta$). Recall the following estimate of Ngô ([12, proposition 5.7.2], based on local results of Goresky, Kottwitz, and MacPherson on the root valuation strata [6]). For each $\delta \geq 0$, there is a number $c_\delta \geq 0$, such that whenever $\deg(D) \geq c_\delta$, we have

$$\operatorname{codim}_{\mathcal{A}^{\mathrm{int}}_{D}}(\mathcal{A}^{\geq\epsilon}_{D}) \geq \epsilon, \quad \forall 1 \leq \epsilon \leq \delta. \tag{3.5.3}$$

Here we write $\mathcal{A}^{\mathrm{int}}_{D}$ to emphasize the dependence of $\mathcal{A}^{\mathrm{int}}$ on D.

Finally, we prove the smallness.

PROPOSITION 3.5.2
Fix $\delta \geq 0$. *For* $\deg(D) \geq c_\delta$ *and for* $n(\deg(D_0) - g + 1) \geq \delta + g_Y$, *the morphisms*

$$\begin{aligned} f_i^{\leq\delta} &: \mathcal{M}_i^{\leq\delta} = \mathcal{M}_i^{\mathrm{int}}|_{\mathcal{A}^{\leq\delta}} \to \mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times}, \quad \forall -d \leq i \leq d, \\ g^{\leq\delta} &: \mathcal{N}^{\leq\delta} = \mathcal{N}^{\mathrm{int}}|_{\mathcal{A}^{\leq\delta}} \to \mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times} \end{aligned}$$

are small.

Proof
First, by Corollary 3.4.3, for any geometric point $(a, b) \in \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}$ and for any integer i, $\dim \mathcal{M}_{i,a,b} \leq \dim \mathcal{N}_{a,b}$; therefore, the smallness of $g^{\leq\delta}$ implies the smallness of $f_i^{\leq\delta}$.

Now we prove that $g^{\le\delta}$ is small. For each $j \ge 1$, let $(\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})_j$ be the locus where the fiber of g_i^{int} has dimension j. By Corollary 3.5.1, we have

$$(\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})_j \subset \mathcal{A}^{\ge j} \times \mathcal{B}^{\times}.$$

In particular, $(\mathcal{A}^{\le\delta} \times \mathcal{B}^{\times})_j$ is nonempty only if $j \le \delta$.

CLAIM 2

For every geometric point $a \in \mathcal{A}^{\mathrm{int}}(\Omega)$, the morphism $g_a^{\mathrm{int}} : \mathcal{N}_a^{\mathrm{int}} \to \mathcal{B}^{\times} \otimes_k \Omega$ is dominant and generically finite when restricted to every irreducible component *of $\mathcal{N}_a^{\mathrm{int}}$.*

Proof

To simplify the notation, we base change everything to $\mathrm{Spec}(\Omega)$ via $a \in \mathcal{A}^{\mathrm{int}}(\Omega)$, and we omit Ω from the notations.

Let $\mathcal{B}^a \subset \mathcal{B}$ be the open subset consisting of geometric points b such that $\gamma_{a,b} : \mathcal{O}_{Y_a}(-D_0) \to \omega_{Y_a/X}(D_0)$ is an isomorphism at all singular points of Y_a. It is clear from the product formula (3.4.2) that if $b \in \mathcal{B}^a$, then $\mathcal{N}_{a,b}$ is finite and nonempty over Ω (finiteness follows from the discussion in Lemma 2.5.5).

Note that, by the Cartesian diagram (3.3.1), the morphism $r : \mathcal{N}_a^{\mathrm{int}} \to \mathcal{N}_a^{\mathrm{Hit}}$ is smooth with fibers isomorphic to punctured vector spaces; therefore, each irreducible component C of $\mathcal{N}_a^{\mathrm{int}}$ is of the form $r^{-1}(C')$ for some irreducible component C' of $\mathcal{N}_a^{\mathrm{Hit}}$. By [12, corollaire 4.16.3], $\mathcal{N}_a^{\mathrm{Hit}}$ is equidimensional, and therefore so is $\mathcal{N}_a^{\mathrm{int}}$. Therefore, it suffices to show that for each irreducible component C of $\mathcal{N}_a^{\mathrm{int}}$, we have $g_a^{\mathrm{int}}(C)\cap\mathcal{B}^a \neq \varnothing$. Because then $g_a^{\mathrm{int}}|_C$ must be generically finite onto its image; but by the equidimensionality statement, we have $\dim C = \dim \mathcal{N}_a^{\mathrm{int}} = \dim \mathcal{B}^a = \dim \mathcal{B}$, and therefore $g_a^{\mathrm{int}}(C) = \mathcal{B}^{\times}$.

Now we fix an irreducible component $C \subset \mathcal{N}_a^{\mathrm{int}}$ of the form $r^{-1}(C')$, where $C' \subset \mathcal{N}_a^{\mathrm{Hit}}$ is an irreducible component. We argue that $g_a^{\mathrm{int}}(C) \cap \mathcal{B}^a \neq \varnothing$. By [12, proposition 4.16.1], the locus of $(\mathcal{F}', h) \in C'$, where $\mathcal{F}'$ is a line bundle on Y_a', is dense. Let $(\mathcal{F}', h) \in C'$ be such a point. Since Y_a' is embedded in a smooth surface, and hence Gorenstein; $\omega_{Y_a'/X}$ is a line bundle on Y_a', and hence $\mathcal{F}'^{\vee}$ is also a line bundle.

Let Z be the singular locus of Y_a. By the definition of $\delta(a)$, we have

$$\#Z \le \dim H^0(Y_a, \mathcal{O}_{\widetilde{Y}_a}/\mathcal{O}_{Y_a}) = \delta(a),$$

where $\widetilde{Y}_a$ is the normalization of Y_a.

To show that $g_a^{\mathrm{int}}(C)\cap\mathcal{B}^a \neq \varnothing$, we only have to find $\beta' \in \mathrm{Hom}(\mathcal{F}', \omega_{Y_a'/X}(D_0)) = H^0(Y_a', \mathcal{F}'^{\vee}(D_0))$ such that $\mathrm{coker}(\beta')$ avoids the singular locus $Z' = \pi_a^{-1}(Z)$ of Y_a', since the support of $\mathrm{coker}(\gamma)$ is the same as the projection of the support of $\mathrm{coker}(\beta')$ to Y_a.

Consider the evaluation map

$$0 \to \mathcal{K} \to \mathcal{F}'^{\vee}(D_0) \xrightarrow{\oplus \mathrm{ev}(y')} \bigoplus_{y' \in Z'} \Omega(y') \to 0.$$

Note that

$$\chi(Y'_a, \mathcal{K}) = \chi\big(Y'_a, \mathcal{F}'^{\vee}(D_0)\big) - \#Z' \geq 2n\big(\deg(D_0) - g + 1\big) - 2\delta(a) \geq 2g_Y > g'_Y.$$

Therefore, by Grothendieck-Serre duality, $H^1(Y'_a, \mathcal{K}) = \mathrm{Hom}_{Y'_a}(\mathcal{K}, \omega_{Y'_a})^{\vee} = 0$. Hence the evaluation map

$$H^0\big(Y'_a, \mathcal{F}'^{\vee}(D_0)\big) \xrightarrow{\oplus \mathrm{ev}(y')} \bigoplus_{y' \in Z'} \Omega(y')$$

is surjective. In particular, we can find $\beta' \in H^0(Y'_a, \mathcal{F}'^{\vee}(D_0))$, which does not vanish at points in Z'. This proves the claim. □

Applying Claim 2 to the geometric *generic* points of $\mathcal{A}$ (note that $\mathcal{N}^{\mathrm{int}} \to \mathcal{A}^{\mathrm{int}}$ is surjective because $\mathcal{N}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}} \to \mathcal{A}^{\mathrm{int}}$ is), we see that g^{int} restricted to every geometric irreducible component of $\mathcal{N}^{\mathrm{int}}$ is generically finite and surjective.

Using the above Claim 2 again, we see that for any geometric point $a \in \mathcal{A}^{\mathrm{int}}$ and $j \geq 1$, the locus of $b \in \mathcal{B}^{\times}$, where $\dim \mathcal{N}_{a,b} = j$, has codimension at least $j+1$ in $\mathcal{B}^{\times}$. Therefore, we have

$$\mathrm{codim}_{\mathcal{A}^{\geq j} \times \mathcal{B}^{\times}}(\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})_j \geq j + 1.$$

Since $\deg(D) \geq c_\delta$ and $j \leq \delta$, we have

$$\begin{aligned}\mathrm{codim}_{\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times}}(\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})_j &\geq \mathrm{codim}_{\mathcal{A}^{\mathrm{int}}}(\mathcal{A}^{\geq j}) \\ &\quad + \mathrm{codim}_{\mathcal{A}^{\geq j} \times \mathcal{B}^{\times}}(\mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times})_j \geq 2j + 1.\end{aligned}$$

This proves the smallness. □

4. Global formulation: Matching of perverse sheaves

4.1. A local system on $\mathcal{M}_i^{\mathrm{int}}$

Consider the morphism

$$\nu : \mathcal{M}_i^{\mathrm{int}} \xrightarrow{r_\beta} \mathrm{Quot}^{d-i}\big(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}}\big) \xrightarrow{\mathfrak{N}_{Y/\mathcal{A}^{\mathrm{int}}}} \mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{int}}) \to \mathrm{Sym}^{d-i}(X/k), \tag{4.1.1}$$

where $\mathfrak{N}_{Y/\mathcal{A}^{\text{int}}}$ is the norm map defined by Grothendieck in [8]. Here $\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\text{int}})$ (resp., $\mathrm{Sym}^{d-i}(X/k)$) is the $(d-i)$th symmetric power of Y over $\mathcal{A}^{\text{int}}$ (resp., X over k), constructed as the geometric invariant theory (GIT) quotient of the fibered power $(Y/\mathcal{A}^{\text{int}})^{d-i}$ (resp., $(X/k)^{d-i}$) by the obvious action of the symmetric group Σ_{d-i}. The morphism $\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\text{int}}) \to \mathrm{Sym}^{d-i}(X/k)$ above is induced from the natural projection $(Y/\mathcal{A}^{\text{int}})^{d-i} \to (X/k)^{d-i}$.

We will construct an étale double cover of $\mathrm{Sym}^{d-i}(X/k)$, which will give us a local system of rank one and order two on $\mathrm{Sym}^{d-i}(X/k)$, and hence on $\mathcal{M}_i^{\text{int}}$.

The groups $(\mathbb{Z}/2)^{d-i}$ and Σ_{d-i} both act on $(X'/k)^{d-i}$, and together they give an action of the semidirect product $\Gamma = (\mathbb{Z}/2)^{d-i} \rtimes \Sigma_{d-i}$ on $(X'/k)^{d-i}$. Then $\mathrm{Sym}^{d-i}(X/k)$ is the GIT quotient of $(X'/k)^{d-i}$ by Γ. We have a canonical surjective homomorphism $\epsilon : \Gamma \to \mathbb{Z}/2$ sending $(v_1, \ldots, v_{d-i}, s) \mapsto v_1 + \cdots + v_{d-i}$, where $v_j \in \mathbb{Z}/2$ and $s \in \Sigma_{d-i}$. Let Γ_0 be the kernel of ϵ.

LEMMA 4.1.1

Let Z^{d-i} be the GIT quotient of $(X'/k)^{d-i}$ by Γ_0. Then the natural morphism ζ : $Z^{d-i} \to \mathrm{Sym}^{d-i}(X/k)$ is an étale double cover.

Proof

Let $t = x_1 + \cdots + x_{d-i}$ be a geometric point of $\mathrm{Sym}^{d-i}(X/k)$, and let $t' = (x'_1, \ldots, x'_{d-i})$ be a geometric point of $(X'/k)^{d-i}$ over t. Then the point $t'' = (\sigma(x'_1), x'_2, \ldots, x'_{d-i})$ is another geometric point of $(X'/k)^{d-i}$ which does not lie in the Γ_0-orbit of t'. Therefore t' and t'' have different images z' and z'' in Z^{d-i}. In other words, the reduced structure of $\zeta^{-1}(t)$ consists of (at least) two points z', z''.

Consider the maps

$$\xi : (X'/k)^{d-i} \xrightarrow{\eta} Z^{d-i} \xrightarrow{\zeta} \mathrm{Sym}^{d-i}(X/k),$$

where ξ is finite flat of degree $2^{d-i}(d-i)!$. The degree of the geometric fibers $\eta^{-1}(z')$ and $\eta^{-1}(z'')$ are *at least* $2^{d-i-1}(d-i)!$ (because this is the generic degree). This forces the two degrees to be equal to $2^{d-i-1}(d-i)!$. This being true for any geometric point of Z^{d-i}, we conclude that the quotient map $(X'/k)^{d-i} \to Z^{d-i}$ is flat, hence faithfully flat. Therefore $\zeta : Z^{d-i} \to \mathrm{Sym}^{d-i}(X/k)$ is also flat (of degree 2). Since every fiber $\zeta^{-1}(x)$ already consists of two distinct points, these two points must be reduced. Therefore ζ is finite flat of degree 2 and unramified, and hence an étale double cover. □

Let L^X_{d-i} be the local system of rank one on $\mathrm{Sym}^{d-i}(X/k)$ associated to the étale double cover $Z^{d-i} \to \mathrm{Sym}^{d-i}(X/k)$ (see Notation 1.6.4). We define

$$L_{d-i} := \nu^* L^X_{d-i}$$

to be the pullback local system on $\mathcal{M}_i^{\mathrm{int}}$.

We describe the stalks of the local system L_{d-i}^X in more concrete terms. Let L be the local system of rank one on X associated to the étale double cover $\pi : X' \to X$. Let $C_0 = \ker((\mathbb{Z}/2)^{d-i} \xrightarrow{\epsilon} \mathbb{Z}/2)$. Then $(X'/k)^{d-i}/C_0$ is an étale double cover of $(X/k)^{d-i}$.

LEMMA 4.1.2

(1) *We have a Cartesian diagram,*

$$\begin{tikzcd} (X'/k)^{d-i}/C_0 \arrow[r] \arrow[d] & Z^{d-i} \arrow[d] \\ (X/k)^{d-i} \arrow[r, "s_{d-i}"] & \mathrm{Sym}^{d-i}(X/k), \end{tikzcd}$$

where the maps are all natural quotient maps.

(2) *We have $s_{d-i}^* L_{d-i}^X \cong L^{\boxtimes(d-i)}$.*

Proof

Part (1) follows from the fact that both vertical maps are étale (see Lemma 4.1.1). In part (2), the local system of rank one associated to the étale double cover $(X'/k)^{d-i}/C_0 \to (X/k)^{d-i}$ is clearly $L^{\boxtimes(d-i)}$. Therefore, (2) follows from (1). □

4.2. The incidence correspondence

Recall that $\mathcal{A}^{\mathrm{sm}}$ is the open locus of $a \in \mathcal{A}^{\mathrm{int}}$, where Y_a is smooth (equivalently, the locus where $\delta(a) = 0$). We assume that $\deg(D)$ is large enough ($\geq c_1$) so that $\mathcal{A}^{\mathrm{sm}}$ is nonempty, and hence dense in $\mathcal{A}^{\mathrm{int}}$. The norm map $\mathfrak{N}_{Y/\mathcal{A}^{\mathrm{int}}}$ is an isomorphism over $\mathcal{A}^{\mathrm{sm}}$.

For each $-d \leq i \leq d$, consider the incidence correspondence

$$I^{d-i,2d} \subset \mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}}) \times \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$$

whose geometric fiber over $a \in \mathcal{A}^{\mathrm{sm}}$ classifies pairs of divisors $T \leq T'$ on Y_a, where $\deg(T) = d - i$, and $\deg(T') = 2d$. Let τ, τ' be the projections of $I^{d-i,2d}$ to $\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}})$ and $\mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$.

LEMMA 4.2.1

The incidence correspondence $I^{d-i,2d}$ is smooth over $\mathcal{A}^{\mathrm{sm}}$.

Proof

Since $Y \to \mathcal{A}^{\mathrm{sm}}$ is a smooth family of curves, we may identify the symmetric power $\mathrm{Sym}^j(Y/\mathcal{A}^{\mathrm{sm}})$ with the Hilbert scheme $\mathrm{Hilb}^j(Y/\mathcal{A}^{\mathrm{sm}})$. We apply the infinitesimal lifting criterion to prove the smoothness. Suppose that R_0 is a local Artinian ring and that R is a thickening of R_0. Let $a \in \mathcal{A}^{\mathrm{sm}}(R)$ with image $a_0 \in \mathcal{A}^{\mathrm{sm}}(R_0)$. Let $T_0 \subset T_0' \subset Y_{a_0}$ be flat subschemes of degree $d-i$ and $2d$ over R_0. We want to find subschemes $T \subset T' \subset Y_a$, flat of degree $d-i$ and $2d$ over R, whose reductions to R_0 are precisely $T_0 \subset T_0'$. We may assume that T_0' is contained in an affine open subset $U \subset Y_a$. Let $U_0 = U \cap Y_{a_0}$. Since Y_{a_0} is a smooth family of curves over R_0, the subschemes T_0 and T_0' are defined by the vanishing of functions f_0 and $f_0' \in \Gamma(U_0, \mathcal{O}_{U_0})$, respectively. Since $T_0 \subset T_0'$, we have $f_0' = f_0 g_0$ for some function $g_0 \in \Gamma(U_0, \mathcal{O}_{U_0})$. Let f, g be arbitrary liftings of f_0, g_0 to $\Gamma(U, \mathcal{O}_U)$ (which exist because U is affine), and define T and T' to be the zero loci of f and fg, respectively. Then it is easy to check that $T \subset T' \subset Y_a$ are flat over R of the correct degree. This proves the smoothness of $I^{d-i,2d}$ over $\mathcal{A}^{\mathrm{sm}}$. $\square$

We define a morphism,

$$\mathrm{div} : \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times} \xrightarrow{\mathrm{coker}(\gamma)} \mathrm{Quot}^{2d}\big(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}}\big) \xrightarrow{\mathfrak{N}_{Y/\mathcal{A}^{\mathrm{int}}}} \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{int}}),$$

which sends (a, b) to the cycle of $\mathrm{coker}(\gamma_{a,b})$ in Y_a.

LEMMA 4.2.2
We have a Cartesian diagram:

$$\begin{tikzcd} \mathcal{M}_i^{\mathrm{sm}} \arrow[r, "\widetilde{\mathrm{div}}"] \arrow[d, "f_i^{\mathrm{sm}}"] & I^{d-i,2d} \arrow[d, "\tau'"] \\ \mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^{\times} \arrow[r, "\mathrm{div}"] & \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}}). \end{tikzcd} \tag{4.2.1}$$

Here the morphism $\widetilde{\mathrm{div}} : \mathcal{M}_i^{\mathrm{sm}} \to I^{d-i,2d}$ *over a point* $a \in \mathcal{A}^{\mathrm{sm}}$ *sends* $(\mathcal{F}, \alpha, \beta)$ *to the pair of divisors* $\mathrm{div}(\beta) \subset \mathrm{div}(a, b)$ *of* Y_a.

Proof
We abbreviate $\mathrm{Quot}^{2d}(\omega_{Y/X}(D_0)/Y/\mathcal{A}^{\mathrm{int}})$ by Quot^{2d}. Let $\mathcal{Q}$ be the universal quotient sheaf on $Y \times_{\mathcal{A}^{\mathrm{int}}} \mathrm{Quot}^{2d}$. Then by the moduli interpretation given in Lemma 3.2.3, we have a Cartesian diagram:

$$\begin{tikzcd} \mathcal{M}_i^{\mathrm{int}} \arrow[r, "r_\beta"] \arrow[d, "f_i^{\mathrm{int}}"] & \mathrm{Quot}^{d-i}\big(\mathcal{Q}/(Y \times_{\mathcal{A}^{\mathrm{int}}} \mathrm{Quot}^{2d})/\mathrm{Quot}^{2d}\big) \arrow[d] \\ \mathcal{A}^{\mathrm{int}} \times \mathcal{B}^{\times} \arrow[r, "\mathrm{coker}(\gamma)"] & \mathrm{Quot}^{2d}. \end{tikzcd}$$

Restricting this diagram to $\mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^{\times}$, we have

$$\mathrm{Quot}^{2d}|_{\mathcal{A}^{\mathrm{sm}}} \cong \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}}),$$
$$\mathrm{Quot}^{d-i}\left(\mathcal{Q}/(Y \times_{\mathcal{A}^{\mathrm{int}}} \mathrm{Quot}^{2d})/\mathrm{Quot}^{2d}\right)|_{\mathcal{A}^{\mathrm{sm}}} \cong I^{d-i,2d}$$

because the norm maps $\mathfrak{N}_{Y/\mathcal{A}^{\mathrm{sm}}}$ are isomorphisms. Therefore the diagram (4.2.1) is Cartesian. □

Let L^Y_{d-i} be the pullback of the local system L^X_{d-i} via $\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}}) \to \mathrm{Sym}^{d-i}(X/k)$. Define

$$K^Y_{d-i} := \tau'_*\tau^* L^Y_{d-i} \in D^b_c\big(\mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}}), \overline{\mathbb{Q}}_\ell\big).$$

LEMMA 4.2.3 (Binomial expansion)
Let $\pi^Y_{2d} : \mathrm{Sym}^{2d}(Y'/\mathcal{A}^{\mathrm{sm}}) \to \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$ *be the natural projection. Then there is a natural isomorphism:*

$$\pi^Y_{2d,*}\overline{\mathbb{Q}}_\ell \cong \bigoplus_{i=-d}^{d} K^Y_{d-i}. \tag{4.2.2}$$

Proof
Since the morphism $\tau' : I^{d-i,2d} \to \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$ is finite and since $I^{d-i,2d}$ is smooth over k by Lemma 4.2.1, we conclude that K^Y_{d-i}, being the direct image of a local system under τ', is a middle extension on $\mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$. Similarly, $\pi^Y_{2d,*}\overline{\mathbb{Q}}_\ell$ is also a middle extension on $\mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$. Therefore, to establish the isomorphism (4.2.2), it suffices to establish such a natural isomorphism over a dense open subset of $\mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$.

Now we consider the dense open subset $U \subset \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}})$ consisting of those divisors which are multiplicity-free. Let U' (resp., $\widetilde{U}$; resp., $\widetilde{U}'$) be the preimage of U in $\mathrm{Sym}^{2d}(Y'/\mathcal{A}^{\mathrm{sm}})$ (resp., $(Y/\mathcal{A}^{\mathrm{sm}})^{2d}$; resp., $(Y'/\mathcal{A}^{\mathrm{sm}})^{2d}$). Then $s_{2d} : \widetilde{U} \to U$ is an étale Galois cover with Galois group Σ_{2d}, and we have a Cartesian diagram:

$$\begin{array}{ccc} \widetilde{U}' & \xrightarrow{s'_{2d}} & U' \\ \Big\downarrow{\scriptstyle (\pi^Y)^{2d}} & & \Big\downarrow{\scriptstyle \pi^Y_{2d}} \\ \widetilde{U} & \xrightarrow{s_{2d}} & U. \end{array}$$

Therefore, we have

$$(s^*_{2d}\pi^Y_{2d,*}\overline{\mathbb{Q}}_\ell)|_{\widetilde{U}} \cong \big((\pi^Y)^{2d}_*\overline{\mathbb{Q}}_\ell\big)|_{\widetilde{U}} = (\overline{\mathbb{Q}}_\ell \oplus L^Y)^{\boxtimes 2d}|_{\widetilde{U}},$$

where L^Y is the pullback of L to Y.

On the other hand, we have a Cartesian diagram,

$$\begin{tikzcd} \coprod_{J\subset\{1,2,\dots,2d\},\#J=d-i}\widetilde{U}_J \arrow[r] \arrow[d] & I^{d-i,2d}|_U \arrow[d, "\tau'"] \\ \widetilde{U} \arrow[r, "s_{2d}"] & U, \end{tikzcd}$$

where for each $J\subset\{1,2,\dots,2d\}$, $\widetilde{U}_J\subset\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}})\times\widetilde{U}$ is the graph of the morphism $\tau_J:(y_1,\dots,y_{2d})\mapsto\sum_{j\in J}y_j$. Therefore, we have

$$(s_{2d}^*K_{d-i}^Y)|_{\widetilde{U}}\cong(s_{2d}^*\tau'_*\tau^*L_{d-i}^X)|_{\widetilde{U}}\cong\bigoplus_{J\subset\{1,2,\dots,2d\},\#J=d-i}\tau_J^*L_{d-i}^Y.$$

For each $J\subset\{1,2,\dots,2d\}$ of cardinality $d-i$, let $p_J:\widetilde{U}\to(Y/\mathcal{A}^{\mathrm{sm}})^J$ be the projection to those coordinates indexed by J. Then we have a factorization

$$\tau_J:\widetilde{U}\xrightarrow{p_J}(Y/\mathcal{A}^{\mathrm{sm}})^J\xrightarrow{s_J}\mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}}).$$

By Lemma 4.1.2, we have $s_J^*L_{d-i}^Y\cong L^{Y,\boxtimes(d-i)}$. Therefore

$$\tau_J^*L_{d-i}^Y\cong p_J^*s_J^*L_{d-i}^Y\cong p_J^*\big((L^Y)^{\boxtimes(d-i)}\big).$$

Finally, we have an Σ_{2d}-equivariant isomorphism of local systems on $\widetilde{U}$:

$$\bigoplus_{i=-d}^{d}(s_{2d}^*K_{d-i}^Y)|_{\widetilde{U}}\cong\bigoplus_{J\subset\{1,2,\dots,2d\}}p_J^*\big((L^Y)^{\boxtimes\#J}\big)\cong(\overline{\mathbb{Q}}_\ell\oplus L^Y)^{\boxtimes 2d}|_{\widetilde{U}}.$$

The last isomorphism justifies the nickname "binomial expansion" of this lemma. This Σ_{2d}-equivariant isomorphism descends to an isomorphism,

$$\bigoplus_{i=-d}^{d}K_{d-i}^Y|_U\cong(\pi_{2d,*}^Y\overline{\mathbb{Q}}_\ell)|_U,$$

which proves the lemma. □

4.3. A decomposition of $g_*^{\mathrm{int}}\overline{\mathbb{Q}}_\ell$

To state the next result, we need to define a technical notion.

Definition 4.3.1
A commutative diagram of schemes

$$\begin{tikzcd} X' \arrow[r, "\alpha"] \arrow[d, "f'"] & X \arrow[d, "f"] \\ Y' \arrow[r, "\beta"] & Y \end{tikzcd} \tag{4.3.1}$$

is said to be *pointwise Cartesian* if, for any algebraically closed field Ω, the corresponding diagram of Ω-points is Cartesian.

For pointwise Cartesian diagrams with reasonable finiteness conditions, the proper base change theorem also holds.

LEMMA 4.3.2
Suppose that we have a pointwise Cartesian diagram (4.3.1), where all maps are of finite type and where f, f' are proper. Let $F \in D_c^b(X, \overline{\mathbb{Q}}_\ell)$ be a constructible $\overline{\mathbb{Q}}_\ell$-complex on X. Then we have a quasi-isomorphism

$$\beta^* f_* F \cong f'_* \alpha^* F.$$

Proof
Let $X'' = Y' \times_Y X$, and let $f'' : X'' \to Y'$ be the projection. Let $\xi : X' \to X''$ be the natural map over Y', which is also proper. By the usual proper base change for Cartesian diagrams, we reduce to showing that for any $G \in D_c^b(X'', \overline{\mathbb{Q}}_\ell)$,

$$f''_* G \cong f'_* \xi^* G. \tag{4.3.2}$$

But, we have

$$f'_* \xi^* G = f''_* \xi_* \xi^* G \cong f''_* (G \otimes \xi_* \overline{\mathbb{Q}}_\ell).$$

Therefore to show (4.3.2), it suffices to show that the natural map $\iota : \overline{\mathbb{Q}}_\ell \to \xi_* \overline{\mathbb{Q}}_\ell$ is a quasi-isomorphism. Since both $\xi_* \overline{\mathbb{Q}}_\ell$ is constructible, it suffices to show that ι is an isomorphism on the stalks of every geometry point $x'' \in X''(\Omega)$; that is,

$$\iota_{x''} : \overline{\mathbb{Q}}_\ell \to H^*\big(\xi^{-1}(x''), \overline{\mathbb{Q}}_\ell\big) \tag{4.3.3}$$

is an isomorphism. By Definition 4.3.1, $\xi^{-1}(x'')(\Omega)$ is a singleton. Therefore, the reduced structure of $\xi^{-1}(x'')$ is $\operatorname{Spec} \Omega$, and (4.3.3) obviously holds. □

Consider the norm map

$$\mathfrak{N}_{Y'/\mathcal{A}^{\mathrm{int}}} : \mathrm{Quot}^{2d}\big(\omega_{Y'/X\times\mathcal{A}}(D_0)/Y'/\mathcal{A}^{\mathrm{int}}\big) \to \mathrm{Sym}^{2d}(Y'/\mathcal{A}^{\mathrm{int}}).$$

LEMMA 4.3.3
The diagram

$$\begin{tikzcd} \mathcal{N}^{\mathrm{sm}} \arrow[r, "\mathfrak{N}_{Y'/\mathcal{A}^{\mathrm{sm}}}\circ r'_\beta"] \arrow[d, "g^{\mathrm{sm}}"] & \mathrm{Sym}^{2d}(Y'/\mathcal{A}^{\mathrm{sm}}) \arrow[d, "\pi^Y_{2d}"] \\ \mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^{\times} \arrow[r, "\mathrm{div}"] & \mathrm{Sym}^{2d}(Y/\mathcal{A}^{\mathrm{sm}}) \end{tikzcd}$$

is pointwise Cartesian.

Proof
Since $\mathcal{N}^{\mathrm{sm}}$ is reduced, to check the commutativity of the diagram, it suffices to check on geometric points. Therefore, we fix a geometric point $(a, b) \in \mathcal{A}^{\mathrm{sm}}(\Omega) \times \mathcal{B}^{\times}(\Omega)$, and we prove that the diagram is commutative and pointwise Cartesian at the same time. Again we omit Ω in the rest of the proof.

Let $\mathrm{div}(a, b) = \sum_{t=1}^{r} m_t y_t$ with $\{y_t\}$ distinct points on Y_a. Recall that $\pi_a : Y'_a \to Y_a$ is the étale double cover induced from $\pi : X' \to X$. Let $\pi_a^{-1}(y_t) = \{y'_t, y''_t\}$. Then a point $\mathcal{O}_{Y'_a}(-D_0) \subset \mathcal{F}' \subset \omega_{Y'_a/X}(D_0)$ is determined by the torsion sheaf $\mathcal{Q}' = \omega_{Y'_a/X}(D_0)/\mathcal{F}'$, which is a quotient of $\omega_{Y'_a/X}(D_0)/\mathcal{O}_{Y'_a}(-D_0)$. Since Y'_a is smooth, $\mathcal{Q}'$ is, in turn, determined by its divisor $\sum_{t=1}^{r} m'_t y'_t + m''_t y''_t \le \pi_a^{-1}(\mathrm{div}(a, b))$. The line bundle $\mathcal{F}'$ is self-dual if and only if $m'_t + m''_t = m_t$ (see the proof of Lemma 2.5.5). Therefore, the image of the divisor of $\mathcal{Q}$ is $\mathrm{div}(a, b)$. This proves the commutativity of the diagram. But this also shows that the map $\mathcal{N}_{a,b} \to \pi_{2d}^{Y,-1}(\mathrm{div}(a, b))$ is a bijection on Ω-points. This completes the proof. □

LEMMA 4.3.4
Fix $\delta \ge 1$, $\deg(D) \ge c_\delta$, *and* $n(\deg(D_0) - g + 1) \ge \delta + g_Y$. *Then we have the following.*

(1) *For each* $-d \le i \le d$, $\mathrm{div}^* K^Y_{d-i}[\dim \mathcal{A} + \dim \mathcal{B}]$ *is a perverse sheaf.*

(2) *We define*

$$K_{d-i} := j^{\mathrm{sm}}_{!*}(\mathrm{div}^* K^Y_{d-i}[\dim \mathcal{A} + \dim \mathcal{B}])[-\dim \mathcal{A} - \dim \mathcal{B}],$$

where $j^{\mathrm{sm}} : \mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^{\times} \hookrightarrow \mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times}$ *is the open inclusion. Then we have a natural decomposition:*

$$g_*^{\leq\delta}\overline{\mathbb{Q}}_\ell \cong \bigoplus_{i=-d}^{d} K_{d-i}.$$

Proof

(1) By Lemma 4.3.3 and Lemma 4.3.2, we have

$$\mathrm{div}^*\pi_{2d,*}^{Y}\overline{\mathbb{Q}}_\ell \cong g_*^{\mathrm{sm}}\overline{\mathbb{Q}}_\ell.$$

Since g^{sm} is finite and since $\mathcal{N}^{\mathrm{sm}}$ is smooth, $g_*^{\mathrm{sm}}\overline{\mathbb{Q}}_\ell[\dim\mathcal{A} + \dim\mathcal{B}]$ is a perverse sheaf. But by Lemma 4.2.3, K_{d-i}^{Y} is a direct summand of $\pi_{2d,*}^{Y}\overline{\mathbb{Q}}_\ell$, and hence $\mathrm{div}^*K_{d-i}^{Y}$ is a direct summand of $\mathrm{div}^*\pi_{2d,*}^{Y}\overline{\mathbb{Q}}_\ell \cong g_*^{\mathrm{sm}}\overline{\mathbb{Q}}_\ell$. Therefore, $\mathrm{div}^*K_{d-i}^{Y}[\dim\mathcal{A} + \dim\mathcal{B}]$ is also a perverse sheaf.

(2) It follows from the smallness of $g^{\leq\delta}$ proved in Proposition 3.5.2. □

4.4. The global matching theorem

The global part of the main theorem of the article is the following.

THEOREM 4.4.1

Fix $\delta \geq 1$, $\deg(D) \geq c_\delta$, *and* $n(\deg(D_0) - g + 1) \geq \delta + g_Y$. *Then for* $-d \leq i \leq d$, *there is a natural isomorphism in* $D_c^b(\mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times})$,

$$f_{i,*}^{\leq\delta}L_{d-i} \cong K_{d-i}, \tag{4.4.1}$$

and hence an isomorphism:

$$\bigoplus_{i=-d}^{d} f_{i,*}^{\leq\delta}L_{d-i} \cong g_*^{\leq\delta}\overline{\mathbb{Q}}_\ell.$$

Proof

By Propositions 3.2.6 and 3.3.2, $\mathcal{M}_i^{\mathrm{int}}$ and $\mathcal{N}^{\mathrm{int}}$ are smooth. Moreover, by Proposition 3.5.2, the morphisms $f_i^{\leq\delta}$ are small. Therefore both $f_{i,*}^{\leq\delta}L_{d-i}$ and K_{d-i} are middle extensions on $\mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times}$. Hence it suffices to establish the isomorphism (4.4.1) on the dense open subset $\mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^{\times} \subset \mathcal{A}^{\leq\delta} \times \mathcal{B}^{\times}$.

The morphism ν (see (4.1.1)) restricted on $\mathcal{M}_i^{\mathrm{sm}}$ factors as

$$\nu^{\mathrm{sm}} : \mathcal{M}_i^{\mathrm{sm}} \xrightarrow{\widetilde{\mathrm{div}}} I^{d-i,2d} \xrightarrow{\tau} \mathrm{Sym}^{d-i}(Y/\mathcal{A}^{\mathrm{sm}}) \to \mathrm{Sym}^{d-i}(X/k).$$

Therefore, we have

$$L_{d-i}|_{\mathcal{M}_i^{\mathrm{sm}}} = \nu^{\mathrm{sm},*} L_{d-i}^X \cong \widetilde{\mathrm{div}}^* \tau^* L_{d-i}^Y.$$

Applying proper base change to the Cartesian diagram (4.2.1), we get

$$\begin{aligned} f_{i,*}^{\mathrm{sm}} L_{d-i} \cong f_{i,*}^{\mathrm{sm}} \widetilde{\mathrm{div}}^* \tau^* L_{d-i}^Y &\cong \mathrm{div}^* \tau'_* \tau^* L_{d-i}^Y, \\ &\cong \mathrm{div}^* K_{d-i}^Y = K_{d-i}|_{\mathcal{A}^{\mathrm{sm}} \times \mathcal{B}^\times}. \end{aligned}$$ □

5. Proof of the local main theorem

We prove Theorem 2.7.1 in this section. Suppose that $\mathrm{char}(F) = \mathrm{char}(k) = p > n$. We are left with the case k'/k nonsplit, and we assume this throughout the remainder of the article. We fix a collection of invariants (a^0, b^0) with $a_i^0, b_i^0 \in \mathcal{O}_E^{\sigma=(-1)^i}$ which is strongly regular semisimple. Suppose that $\mathrm{val}_F(\Delta_{a_0,b_0})$ is even. Let $\delta(a^0)$ be the local Serre invariant associated to the algebra R_{a^0}; that is, $\delta(a^0) = \dim_k(\widetilde{R}_{a^0}/R_{a^0})$, where $\widetilde{R}_{a^0}$ is the normalization of R_{a^0}.

5.1. Local constancy of the local moduli spaces

In this section, we prove an analogous statement to [12, proposition 3.5.1] in our situation. This is a geometric interpretation of Harish-Chandra's theorem on local constancy of orbital integrals.

PROPOSITION 5.1.1

There is an integer $N \geq 1$ (depending on (a^0, b^0)) such that, for any field $\Omega \supset k$ and for any collection of invariants (a, b) with $a_i, b_i \in (\mathcal{O}_E \otimes_k \Omega)^{\sigma=(-1)^i}$, if $(a, b) \equiv (a^0, b^0)$ mod ϖ^N, then

(1) *(a, b) is strongly regular semisimple;*

(2) *$\delta(a) \leq \delta(a^0) + n/2$;*

(3) *there are canonical isomorphisms of schemes over Ω:*

$$\mathcal{M}_{i,a,b}^{\mathrm{loc}} \otimes_k \Omega \cong \mathcal{M}_{i,a^0,b^0}^{\mathrm{loc}} \otimes_k \Omega,$$
$$\mathcal{N}_{a,b}^{\mathrm{loc}} \otimes_k \Omega \cong \mathcal{N}_{a^0,b^0}^{\mathrm{loc}} \otimes_k \Omega.$$

Proof

We stick to the case $\Omega = k$ (the general case is argued in the same way). First of all, by Lemma 2.2.4, the strong regular semisimplicity of (a, b) is checked by the nonvanishing of polynomial equations with $\mathcal{O}_F$-coefficients in a_i, b_i: the discriminant $\mathrm{Disc}(P_a)$ of the polynomial $P_a(t) = t^n - a_1 t^{n-1} + \cdots + (-1)^n a_n$ and the Δ-invariant

$\Delta_{a,b}$. Whenever $(a, b) \equiv (a^0, b^0) \mod \varpi^N$, we have

$$\mathrm{Disc}(P_a) \equiv \mathrm{Disc}(P_{a^0}) \mod \varpi^N,$$
$$\Delta_{a,b} \equiv \Delta_{a^0,b^0} \mod \varpi^N.$$

If we choose $N > \max\{\mathrm{val}_F(\mathrm{Disc}(P_{a^0})), \mathrm{val}_F(\Delta_{a^0,b^0})\}$, then whenever $(a, b) \equiv (a^0, b^0) \mod \varpi^N$, $\mathrm{Disc}(P_a)$ and $\Delta_{a,b}$ are nonzero; hence (a, b) are strongly regular semisimple.

Now fix this choice of N and any (a, b) such that $a_i, b_i \in (\mathcal{O}_E)^{\sigma=(-1)^i}$ and such that $(a, b) \equiv (a^0, b^0) \mod \varpi^N$. Let $\gamma^0 = \gamma_{a^0,b^0}$ and let $\gamma = \gamma_{a,b}$.

By the formula for $\delta(a)$ (see [3] and [12, section 3.7]), we have

$$\delta(a) \leq \frac{\mathrm{val}_F(\mathrm{Disc}(P_a))}{2} = \frac{\mathrm{val}_F(\mathrm{Disc}(P_{a^0}))}{2} \leq \delta(a^0) + \frac{n}{2}.$$

Since $N \geq \mathrm{val}_F(\Delta_{a^0,b^0}) = \mathrm{val}_F(\Delta_{a,b})$, we have

$$R_a^\vee/\gamma(R_a) = (R_a^\vee/\varpi^N R_a^\vee)/\gamma(R_a/\varpi^N R_a), \tag{5.1.1}$$
$$R_{a^0}^\vee/\gamma^0(R_{a^0}) = (R_{a^0}^\vee/\varpi^N R_{a^0}^\vee)/\gamma^0(R_{a^0}/\varpi^N R_{a^0}). \tag{5.1.2}$$

We prove that $\mathcal{M}_{i,a,b}^{\mathrm{loc}}$ and $\mathcal{M}_{i,a^0,b^0}^{\mathrm{loc}}$ are canonically isomorphic. First, we have a canonical isomorphism of $\mathcal{O}_F/\varpi^N\mathcal{O}_F$-algebras

$$\iota : R_a/\varpi^N R_a \cong R_{a^0}/\varpi^N R_{a^0}.$$

We also have a commutative diagram

$$\begin{tikzcd}
R_a/\varpi^N R_a \arrow[r, "\gamma"] \arrow[d, "\iota"] & R_a^\vee/\varpi^N R_a^\vee \\
R_{a^0}/\varpi^N R_{a^0} \arrow[r, "\gamma^0"] & R_{a^0}^\vee/\varpi^N R_{a^0}^\vee, \arrow[u, "\iota^\vee"']
\end{tikzcd}$$

because $\gamma \mod \varpi^N$ depends only on $(a, b) \mod \varpi^N$. Therefore, from (5.1.1) and (5.1.2), we conclude that $R_a^\vee/\gamma(R_a)$ as an $R_a/\varpi^N R_a$-module is canonically isomorphic to $R_{a^0}^\vee/\gamma^0(R_{a^0})$ as an $R_{a^0}/\varpi^N R_{a^0}$-module. Looking back into the definition of $\mathcal{M}_{i,a,b}^{\mathrm{loc}}$, we observe that this scheme canonically depends only on $R_a^\vee/\gamma(R_a)$ as an R_a-module. Therefore we get a canonical isomorphism $\mathcal{M}_{i,a,b}^{\mathrm{loc}} \cong \mathcal{M}_{i,a^0,b^0}^{\mathrm{loc}}$.

The argument for the other isomorphism $\mathcal{N}_{a,b}^{\mathrm{loc}} \cong \mathcal{N}_{a^0,b^0}^{\mathrm{loc}}$ is the same. □

5.2. Preparations

We fix a smooth, projective, and geometrically connected curve X over k of genus g with a k-point x_0. Also fix an étale double cover $\pi : X' \to X$, also geometrically connected, with only one closed point x_0' above x_0. We choose identifications $\mathcal{O}_F \xrightarrow{\sim} \mathcal{O}_{X,x_0}$ and $\mathcal{O}_E \cong \mathcal{O}_{X',x_0'}$.

Fix an integer $\delta \geq \delta(a^0) + n/2$. Fix effect divisors $D = 2D'$ and D_0 on X, disjoint from x_0, such that

$$\begin{aligned} &\deg(D) \geq \max\{c_\delta, 2g + Nn + 1\},\\ &\deg(D_0) \geq \frac{n-1}{2}\deg(D) + 2g + \max\left\{\frac{\delta}{n}, \left(1 + \frac{1}{n}\right)N\right\}. \end{aligned}$$

These numerical assumptions will make sure that all the numerical conditions in the propositions or lemmas of the article (including those which we are about to prove) are satisfied.

For each closed point $x : \operatorname{Spec} k(x) \to X$, let ϖ_x be a uniformizing parameter of $\mathcal{O}_{X,x}$, and let F_x be the field of fractions of $\mathcal{O}_{X,x}$. Let Frob_x be the geometric Frobenius element in $\mathrm{Gal}(\overline{k(x)}/k(x))$. Let E_x be the ring of total fractions of $\mathcal{O}_{X',x}$. Let η_x be the quadratic character of $F_x^\times$ associated to the quadratic extension E_x/F_x.

Recall that the double cover $\pi : X' \to X$ gives a local system L according to Notation 1.6.4. Let $L_x = x^*L$ be the rank one local system on $\operatorname{Spec} k(x)$ given by the pullback of L via $x : \operatorname{Spec} k(x) \to X$. Then we have

$$\eta_x(\varpi_x) = \mathrm{Tr}(\mathrm{Frob}_x, L_x). \tag{5.2.1}$$

Let $\mathcal{A}_0 \times \mathcal{B}_0 \subset \mathcal{A} \times \mathcal{B}$ be the affine subspace consisting of (a, b) such that

$$(a, b) \equiv (a^0, b^0) \mod \varpi^N.$$

LEMMA 5.2.1

Let $\mathcal{A}_0' \subset \mathcal{A}_0 \cap \mathcal{A}^{\leq\delta}$ be the open locus of a such that Y_a is smooth away from $p_a^{-1}(x_0)$. If $\deg(D) \geq 2g + 2Nn + 1$ and if $2\deg(D_0) \geq 2g + N - 1$, then $\mathcal{A}_0'$ and $\mathcal{B}_0$ are nonempty.

Proof

First, we have to make sure that $\mathcal{A}_0$ and $\mathcal{B}_0$ are nonempty. For this it suffices to show that the following evaluation maps at an Nth infinitesimal neighborhood of x are surjective:

$$H^0\big(X, \mathcal{L}(D)^{\otimes i}\big) \to \mathcal{O}_F/\varpi^N, \quad 1 \le i \le n,$$
$$H^0\big(X, \mathcal{O}_X(2D_0) \otimes \mathcal{L}(D)^{\otimes i}\big) \to \mathcal{O}_F/\varpi^N, \quad 0 \le i \le n-1.$$

This is guaranteed as long as $\deg(D) \ge 2g + N - 1$ and $2\deg(D_0) \ge 2g + N - 1$.

Next, we make sure that $\mathcal{A}_0 \cap \mathcal{A}^{\mathrm{int}}$ is nonempty. By Proposition 5.1.1, any $a \in \mathcal{A}_0$ is strongly regular semisimple at x_0, hence, in particular, R_a is reduced. This implies that $\mathcal{A}_0 \subset \mathcal{A}^{\heartsuit}$. By Lemma 3.2.1, we have

$$\mathrm{codim}_{\mathcal{A}^{\heartsuit}}(\mathcal{A}^{\heartsuit} - \mathcal{A}^{\mathrm{int}}) \ge \deg(D) > Nn \ge \mathrm{codim}_{\mathcal{A}^{\heartsuit}}(\mathcal{A}_0).$$

Therefore, $\mathcal{A}_0 \cap \mathcal{A}^{\mathrm{int}} \neq \varnothing$.

Finally, we prove that $\mathcal{A}'_0$ is nonempty. We base change the whole situation to $\overline{k}$. We use the argument for Bertini's theorem (for details, we refer to [12, proposition 4.7.1]). We only point out that for $\deg(D) \ge 2g + N + 1$, the evaluation maps at both x_0 and any other $x \in X(\overline{k})$,

$$\mathrm{ev}(x_0) \oplus \mathrm{ev}(x) : H^0\big(X \otimes_k \overline{k}, \mathcal{L}(D)^{\otimes i}\big) \to \mathcal{O}_F \oplus_k \overline{k}/\varpi^N \bigoplus \mathcal{O}_{X\otimes_k \overline{k}, x}/\varpi_x^2,$$

are also surjective. This is all we need to apply the Bertini argument.

We still have to check that for each $a \in \mathcal{A}'_0$, $\delta(a) \le \delta$. But since Y_a is smooth away from $p_a^{-1}(x_0)$, we have $\delta(a) = \delta(a, x) \le \delta(a^0) + n/2 \le \delta$ by Proposition 5.1.1(2). □

Since $\mathcal{A}'_0$ is an open subset of an affine space and nonempty, it contains a k_m-point for every $m \ge m_0$. Here k_m is the degree m extension of k. Now we fix $m \ge m_0$ and fix a point $a \in \mathcal{A}'_0(k_m)$. We base change the whole situation from k to k_m (in particular, we let $R_{a,m} = R_a \otimes_k k_m$, $E_m = E \otimes_k k_m$, $X_m = X \otimes_k k_m$ and so on). Since D and D_0 are disjoint from x_0, the trivializations that we fixed allow us to get σ-equivariant isomorphisms:

$$R_a(\mathcal{O}_{E_m}) \cong \mathcal{O}_{Y'_a, x_0} \cong \mathcal{O}_{Y'_a}(-D_0)|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}}, \tag{5.2.2}$$

$$R_a^{\vee}(\mathcal{O}_{E_m}) \cong \omega_{Y'_a/X}|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \cong \omega_{Y'_a/X}(D_0)|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}}. \tag{5.2.3}$$

PROPOSITION 5.2.2

If $n(\deg(D_0) - g + 1) \ge (n+1)N + g_Y$, then there exists $b \in \mathcal{B}_0(k_m)$ such that, for each closed point $x \neq x_0$ of X_m, $\mathcal{N}^x_{a_x, b_x}(k_m(x)) \neq \varnothing$.

Proof

We first choose any $b \in \mathcal{B}_0(k_m)$ (which exists because $\mathcal{B}_0 \neq \varnothing$ is an affine space over k). The pair (a, b) determines $\gamma = \gamma_{a,b} : R_{a,m} \hookrightarrow R^{\vee}_{a,m}$.

CLAIM 3
There exists $\mathcal{F}' \in \mathrm{Pic}(Y'_a)$ *and a homomorphism* $h : \mathcal{F}' \hookrightarrow \sigma^*\mathcal{F}'^\vee$ *such that*

(1) $\sigma^* h^\vee = h$;

(2) *h is an isomorphism away from* $p'^{-1}_a(x'_0)$;

(3) *there is an* $R_a(\mathcal{O}_{E_m})$*-linear isomorphism* $\iota : \mathcal{F}'|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \cong R_a(\mathcal{O}_{E_m})$ *(compatible with the identification (5.2.2) such that the following diagram is commutative:*

$$\begin{tikzcd}
\mathcal{F}'|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \arrow[r, "h"] \arrow[d, "\iota"] & \sigma^*\mathcal{F}'^\vee|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \\
R_a(\mathcal{O}_{E_m}) \arrow[r, "\sigma_R\circ\gamma"] & R_a^\vee(\mathcal{O}_{E_m}). \arrow[u, "\sigma^*\iota^\vee"']
\end{tikzcd}$$

Proof
Condition (1) in fact follows from conditions (2) and (3). Since Y'_a is geometrically irreducible, h and $\sigma^*h^\vee$ agree up to a scalar and this scalar must be 1 by (3).

Suppose that $\mathcal{F}'_0$ is a line bundle on Y'_a with an isomorphism $h_0 : \mathcal{F}'_0 \xrightarrow{\sim} \sigma^*\mathcal{F}'^\vee_0$ satisfying $\sigma^*h_0^\vee = h_0$. Such a line bundle exists because we have a Kostant section $\mathcal{A}^{\mathrm{int}} \to \mathcal{N}^{\mathrm{Hit}}|_{\mathcal{A}^{\mathrm{int}}}$ by [10, Section 2.3] when $D = 2D'$. The existence of the Kostant section requires that $\mathrm{char}(k) > n$. Choose an $R_a(\mathcal{O}_{E_m})$-linear isomorphism $\iota_0 : \mathcal{F}'|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \cong R_a(\mathcal{O}_{E_m})$ compatible with the identification (5.2.2). Let $U' = Y'_a - p'^{-1}(x'_0)$, and let U'_{x_0} be the punctured formal neighborhood of $p'^{-1}_a(x'_0)$ in Y'_a. Since $R^\vee_{a,m}$ is an invertible $R_{a,m}$-module, there is a unique invertible element $\beta \in R_a(E_m)$ such that the following diagram is commutative:

$$\begin{tikzcd}
\mathcal{F}'_0|_{U'_{x_0}} \arrow[r, "h_0"] \arrow[d, "\iota_0"] & \sigma^*\mathcal{F}'^\vee_0|_{U'_{x_0}} \\
R_a(\mathcal{O}_{E_m}) \arrow[r, "(\sigma_R\circ\gamma)\cdot\beta^{-1}"] & R_a^\vee(\mathcal{O}_{E_m}). \arrow[u, "\sigma^*\iota_0^\vee"']
\end{tikzcd} \tag{5.2.4}$$

Since both h_0 and γ are Hermitian, we actually have $\beta \in R_a(F_m)$.

Suppose that we can choose a $\mathcal{G} \in \mathrm{Pic}(Y'_a)$ with a map $s : \mathcal{G} \xrightarrow{\sim} \sigma^*\mathcal{G}^{-1}$, which is an isomorphism over U', together with an $R_a(\mathcal{O}_{E_m})$-linear isomorphism,

$$\rho : \mathcal{G}|_{\mathrm{Spec}\,\mathcal{O}_{Y'_a,x_0}} \cong R_a(\mathcal{O}_{E_m}),$$

such that the following diagram is commutative:

$$\begin{array}{ccc} \mathscr{G}|_{U'_{x_0}} & \xrightarrow{s} & \sigma^*\mathscr{G}^{-1}|_{U'_{x_0}} \\ \Big\downarrow \rho & & \Big\uparrow \sigma^*\rho^\vee \\ R_a(E_m) & \xrightarrow{\sigma_R\circ\beta} & R_a^\vee(E_m). \end{array} \tag{5.2.5}$$

Then the triple $(\mathscr{F}' = \mathscr{F}'_0 \otimes \mathscr{G}, h_0 \otimes s, \iota_0 \otimes \rho)$ satisfies the requirement of the claim. Hence our task is to find the triple $(\mathscr{G}, s, \rho)$ such that the diagram (5.2.5) is commutative. We can translate this problem into the language of ideles. Let K (resp., K') be the function field of the geometrically connected curve Y_a (resp., Y'_a) over k_m. Let $\mathbb{A}_K$ (resp., $\mathbb{A}_{K'}$) be the ring of adeles of K (resp., K'). Let $\mathbb{O}_{Y'_a} \subset \mathbb{A}_{K'}$ and $\mathbb{O}_{Y_a} \subset \mathbb{A}_K$ be the product of completions of local rings of Y'_a and Y_a, respectively. Let $\mathrm{Nm} : \mathbb{A}_{K'}^\times \to \mathbb{A}_K^\times$ be the norm map. Then $R_a(E_m)$ is the product of local fields corresponding to places of K over x_0. Thus we get a canonical embedding $R_a(E_m) \subset \mathbb{A}_K$. In particular, we can identify β with an idèle $(\beta_v) \in \mathbb{A}_K^\times$ which is nontrivial only at places $v|x_0$.

A choice of the triple $(\mathscr{G}, s, \rho)$ as above (up to isomorphism) is the same as the choice of an idèle class $\theta \in K'^\times\backslash\mathbb{A}_{K'}^\times/\mathbb{O}_{Y'_a}^\times = \mathrm{Pic}(Y'_a)(k_m)$ such that $\mathrm{Nm}(\theta) = \theta\cdot\sigma\theta = \beta$ as an idèle class in $K^\times\backslash\mathbb{A}_K^\times/\mathbb{O}_{Y_a}^\times$. Let $W_{K'}$ and W_K be the Weil groups of K' and K, respectively. By class field theory, we have the following commutative diagram,

$$\begin{array}{ccccccc} K'^\times\backslash\mathbb{A}_{K'}^\times & \xrightarrow{\mathrm{Nm}} & K^\times\backslash\mathbb{A}_K^\times & \xrightarrow{\partial} & \mathbb{Z}/2\mathbb{Z} & \longrightarrow & 0 \\ \wr\Big\downarrow \mathrm{Art}_{K'} & & \wr\Big\downarrow \mathrm{Art}_K & & \Big\downarrow\wr & & \\ W_{K'}^{\mathrm{ab}} & \longrightarrow & W_K^{\mathrm{ab}} & \longrightarrow & \mathrm{Gal}(K'/K) & \longrightarrow & 1, \end{array}$$

where the map ∂ is defined by

$$\partial(\xi) = \sum_{v \text{ nonsplit}} \mathrm{val}_{K_v}(\xi_v) \mod 2,$$

where v runs over all places of K (which is nonsplit in K') and where k_v is the residue field of $\mathcal{O}_{K_v}$. Now to solve the equation $\mathrm{Nm}(\theta) = \beta$; the only obstruction is $\partial(\beta)$. Since x_0 is nonsplit in X'_m, a place v over x_0 is nonsplit in K' if and only if

$[k_v : k_m]$ is odd. Therefore, we have

$$\partial(\beta) \equiv \sum_{v|x_0,[k_v:k_m]\text{ odd}} \mathrm{val}_{K_v}(\beta_v) \equiv \sum_{v|x_0} \mathrm{val}_{K_v}(\beta_v)[k_v : k_m] \mod 2.$$

From the diagram (5.2.4), we see that $\gamma \cdot \beta^{-1} : R_a(E_m) \xrightarrow{\sim} R_a^\vee(E_m)$ sends $R_a(\mathcal{O}_{E_m})$ isomorphically to $R_a^\vee(\mathcal{O}_{E_m})$. Therefore

$$\begin{aligned}\mathrm{leng}_{\mathcal{O}_{F_m}}\left(R_{a,m} : \beta(R_{a,m})\right) &= \mathrm{leng}_{\mathcal{O}_{F_m}}\left(R_{a,m}^\vee : \gamma(R_{a,m})\right)\\ &= \mathrm{val}_{F_m}(\Delta_{a,b}) = \mathrm{val}_F(\Delta_{a^0,b^0})\end{aligned}$$

is even. Hence

$$\sum_{v|x_0} \mathrm{val}_{K_v}(\beta_v)[k_v : k_m] = \mathrm{leng}_{\mathcal{O}_{F_m}}\left(R_{a,m} : \beta(R_{a,m})\right) \tag{5.2.6}$$

is also even. This shows the vanishing of $\partial(\beta)$ in $\mathbb{Z}/2\mathbb{Z}$. Therefore we can always find $\theta \in K'^\times\backslash\mathbb{A}_{K'}^\times$ such that $\mathrm{Nm}(\theta) = \beta$. Translating back into geometry, we have found the desired $(\mathcal{G}, s, \rho)$, and hence the desired $(\mathcal{F}', h, \iota)$. □

Now we pick such a triple $(\mathcal{F}', h, \iota)$ from Claim 3. By construction, we have

$$\chi(Y_a', \mathcal{F}') = -2n(g-1) - \frac{\mathrm{val}_{F_m}(\Delta_{a,b})}{2}.$$

CLAIM 4
There is a homomorphism $\alpha' : \mathcal{O}_{Y_a'}(-D_0) \to \mathcal{F}'$ such that the composition

$$\iota \circ \alpha' : R_a(\mathcal{O}_{E_m}) \cong \mathcal{O}_{Y_a',x_0} \xrightarrow{\alpha'} \mathcal{F}'|_{\mathrm{Spec}\,\mathcal{O}_{Y_a',x_0}} \xrightarrow{\iota} R_a(\mathcal{O}_{E_m})$$

is the identity modulo ϖ^N.

Proof
Consider the following evaluation map at the Nth infinitesimal neighborhood of $p_a'^{-1}(x_0') \subset Y_a'$:

$$\mathrm{ev} : \mathcal{F}'(D_0) \to \mathcal{F}'(D_0)|_{\mathrm{Spec}\,\mathcal{O}_{Y_a',x_0}} \xrightarrow{\iota} R_a(\mathcal{O}_{E_m}) \otimes_{\mathcal{O}_F} (\mathcal{O}_F/\varpi^N).$$

Let $\mathcal{K}$ be the kernel of ev, which is a coherent sheaf on Y_a'. By Grothendieck-Serre duality, we have

$$H^1(Y_a', \mathcal{K}) \cong \mathrm{Hom}_{Y_a'}(\mathcal{K}, \omega_{Y_a'})^\vee.$$

But since

$$\begin{aligned}\chi(Y'_a,\mathcal{K}) &= \chi\big(Y'_a,\mathcal{F}'(D_0)\big)-2nN\\ &= \frac{-\operatorname{val}_{F_m}(\Delta_{a,b})}{2}+2n\big(\deg(D_0)-g+1\big)-2nN\\ &\geq 2n\big(\deg(D_0)-g+1\big)-(2n+1)N\\ &\geq 2g_Y > g'_Y-1=\chi(Y'_a,\omega_{Y'_a}),\end{aligned}$$

we must have $\operatorname{Hom}(\mathcal{K},\omega_{Y'_a})=0$. Therefore, $H^1(Y'_a,\mathcal{K})=0$. This implies that

$$\operatorname{ev}:\operatorname{Hom}_{Y'_a}\big(\mathcal{O}_{Y'_a}(-D_0),\mathcal{F}'\big)=H^0\big(Y'_a,\mathcal{F}'(D_0)\big)\to R_a(\mathcal{O}_{E_m})\otimes_{\mathcal{O}_F}(\mathcal{O}_F/\varpi^N)$$

is surjective. Hence there exists $\alpha'\in\operatorname{Hom}_{Y'_a}(\mathcal{O}_{Y'_a}(-D_0),\mathcal{F}')$ such that $\iota\circ\operatorname{ev}(\alpha')\equiv 1 \mod \varpi^N$. □

Now let γ' be the composition

$$\mathcal{O}_{Y'_a}(-D_0)\xrightarrow{\alpha'}\mathcal{F}'\xrightarrow{h}\sigma^*\mathcal{F}'^{\vee}\xrightarrow{\sigma^*\alpha'^{\vee}}\omega_{Y'_a/X}(D_0).$$

Then γ' gives back another $b'\in\mathcal{B}(k_m)$. From the construction it is clear that $b'\equiv b \mod \varpi^N$, and therefore $b'\in\mathcal{B}_0(k_m)$. Now for each closed point $x\neq x_0$ of X_m, the local moduli space $\mathcal{N}^x_{a_x,b'_x}(k_m(x))$ is nonempty because it contains a point given by $\mathcal{F}'|_{\operatorname{Spec}\mathcal{O}_{Y'_a,x}}$. (This is self-dual because h is an isomorphism over $x\neq x_0$ by construction). Therefore, the pair (a,b') satisfies the requirement of Proposition 5.2.2. □

5.3. The proof

Now for each $m\geq m_0$, we have a pair $(a,b)\in\mathcal{A}'_0(k_m)\times\mathcal{B}_0(k_m)$ such that the condition in Proposition 5.2.2 holds. Using Theorem 4.4.1 (taking the stalks of the two complexes in (4.4.1) at the point (a,b)), we get an isomorphism of graded Frob_k^m modules:

$$\bigoplus_{i=-d}^{d}H^\bullet(\mathcal{M}_{i,a,b}\otimes_{k_m}\overline{k},L_{d-i})\cong H^\bullet(\mathcal{N}_{a,b}\otimes_{k_m}\overline{k},\overline{\mathbb{Q}}_\ell). \tag{5.3.1}$$

By the product formulas (3.4.1) and (3.4.2), we can rewrite (5.3.1) as

$$\begin{aligned}&\bigotimes_{x\in|X_m|}\Big(\bigoplus_{i_x}H^\bullet\big((\operatorname{Res}_{k_m(x)/k_m}\mathcal{M}^x_{i_x,a_x,b_x})\otimes_{k_m}\overline{k},\operatorname{Res}_{k_m(x)/k_m}L_x^{\otimes i_x}\big)\Big)\\ &\quad\cong\bigotimes_{x\in|X_m|}H^\bullet\big((\operatorname{Res}_{k_m(x)/k_m}\mathcal{N}^x_{a_x,b_x})\otimes_{k_m}\overline{k},\overline{\mathbb{Q}}_\ell\big).\end{aligned}$$

Here $\mathrm{Res}_{k_m(x)/k_m} L_x$ is the local system of rank one on $\mathrm{Res}_{k_m(x)/k_m} \mathrm{Spec}\, k_m(x)$ induced from L_x. We have such a tensor product decomposition on the left-hand side because the local system L_{d-i} on $\mathcal{M}_{i,a,b}$, when pulled back via the isomorphism (3.4.1), becomes $\boxtimes_{x\in|X_m|} \mathrm{Res}_{k_m(x)/k_m} L_x^{\otimes i_x}$ on $\prod_{x\in|X_m|} \mathrm{Res}_{k_m(x)/k_m} \mathcal{M}^x_{i_x,a_x,b_x}$ (see Lemma 4.1.2).

We use the abbreviations

$$\begin{aligned} M_x^j &:= \bigoplus_{i_x} H^j\big((\mathrm{Res}_{k_m(x)/k_m} \mathcal{M}^x_{i_x,a_x,b_x}) \otimes_{k_m} \overline{k}, \mathrm{Res}_{k_m(x)/k_m} L_x^{\otimes i_x}\big), \\ N_x^j &:= H^j\big((\mathrm{Res}_{k_m(x)/k_m} \mathcal{N}^x_{a_x,b_x}) \otimes_{k_m} \overline{k}, \overline{\mathbb{Q}}_\ell\big), \\ M_0^j &:= \bigoplus_{i=0}^{\mathrm{val}_F(\Delta_{a^0,b^0})} H^j\big(\mathcal{M}^{\mathrm{loc}}_{i,a^0,b^0} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell(\eta_{k'/k})^{\otimes i}\big), \\ N_0^j &:= H^j(\mathcal{N}^{\mathrm{loc}}_{a^0,b^0} \otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell), \end{aligned}$$

where the first two are Frob_k^m-modules and where the last two are Frob_k-modules.

By Proposition 5.1.1, we have

$$M^j_{x_0} \cong M^j_0 \quad \text{and} \quad N^j_{x_0} \cong N^j_0$$

as Frob_k^m-modules.

On the other hand, since we have assumed that Y_a is smooth away from $p_a^{-1}(x_0)$ for any $x \neq x_0$, the local moduli spaces $\mathcal{M}^x_{i_x,a_x,b_x}$ and $\mathcal{N}^x_{a_x,b_x}$ are zero-dimensional, and hence have no higher cohomology. For $x \neq x_0$, we write M_x (resp., N_x) for M_x^0 (resp., N_x^0). Therefore, for each j, we get an isomorphism of Frob_k^m-modules

$$M_0^j \otimes \Big(\bigotimes_{x_0\neq x\in|X_m|} M_x\Big) \cong N_0^j \otimes \Big(\bigotimes_{x_0\neq x\in|X_m|} N_x\Big).$$

Taking the traces of Frob_k^m and using the Lefschetz trace formula for $\mathcal{M}^x_{i_x,a_x,b_x}$ and $\mathcal{N}^x_{a_x,b_x}$, we get, for any $j \geq 0$,

$$\mathrm{Tr}(\mathrm{Frob}_k^m, M_0^j) \prod_{x_0\neq x\in|X_m|} \Big(\sum_{i_x} \eta_x(\varpi_x)^{i_x} \#\mathcal{M}^x_{i_x,a_x,b_x}\big(k_m(x)\big)\Big) \tag{5.3.2}$$

$$= \mathrm{Tr}(\mathrm{Frob}_k^m, N_0^j) \prod_{x_0\neq x\in|X_m|} \#\mathcal{N}^x_{a_x,b_x}\big(k_m(x)\big). \tag{5.3.3}$$

Here we used (5.2.1). Since Y_a is smooth away from $p_a^{-1}(x_0)$, for $x \neq x_0$, $R_{a_x} \cong \mathcal{O}_{Y_a,x}$ is a product of DVRs, and we can apply Lemma 2.5.5 to conclude that

$$\sum_{i_x} \eta_x(\varpi_x)^{i_x} \#\mathcal{M}^x_{i_x,a_x,b_x}\big(k_m(x)\big) = \#\mathcal{N}^x_{a_x,b_x}\big(k_m(x)\big), \quad \forall x \neq x_0.$$

Moreover, the right-hand side for each $x \neq x_0$ is nonzero because (a, b) satisfies the condition in Proposition 5.2.2. Therefore (5.3.2) implies that

$$\mathrm{Tr}(\mathrm{Frob}_k^m, M_0^j) = \mathrm{Tr}(\mathrm{Frob}_k^m, N_0^j), \quad \forall j \geq 0.$$

Since this is true for any $m \geq m_0$, we get an isomorphism of semisimplified Frob_k-modules:

$$M_0^{j,\mathrm{ss}} \cong N_0^{j,\mathrm{ss}}.$$

But by Lemma 2.7.3, M_0^j and N_0^j are isomorphic as Frob_k^2-modules. Therefore we can conclude that $M_0^j \cong N_0^j$ as Frob_k-modules since the unipotent part of the Frob_k action is uniquely determined by that of Frob_k^2 by taking the square root. This proves the main theorem.

Acknowledgments. The author would like to thank Xinyi Yuan and Wei Zhang for suggesting the problem and for helpful discussions. He also thanks Joseph Bernstein, Julia Gordon, Hervé Jacquet, Yifeng Liu, Shou-Wu Zhang, and the referees for useful suggestions. Different parts of this work were conceived during the workshop on Trace Formula held at Columbia University in December 2008, and the AIM workshop on Relative Trace Formula and Periods of Automorphic Forms in August 2009. The author would like to thank the organizers of these workshops.

References

[1] A. B. ALTMAN and S. L. KLEIMAN, *Compactifying the Picard scheme*, Adv. Math. **35** (1980), 50 – 112. MR 0555258

[2] A. BEAUVILLE, M. S. NARASIMHAN, and S. RAMANAN, *Spectral curves and the generalized theta divisor*, J. Reine Angew. Math. **398** (1989), 169 – 179. MR 0998478

[3] R. BEZRUKAVNIKOV, *The dimension of the fixed point set on affine flag manifolds*, Math. Res. Lett. **3** (1996), 185 – 189. MR 1386839

[4] I. BISWAS and S. RAMANAN, *An infinitesimal study of the moduli of Hitchin pairs*, J. London Math. Soc. **49** (1994), 219 – 231. MR 1260109

[5] M. GORESKY, R. KOTTWITZ, and R. MACPHERSON, *Homology of affine Springer fibers in the unramified case*, Duke Math. J. **121** (2004), 509 – 561. MR 2040285

[6] ———, *Codimensions of root valuation strata*, Pure Appl. Math. Q. **5** (2009), 1253 – 1310. MR 2560317

[7] M. GORESKY and R. MACPHERSON, *Intersection homology, II*, Invent. Math. **72** (1983), 77 – 129. MR 0696691

[8] A. GROTHENDIECK, "Techniques de construction et théorèmes d'existence en géométrie algébrique, IV: Les schémas de Hilbert," *Séminaire Bourbaki, Vol. 6,* no. 221, Soc. Math. France, Paris, 1995, 249 – 276. MR 1611822

[9] H. JACQUET and S. RALLIS, *On the Gross-Prasad conjecture for the unitary group in 3 variables*, http://www.math.columbia.edu/hj.

[10] G. LAUMON and B.-C. NGÔ, *Le lemme fondamental pour les groupes unitaires*, Ann. of Math. (2) **168** (2008), 477 – 573. MR 2434884

[11] B.-C. NGÔ, *Fibration de Hitchin et endoscopie*, Invent. Math. **164** (2006), 399 – 453. MR 2218781

[12] ———, *Le lemme fondamental pour les algebres de Lie*, Publ. Math. Inst. Hautes Études Sci. **111** (2010), 1 – 169. MR 2653248

[13] W. ZHANG, *On the refined Gross-Prasad conjecture for unitary groups in three variables*, preprint, 2009.

Appendix. Transfer to characteristic zero

JULIA GORDON

The purpose of this appendix is to point out that the work of Cluckers, Hales, and Loeser [1] implies that the transfer principle of Cluckers and Loeser [4] applies to the version of the fundamental lemma proved in this article. Thus, Conjectures 1.1.1 and 1.1.2 are true when F is a local field of characteristic zero with sufficiently large residue characteristic.

We need to emphasize that even though these conjectures in the equal characteristic case are proved here for the fields F of characteristic larger than n, the transfer principle leads to a slightly weaker result for the fields of characteristic zero—namely, that there exists (an algorithmically computable) constant M such that the conjectures hold for the characteristic zero local fields F of residue characteristic larger than M. However, we hope that such a result is sufficient for some applications.

Since this appendix is of an expository nature, the references often point not to the original sources, but to more expository articles. All the references to specific sections, conjectures, and definitions that do not mention a source are to Yun's article, for which this appendix is written.

A.1. Denef-Pas language

The idea behind the approach to transfer described here is to express everything involved in the statement of the fundamental lemma by means of formulas in a certain first-order language of logic (called the *Denef-Pas language*) $\mathcal{L}_{\mathrm{DP}}$ (see, e.g., [1, Section 1.6] for the detailed definition), and then work with these formulas directly instead of with the sets and functions described by them. Denef-Pas language is designed for valued fields. It is a *three-sorted language*, meaning that it has three sorts of variables. Variables of the first sort run over the valued field; variables of the second sort run over the value group (for simplicity, we shall assume that the value group is $\mathbb{Z}$); and variables of the third sort run over the residue field.

Let us describe the set of symbols that, along with parentheses, the binary relation symbol $=$ in every sort, the standard logical symbols for conjunction, disjunction, and negation, and the quantifiers, are used to build formulas in Denef-Pas language.

- In the valued field sort, there are constant symbols 0 and 1, and the symbols $+$ and $\times$ for the binary operations of addition and multiplication. Additionally, there are symbols for two functions from the valued field sort: $\mathrm{ord}(\cdot)$ to denote a function from the valued field sort to the $\mathbb{Z}$-sort, and $\overline{\mathrm{ac}}(\cdot)$ to denote a function from the valued field sort to the residue field sort. These functions are called the *valuation map* and the *angular component map*, respectively.
- In the residue field sort, there are constant symbols 0 and 1, and the binary operations symbols $+$ and $\times$ (thus, restricted to the residue field sort, this is the language of rings).
- In the $\mathbb{Z}$-sort, there are 0 and 1, and the operation $+$. Additionally, for each $d = 2, 3, 4, \ldots$, there is a symbol $\equiv_d$ to denote the binary relation $x \equiv y \bmod d$. Finally, there is a binary relation symbol $\geq$. (This is Presburger language for the integers).

Given a discretely valued field K with a uniformizer of the valuation ϖ, the functions $\mathrm{ord}(\cdot)$ and $\overline{\mathrm{ac}}(\cdot)$ are interpreted as follows. The function $\mathrm{ord}(x)$ stands for the valuation of x. It is in order to provide the interpretation for the symbol $\overline{\mathrm{ac}}(x)$ that a choice of the uniformizing parameter ϖ (so that $\mathrm{ord}(\varpi) = 1$) is needed. If $x \in \mathcal{O}_K^*$ is a unit, then there is a natural definition of $\overline{\mathrm{ac}}(x)$; it is the reduction of x modulo the ideal (ϖ). For $x \neq 0$ in K, $\overline{\mathrm{ac}}(x)$ is defined by $\overline{\mathrm{ac}}(x) = \overline{\mathrm{ac}}(\varpi^{-\mathrm{ord}(x)}x)$ and, by definition, $\overline{\mathrm{ac}}(0) = 0$.

A formula φ in $\mathcal{L}_{\mathrm{DP}}$ can be interpreted in any discretely valued field (once a uniformizer of the valuation is chosen) in the sense that, given a valued field K with a uniformizer ϖ and the residue field k_K, one can allow the free variables of φ to range over K, k_K, and $\mathbb{Z}$, respectively, according to their sort (naturally, the variables bound by a quantifier then also range over K, k_K, and $\mathbb{Z}$, respectively). Thus, any discretely valued field is a structure for Denef-Pas language.

A.2. Constructible motivic functions

In the foundational articles [3] and [4], Cluckers and Loeser developed the theory of motivic integration for functions defined by means of formulas in Denef-Pas language, and proved a very general transfer principle. We refer to [1] and [2] for the introduction to this subject and all definitions (note that the article [1] is self-contained and essentially covers everything in this appendix).

Here we need to use the terms *definable subassignment* and *constructible motivic function*. Let $h[m, n, r]$ be the functor from the category of fields to the category of sets defined by

$$h[m, n, r](K) = K((t))^m \times K^n \times \mathbb{Z}^r.$$

The term *subassignment* was first introduced in [5]. Given a functor F from some category C to Sets, a subassignment X of F is a collection of subsets $X(A) \subset F(A)$ for every object A of C. A definable set is a set that can be described by a formula in Denef-Pas language, and a subassignment X of the functor $h[m, n, r]$ is called *definable* if there exists a formula φ in Denef-Pas language with m free variables of the valued field sort, n free variables of the residue field sort, and r free variables of the value group sort such that, for every field K, the set $X(K)$ is exactly the set of points in $K((t))^m \times K^n \times \mathbb{Z}^r$ where φ takes the value "true." Note that there are slightly different variants of Denef-Pas language, depending on the sets of coefficients for a formula φ allowed in every sort (the smallest set of coefficients is $\mathbb{Z}$ in every sort; however, one can add constant symbols that can later be used as coefficients—one such variant will be discussed below). We emphasize, however, that regardless of the variant, the coefficients come from a fixed set and are independent of K. Definable subassignments form a Boolean algebra in a natural way, and this algebra is the replacement, in the theory of motivic integration, for the Boolean algebra of measurable sets in the traditional measure theory.

For a definable subassignment X, the ring of the so-called constructible motivic functions on X, denoted by $\mathcal{C}(X)$, is defined in [3]. The elements of $\mathcal{C}(X)$ are, essentially, formal constructions defined using the language $\mathcal{L}_{\mathrm{DP}}$. The main feature of constructible motivic functions is specialization to functions on discretely valued fields. Namely, let $f \in \mathcal{C}(X)$. Let F be a non-Archimedean local field (either of characteristic zero or of positive characteristic). Let ϖ be the uniformizer of the valuation on F. Given these data, one gets a specialization X_F of the subassignment X to F, which is a definable subset of $F^m \times k_F^n \times \mathbb{Z}^r$ for some m, n, r, and the constructible motivic function f specializes to a $\mathbb{Q}$-valued function f_F on X_F, for all fields F of residue characteristic bigger than a constant that depends only on the $\mathcal{L}_{\mathrm{DP}}$-formulas defining f and X. As explained in [1, Section 2.9], one can tensor the ring $\mathcal{C}(X)$ with $\mathbb{C}$, and then the specializations f_F of elements of $\mathcal{C}(X) \otimes \mathbb{C}$ form a $\mathbb{C}$-algebra of functions on X_F.

A.3. Integration and transfer principle

In [3], Cluckers and Loeser defined a class $\mathrm{I}C(X)$ of integrable constructible motivic functions closed under integration with respect to parameters (where integration is with respect to the motivic measure). Given a local field F with a choice of the uniformizer, these functions specialize to integrable (in the classical sense) functions on X_F, and motivic integration specializes to the classical integration with respect to an appropriate Haar measure when the residue characteristic of F is sufficiently large.

From now on, we will use the variant of the theory of motivic integration with coefficients in the ring of integers of a given global field. Let Ω be a global field with the ring of integers $\mathcal{O}$. Following [2], we denote by $\mathcal{A}_{\mathcal{O}}$ the collection of all p-adic completions of all finite extensions of Ω, and we denote by $\mathcal{B}_{\mathcal{O}}$ the set of all local fields of positive characteristic that are $\mathcal{O}$-algebras. Let $\mathcal{A}_{\mathcal{O},M}$ (resp., $\mathcal{B}_{\mathcal{O},M}$) be the set of all local fields F in $\mathcal{A}_{\mathcal{O}}$ (resp., $\mathcal{B}_{\mathcal{O}}$) such that the residue field k_F has characteristic larger than M. Let $\mathcal{L}_{\mathcal{O}}$ be the variant of Denef-Pas language with coefficients in $\mathcal{O}[[t]]$ (see [2, Section 6.7] for the precise definition). This means, roughly, that a constant symbol for every element of $\mathcal{O}[[t]]$ is added to the valued field sort so that a formula in $\mathcal{L}_{\mathcal{O}}$ is allowed to have coefficients in $\mathcal{O}[[t]]$ in the valued field sort, coefficients in Ω in the residue field sort, and coefficients in $\mathbb{Z}$ in the value group sort.

Then the transfer principle can be stated as follows.

THEOREM A.1 (Abstract transfer principle; see [1, Theorem 2.7.2.])
Let X be a definable subassignment, and let φ be a constructible (with respect to the language $\mathcal{L}_{\mathcal{O}}$) motivic function on X. Then there exists $M > 0$ such that, for every $K_1, K_2 \in \mathcal{A}_{\mathcal{O},M} \cup \mathcal{B}_{\mathcal{O},M}$ with $k_{K_1} \simeq k_{K_2}$,

$$\varphi_{K_1} = 0 \quad \text{if and only if} \quad \varphi_{K_2} = 0.$$

Remark A.2
In fact, the transfer principle is proved in [4] for an even richer class of functions, called *constructible motivic exponential functions*, that contain additive characters of the field along with the constructible motivic functions. However, we do not discuss it here since the characters are not needed in the present setting.

The goal of this appendix is to check that the conjectures proved in this article can be expressed as equalities between specializations of constructible motivic functions. We emphasize that all the required work is actually done in [1]; here we just check that it indeed applies in the present situation.

A.4. Definability of all the ingredients

Here we go through Section 2.1 and check that every object appearing in it is definable.

A.4.1. The degree 2 algebra E/F
Following [1, Section 4], we fix, once and for all, a $\mathbb{Q}$-vector space V of dimension n, and we fix the basis $e_0, \ldots, e_{n-1}$ of V over $\mathbb{Q}$.

As in [1, Section 3.2], we introduce a parameter (which we denote by ϵ) that will appear in all the formulas that involve an unramified quadratic extension of the base field. We think of ϵ as a nonsquare unit, and we denote by Λ be the subassignment of $h[1, 0, 0]$ defined by the formula $\mathrm{ord}(\epsilon) = 0 \wedge \nexists x : x^2 = \epsilon$. From now on, we only consider the relative situation: all the subassignments we consider will come with a fixed projection morphism to Λ (in short, we are considering the category of definable subassignments over Λ; see [3, Section 2.1]). That is, we replace all the constructions that depend on an unramified quadratic extension E/F (such as the unitary group), with the family of isomorphic objects parameterized by a nonsquare unit ϵ in F. Now imagine that we fixed the basis $(1, \sqrt{\epsilon})$ for the quadratic extension E. Then E can be identified with F^2 via this basis, so from now on we shall think of the elements of E as pairs of variables that range over F. The nontrivial Galois automorphism σ of E over F now can be expressed as a (2×2)-matrix with entries in F and can be used in the expressions in Denef-Pas language.

The nontrivial quadratic character $\eta_{E/F}$ can be expressed by a Denef-Pas formula $\eta_{E/F}(x) = 1 \Leftrightarrow \exists (a, b) \in F^2 : (a^2 + \epsilon b^2 = x)$, or simply by $\eta_{E/F}(x) = 1 \Leftrightarrow \mathrm{ord}(x) \equiv 0 \mod 2$.

In the case E/F split, we just treat elements of E as pairs of elements of F.

A.4.2. The groups and their Lie algebras

In Sections 2.1 and 2.2, one starts out with free $\mathcal{O}_F$-modules W and V and then proceeds to choose a basis vector e_0 with certain properties. We shall reverse the thinking here. We fix a basis $e_0, \dots, e_{n-1}$, and we fix the dual basis $e_0^*, \dots, e_{n-1}^*$ such that $e_0^*(e_0) = 1$ and such that the Hermitian form $(\cdot, \cdot)$ on V, with respect to this basis, corresponds to a matrix with entries in the set $\{0, \pm 1\}$, and where $(e_0, e_0) = 1$ and $(e_0, e_j) = 0$ for $1 \le j \le n-1$. We let W be the span of the vectors $e_1, \dots, e_{n-1}$. With this choice of basis, we think of the elements of $\mathfrak{gl}_n$ as n^2-tuples of variables $A = (a_{ij})$. (Formally speaking, we identify $\mathfrak{gl}_n$ with the definable subassignment $h[n^2, 0, 0]$.) All the split algebraic groups are naturally defined by polynomial equations in these variables and thus can be replaced with definable subassignments of $h[n^2, 0, 0]$. The embedding $\mathrm{GL}_{n-1} \hookrightarrow \mathrm{GL}_n$, where

$$A \mapsto \begin{pmatrix} A & \\ & 1 \end{pmatrix},$$

is clearly definable.

To find the definable subassignments that specialize to $\mathfrak{s}_n$, $\mathfrak{u}_n$, and U_n, we introduce the parameter ϵ as in Section A.4.1. Then $\mathfrak{s}_n$ naturally becomes a definable subassignment of $h[2n^2, 0, 0] \times \Lambda \subset h[2n^2 + 1, 0, 0]$. Indeed, as discussed above, the Galois automorphism σ can be used in $\mathcal{L}_{\mathrm{DP}}$-expressions when we think of the elements of E as pairs of F-variables: we replace each variable a_{ij} ranging over E

with a pair of variables (x_{ij}, y_{ij}) ranging over F. The Hermitian form that is used to define the unitary group, given the choice of the basis, gives rise to polynomial equations in (x_{ij}, y_{ij}) that define the unitary group. Hence, $\mathfrak{u}_n$ and U_n can also be replaced with definable subassignments of $h[2n^2+1, 0, 0]$.

A.4.3. The invariants

By definition, $a_i(A)$ are the coefficients of the characteristic polynomial of A. In particular, they are polynomial expressions in the matrix entries of A, and therefore they are given by terms in $\mathcal{L}_{\mathrm{DP}}$, and the map $A \mapsto (a_i(A))_{1\le i\le n}$ is definable (recall that a function is called *definable* if its graph is a definable set).

First, let us consider the case when E/F is a field extension.

The linear functional e_0^* on V (defined in Section 2.1) with our choice of the bases is just the covector $(1, 0, \dots, 0)$. Then the invariants $b_i(A)$ of Section 2.2 are also given by terms in $\mathcal{L}_{\mathrm{DP}}$.

The vectors $A^i e_0$ are clearly just columns of polynomial expressions in the matrix entries (x_{ij}, y_{ij}) of A. The condition that a collection of vectors forms a basis of a given vector space is a predicate in $\mathcal{L}_{\mathrm{DP}}$. Hence, the set of semisimple elements in $\mathfrak{gl}_n(E)$ that are strongly regular with respect to $\mathrm{GL}_{n-1}(E)$-action (in the sense of Definition 2.2.1) is a specialization (to F) of a definable subassignment of $h[2n^2+1, 0, 0]$.

We observe that $\Delta_{a,b} = \det(e_0^* A^{i+j} e_0)_{0\le i,j,\le n-1}$ (of Definition 2.2.3) is also a polynomial expression in (x_{ij}, y_{ij}).

Recall the subassignment Λ from Section A.4.2 that specializes to the domain for a parameter ϵ defining the extension E. Since the image of a definable subassignment under a definable morphism is a definable subassignment, we have the definable subassignment $\mathcal{P}$ over Λ, which we denote by $\mathcal{P} \to \Lambda$, that corresponds to the set of pairs $(a, b) \in E^{2n}$ that are invariants of some strongly regular element of $\mathfrak{gl}_n(E)$. More precisely, $\mathcal{P}$ is a subassignment of $\Lambda \times h[4n, 0, 0]$ that satisfies the condition that there exists $N > 0$ such that, for every local field $F \in \mathcal{A}_{\mathcal{O},N} \cup \mathcal{B}_{\mathcal{O},N}$ and for every $\epsilon \in \Lambda_F$, the fiber $\mathcal{P}_\epsilon$ of $\mathcal{P}$ at ϵ specializes to the set of pairs (a, b) that are invariants of some $A \in \mathfrak{gl}_n(E)$, strongly regular with respect to $\mathrm{GL}_{n-1}(E)$-action (in the sense of Definition 2.2.1), where E is the field extension corresponding to ϵ.

If E/F is split, the same argument works, except that there is no need to consider the relative situation over Λ.

Since we have a symbol for the F-valuation in $\mathcal{L}_{\mathrm{DP}}$, the parameter $\nu(A)$ of Definition 2.2.2 is also an expression in $\mathcal{L}_{\mathrm{DP}}$.

A.4.4. The orbital integrals

Since the quadratic character $\eta_{E/F}$ only takes the values ± 1, we can break the orbital integral $\mathrm{O}_A^{\mathrm{GL}_{n-1},\eta}\left(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}\right)$ into the difference of two integrals:

$$\mathrm{O}_A^{\mathrm{GL}_{n-1},\eta}\left(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}\right) = \int_{\mathrm{GL}_{n-1}(F)\cap\{g\mid\eta_{E/F}(\det g)=1\}} \mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}(g^{-1}Ag)\,dg \\ - \int_{\mathrm{GL}_{n-1}(F)\cap\{g\mid\eta_{E/F}(\det g)=-1\}} \mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}(g^{-1}Ag)\,dg.$$

By the remarks in Section A.4.1 above, both domains of integration are definable sets. For each point A in the subassignment of strongly regular elements, $\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}(g^{-1}Ag)$ is, by Section A.4.2 above, a specialization of a constructible motivic function of g. We need to briefly discuss the normalization of the measures. The p-adic measure to which the motivic measure specializes is the so-called Serre-Oesterlé measure (defined in [7]). The Serre-Oesterlé measure on a classical group G is the Haar measure such that the volume of the maximal compact subgroup is $q^{\dim G}$. Hence, the Haar measure dg differs from the Serre-Oesterlé measure by a factor of $q^{-(n-1)^2}$, where q is the cardinality of the residue field, since, as in [6], the Haar measures here are chosen so that the standard maximal compact subgroups have volume 1. This factor is the specialization of the (constant) constructible motivic function $\mathbb{L}^{-(n-1)^2}$ (see, e.g., [1, Section 2.3] for the discussion of the symbol $\mathbb{L}$). We conclude that $\mathrm{O}_A^{\mathrm{GL}_{n-1},\eta}\left(\mathbf{1}_{\mathfrak{s}_n(\mathcal{O}_F)}\right)$ is a specialization of a constructible motivic function of A.

By a similar inspection, we see that the integral $\mathrm{O}_{A'}^{U_{n-1}}\left(\mathbf{1}_{\mathfrak{u}_n(\mathcal{O}_F)}\right)$ is a specialization of a constructible motivic function of A', and thus so is the right-hand side of Conjecture 1.1.1(1).

Finally, recall the subassignment $\mathcal{P}$ from Section A.4.3 that specializes to the set of invariants. Consider the subassignment $\mathcal{X}$ of $\mathfrak{s}_n \times \mathfrak{u}_n$ defined by $(A, A') \in \mathcal{X}$ if and only if A and A' have the same invariants. Since as we discussed above, the map that maps A to its collection of invariants is a definable map, this is a definable subassignment (note that it has a natural projection to $\mathcal{P}$). We have shown that the difference of the left-hand side and the right-hand side of equation (1) in Conjecture 1.1.1 is a constructible motivic function on $\mathcal{X}$. Therefore, the transfer principle applies to it.

By inspection, all the ingredients of all the other variants of Conjecture 1.1.1 and Conjecture 1.1.2 are definable in the language $\mathcal{L}_{\mathcal{O}}$, and hence the transfer principle applies in all these cases.

Acknowledgments. This appendix emerged as a result of the AIM workshop on Relative Trace Formula and Periods of Automorphic Forms in August 2009. It is a pleasure to thank the organizers and participants of this workshop. I would like to emphasize that all the original ideas used and described here appear in the works of R. Cluckers, T. C. Hales, and F. Loeser. I am grateful to R. Cluckers for a careful reading.

Appendix references

[1] R. CLUCKERS, T. HALES, and F. LOESER, *Transfer principle for the fundamental lemma*, preprint, arXiv:0712.0708v1 [math.RT]

[2] R. CLUCKERS and F. LOESER, "Ax-Kochen-Eršov theorems for p-adic integrals and motivic integration" in *Geometric Methods in Algebra and Number Theory*, Progr. Math. **235**, Birkhaüser, Boston, 2005, 109 – 137. MR 2159379

[3] ———, *Constructible motivic functions and motivic integration*, Invent. Math. **173**, (2008) 23 – 121. MR 2403394

[4] ———, *Constructible exponential functions, motivic Fourier transform and transfer principle*, Ann. of Math. (2) **171** (2010), 1011 – 1065. MR 2630060

[5] J. DENEF, and F. LOESER, *Definable sets, motives and p-adic integrals*, J. Amer. Math. Soc. **14** (2001), 429 – 469. MR 1815218

[6] H. JACQUET and S. RALLIS, *On the Gross-Prasad conjecture for the unitary group in* 3 *variables*, http://www.math.columbia.edu/hj.

[7] J. OESTERLÉ, *Réduction modulo p^n des sous-ensembles analytiques fermés de $\mathbb{Z}_p^N$*, Invent Math. **66** (1982), 325 – 341. MR 0656627

Yun
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; zyun@math.mit.edu

Gordon
Department of Mathematics, University of British Columbia, Vancouver, British Columbia V6T 1Z2, Canada; gor@math.ubc.ca