More Arithmetic Fundamental Lemma conjectures: the case of Bessel subgroups

Wei Zhang[*]

Abstract: We define some formal moduli spaces of quasi-isogenies of isoclinic p-divisible groups with a non-reductive group as the "structure group". We then formulate new Arithmetic Fundamental Lemma conjectures for Bessel subgroups in the context of the arithmetic Gan–Gross–Prasad conjectures. Some (very limited) evidence is presented.

Keywords: Local Shimura variety, Arithmetic Gan–Gross–Prasad conjecture, Arithmetic Fundamental Lemma, Bessel cycle, orbital integral.

1. Introduction

For a (smooth projective) algebraic variety over a number field, the vanishing of its Hasse–Weil L-function at the central point is conjectured to be ac-

counted for by the non-triviality of its Chow group of (homologically trivial) algebraic cycles. A notable example is the conjecture of Birch and Swinnerton-Dyer for elliptic curves. There is now more evidence in higher dimensional cases provided by special cycles on Shimura varieties. In this context, the first example is provided by the Gross–Zagier formula [7, 27], which relates the Néron–Tate heights of Heegner divisors on modular curves to the first central derivative of the relevant L-functions. The arithmetic Gan–Gross–Prasad (GGP) conjecture [4, 18, 29, 30] is one of the generalizations of the Gross–Zagier formula to higher-dimensional Shimura varieties. The relevant Shimura variety is associated to the product unitary group $G = \mathrm{U}(1, n-2) \times \mathrm{U}(1, n-1)$ or orthogonal group $G = \mathrm{SO}(2, n - 2) \times \mathrm{SO}(2, n - 1)$, on which there is the so-called arithmetic diagonal cycle associated to the diagonally embedded $H = \mathrm{U}(1, n - 2)$ or $H = \mathrm{SO}(2, n - 2)$. Then the case $n = 2$ (for both unitary and orthogonal groups) essentially recovers Heegner divisors on modular curves.

The arithmetic Gan–Gross–Prasad conjecture is inspired by the (usual) Gan–Gross–Prasad conjecture, which relates automorphic period integrals on classical groups to the special values of Rankin–Selberg tensor product L-functions (cf. [2, 3] for recent advances in this direction). The latter conjecture makes sense for more general groups, such as the product group $G = \mathrm{U}(m) \times \mathrm{U}(n)$ and $G = \mathrm{SO}(m) \times \mathrm{SO}(n)$ when $m - n$ is odd, while the arithmetic conjecture only exists for the case $|n - m| = 1$[1]. A natural question is how to formulate the arithmetic conjecture for the general product groups. This paper aims to give a partial answer to this question. In short, there is no global cycle to formulate a global conjecture but there is a local cycle and one can formulate a local conjecture. In other words, there is no Shimura variety with the desired special algebraic cycle when $|n - m| > 1$, but there is a local Shimura variety ([20, 21], a formal scheme in this paper) over a p-adic integer ring with the desired cycle (a closed formal subscheme) to formulate an Arithmetic Fundamental Lemma (AFL) conjecture.

To be more precise let us recall that Jacquet and Rallis formulated a relative trace formula (RTF) approach to the (usual) Gan–Gross–Prasad conjecture for $\mathrm{U}(n-1) \times \mathrm{U}(n)$ [9]. Inspired by their approach, in [30] the author proposed a relative trace formula approach to the arithmetic Gan–Gross–Prasad conjecture. A key local statement of this approach is the Arithmetic Fundamental Lemma (AFL) conjecture, which relates the central derivatives of

---

[1]Yifeng Liu [14] discovered how to formulate the arithmetic conjecture in the equal rank unitary case $\mathrm{U}(n) \times \mathrm{U}(n)$; however we will not discuss the even $n - m$ case in this paper.

certain orbital integrals to the arithmetic intersection numbers of "local arithmetic diagonal cycles" on a Rapoport–Zink (RZ) space for unitary groups,

$$(1.1) \qquad \partial \mathrm{Orb}\big(\gamma, \mathbf{1}_{S_n(O_{F_0})}\big) = -\mathrm{Int}(g) \cdot \log q,$$

where F_0 is a p-adic local field with *odd* residue cardinality q, and $S_n \simeq \mathrm{GL}_{n,F}/\mathrm{GL}_{n,F_0}$ is the symmetric space with respect to an unramified quadratic extension F of F_0. See §7.2 for the precise formulation. This conjecture is largely proved by the author [31] when $F_0 = \mathbb{Q}_p$, and by Mihatsch and the author in [17] for general p-adic fields F_0, provided $q \geq n$ of F_0 in both cases. (Very recently Zhiyu Zhang has removed the restriction $q \geq n$ in [32].)

In [13], Liu generalized the RTF construction of Jacquet and Rallis to all unitary groups $\mathrm{U}(m) \times \mathrm{U}(n)$ (the Bessel case when $n - m$ is odd, and Fourier–Jacobi case when $n - m$ is even). In particular, he has essentially generalized the left hand side of the AFL conjecture (1.1) above to the general case of unitary groups; we will recall the Bessel case of his orbital integrals in §6. We will also give a lattice counting interpretation of the orbital integral in §6.6 and prove the conjecture for some special elements (by reduction to Jacquet–Rallis FL). Then we will generalize the local arithmetic diagonal cycle to the general Bessel case, see §7. The key new input is a class of RZ spaces with non-reductive groups as the "structure groups", which may be of independent interest. The RZ spaces for reductive groups (in the EL or PEL cases) parameterize p-divisible groups with additional structures, cf. §2.1. Here we add a filtration by p-divisible groups (respecting the additional structure) to define the RZ spaces for certain non-reductive groups, see §3.

There are vast generalizations of the GGP conjecture in terms of period integrals on (homogeneous) *spherical varieties* G/H for a pair of algebraic groups $(G, H), H \subset G$, over number fields; see the series of work by Sakellaridis (e.g. [22]) and a version of the general conjecture of Ichino–Ikeda type, formulated by Sakellaridis and Venkatesh [23]. The idea in this paper seems applicable to more spherical pairs (G, H) where G/H is non-affine (equivalently, by a theorem of Matsushima, H is non-reductive; here G is always assumed to be reductive). In fact, there are many more non-affine (homogeneous) spherical varieties than affine ones, and it would be interesting to classify all such spherical varieties which behaves like the cases in this paper. For all of the non-affine ones, it is currently hopeless to construct the global cycles but we expect to find the desired local cycles over p-adic fields, at least assuming the existence of a "good" theory of (integral models of) local Shimura varieties (i.e., one needs a suitable intersection theory on them). Along this direction, we hope to investigate the arithmetic GGP conjecture

for $G = \mathrm{SO}(m) \times \mathrm{SO}(n)$ when $m - n$ is odd, which involves local Shimura varieties of Hodge type (GSpin RZ spaces). In the non-affine case, the arguably "simplest" example is the Ginzburg–Rallis subgroup, which only involves RZ spaces of EL type and is a natural generalization of the Lubin–Tate deformation space; we will discuss this briefly in §4. There are also non-affine spherical varieties for exceptional groups (e.g., certain cases in the work of Wan–Zhang [26]), which are more challenging at this moment, due to the lack of a "good" theory of integral model of local Shimura varieties.

2. RZ spaces: the reductive case

Rapoport and Zink have constructed formal moduli spaces of p-divisible groups with EL or PEL structures (see [21, §3]). These moduli spaces admit actions by (the $\mathbb{Q}_p$-points of) certain reductive groups over $\mathbb{Q}_p$. We recall their construction in this section.

2.1. The RZ datum: the EL and PEL case

The formal moduli spaces of Rapoport and Zink, abbreviated as RZ spaces in this paper, depend on some linear algebraic data, called "RZ data". In this paper, we will only consider *simple* RZ data, cf. [20, §4.1]. A *simple rational RZ datum* in the EL case is a tuple

$$\mathcal{D} = (F, B, V, \{\mu\}, [b]),$$

where

- F is a finite field extension of $\mathbb{Q}_p$,

- B is a central *division* F-algebra,

- V is a finite dimensional left B-module,

- $\{\mu\}$ is a conjugacy class of minuscule cocharacters $\mu : \mathbb{G}_m \to G_{\overline{\mathbb{Q}}_p}$ where we define $G = \mathrm{Res}_{F/\mathbb{Q}_p} \mathrm{GL}_B(V)$ (an algebraic group over $\mathbb{Q}_p$),

- $[b] \in B(G)$, where $B(G)$ is the set of σ-conjugacy classes of $G(\breve{\mathbb{Q}}_p)$ (cf. [20, §2.1]).

Here, we denote the completion of a maximal unramified extension of a p-adic field F by $\breve{F}$. Moreover, $\sigma \in \mathrm{Aut}(\breve{\mathbb{Q}}_p/\mathbb{Q}_p)$ denotes the Frobenius automorphism.

A *simple rational RZ datum* in the PEL case is a tuple

$$\mathcal{D} = (F, B, V, (\cdot, \cdot), *, \{\mu\}, [b])$$

where $F, B, V, \{\mu\}, [b]$ are as in the EL case (but for the structure group G defined below) and

- $(\cdot, \cdot) : V \times V \to \mathbb{Q}_p$ is a non-degenerate alternating $\mathbb{Q}_p$-bilinear form,

- $* : B \to B$ is an involution such that $(ax, y) = (x, a^*y)$ holds for $a \in B$ and $x, y \in V$,

- G is the subgroup of $\mathrm{Res}_{F/\mathbb{Q}_p} \mathrm{GL}_B(V)$ such that, for any $\mathbb{Q}_p$-algebra R, its R-points are

$$G(R) = \left\{ g \in \mathrm{GL}_B(V \otimes_{\mathbb{Q}_p} R) \mid (gx, gy) = c(g)(x, y), c(g) \in R^\times \right\}.$$

In both cases, we imposes the following condition on $\{\mu\}$. Any cocharacter μ induces a weight decomposition on $V \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p$. We require that for any $\mu \in \{\mu\}$, only weights 0 and 1 occur in the decomposition, i.e.

(2.1) $$V \otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p = V_0 \oplus V_1.$$

In the PEL case, we further assume that the prime p is odd throughout the paper, and require that for any $\mu \in \{\mu\}$, the composition with the similitude character c

$$\mathbb{G}_m \xrightarrow{\mu} G_{\overline{\mathbb{Q}}_p} \xrightarrow{c} \mathbb{G}_{m, \overline{\mathbb{Q}}_p}$$

is the identity.

Now we turn to the integral datum. A *simple integral RZ datum* $\mathcal{D}_{\mathbb{Z}_p}$ in the EL case consists, in addition to the rational data $\mathcal{D}$, of a maximal order O_B in B and an O_B-stable lattice Λ in V. A *simple integral RZ datum* $\mathcal{D}_{\mathbb{Z}_p}$ in the PEL case consists, in addition to the rational data $\mathcal{D}$, of a maximal order O_B in B which is stable under the involution $*$ and an O_B-stable lattice Λ in V such that $\Lambda \subset \Lambda^\vee \subset \varpi^{-1}\Lambda$. Here ϖ denotes a uniformizer of O_B, and

$$\Lambda^\vee = \{x \in V \mid (x, \Lambda) \subset \mathbb{Z}_p\}$$

denotes the "dual lattice" of Λ[2].

---

[2]We may call a lattice Λ a *vertex lattice* if it satisfies $\Lambda \subset \Lambda^\vee \subset \varpi^{-1}\Lambda$.

Remark 2.1. More general RZ data (still in the EL and PEL case) can be found in [21, §3], where B could be a semisimple $\mathbb{Q}_p$-algebra and Λ is replaced by a "periodic lattice chain".

For simplicity we will assume that $[b]$ is *basic* for the rest of this paper, cf. [20, §2.1]. Notice that there is a unique basic element in the subset $B(G, \mu)$ of μ-neutral acceptable elements in $B(G)$ (see [20, Def. 2.3]). In the EL and PEL case, the basic element can be characterized by the associated isocrystal (given by $b(1_V \otimes \sigma)$) on $V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p$ being *isoclinic* (i.e., only a single slope appears).

We let $E = E_{\{\mu\}}$ denote the reflex field (i.e., the field of definition of $\{\mu\}$). To each $b \in G(\breve{\mathbb{Q}}_p)$ one may associate an algebraic group over $\mathbb{Q}_p$, denoted by J_b, such that for any $\mathbb{Q}_p$-algebra R,

$$(2.2) \qquad J_b(R) = \{g \in G(R \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p) \mid g(b\sigma) = (b\sigma)g\}.$$

Up to isomorphism the group J_b depends only on the class $[b] \in B(G)$. For $[b]$ basic, the group J_b is an inner form of G.

We also single out the (simple) *unramified* case [21, §3.82], which is of particular interest in this paper. We recall that this means that $B = F$ is an unramified field extension of $\mathbb{Q}_p$, and in the PEL case, $\Lambda = \Lambda^\vee$ (i.e., a self-dual lattice with respect to the given alternating form on V).

We will mostly concentrate on two (families of) examples.

Example 2.2 (EL case)**.** All of the simple unramified integral RZ data $\mathcal{D}_{\mathbb{Z}_p}$ in the EL case arise as follows. Let $B = F$ is an unramified field extension of $\mathbb{Q}_p$, $V = F^n$ is the standard n-dimensional F-vector space, $\Lambda \subset V$ an O_F-lattice. Then $G = \operatorname{Res}_{F/\mathbb{Q}_p} \operatorname{GL}_{n,F}$. A cocharacter

$$\mu : \mathbb{G}_m \longrightarrow G_{\overline{\mathbb{Q}}_p} \simeq \prod_{\varphi \in \operatorname{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)} \operatorname{GL}_{n, \overline{\mathbb{Q}}_p}$$

consists of a tuple $(\mu_\varphi)_{\varphi \in \operatorname{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ of cocharacters $\mu_\varphi : \mathbb{G}_m \to \operatorname{GL}_{n, \overline{\mathbb{Q}}_p}$. We consider (the conjugacy class of)

$$\mu_\varphi(z) = \begin{bmatrix} \mathbf{1}_{r_\varphi} & \\ & z \cdot \mathbf{1}_{s_\varphi} \end{bmatrix},$$

where $(r_\varphi, s_\varphi)_{\varphi \in \operatorname{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ is a tuple of non-negative integers such that $r_\varphi + s_\varphi = n$. The reflex field E is then the subfield of $\overline{\mathbb{Q}}_p$ corresponding to the

open subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ fixing the tuple $(r_\varphi, s_\varphi)_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ (under the obvious action). Let $[b]$ be the unique basic element in $B(G, \mu)$.

Example 2.3 (unitary PEL case). We consider the following simple integral RZ data $\mathcal{D}_{\mathbb{Z}_p}$ in the PEL case, related to unitary groups. Let F/F_0 be an unramified quadratic extension of p-adic fields with the non-trivial Galois involution denoted by $a \mapsto \overline{a}$. Let $B = F$, and V a non-degenerate F/F_0-Hermitian space of F-dimension n. Upon choosing a unit in the "imaginary" F_0-line $F^- := \{a \in F \mid \overline{a} = -a\}$, the Hermitian form induces an alternating $\mathbb{Q}_p$-bilinear form on V. Then $G = \mathrm{GU}(V, (\cdot, \cdot))$ is the unitary similitude group. Let $\Lambda \subset V$ be an O_F-lattice such that $\Lambda \subset \Lambda^\vee \subset \varpi^{-1}\Lambda$. We consider the conjugacy class of $\mu : \mathbb{G}_m \to G_{\overline{\mathbb{Q}}_p}$ corresponding to the "signature" $(r_\varphi, s_\varphi)_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ such that $r_\varphi + s_\varphi = n$ and $r_\varphi = s_{\overline{\varphi}}$, where r_φ denotes the dimension of $V_{0,\varphi}$. Let $[b]$ be the unique basic element in $B(G, \mu)$.

2.2. The formal moduli spaces in the reductive case

Let O_E and $O_{\breve{E}}$ denote the ring of integers in E and $\breve{E}$ respectively. Let $\mathbb{F}$ be the residue field of $O_{\breve{E}}$. Let $\mathtt{Nilp} = \mathtt{Nilp}_{O_{\breve{E}}}$ denote the category of $O_{\breve{E}}$-schemes S on which p is locally nilpotent. For $S \in \mathtt{Nilp}$, we denote by $\overline{S}$ the closed subscheme of S defined by $p\mathcal{O}_S$, i.e., $\overline{S} := S \times_{\mathrm{Spec}\, O_{\breve{E}}} \mathrm{Spec}\, O_{\breve{E}}/(p)$. Note that $\mathrm{Spec}\, O_{\breve{E}}/(p)$ is naturally a scheme over $\mathrm{Spec}\, W(\mathbb{F})/(p) \simeq \mathrm{Spec}\, \mathbb{F}$ and hence we can view $\overline{S}$ as a scheme over $\mathrm{Spec}\, \mathbb{F}$.

Fix a simple integral RZ data $\mathcal{D}_{\mathbb{Z}_p}$. We will consider pairs (X, ι) in the EL case, and triples (X, ι, λ) in the PEL case, where

- X is a p-divisible group over $S \in \mathtt{Nilp}$,

- $\iota : O_B \to \mathrm{End}(X)$ is an action of O_B on X,

- In the PEL case, a polarization $\lambda : X \to X^\vee$.

We impose the Kottwitz condition, i.e., the equality of characteristic polynomials for all $a \in O_B$

$$\mathrm{char}\big(\iota(a); \mathrm{Lie}(X)\big) = \mathrm{char}\big(a; V_0\big),$$

where V_0 is the weight zero subspace of V, cf. (2.1). In the PEL case, we further impose the condition that $\ker(\lambda) \subset X[\varpi]$ has order equal to $\#(\Lambda^\vee/\Lambda)$, and that the Rosati involution on O_B coincides with $*$.

To define the moduli space, besides an integral RZ datum $\mathcal{D}_{\mathbb{Z}_p}$, we also need to fix a *framing object*, i.e., a pair $(\mathbb{X}, \iota_{\mathbb{X}})$ over $\mathrm{Spec}\, \mathbb{F}$ in the EL case

(resp. a triple $(\mathbb{X}, \iota_{\mathbb{X}}, \lambda_{\mathbb{X}})$ in the PEL case) as above. We assume that the induced isocrystal with the O_B-action (and the alternating form induced by the polarization in the PEL case) is isomorphic to $(V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, b\sigma)$ (in the PEL case, $(V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, (\cdot, \cdot), b\sigma)$, the alternating form preserved up to a $\breve{\mathbb{Q}}_p^{\times}$-factor). Then the group of self quasi-isogenies of $\mathbb{X}$ respecting the additional structure is isomorphic to the group $J_b(\mathbb{Q}_p)$.

We are ready to state the definition of the moduli functor of Rapoport–Zink in the EL case:

$$\mathcal{M} : \mathtt{Nilp} \longrightarrow Sets \,.$$

It associates to $S \in \mathtt{Nilp}$ the set of isomorphism classes (X, ι, ρ) where (X, ι) is as above and

$$\rho : X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S}$$

is an O_B-linear quasi-isogeny [21, §2]. In the PEL case, $\mathcal{M}(S)$ is the set of isomorphism classes $(X, \iota, \lambda, \rho)$, where (X, ι, λ) is as above and

$$\rho : X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S}$$

is an O_B-linear quasi-isogeny that preserves the polarizations up to a factor in $\mathbb{Q}_p^{\times}$, locally on $\overline{S}$.

Here the datum ρ is called a *framing* of (X, ι) or (X, ι, λ). The group $J_b(\mathbb{Q}_p)$ acts on $\mathcal{M}$ by changing the framing ρ.

Theorem 2.4 (Rapoport–Zink)**.** *The functor $\mathcal{M}$ is (pro-)representable by a formal scheme, formally locally of finite type over* $\mathrm{Spf}\,O_{\breve{E}}$*.*

See [21, Thm. 3.25] for a more general statement (cf. [21, Def. 3.21] and the remarks in §3.23 of *loc. cit.* for the equivalence of the definitions).

2.3. Local models

The local structure of $\mathcal{M}$ is governed by the so-called *local model* [21, §3.26–3.35]. For brevity we recall the definition in the unramified $\mathcal{D}_{\mathbb{Z}_p}$ case.

We consider the following functor from the category $\mathtt{Sch}_{O_E}$ of schemes over O_E

$$\mathbf{M}^{\mathrm{loc}} : \mathtt{Sch}_{O_E} \longrightarrow Sets$$

which associates to $S \in \mathtt{Sch}_{O_E}$ the set of isomorphism classes of the following data:

- A locally free $\mathcal{O}_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-module $\mathcal{F}$ on S:

- A homomorphism of $\mathcal{O}_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules

$$t : \Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S \longrightarrow \mathcal{F}.$$

We require the following conditions to hold:

- the action of O_B satisfies the analog of Kottwitz condition

$$\mathrm{char}(\iota(a); \mathcal{F}) = \mathrm{char}(a; V_0), \quad \forall a \in O_B.$$

- The homomorphisms t is surjective.

- In the PEL case, the isomorphism $\Lambda^* \simeq \Lambda$ induces the following commutative diagram

$$
\begin{array}{ccc}
\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \overset{\sim}{\longrightarrow} & (\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S)^* \\
\downarrow{\scriptstyle t} & & \downarrow \\
\mathcal{F} & \overset{\sim}{\longrightarrow} & (\ker(t))^*.
\end{array}
$$

(Note that the polarization is assumed to be principal.) Here $*$ denotes the $\mathcal{O}_S$-linear dual.

It is easy to see that the functor $\mathbf{M}^{\mathrm{loc}}$ is represented by a projective scheme over $\mathrm{Spec}\, O_E$, being a closed subscheme of a product of Grassmannians. In our unramified case, it is smooth.

Let $\hat{\mathbf{M}}^{\mathrm{loc}}$ denote the p-adic completion of the base change $\mathbf{M}^{\mathrm{loc}} \times_{\mathrm{Spec}\, O_E} \mathrm{Spec}\, O_{\breve{E}}$; it is a formal scheme over $\mathrm{Spf}\, O_{\breve{E}}$. Then Grothendieck–Messing theory implies that any point of $\mathcal{M}$ has an étale neighborhood which is formally étale over $\hat{\mathbf{M}}^{\mathrm{loc}}$ (see [21, §3.32]). Therefore, the smoothness of the local model $\mathbf{M}^{\mathrm{loc}} \to \mathrm{Spec}\, O_E$ implies that $\mathcal{M}$ is formally smooth over $\mathrm{Spf}\, O_{\breve{E}}$. In the two examples 2.2 and 2.3, if the RZ data are unramified, then relative dimension of $\mathcal{M}$ over $\mathrm{Spf}\, O_{\breve{E}}$ is given by

$$(2.3) \qquad \sum_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)} r_\varphi s_\varphi$$

in the EL case, and

$$(2.4) \qquad \frac{1}{2} \sum_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)} r_\varphi s_\varphi = \sum_{\varphi_0 \in \mathrm{Hom}_{\mathbb{Q}_p}(F_0, \overline{\mathbb{Q}}_p)} r_\varphi s_\varphi$$

in the unitary PEL case, where $\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)$ is any extension of $\varphi_0 \in \mathrm{Hom}_{\mathbb{Q}_p}(F_0, \overline{\mathbb{Q}}_p)$.

Example 2.5 (The "totally definite" or "banal" case)**.** In Example 2.2 and 2.3, we call an unramified $\mathcal{D}_{\mathbb{Z}_p}$ *totally definite* if

$$r_\varphi s_\varphi = 0, \quad \forall \varphi \in \operatorname{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p).$$

Then the relative dimension of $\mathcal{M}$ over $\operatorname{Spf} O_{\breve{E}}$ is zero and $J_b \simeq G$. There is an isomorphism of formal schemes

$$\mathcal{M} \simeq \bigsqcup_{G(\mathbb{Q}_p)/K} \operatorname{Spf} O_{\breve{E}}$$

where K is the stabilizer of Λ.

3. RZ spaces: the non-reductive case

In this section, we construct some generalized RZ spaces in the *basic* case, which may be viewed as RZ spaces with the structure group being non-reductive. The new piece of datum defining the moduli problem is a filtration by p-divisible (sub)groups. These formal moduli spaces seem new and they appear to have interesting structure (e.g., even the connected components seem not understood to the author). The moduli spaces of "filtered p-divisible groups" (with EL or PEL structures) in the *non-basic* case have appeared in Mantovan's work [15]. Just as in *loc. cit.*, here we can establish preliminary properties (representability, local model) of our moduli spaces using the strategy very similar to [21, §3].

3.1. The filtered RZ datum: the EL and PEL case

A simple rational *filtered* RZ datum in the EL case is a tuple

$$\mathcal{D} = (F, B, V, \operatorname{Fil}^\bullet, \{\mu\}, [b]),$$

where F, B, V are as in the simple rational RZ datum §2.1 and

• $\operatorname{Fil}^\bullet$ is a filtration of finite length

$$0 = \operatorname{Fil}^0 V \subset \operatorname{Fil}^1 V \subset \cdots \subset \operatorname{Fil}^\ell V = V$$

by B-stable submodules $\operatorname{Fil}^i V$,

• $\{\mu\}$ is a conjugacy class of minuscule cocharacters $\mu : \mathbb{G}_m \to H_{\overline{\mathbb{Q}}_p}$ where $H := \operatorname{GL}_B(V, \operatorname{Fil}^\bullet)$ is the stabilizer of the filtration $\operatorname{Fil}^\bullet$, hence a parabolic subgroup of $G = \operatorname{GL}_B(V)$.

- $[b]$ is a σ-conjugacy class of $H(\breve{\mathbb{Q}}_p)$.

The filtered RZ datum $\mathcal{D}$ induces (unfiltered) RZ data

$$\mathrm{Fil}^i\mathcal{D} = (F, B, \mathrm{Fil}^i V, \{\mathrm{Fil}^i\mu\}, [\mathrm{Fil}^i b]), \quad i = 1, \cdots, \ell,$$

where $\mathrm{Fil}^i G := \mathrm{GL}_B(\mathrm{Fil}^i V)$, $\mathrm{Fil}^i\mu : \mathbb{G}_m \to \mathrm{Fil}^i G_{\overline{\mathbb{Q}}_p}$, and $[\mathrm{Fil}^i b] \in B(\mathrm{Fil}^i G)$ are induced by $\{\mu\}$ and $[b]$. Similarly, we have induced RZ data

$$\mathrm{gr}^i\mathcal{D} = (F, B, \mathrm{gr}^i V, \{\mathrm{gr}^i\mu\}, [\mathrm{gr}^i b]), \quad i = 1, \cdots, \ell,$$

where $\mathrm{gr}^i V := \mathrm{Fil}^i V/\mathrm{Fil}^{i-1} V$ is the i-th graded piece, $\mathrm{gr}^i G := \mathrm{GL}_B(\mathrm{gr}^i V)$, $\mathrm{gr}^i\mu : \mathbb{G}_m \to \mathrm{gr}^i G_{\overline{\mathbb{Q}}_p}$, and $[\mathrm{gr}^i b] \in B(\mathrm{gr}^i G)$ are induced by $\{\mu\}$ and $[b]$.

A simple *filtered* rational RZ datum in the PEL case is a tuple

$$\mathcal{D} = (F, B, V, \mathrm{Fil}^\bullet, (\cdot, \cdot), *, \{\mu\}, [b]),$$

where $F, B, V, \mathrm{Fil}^\bullet, \{\mu\}, [b]$ are as in the filtered EL case (but for H defined below), $(\cdot, \cdot), *$ are as in the (unfiltered) PEL case, such that

- $\mathrm{Fil}^\bullet$ is *self-dual*, in the sense that, for every $i = 0, \cdots, \ell$, the space $\mathrm{Fil}^i V$ is the exact annihilator of $\mathrm{Fil}^{\ell-i} V$ under the alternating form $(\cdot, \cdot)$ on V,

- $H := \mathrm{GL}_B(V, \mathrm{Fil}^\bullet, (\cdot, \cdot))$, a parabolic subgroup of $G = \mathrm{GL}_B(V, (\cdot, \cdot))$.

We have induced RZ data of PEL type

$$\mathrm{Fil}^{i,\ell-i}\mathcal{D} = (F, B, \mathrm{Fil}^{i,\ell-i} V, (\cdot, \cdot), *, \{\mathrm{Fil}^{i,\ell-i}\mu\}, [\mathrm{Fil}^{i,\ell-i} b]), \quad i = 1, \cdots, [\ell/2],$$

with the induced alternating pairing $(\cdot, \cdot)$ on $\mathrm{Fil}^{i,\ell-i} V := \mathrm{Fil}^{\ell-i} V/\mathrm{Fil}^i V$. We also have RZ data of EL type

$$\mathrm{gr}^i\mathcal{D} = (F, B, \mathrm{gr}^i V, \{\mathrm{gr}^i\mu\}, [\mathrm{gr}^i b]), \quad i = 1, \cdots, [\ell/2],$$

and, if ℓ is odd, an RZ datum of PEL type

$$\mathrm{gr}^{[(\ell+1)/2]}\mathcal{D} = (F, B, \mathrm{gr}^{[(\ell+1)/2]} V, (\cdot, \cdot), *, \{\mathrm{gr}^{[(\ell+1)/2]}\mu\}, [\mathrm{gr}^{[(\ell+1)/2]} b]).$$

Now we turn to the integral datum. A simple integral filtered RZ datum $\mathcal{D}_{\mathbb{Z}_p}$ in the EL case consists, in addition to the rational data $\mathcal{D}$, of a maximal order O_B in B and an O_B-stable lattice Λ in V. A simple integral filtered RZ datum $\mathcal{D}_{\mathbb{Z}_p}$ in the PEL case consists, in addition to the rational data $\mathcal{D}$, of a maximal order O_B in B which is stable under the involution $*$ and

an O_B-stable lattice Λ in V such that $\Lambda \subset \Lambda^\vee \subset \varpi^{-1}\Lambda$. In the EL case, the O_B-stable Λ induces O_B-stable lattice $\mathrm{Fil}^i\Lambda := \Lambda \cap \mathrm{Fil}^iV \subset \mathrm{Fil}^iV$ and $\mathrm{gr}^i\Lambda := \mathrm{Fil}^i\Lambda/\mathrm{Fil}^{i-1}\Lambda \subset \mathrm{gr}^iV$. Therefore, we obtain induced simple integral RZ data $\mathrm{Fil}^i\mathcal{D}_{\mathbb{Z}_p}$ and $\mathrm{gr}^i\mathcal{D}_{\mathbb{Z}_p}$ of EL type. Similarly, in the PEL case, we define $\mathrm{Fil}^{i,\ell-i}\Lambda := \mathrm{Fil}^{\ell-i}\Lambda/\mathrm{Fil}^i\Lambda \subset \mathrm{Fil}^{i,\ell-i}V$. Then, by the following lemma, we obtain simple integral RZ data $\mathrm{Fil}^{i,\ell-i}\mathcal{D}_{\mathbb{Z}_p}$ and $\mathrm{gr}^i\mathcal{D}_{\mathbb{Z}_p}$. Moreover, if $\mathcal{D}_{\mathbb{Z}_p}$ is unramified, so are all of the induced data.

Lemma 3.1. *In the PEL case, let $\Lambda \subset V$ be an O_B-stable lattice such that $\Lambda \subset \Lambda^\vee \subset \varpi^{-1}\Lambda$. Then, for all $i = 1, \cdots, [\ell/2]$, $\mathrm{Fil}^{i,\ell-i}\Lambda$ is an O_B-stable lattice in $\mathrm{Fil}^{i,\ell-i}V$ such that*

$$\mathrm{Fil}^{i,\ell-i}\Lambda \subset (\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee \subset \varpi^{-1}\mathrm{Fil}^{i,\ell-i}\Lambda.$$

Moreover, if $\Lambda = \Lambda^\vee$, then $\mathrm{Fil}^{i,\ell-i}\Lambda = (\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee$.

Proof. For a $\mathbb{Z}_p$-lattice L we will denote $L^* = \mathrm{Hom}_{\mathbb{Z}_p}(L, \mathbb{Z}_p)$ the $\mathbb{Z}_p$-linear dual. If L is an O_B-lattice, we endow L^* the twist of the O_B-action by the involution $*$. Then the alternating form $(\cdot, \cdot)$ induces an O_B-linear map $\alpha : \Lambda \to \Lambda^*$ and an O_B-linear isomorphism $\Lambda^* \simeq \Lambda^\vee$. Since Fil^iV is totally isotropic, α induces an O_B-linear map $\beta : \mathrm{Fil}^i\Lambda \to (\Lambda/\mathrm{Fil}^i\Lambda)^*$, which then induces an O_B-linear map $\beta^* : (\Lambda/\mathrm{Fil}^i\Lambda)^* \to ((\mathrm{Fil}^i\Lambda)^*)^* \simeq \mathrm{Fil}^i\Lambda$. We obtain a commutative diagram:

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \mathrm{Fil}^i\Lambda & \longrightarrow & \Lambda & \longrightarrow & \Lambda/\mathrm{Fil}^i\Lambda & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta^*} & & \\
0 & \longrightarrow & (\Lambda/\mathrm{Fil}^i\Lambda)^* & \longrightarrow & \Lambda^* & \longrightarrow & (\mathrm{Fil}^i\Lambda)^* & \longrightarrow & 0.
\end{array}
$$

Now note that $\ker(\beta^*) = \mathrm{Fil}^{\ell-i}\Lambda/\mathrm{Fil}^i\Lambda = \mathrm{Fil}^{i,\ell-i}\Lambda$ and hence $\mathrm{coker}(\beta) \simeq (\mathrm{Fil}^{i,\ell-i}\Lambda)^* \simeq (\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee$. By the snake lemma, we obtain a long exact sequence of O_B-modules

$$0 \longrightarrow \mathrm{Fil}^{i,\ell-i}\Lambda \longrightarrow (\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee \longrightarrow \Lambda^\vee/\Lambda \longrightarrow \mathrm{coker}(\beta^*) \longrightarrow 0,$$

or equivalently,

$$0 \longrightarrow (\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee/\mathrm{Fil}^{i,\ell-i}\Lambda \longrightarrow \Lambda^\vee/\Lambda \longrightarrow \mathrm{coker}(\beta^*) \longrightarrow 0.$$

If $\Lambda^\vee/\Lambda$ is annihilated by $\varpi \in O_B$, so is $(\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee/\mathrm{Fil}^{i,\ell-i}\Lambda$. If $\Lambda^\vee/\Lambda$ vanishes, so does $(\mathrm{Fil}^{i,\ell-i}\Lambda)^\vee/\mathrm{Fil}^{i,\ell-i}\Lambda$. $\qquad\square$

In both cases, we will again assume that the induced σ-conjugacy class of $G(\breve{\mathbb{Q}}_p)$ is *basic*. Since this means (in our EL and PEL case) that the associated isocrystal is isoclinic, the induced isocrystals on $\mathrm{Fil}^{i,\ell-i}V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p$ and $\mathrm{gr}^i V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p$ are automatically isoclinic and hence the induced σ-conjugacy classes $\mathrm{Fil}^{i,\ell-i}b$ and $\mathrm{gr}^i b$ are all basic.

Similar to (2.2) we can define a functor J_b. Though we have not checked the detail, it is reasonable to expect that J_b is represented by a smooth affine group scheme over $\mathbb{Q}_p$ and it is a parabolic subgroup of $J_{\mathrm{Fil}^{0,\ell}b}$. Nevertheless, one can verify this assertion in all the examples below 3.2 and 3.3.

Example 3.2 (EL case)**.** Let $B = F, V, G, \Lambda$ be as in Example 2.2. Let $\mathrm{Fil}^\bullet V$ be any filtration of V and let $H = \mathrm{GL}_B(V, \mathrm{Fil}^\bullet)$. Then we may arrange a cocharacter of G to factor through $\mu : \mathbb{G}_m \to H$. Let $(r^i_\varphi, s^i_\varphi)_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ be the invariants associated to the induced data on $\mathrm{gr}^i V$. In particular, we have

$$r_\varphi = \sum_i r^i_\varphi, \quad s_\varphi = \sum_i s^i_\varphi, \quad \forall \varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p).$$

Example 3.3 (unitary PEL case)**.** Let $B = F, V, (\cdot, \cdot), G, \Lambda$ be as in Example 2.3. Let $\mathrm{Fil}^\bullet V$ be a self-dual filtration of V and let $H = \mathrm{GL}_B(V, (\cdot, \cdot), \mathrm{Fil}^\bullet) \subset G = \mathrm{GL}_B(V, (\cdot, \cdot))$. Let $\mu : \mathbb{G}_m \to H_{\overline{\mathbb{Q}}_p}$ be a cocharacter. Let $(r^i_\varphi, s^i_\varphi)_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)}$ be the invariants associated to the induced data on $\mathrm{gr}^i V$. Then for all $\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p)$

$$r_\varphi = \sum_i r^i_\varphi, \quad s_\varphi = \sum_i s^i_\varphi,$$

and

$$r^i_\varphi = s^{\ell+1-i}_{\overline{\varphi}}.$$

We remark that there are some restrictions to the existence of above Fil and μ due to the restriction to the weights and the compatibility with the similitude character. For example, when $F_0 = \mathbb{Q}_p$ and $\dim_F V = 2$ there does not exist any nontrivial filtered data (besides the "unfiltered" one) satisfying all of the conditions.

3.2. The formal moduli spaces in the non-reductive case

We fix a simple integral filtered RZ data $\mathcal{D}_{\mathbb{Z}_p}$ for the rest of this section. In view of the possible complexity of the induced data, as illustrated by Lemma 3.1, we assume that $\mathcal{D}_{\mathbb{Z}_p}$ is *unramified*. In particular, $B = F$ is an

unramified field extension of $\mathbb{Q}_p$, the reflex field E is unramified over $\mathbb{Q}_p$ and hence $\breve{E} = \breve{\mathbb{Q}}_p$. We consider tuples $(X, \mathrm{Fil}^\bullet X, \iota)$ in the EL case, and $(X, \mathrm{Fil}^\bullet X, \iota, \lambda)$ in the PEL case, where the unfiltered data are the same as before, and

- $\mathrm{Fil}^\bullet X$ is a filtration of X by O_B-stable p-divisible subgroups

$$0 = \mathrm{Fil}^0 X \subset \mathrm{Fil}^1 X \subset \cdots \subset \mathrm{Fil}^\ell X = X.$$

- In the PEL case, a *principal* polarization $\lambda : X \to X^\vee$ such that the Rosati involution on O_B coincides with $*$. (The requirement of the polarization being principal is due to our assumption on the unramifiedness of the integral datum.)

We impose the Kottwitz condition for every $i = 0, \cdots, \ell$,

$$\mathrm{char}\big(\iota(a); \mathrm{Lie}(\mathrm{Fil}^i X)\big) = \mathrm{char}\big(a; (\mathrm{Fil}^i V)_0\big), \quad \forall a \in O_B,$$

where $(\mathrm{Fil}^i V)_0$ is the weight zero subspace of $\mathrm{Fil}^i V$. In the PEL case, we require that the filtration $\mathrm{Fil}^\bullet$ is *self-dual*, in the sense that, for every $i = 0, \cdots, \ell$, we have an exact sequence (of fppf abelian sheaves) induced by the polarization λ:

$$0 \longrightarrow \mathrm{Fil}^{\ell-i} X \longrightarrow X \simeq X^\vee \longrightarrow (\mathrm{Fil}^i X)^\vee \longrightarrow 0.$$

Next we fix a *framing object* $(\mathbb{X}, \mathrm{Fil}^\bullet \mathbb{X}, \iota_{\mathbb{X}})$ over $\mathrm{Spec}\,\mathbb{F}$ in the EL case (resp. a triple $(\mathbb{X}, \mathrm{Fil}^\bullet \mathbb{X}, \iota_{\mathbb{X}}, \lambda_{\mathbb{X}})$ in the PEL case). We assume that the induced isocrystal with the filtration, with the O_B-action (and the alternating form induced by the polarization in the PEL case) is isomorphic to $(V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, \mathrm{Fil}^\bullet V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, b\sigma)$ (in the PEL case, $(V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, \mathrm{Fil}^\bullet V \otimes_{\mathbb{Q}_p} \breve{\mathbb{Q}}_p, (\cdot, \cdot), b\sigma)$ with the alternating form preserved up to a $\breve{\mathbb{Q}}_p^\times$-factor).

We now state the definition of the moduli functor analogous to that of Rapoport–Zink, in the EL case:

$$\mathcal{M} : \mathtt{Nilp} \longrightarrow Sets,$$

which associates to $S \in \mathtt{Nilp}$ the set of isomorphism classes $(X, \mathrm{Fil}^\bullet X, \iota, \rho)$ where $(X, \mathrm{Fil}^\bullet X, \iota)$ is as above and

$$\rho : X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S}$$

is an O_B-linear quasi-isogeny that preserves the filtration $\mathrm{Fil}^\bullet X$ and $\mathrm{Fil}^\bullet \mathbb{X}$. In the PEL case, $\mathcal{M}(S)$ is the set of isomorphism classes $(X, \mathrm{Fil}^\bullet X, \iota, \lambda, \rho)$, where $(X, \mathrm{Fil}^\bullet X, \iota, \lambda)$ is as above and

$$\rho : X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S}$$

is an O_B-linear quasi-isogeny that preserves the filtration $\mathrm{Fil}^\bullet X$ and $\mathrm{Fil}^\bullet \mathbb{X}$ and preserves the polarizations up to a factor in $\mathbb{Q}_p^\times$, locally on $\overline{S}$.

The following result follows essentially from the representability theorems in [21, §2, §3]. Note that we have assumed that $\mathcal{D}_{\mathbb{Z}_p}$ is unramified.

Theorem 3.4. *The functor $\mathcal{M}$ is (pro-)representable by a formal scheme, formally locally of finite type, and formally smooth over $\mathrm{Spf}\,O_{\breve{E}}$.*

Proof. We will defer the proof of the formal smoothness after we introduce the local model.

We consider the EL case and the same argument applies to the PEL case. We consider the RZ moduli functors $\mathcal{M}_i$ for the framing objects $(\mathrm{Fil}^i \mathbb{X}, \iota_{\mathrm{Fil}^i \mathbb{X}})$, $i = 1, \cdots, \ell$. In view of the representability theorem 2.4 of Rapoport–Zink, it suffices to consider the relatively representability of the natural map

$$(3.1) \qquad\qquad j : \mathcal{M} \longrightarrow \prod_{i=1}^{\ell} \mathcal{M}_i.$$

The map j factorizes through $\mathcal{M}^\bullet$ which parameterize similar data as $\mathcal{M}$ except that $\mathrm{Fil}^i X \to \mathrm{Fil}^{i+1} X$ is not required to be injective (as a morphism of fppf abelian sheaves). Then $\mathcal{M}^\bullet$ is the locus where the quasi-homomorphisms $\mathrm{Fil}^i \mathbb{X} \to \mathrm{Fil}^{i+1} \mathbb{X}$ pull back to homomorphisms $\mathrm{Fil}^i X \to \mathrm{Fil}^{i+1} X$. By [21, Prop. 2.9][3], $\mathcal{M}^\bullet$ is a closed formal subscheme of $\prod_{i=1}^{\ell} \mathcal{M}_i$. Finally, the injectivity imposes an open condition and hence $\mathcal{M}$ is an open formal subscheme of $\mathcal{M}^\bullet$. The desired result follow. $\square$

Remark 3.5. Unlike the reductive case of RZ spaces, the irreducible components of the reduced scheme of $\mathcal{M}$ are not necessarily proper. The moduli functor $\mathcal{M}^\bullet$ may be viewed as a partial "compactification" of $\mathcal{M}$.

Remark 3.6. Closely related to the above constructions, one may also consider moduli functors in the *degenerate* PEL case, i.e., the alternative form on V is degenerate. In fact, from the filtered rational RZ datum $\mathcal{D}$ of PEL type, one has an induced datum from $\mathrm{Fil}^{\ell-1} V$, which carries a degenerate alternating form.

---

[3]Note that [21, Prop. 2.9] is stated only for quasi-isogenies; but the proof applies verbatim to quasi-homomorphisms.

3.3. Connected components

As a natural question, what is the set of connected components of $\mathcal{M}$ defined in the last subsection? Similar questions for RZ spaces (and generalizations to local Shimua varieties in [20] with parahoric level structure) have received considerable attention in recent years. For the moduli space defined here, it seems unclear what the answer should be. Nevertheless, we can partly answer the question in some special examples related to the unitary RZ spaces, where we have a very explicit description of the reduced scheme (via the BT-stratification); see Example 5.10 in §5.6.

3.4. Local models

Analogous to §2.3, the local model construction in [21, §3.26] can be carried over to the fixed unramified integral filtered RZ data $\mathcal{D}_{\mathbb{Z}_p}$. We denote $\Lambda^i = \mathrm{Fil}^i\Lambda := \mathrm{Fil}^iV \cap \Lambda$, an O_B-stable lattice. In the PEL case, the unramified hypothesis implies that there is a short exact sequence of O_B-modules

$$(3.2) \qquad 0 \longrightarrow \Lambda^{\ell-i} \longrightarrow \Lambda \simeq \Lambda^* \longrightarrow (\Lambda^i)^* \longrightarrow 0 .$$

We consider the following functor from the category $\mathtt{Sch}_{O_E}$ of schemes over O_E

$$\mathbf{M}^{\mathrm{loc}} : \mathtt{Sch}_{O_E} \longrightarrow Sets ,$$

which associates to $S \in \mathtt{Sch}_{O_E}$ the set of isomorphism classes of the following data:

- A flag of locally free $\mathcal{O}_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules $\mathcal{F}^\bullet$ on S:

$$0 = \mathcal{F}^0 \subset \mathcal{F}^1 \subset \cdots \subset \mathcal{F}^\ell = \mathcal{F}.$$

- Homomorphisms of $\mathcal{O}_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules

$$t_i : \Lambda^i \otimes_{\mathbb{Z}_p} \mathcal{O}_S \longrightarrow \mathcal{F}^i, \quad i = 1, \cdots, \ell.$$

We require the following conditions to hold:

- all $\mathcal{F}^i$ are subbundles of $\mathcal{F}$ (i.e., $\mathcal{F}/\mathcal{F}^i$ are locally free), and the action of O_B satisfies the analog of Kottwitz condition

$$\mathrm{char}(a; \mathcal{F}_i) = \mathrm{char}(a; (\mathrm{Fil}^iV)_0), \quad \forall a \in O_B,$$

- The homomorphisms t_i are all surjective and the following diagram commutes

$$
\begin{array}{ccccccc}
\Lambda^1 \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \hookrightarrow \cdots \hookrightarrow & \Lambda^i \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \hookrightarrow \cdots \hookrightarrow & \Lambda^\ell \otimes_{\mathbb{Z}_p} \mathcal{O}_S \\
\downarrow{\scriptstyle t_1} & & \downarrow{\scriptstyle t_i} & & \downarrow{\scriptstyle t_\ell} \\
\mathcal{F}^1 & \hookrightarrow \cdots \hookrightarrow & \mathcal{F}^i & \hookrightarrow \cdots \hookrightarrow & \mathcal{F}^\ell = \mathcal{F}.
\end{array}
$$

In particular, when $i \le j$, $\mathcal{F}^i$ the image of $\Lambda^i \otimes_{\mathbb{Z}_p} \mathcal{O}_S$ under t_j, and hence $(t_i, \mathcal{F}^i)$ is determined by $(t_j, \mathcal{F}^j)$.

- In the PEL case, the isomorphism $\Lambda^* \simeq \Lambda$ induces the following commutative diagram

$$
\begin{array}{ccc}
\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \xrightarrow{\sim} & (\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S)^* \\
\downarrow{\scriptstyle t_\ell} & & \downarrow \\
\mathcal{F} & \xrightarrow{\sim} & (\ker(t_\ell))^*.
\end{array}
$$

(Note that our polarization is assumed to be principal.) Here $*$ denotes the $\mathcal{O}_S$-linear dual. Together with (3.2) we have a commutative diagram for all $i = 1, \cdots, \ell$:

(3.3)

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \Lambda^{\ell-i} \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & \Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S \simeq (\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S)^* & \longrightarrow & (\Lambda^i \otimes_{\mathbb{Z}_p} \mathcal{O}_S)^* & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \mathcal{F}^{\ell-i} & \longrightarrow & \mathcal{F} \simeq (\ker(t_\ell))^* & \longrightarrow & (\ker(t_i))^* & \longrightarrow & 0
\end{array}
$$

where the bottom row is required to be exact.

The functor $\mathbf{M}^{\mathrm{loc}}$ is represented by a quasi-projective scheme over $\operatorname{Spec} O_E$. Unlike the reductive case in [21] (cf. §2.3), it is *not necessarily* projective, as we will see from the proof of Lemma 3.7 below.

We consider the (unfiltered) RZ data $\operatorname{gr}^i \mathcal{D}_{\mathbb{Z}_p}$ associated to $\operatorname{gr}^i V$, $1 \le i \le \ell$ in the EL case and $1 \le i \le [\frac{\ell+1}{2}]$ in the PEL case (cf. §3.1). Let $\mathbf{M}^{\mathrm{loc}}_i$ be the corresponding local models as defined in §2.3. Then we have a natural morphism

$$
\mathbf{M}^{\mathrm{loc}} \longrightarrow \prod_i \mathbf{M}^{\mathrm{loc}}_i.
$$

Note that each factor in the right hand side is the usual local model in [21] and is smooth over $\operatorname{Spec} O_E$ under our assumption of unramifiedness of the datum $\mathcal{D}_{\mathbb{Z}_p}$.

Lemma 3.7. *The morphism* $\mathbf{M}^{\mathrm{loc}} \to \prod_i \mathbf{M}_i^{\mathrm{loc}}$ *is smooth and quasi-projective. In particular, the scheme* $\mathbf{M}^{\mathrm{loc}}$ *is smooth over* $\operatorname{Spec} O_E$.

Proof. Similar to $\mathbf{M}^{\mathrm{loc}}$ we define a functor $\mathbf{M}^{\flat,\mathrm{loc}}$, only recording the first $\ell-1$ steps in $\mathcal{F}^{\bullet}$ in the EL case, and the quotients $\mathcal{F}^i/\mathcal{F}^1, i = 1, 2, \cdots, \ell-1$ in the PEL case. Then we have a natural morphism

$$(3.4) \qquad\qquad \pi : \mathbf{M}^{\mathrm{loc}} \longrightarrow \mathbf{M}^{\flat,\mathrm{loc}} \times \mathbf{M}_?^{\mathrm{loc}}$$

where $? = \ell$ in the EL case, and $? = 1$ in the PEL case.

In the EL case, by induction on ℓ, it suffices to show that π is representable, quasi-projective and smooth. Consider the commutative diagram

$$(3.5) \quad \begin{array}{ccccccccc} 0 & \longrightarrow & \Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & \Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & (\Lambda/\Lambda^{\ell-1}) \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle t_{\ell-1}} & & \downarrow{\scriptstyle t_\ell} & & \downarrow{\scriptstyle t} & & \\ 0 & \longrightarrow & \mathcal{F}^{\ell-1} & \longrightarrow & \mathcal{F} & \longrightarrow & \mathcal{F}/\mathcal{F}^{\ell-1} & \longrightarrow & 0 \end{array}$$

By the snake lemma we obtain a short exact sequence of locally free $O_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules:

$$0 \longrightarrow \ker(t_{\ell-1}) \longrightarrow \ker(t_\ell) \longrightarrow \ker(t) \longrightarrow 0 .$$

Moreover, we have an induced diagram of locally free $O_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules

$$(3.6) \quad \begin{array}{ccccccccc} 0 & \longrightarrow & \dfrac{\Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S}{\ker(t_{\ell-1})} & \longrightarrow & \mathcal{E} & \longrightarrow & \ker(t) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle =} & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \dfrac{\Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S}{\ker(t_{\ell-1})} & \longrightarrow & \dfrac{\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S}{\ker(t_{\ell-1})} & \longrightarrow & (\Lambda/\Lambda^{\ell-1}) \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & 0 \end{array}$$

where the top row is the pull-back of the bottom one via the right vertical map. Now, for any given $t_{\ell-1}$ and t, the datum of t_ℓ is equivalent to the datum of an $O_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-subbundle $\ker(t_\ell)/\ker(t_{\ell-1})$ of $\mathcal{E}$, required to have trivial intersection with $\frac{\Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S}{\ker(t_{\ell-1})}$. Note that the quotient bundle of $\mathcal{E}$ by $\ker(t_\ell)/\ker(t_{\ell-1})$ is isomorphic to $\mathcal{F}^{\ell-1}$, hence the characteristic polynomial for the action of $a \in O_B$ is given by $\operatorname{char}(a; (\operatorname{Fil}^{\ell-1}V)_0)$.

Locally (on $\mathbf{M}^{\flat,\mathrm{loc}} \times \mathbf{M}_\ell^{\mathrm{loc}}$, considering the universal bundles) we may trivialize both $\mathcal{E}$ and its subbundle $\mathcal{E}_0 := \frac{\Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S}{\ker(t_{\ell-1})}$. Then the fiber of π

is isomorphic to (an open subscheme of) the Hilbert/Quot scheme, denoted by $\mathrm{Hilb}_{\ker(t),(\mathrm{Fil}^{\ell-1}V)_0}$ (to indicate the action of O_B on the kernel and the cokernel of the map λ below), parameterizing surjective homomorphisms of $\mathcal{O}_B \otimes_{\mathbb{Z}_p} \mathcal{O}_S$-modules

$$(3.7) \qquad\qquad \lambda : \mathcal{E} \longrightarrow \mathcal{L},$$

such that $\mathrm{char}(a; \mathcal{L}) = \mathrm{char}(a; (\mathrm{Fil}^{\ell-1}V)_0), \forall a \in O_B$, with the additional condition that

$$(\heartsuit) \quad \lambda|_{\mathcal{E}_0} \text{ is an isomorphism.}$$

Without the condition $(\heartsuit)$, the Hilbert scheme is a product of Grassmannians, hence projective and smooth over $\mathrm{Spec}\, O_E$. The condition $(\heartsuit)$ is an open condition. Therefore, the map π is smooth and quasi-projective. Moreover, π is projective precisely when $\mathrm{rank}\,\mathcal{F}^{\ell-1} = 0$ or $\mathrm{rank}\,\ker(t) = 0$ (e.g., the two factors in the target of π are in the definite case, see Example 2.5).

We now turn to the PEL case. Note that $\mathbf{M}^{\flat,\mathrm{loc}}$ is the local model for the filtered RZ datum induced by $\mathrm{Fil}^{\ell-1}/\mathrm{Fil}^1 V$, hence also in the PEL case. By induction it suffices to show that the map (3.4) $\pi : \mathbf{M}^{\mathrm{loc}} \to \mathbf{M}^{\flat,\mathrm{loc}} \times \mathbf{M}_1^{\mathrm{loc}}$ is representable, quasi-projective and smooth. Consider the analogous functor $\mathbf{N}^{\mathrm{loc}}$, recording the first $\ell - 1$ steps in $\mathcal{F}^\bullet$. Then π factors as

$$(3.8) \qquad \mathbf{M}^{\mathrm{loc}} \xrightarrow{\ \pi_1\ } \mathbf{N}^{\mathrm{loc}} \xrightarrow{\ \pi_2\ } \mathbf{M}^{\flat,\mathrm{loc}} \times \mathbf{M}_1^{\mathrm{loc}}$$

and it suffices to show that both π_1 and π_2 are representable, quasi-projective and smooth. The case for π_2 is similar to the EL case and its fiber is locally a Hilbert scheme $\mathrm{Hilb}_{(\mathrm{Fil}^{\ell-1}/\mathrm{Fil}^1 V)_1,(\mathrm{Fil}^1 V)_0}$, where $(\mathrm{Fil}^{\ell-1}/\mathrm{Fil}^1 V)_1$ stipulates the action of O_B on $\ker(t)$ for t in the diagram:

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \Lambda^1 \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & \Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & (\Lambda^{\ell-1}/\Lambda^1) \otimes_{\mathbb{Z}_p} \mathcal{O}_S & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle t_1} & & \downarrow{\scriptstyle t_{\ell-1}} & & \downarrow{\scriptstyle t} & & \\
0 & \longrightarrow & \mathcal{F}^1 & \longrightarrow & \mathcal{F}^{\ell-1} & \longrightarrow & \mathcal{F}^{\ell-1}/\mathcal{F}^1 & \longrightarrow & 0.
\end{array}
$$

For the morphism π_1, we denote by $(\ker(t_i))^\perp$ the kernel of the composite map

$$(3.9) \qquad \Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S \xrightarrow{\ \sim\ } (\Lambda \otimes_{\mathbb{Z}_p} \mathcal{O}_S)^* \longrightarrow \ker(t_i)^*.$$

Note that $(\ker(t_\ell))^\perp = \ker(t_\ell)$ by the condition in the definition of local model. By (3.5), $\ker(t_{\ell-1})$ is contained in $\ker(t_\ell)$ and hence we have

$$\ker(t_{\ell-1}) \subset \ker(t_\ell) = (\ker(t_\ell))^\perp \subset (\ker(t_{\ell-1}))^\perp.$$

Consider the quotient $\mathcal{E} := (\ker(t_{\ell-1}))^\perp / \ker(t_{\ell-1})$ with the induced perfect alternating form. Then the datum of t_ℓ is equivalent to a subbundle $\mathcal{L} \subset \mathcal{E}$, which is a Lagrangian (i.e., the composite map $\mathcal{L} \to \mathcal{E} \to \mathcal{E}^* \to \mathcal{L}^*$ is zero), such that O_B acts on the quotient $\mathcal{E}/\mathcal{L}$ according to the action of O_B on $\ker(t_\ell)/\ker(t_{\ell-1}) \simeq \mathcal{F}_1^*$ and such that $\mathcal{L}$ intersects trivially with $(\Lambda^{\ell-1} \otimes_{\mathbb{Z}_p} \mathcal{O}_S)/\ker(t_{\ell-1})$. Now it is easy to see that π_1 is representable, quasi-projective and smooth. This completes the proof. $\qquad\square$

Let $\hat{\mathbf{M}}^{\mathrm{loc}}$ denote the p-adic completion of the base change $\mathbf{M}^{\mathrm{loc}} \times_{\mathrm{Spec}\, O_E} \mathrm{Spec}\, O_{\breve{E}}$. The Grothendieck–Messing deformation theory implies that any point of $\mathcal{M}$ has an étale neighborhood which is formally étale over $\hat{\mathbf{M}}^{\mathrm{loc}}$ (the argument of [21, §3.32] applies verbatim). Hence the smoothness claimed in Theorem 3.4 follows from Lemma 3.7.

The proof of Lemma 3.7 also tells us how to compute the relative dimension of $\mathbf{M}^{\mathrm{loc}}$ and hence $\mathcal{M}$. In Example 3.2 (the EL case), the relative dimension of the map (3.4) is given by that of the Hilbert scheme $\mathrm{Hilb}_{\ker(t),(\mathrm{Fil}^{\ell-1}V)_0}$ (cf. (2.3)):

$$(3.10) \qquad \sum_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F,\overline{\mathbb{Q}}_p)} \left(s_\varphi^\ell \sum_{i \leq \ell-1} r_\varphi^i \right).$$

It follows that the relative dimension of $\mathcal{M}$ over $\mathrm{Spf}\, O_{\breve{E}}$ is

$$(3.11) \qquad \sum_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F,\overline{\mathbb{Q}}_p)} \sum_{j=1}^{\ell} \left(s_\varphi^j \sum_{1 \leq i \leq j} r_\varphi^i \right).$$

In Example 3.3 (the unitary PEL case), similarly we get the relative dimension of the map (3.4) as the sum of that of the map π_1 and π_2 in (3.8). The relative dimension of the map π_2 is

$$\sum_{\varphi \in \mathrm{Hom}_{\mathbb{Q}_p}(F,\overline{\mathbb{Q}}_p)} \left(r_\varphi^1 \sum_{2 \leq i \leq \ell-1} s_\varphi^i \right).$$

By the total definiteness of $\mathrm{gr}^i, i \leq [(\ell-1)/2]$, the relative dimension of π_1 is zero.

Example 3.8. We specialize to the unitary PEL case: $F_0 = \mathbb{Q}_p$ and $\dim V = n$. Let $\mathrm{Hom}_{\mathbb{Q}_p}(F, \overline{\mathbb{Q}}_p) = \{\varphi, \overline{\varphi}\}$. Suppose

- $(r_\varphi, s_\varphi) = (r, n - r)$. Then $(r_{\overline{\varphi}}, s_{\overline{\varphi}}) = (n - r, r)$.

- $\ell = 2j + 1$ is odd, $\dim_F \mathrm{gr}^i V = 1$ for all $i \leq j$. Then $\dim_F \mathrm{gr}^{j+1} V = n - 2j$

- $(r_\varphi^i, s_\varphi^i) = (0, 1)$ for all $i \leq j$. Then $(r_\varphi^{j+1}, s_\varphi^{j+1}) = (r, n - r - 2j)$. (Implicitly $n \geq r + 2j$.)

Note that $r_{\overline{\varphi}}^i = s_\varphi^i$. By the formula above we find the relative dimension of π is equal to r and hence the relative dimension of $\mathbf{M}^{\mathrm{loc}}$ over $\mathrm{Spec}\, O_E$ is

$$(3.12) \qquad rj + r(n - r - 2j) = r(n - r) - rj.$$

As we will see in §5, this is the same as the expected dimension of KR cycle attached to a totally isotropic rank-j sublattice. In fact, the filtered RZ space in this case is the "smooth locus" of the corresponding KR cycle, at least when $r = 1$, cf. Prop. 5.3.

4. Ginzburg–Rallis cycle

A particularly interesting example in the EL case arises from the Ginzburg–Rallis period [5], which is (conjecturally) related to the exterior cube L-function on GL_6.

Consider the integral RZ datum $\mathcal{D}_{G, \mathbb{Z}_p}$ where $F = \mathbb{Q}_p, V = F^{\oplus 6}$ with a standard basis $\{e_1, \cdots, e_6\}$, $\mu = (0, 0, 0, 1, 1, 1)$ (i.e., the cocharacter $z \mapsto \mathrm{diag}(1, 1, 1, z, z, z)$), b basic, $\Lambda = \langle e_1, \cdots, e_6 \rangle$. Henceforth, we denote by $\langle e_1, \cdots, e_i \rangle$ (resp. $\langle e_1, \cdots, e_i \rangle_F$) the lattice (resp. the vector space) spanned by the specified vectors. To formulate the moduli space, we fix the framing object

$$\mathbb{X} = \mathbb{E} \times \mathbb{E} \times \mathbb{E},$$

where $\mathbb{E}$ is the (unique up-to-isomorphism) supersingular p-divisible group over $\mathbb{F}$ of dimension one and height 2. In terms of the notation in Example 2.2, we have $r = s = 3$. Then the RZ space $\mathcal{M}_G$ is formally smooth over $\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}$ of relative dimension $3 \times 3 = 9$ by the formula (2.3).

Consider the filtration

$$0 \subset \langle e_1, e_2 \rangle_F \subset \langle e_1, e_2, e_3, e_4 \rangle_F \subset V$$

Its stabilizer is the parabolic subgroup P with Levi subgroup isomorphic to GL_2^3. Denote the filtered RZ datum by $\mathcal{D}_{P,\mathbb{Z}_p}$ and the RZ space by $\mathcal{M}_P$. In terms of Example 3.2, we have $\ell = 3$ and

$$r^i = s^i = 1, \quad 1 \le i \le 3.$$

Then the RZ space $\mathcal{M}_P$ is formally smooth over $\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}$ of relative dimension, by the formula (3.11),

$$1 \times 1 + 1 \times 2 + 1 \times 3 = 6.$$

Note that $\mathcal{M}_P$ comes with two natural morphisms

$$
\begin{array}{ccc}
\mathcal{M}_P & \longrightarrow & \mathcal{M}_G \\
\downarrow & & \\
\mathbf{LT} \times_{\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}} \mathbf{LT} \times_{\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}} \mathbf{LT} & &
\end{array}
$$

where $\mathbf{LT}$ is the RZ space for $\mathrm{GL}_2, \mu = (0,1)$ and b basis (i.e., each connected component is isomorphic to the Lubin–Tate deformation space in the height two case).

The Ginzburg–Rallis subgroup of $G = \mathrm{GL}_6$ is defined by

$$H = P \times_{\mathrm{GL}_2^3, \Delta} \mathrm{GL}_2,$$

where $\Delta : \mathrm{GL}_2 \to \mathrm{GL}_2^3$ is the diagonal embedding. Then $H \simeq U \rtimes \mathrm{GL}_2$ where U denotes the unipotent radical. We define the (local) Ginzburg–Rallis cycle $\mathcal{M}_H$ as the fiber product

$$
\begin{array}{ccc}
\mathcal{M}_H & \longrightarrow & \mathcal{M}_P \\
\downarrow & & \downarrow \\
\mathbf{LT} & \xrightarrow{\ \Delta\ } & \mathbf{LT} \times_{\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}} \mathbf{LT} \times_{\mathrm{Spf}\, O_{\breve{\mathbb{Q}}_p}} \mathbf{LT}.
\end{array}
$$

By the proof of Lemma 3.7 (cf. (3.10)), the vertical map is smooth of relative dimension 3. In particular, we have the total dimension $\dim \mathcal{M}_H = 5$, exactly half of the total dimension $\dim \mathcal{M}_G = 10$:

$$i_{H,G} : \mathcal{M}_H \longrightarrow \mathcal{M}_G.$$

Remark 4.1. The drawback is that the morphism $i_{H,G}$ is not proper. One could consider a Laumon style compactification, but it is rather complicated and we hope to investigate its structure in the future. In fact, even the reduced scheme of $\mathcal{M}_G$ seems not understood.

Remark 4.2. There is a relative trace formula approach to the Ginzburg–Rallis period. We hope to formulate the Arithmetic Fundamental Lemma conjecture in that context in the future.

5. Connection to KR cycles

In this section, we fix a p-adic field F_0 and an unramified quadratic extension F. Let ϖ be a uniformizer of F_0 and q the residue cardinality of F_0.

5.1. The unitary RZ space

Fix an integer $n \geq 1$. We explicate (a slight variant of) the RZ formal moduli scheme $\mathcal{N}_n = \mathcal{N}_{n,F/F_0}$ associated to (a special case of) the *unramified* integral RZ datum in the unitary PEL case, cf. Example 2.3. For a greater generality, we will consider the relative unitary RZ space defined as follows, cf. [16] for more details.

For $S \in \mathtt{Nilp}$, we consider triples (X, ι, λ), where

- X is a p-divisible group of absolute height $2nd$ and dimension n over S, where $d := [F_0 : \mathbb{Q}_p]$,

- ι is an action of O_F such that the induced action of O_{F_0} on $\operatorname{Lie} X$ is via the structure morphism $O_{F_0} \to \mathcal{O}_S$, and

- λ is a principal polarization.

Hence $(X, \iota|_{O_{F_0}})$ is a formal (strict) O_{F_0}-module of relative height $2n$ and dimension n. We require that the Rosati involution $\operatorname{Ros}_\lambda$ on O_F is the non-trivial Galois automorphism in $\operatorname{Gal}(F/F_0)$, and that the *Kottwitz condition* of signature $(n-1, 1)$ is satisfied:

$$(5.1) \qquad \operatorname{char}(a; \operatorname{Lie} X) = (T-a)^{n-1}(T-\overline{a}) \in \mathcal{O}_S[T] \quad \text{for all} \quad a \in O_F.$$

An isomorphism $(X, \iota, \lambda) \xrightarrow{\sim} (X', \iota', \lambda')$ between two such triples is an O_F-linear isomorphism $\varphi \colon X \xrightarrow{\sim} X'$ such that $\varphi^*(\lambda') = \lambda$.

Over the residue field $\mathbb{F}$ of $O_{\breve{F}}$, we fix a *framing object*, i.e., a triple $(\mathbb{X}, \iota_{\mathbb{X}}, \lambda_{\mathbb{X}})$ such that $\mathbb{X}$ is supersingular, unique up to O_F-linear quasi-isogeny

compatible with the polarization. Then the formal scheme $\mathcal{N}_n$ (pro-)represents the functor

$$\mathcal{N}_n : \mathtt{Nilp} \longrightarrow Sets \,,$$

which associates to each $S \in \mathtt{Nilp}$ the set of isomorphism classes of quadruples $(X, \iota, \lambda, \rho)$ over S, where the final entry is an O_F-linear quasi-isogeny of *height zero*[4]

$$\rho \colon X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S},$$

such that $\rho^*((\lambda_{\mathbb{X}})_{\overline{S}}) = \lambda_{\overline{S}}$.

Remark 5.1. When F_0 is unramified over $\mathbb{Q}_p$, Mihatsch showed that in [16] the above $\mathcal{N}_n$ is isomorphic to a certain RZ space in Example 2.3. For example, when $F_0 = \mathbb{Q}_p$ in Example 2.3, we are in the case $B = F, \dim_F V = n$ and

$$(r_\varphi, s_\varphi) = (n-1, 1), \quad (r_{\overline{\varphi}}, s_{\overline{\varphi}}) = (1, n-1),$$

where $\varphi, \overline{\varphi} : F \hookrightarrow \overline{\mathbb{Q}}_p$ are the two distinct embeddings.

The formal scheme $\mathcal{N}_n$ is connected, formally smooth over $\mathrm{Spf}\, O_{\breve{F}}$ of relative dimension $n - 1$. When $n = 1$, the framing object is the supersingular formal O_{F_0}-module $(\overline{\mathbb{E}}, \iota_{\overline{\mathbb{E}}}, \lambda_{\overline{\mathbb{E}}})$ of (relative) height two and dimension one with signature $(0, 1)$, and $\mathcal{N}_1 \simeq \mathrm{Spf}\, O_{\breve{F}}$ with the universal object being the canonical lifting $(\overline{\mathscr{E}}, \iota_{\overline{\mathscr{E}}}, \lambda_{\overline{\mathscr{E}}}, \rho_{\overline{\mathscr{E}}})$. Here the bar in the notation indicates the action of O_F is through the Galois conjugate of the usual one. We also consider the signature $(1, 0)$ case with the framing object denoted by $(\mathbb{E}, \iota_{\mathbb{E}}, \lambda_{\mathbb{E}})$ and the universal object $(\mathscr{E}, \iota_{\mathscr{E}}, \lambda_{\mathscr{E}}, \rho_{\mathscr{E}})$.

We make an explicit choice of the framing object $\mathbb{X}_n = \mathbb{E}^{n-1} \times \overline{\mathbb{E}}$ with the obvious choice of $\iota_{\mathbb{X}_n}, \lambda_{\mathbb{X}_n}$. Let

$$(5.2) \qquad\qquad\qquad \mathbb{V}_n := \mathrm{Hom}^\circ_{O_F}(\mathbb{E}, \mathbb{X}_n)$$

be the F/F_0-Hermitian space of *special quasi-homomorphisms*, where the Hermitian form is induced by the principal polarizations on $\mathbb{X}_n$ and $\mathbb{E}$: for $x, y \in \mathbb{V}_n$,

$$(x, y) := \lambda_{\mathbb{E}}^{-1} \circ y^\vee \circ \lambda_{\mathbb{X}} \circ x \in \mathrm{End}^\circ_{O_F}(\mathbb{E}) \simeq F.$$

Note that $\mathbb{V}$ is a non-split Hermitian space, i.e., it does not contain a self-dual lattice.

---

[4]In §2.2 on RZ space of PEL type, there was no restriction on the height of ρ. Hence here the space $\mathcal{N}_n$ is a connected component of what's defined earlier.

Let $\mathrm{Aut}^\circ(\mathbb{X}_n, \iota_{\mathbb{X}_n}, \lambda_{\mathbb{X}_n})$ be the group of self-quasi-isogenies of the framing object (here we require self-quasi-isogenies to respect the auxiliary structure $\iota_{\mathbb{X}_n}$ and $\lambda_{\mathbb{X}_n}$). It acts naturally on $\mathbb{V}_n$ and induces an isomorphism

$$(5.3) \qquad \mathrm{Aut}^\circ(\mathbb{X}_n, \iota_{\mathbb{X}_n}, \lambda_{\mathbb{X}_n}) \overset{\sim}{\longrightarrow} \mathrm{U}(\mathbb{V}_n)(F_0) \ ,$$

Then the group $\mathrm{U}(\mathbb{V}_n)(F_0)$ acts via the inverse of (5.3) on $\mathcal{N}_n$ by changing the framing:

$$g \cdot (X, \iota, \lambda, \rho) = (X, \iota, \lambda, g \circ \rho).$$

Remark 5.2. We may also consider the totally definite case (cf. Example 2.5), i.e., the signature $(n, 0)$ case. We may fix the framing object as $\mathbb{X}'_n = \mathbb{E}^n$ with the obvious choice of $\iota_{\mathbb{X}'_n}, \lambda_{\mathbb{X}'_n}$. Then the resulting moduli space, denoted by $\mathcal{N}'_n$, is formally smooth of relative dimension zero. Similarly, define $\mathbb{V}'_n = \mathrm{Hom}^\circ_{O_F}(\mathbb{E}, \mathbb{X}'_n)$, a split Hermitian space (hence isomorphic to V in the RZ datum). Let $G = \mathrm{U}(\mathbb{V}'_n)$ and let K the stabilizer of the lattice $\mathrm{Hom}_{O_F}(\mathbb{E}, \mathbb{X}'_n)$, which is a hyperspecial subgroup of G. Then there is a $G(\mathbb{Q}_p)$-equivariant isomorphism

$$\mathcal{N}'_n \overset{\sim}{\longrightarrow} \bigsqcup_{G(\mathbb{Q}_p)/K} \mathrm{Spf}\, O_{\breve{F}} \ .$$

Here $G(\mathbb{Q}_p)/K$ is naturally bijective to the set of self-dual lattices in $\mathbb{V}'_n$.

5.2. KR cycles

In the sections below, when n is fixed, we will simplify write $\mathbb{V}$ (resp. $\mathbb{X}$) for $\mathbb{V}_n$ (resp. $\mathbb{X}_n$). For any subset $L \subseteq \mathbb{V}$, we recall from [10, Def. 3.2] that the *Kudla–Rapoport cycle* (or *special cycle*) $\mathcal{Z}(L)$ is the closed formal subscheme of $\mathcal{N}_n$, which represents the functor sending each $S \in \mathtt{Nilp}$ to the set of isomorphism classes of tuples $(X, \iota, \lambda, \rho) \in \mathcal{N}_n(S)$ such that any (quasi-homomorphism) $x : \mathbb{E} \to \mathbb{X}$ in L lifts to a homomorphism $\mathscr{E}_S \to X$, in the sense that the quasi-homomorphism

$$\rho^{-1} \circ x \circ \rho_{\mathscr{E}} : \mathscr{E}_S \times_S \overline{S} \xrightarrow{\rho_{\mathscr{E}}} \mathbb{E} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S} \xrightarrow{x} \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S} \xrightarrow{\rho^{-1}} X \times_S \overline{S}$$

extends to a homomorphism $\mathscr{E}_S \to X$. Note that $\mathcal{Z}(L)$ only depends on the O_F-linear span of L in $\mathbb{V}$.

Let $\mathcal{Z}(L)^\dagger$ be the "(formally) smooth locus" of the morphism $\mathcal{Z}(L) \to \mathrm{Spf}\, O_{\breve{F}}$, i.e., the maximal open formal subscheme of $\mathcal{Z}(L)$ which is flat over

$\operatorname{Spf} O_{\breve{F}}$ and whose relative tangent spaces over $\operatorname{Spf} O_{\breve{F}}$ have a constant dimension.

Note that when $L = \{x\}$ consists of a single non-zero element x, we also denote the special cycle by $\mathcal{Z}(x)$. By [10], $\mathcal{Z}(x)$ is a Cartier divisor on $\mathcal{N}_n$, flat over $\operatorname{Spf} O_{\breve{F}}$.

5.3. Filtered RZ space

Let $m, r \in \mathbb{Z}_{\geq 0}$ and $n = m + 1 + 2r$. We consider the filtered RZ datum in Example 3.3, with the self-dual filtration

$$0 = \operatorname{Fil}^0 V \subset \operatorname{Fil}^1 V \subset \cdots \subset \operatorname{Fil}^\ell V = V$$

of length $\ell = 2r + 1$. We further assume that

- The graded quotients have dimensions:

$$\dim \operatorname{gr}^i V = \begin{cases} 1, & i \neq r + 1, \\ n - 2r, & i = r + 1. \end{cases}$$

- The signature $(r_\varphi^i, s_\varphi^i) = (1, 0)$ for all $i \leq r$. (Then the signature of $\operatorname{gr}^{r+1} V$ is $(n - 1 - 2r, 1)$.)

We denote by $P = P_{1^r, m+1, 1^r}$ the parabolic subgroup of G stabilizing the filtration.

We explicate the formal moduli functor. For $S \in \texttt{Nilp}$, we consider tuples $(X, X_\bullet, \iota, \lambda)$, where (X, ι, λ) is as in §5.1, and

$$0 = X_0 \subset X_1 \subset \cdots \subset X_{2r+1} = X$$

is a self-dual filtration (cf. §3.2) such that the following Kottwitz condition holds: $\forall a \in O_F$,

$$\operatorname{char}\big(a; \operatorname{Lie}(X_i)\big) = \begin{cases} (T - a)^i, & i \leq r, \\ (T - a)^{n-1-j}(T - \overline{a}), & i = 2r + 1 - j \geq r. \end{cases}$$

Fix a framing object $(\mathbb{X}, \mathbb{X}_\bullet, \iota_{\mathbb{X}}, \lambda_{\mathbb{X}})$ over $\operatorname{Spec} \mathbb{F}$. Then we define a functor

$$\mathcal{N}_P : \texttt{Nilp} \longrightarrow Sets\,,$$

which associates to $S \in \mathtt{Nilp}$ the set of isomorphism classes $(X, X_\bullet, \iota, \lambda, \rho)$ where $(X, \mathrm{Fil}^\bullet X, \iota, \lambda)$ is as above and

$$\rho : X \times_S \overline{S} \longrightarrow \mathbb{X} \times_{\mathrm{Spec}\,\mathbb{F}} \overline{S}$$

is an O_F-linear quasi-isogeny of *height zero* that preserves the filtration $X_\bullet$ and $\mathbb{X}_\bullet$, and such that $\rho^*((\lambda_{\mathbb{X}})_{\overline{S}}) = \lambda_{\overline{S}}$.

Let $\mathcal{N}_{1^r}$ denote the functor recording only the first r-steps $(X_r, 0 = X_0 \subset X_1 \subset \cdots \subset X_r, \iota|_{X_r}, \rho|_{X_r})$ (without any polarization). It is associated to a filtered RZ datum of EL type, cf. §3.1 and §3.2. To avoid confusion with the unitary RZ space $\mathcal{N}_1$ (the special case $n = 1$ of $\mathcal{N}_n$ in §5.1), we will denote

$$(5.4) \qquad\qquad \mathcal{M}_1 = \mathcal{N}_{1^r}, \quad \text{when } r = 1.$$

Then we have a natural morphism

$$(5.5) \qquad\qquad \mathcal{N}_P \longrightarrow \mathcal{N}_{1^r} \times \mathcal{N}_{n-2r} \ ,$$

where the second factor sends $(X, X_\bullet, \iota, \lambda, \rho)$ to the p-divisible group X_{r+1}/X_r with the induced additional structure. Note that $\mathcal{N}_{1^r}$ is an example of totally definite RZ space of EL type (cf. Example 2.5), and hence has relative dimension zero over $\mathrm{Spf}\,O_{\breve{F}}$. A more explicit description of $\mathcal{N}_{1^r}$ is as follows. Let $E_i = \mathrm{Hom}^\circ_{O_F}(\mathbb{E}, \mathbb{X}_i) \subset \mathbb{V}$ and $E = E_r$. (A notational remark: henceforth the letter E will no longer denote the reflex field of a local Shimura datum.) Then there is a natural bijection

$$(5.6) \qquad\qquad \mathcal{N}_{1^r} \overset{\sim}{\longrightarrow} \coprod_{\mathcal{FL}_E} \mathrm{Spf}\,O_{\breve{F}} \ ,$$

where $\mathcal{FL}_E$ the set of complete flags $\mathcal{E}_\bullet$ of lattices in $E = E_r$

$$0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 \subset \cdots \subset \mathcal{E}_r = \mathcal{E}$$

such that $\mathcal{E}_i = \mathcal{E} \cap E_i$ is of rank i. Upon choosing a basis of E[5], the index set is then isomorphic to $B_r(F)/B_r(O_F)$ where $B_r \subset \mathrm{GL}_{r,F}$ is the Borel subgroup of upper triangular matrices. Note that the set $\mathcal{FL}_E$ is also isomorphic to $\mathrm{GL}_r(F)/\mathrm{GL}_r(O_F)$ via the map sending the chain $\mathcal{E}_\bullet$ to the largest lattice $\mathcal{E} \subset \mathbb{V}_r$. We denote by $\mathcal{N}_{P,\mathcal{E}_\bullet}$ the fiber of the map over the copy of $\mathrm{Spf}\,O_{\breve{F}}$

---

[5]We need less than the basis; it suffices to fix a base point in the set $\{\mathcal{E}_\bullet\}$.

indexed by $\mathcal{E}_\bullet \in \mathcal{FL}_E$:

(5.7)
$$
\begin{array}{ccc}
\mathcal{N}_{P,\mathcal{E}_\bullet} & \longrightarrow & \mathrm{Spf}\, O_{\breve{F}} \\
\downarrow & & \downarrow{\scriptstyle \mathcal{E}_\bullet} \\
\mathcal{N}_P & \longrightarrow & \mathcal{N}_{1^r}
\end{array}
$$

Similar to (5.5) we may further define a natural morphism, recording the graded quotients X_i/X_{i-1} for $i = 1, \cdots, r+1$

(5.8)
$$
\mathcal{N}_P \longrightarrow \mathcal{N}_M := \mathcal{M}_1 \times_{\mathrm{Spf}\, O_{\breve{F}}} \cdots \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{M}_1 \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{N}_{m+1}
$$

where M stands for the Levi of P, which is isomorphic to $\mathrm{Res}_{E/F}(\mathrm{GL}_1)^r \times \mathrm{U}(E^\perp/E)$. Here the induced Hermitian form on $E^\perp/E \simeq \mathbb{V}_{m+1}$ is non-degenerate (also cf. (6.1) below). Then the case $r = 1$ of (5.6) asserts (with the chosen basis)

(5.9)
$$
\mathcal{M}_1 \xrightarrow{\ \sim\ } \coprod_{F^\times/O_F^\times \simeq \mathbb{Z}} \mathrm{Spf}\, O_{\breve{F}} \,.
$$

5.4. Connection to KR cycles

For the rest of this section, we fix a flag of lattices $\mathcal{E}_\bullet \in \mathcal{FL}_E$ as above, and let $\mathcal{E}$ be the largest one in the flag, which is of rank r. Recall from §5.2 that we have $\mathcal{Z}(\mathcal{E})^\dagger$, the "(formally) smooth locus" of the morphism $\mathcal{Z}(\mathcal{E}) \to \mathrm{Spf}\, O_{\breve{F}}$. From the definition of $\mathcal{N}_{P,\mathcal{E}_\bullet}$, it is clear that there is a natural morphism of formal schemes $\mathcal{N}_{P,\mathcal{E}_\bullet} \to \mathcal{Z}(\mathcal{E})$.

Proposition 5.3. *There is a natural isomorphism:*

$$
\mathcal{N}_{P,\mathcal{E}_\bullet} \xrightarrow{\ \sim\ } \mathcal{Z}(\mathcal{E})^\dagger.
$$

Proof. Let $(X, X_\bullet, \iota, \lambda) \in \mathcal{N}_{P,\mathcal{E}_\bullet}(S)$. Since $\mathcal{E}$ is an O_F-lattice of rank r, the formal scheme $\mathcal{Z}(\mathcal{E})$ is an intersection of r Cartier divisors on $\mathcal{N}_n$. Therefore, $\mathcal{Z}(\mathcal{E})^\dagger$ is the open formal subscheme of $\mathcal{Z}(\mathcal{E})$ with underlying reduced scheme consisting of points on the reduced scheme $\mathcal{Z}(\mathcal{E})^{\mathrm{red}}$ where the relative tangent spaces over $\mathrm{Spf}\, O_{\breve{F}}$ have a constant dimension $n - r - 1$ (once we show the non-emptiness of such points, which is clear from the argument below). Hence it suffices to show that the condition on the tangent space is equivalent to the condition that the homomorphism $\mathscr{E}_S \otimes_{O_F} \mathcal{E} \to X$ is a monomorphism, where $\mathscr{E}_S \otimes_{O_F} \mathcal{E}$ is Serre's tensor construction. This equivalence can be checked on

geometric points of S and it suffices to consider the case where $S = \operatorname{Spec} k$ is a geometric point x on $\mathcal{N}_n^{\mathrm{red}}$. Below our argument works for an arbitrary algebraically closed field k in characteristic p; but for notational simplicity we will restrict ourselves to the case $k = \mathbb{F}$.

We recall some preliminary constructions (see [25], [10, §2], [19, §3], [12, §2] for more details). Let $\mathbb{D}(X)$ be the (covariant) Dieudonné module of X with the induced action by

$$O_F \otimes_{O_{F_0}} O_{\breve{F}} \simeq \prod_{\operatorname{Hom}_{O_{F_0}}(O_F, O_{\breve{F}})} O_{\breve{F}}.$$

Let $A = \mathbb{D}(X)_1$ be the eigenspace corresponding to the Galois conjugate of the tautological embedding $O_F \subset O_{\breve{F}}$. Via the framing ρ we may identify the $\breve{F}$-vector space $A_{\breve{F}} := A \otimes_{O_{\breve{F}}} \breve{F}$ with $\mathbb{V}_{\breve{F}} := \mathbb{V} \otimes_F \breve{F}$, respecting the σ-sesquilinear pairing on $A_{\breve{F}}$ induced by the polarization and the extension of the Hermitian paring on $\mathbb{V}$ to $\mathbb{V}_{\breve{F}}$. Then the lattice A is a *special lattice* in $\mathbb{V}_{\breve{F}}$, in the sense that $A^\vee \subset^1 A \subset \varpi^{-1} A^\vee$, where $A^\vee$ is the dual lattice. Then a point $x \in \mathcal{N}_n^{\mathrm{red}}$ lies on $\mathcal{Z}(\mathcal{E})$ if and only if the $\mathcal{E} \subset A^\vee$.

The map $\mathscr{E}_S \otimes_{O_F} \mathcal{E} \to X$ is a monomorphism if and only if the induced $O_{\breve{F}}$-linear map $\mathcal{E} \otimes_{O_F} O_{\breve{F}} \to \mathbb{D}(X)$ is saturated (i.e., the cokernel is torsion free). By the action on Lie algebra we may replace $\mathbb{D}(X)$ by $A = \mathbb{D}(X)_1$. Furthermore, the saturation is equivalent to the condition that the $\mathbb{F}$-linear map

$$(5.10) \qquad \mathcal{E} \otimes_{O_F} \mathbb{F} \longrightarrow A \otimes_{O_{\breve{F}}} \mathbb{F}$$

is injective.

By Grothendieck–Messing theory, the tangent space of $\mathcal{Z}(\mathcal{E}) \times_{\operatorname{Spf} O_{\breve{F}}} \operatorname{Spec} \mathbb{F}$ at x is isomorphic to the cokernel of the $\mathbb{F}$-linear map (cf. [12])

$$(5.11) \qquad \mathcal{E} \otimes_{O_F} \mathbb{F} \longrightarrow A^\vee / \varpi A$$

where the target $A^\vee / \varpi A$ is an $\mathbb{F}$-subspace in $A/\varpi A = A \otimes_{O_{\breve{F}}} \mathbb{F}$ of codimension one. Note that the KR cycle $\mathcal{Z}(\mathcal{E})$ is the intersection of $r = \operatorname{rank} \mathcal{E}$ Cartier divisors. The smoothness at x is then equivalent to the above $\mathbb{F}$-linear map (5.11) is injective, which is equivalent to the injectivity of the map (5.10). This completes the proof. $\qquad \square$

We now determine the reduced scheme $\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}}$ of $\mathcal{Z}(\mathcal{E})^\dagger$, for $\mathcal{E}$ arising from a flag of lattices $\mathcal{E}_\bullet \in \mathcal{FL}_E$.

Proposition 5.4. *Let $\mathcal{E}$ be a totally isotropic lattice of rank r (every such $\mathcal{E}$ arises from a flag $\mathcal{E}_\bullet \in \mathcal{FL}_E$). Then we have*

$$\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} = \mathcal{Z}(\mathcal{E})^{\mathrm{red}} \setminus \cup_{\mathcal{E} \subsetneq \mathcal{E}' \subset E} \mathcal{Z}(\mathcal{E}')^{\mathrm{red}},$$

where the union runs over all O_F-lattices $\mathcal{E}'$ of rank r in $E = \mathcal{E} \otimes_{O_F} F$, and

$$\mathcal{Z}(\mathcal{E})^\dagger = \mathcal{Z}(\mathcal{E}) \setminus \cup_{\mathcal{E} \subsetneq \mathcal{E}' \subset E} \mathcal{Z}(\mathcal{E}'),$$

where the right hand side denotes the open formal subscheme of $\mathcal{Z}(\mathcal{E})$ whose topological space is the open subscheme $\mathcal{Z}(\mathcal{E})^{\mathrm{red}} \setminus \cup_{\mathcal{E} \subsetneq \mathcal{E}' \subset E} \mathcal{Z}(\mathcal{E}')^{\mathrm{red}}$ of $\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}}$.

Proof. It suffices to prove the first assertion on the reduced schemes. We use the criterion of smoothness via the injectivity of (5.10). It is clear that $\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} \subset \mathcal{Z}(\mathcal{E})^{\mathrm{red}} \setminus \cup_{\mathcal{E} \subsetneq \mathcal{E}'} \mathcal{Z}(\mathcal{E}')^{\mathrm{red}}$. Now let $x \in \mathcal{Z}(\mathcal{E})(\mathbb{F})$ and $A \subset \mathbb{V}_{\breve{F}}$ the corresponding special lattice. Then $\mathcal{E} \subset A^\vee$.

It remains to show that, if the map (5.10) is not injective, then x lies on $\mathcal{Z}(\mathcal{E}')^{\mathrm{red}}$ for some $\mathcal{E} \subsetneq \mathcal{E}' \subset E$. Let $\mathcal{E}' = A \cap E$. The non-injectivity of (5.10) implies that $\mathcal{E} \neq \mathcal{E}'$. Since E is totally isotropic, all vectors in $\mathcal{E}'$ have integral Hermitian norms. By the claim below, we have $\mathcal{E} \subset A^\vee$ and hence the point $x \in \mathcal{Z}(\mathcal{E}')$.

Claim: Let u be a vector in $\mathbb{V}$ such that $(u,u) \in O_F$. If $u \in A$, then $u \in A^\vee$.

To show the claim, we denote by $\{\cdot,\cdot\}$ the σ-sesquilinear pairing on $\mathbb{V}_{\breve{F}}$ and denote by Φ the automorphism $\mathrm{id}_\mathbb{V} \otimes \sigma$. The dual lattice $A^\vee$ is characterized by $\{A^\vee, A\} = O_{\breve{F}}$ and we have $(A^\vee)^\vee = \Phi(A)$. It follows that $u = \Phi(u) \in \Phi(A)$ and hence we have $\{u, A^\vee\} \subset O_{\breve{F}}$.

Now suppose that $u \in A \setminus A^\vee$. Since $\dim_\mathbb{F} A/A^\vee = 1$, we must have

$$A = \langle u \rangle_{O_{\breve{F}}} + A^\vee.$$

It follows from $(u,u) = \{u,u\} \in O_F$ that

$$\{u, A\} \subset O_{\breve{F}}$$

and hence $u \in A^\vee$. Contradiction! $\qquad\qquad\square$

Remark 5.5. The assertion fails without the totally isotropic condition on $\mathcal{E}$. For example, consider a rank one $\mathcal{E} = \langle u \rangle$ with a generator u such that its Hermitian norm $(u,u) = \varpi$. The special divisor $\mathcal{Z}(\mathcal{E}) = \mathcal{Z}(u)$ is not formally smooth over $\mathrm{Spf}\, O_{\breve{F}}$ while $\mathcal{Z}(\mathcal{E}')$ is empty for any lattice $\mathcal{E}' \subset E$ such that $\mathcal{E}' \supsetneq \mathcal{E}$.

Example 5.6 (The case rank $\mathcal{E} = 1$). Let $\mathcal{E} = \langle u \rangle$ be a rank one lattice generated by a non-zero $u \in \mathbb{V}$ with Hermitian norm $(u, u) = 0$. Then by Prop. 5.4,

$$\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} = \mathcal{Z}(u)^{\mathrm{red}} \setminus \mathcal{Z}(\varpi^{-1}u)^{\mathrm{red}}, \tag{5.12}$$

and

$$\mathcal{Z}(\mathcal{E})^{\dagger} = \mathcal{Z}(u) \setminus \mathcal{Z}(\varpi^{-1}u), \tag{5.13}$$

the restriction of $\mathcal{Z}(u)$ to the open subscheme $\mathcal{Z}(u)^{\mathrm{red}} \setminus \mathcal{Z}(\varpi^{-1}u)^{\mathrm{red}}$ of $\mathcal{Z}(u)^{\mathrm{red}}$. Note that $\mathcal{Z}(\varpi^{-1}u)$ is a Cartier divisor and is a closed formal subscheme of $\mathcal{Z}(u)$ (though abstractly they are isomorphic!). The difference divisor $\mathcal{D}(u)$ is defined as the difference of Cartier divisors

$$\mathcal{D}(u) := \mathcal{Z}(u) - \mathcal{Z}(\varpi^{-1}u), \tag{5.14}$$

i.e., $\mathcal{D}(u)$ is locally defined by $f_u/f_{\varpi^{-1}u} = 0$, where f_u and $f_{\varpi^{-1}u}$ are respectively the local equations defining $\mathcal{Z}(u)$ and $\mathcal{Z}(\varpi^{-1}u)$. Then $\mathcal{Z}(\mathcal{E})^{\dagger}$ is an open formal subscheme of $\mathcal{D}(u)$ and hence we may view $\mathcal{D}(u)$ as a compactification of $\mathcal{Z}(\mathcal{E})^{\dagger}$. It is analogous to the Laumon style compactification of $\mathrm{Bun}_P \to \mathrm{Bun}_{\mathrm{GL}_n}$ for a parabolic P of GL_n.

5.5. The Bruhat–Tits stratification of $\mathcal{Z}(\mathcal{E})^{\dagger,\mathbf{red}}$

We first recall the Bruhat–Tits stratification of the reduced scheme $\mathcal{N}_n^{\mathrm{red}}$ of $\mathcal{N}_n$, following the theorem of Vollaard–Wedhorn [25][6], see also [11, §2.7].

Let $\mathrm{Vert}(\mathbb{V})$ denote the set of all vertex lattices $\Lambda \subseteq \mathbb{V}$, i.e., lattices (of full rank) such that $\Lambda \subset \Lambda^{\vee} \subset \varpi^{-1}\Lambda$. The reduced subscheme of $\mathcal{N}_n$ is a union of closed subschemes $\mathcal{V}(\Lambda)$

$$\mathcal{N}_n^{\mathrm{red}} = \bigcup_{\Lambda \in \mathrm{Vert}(\mathbb{V})} \mathcal{V}(\Lambda).$$

Here $\mathcal{V}(\Lambda)$ is the generalized Deligne–Lusztig variety [11, §2.5], a smooth projective variety over $\mathbb{F}$ of dimension $\frac{t(\Lambda)-1}{2}$. Here

$$t(\Lambda) := \dim \Lambda^{\vee}/\Lambda$$

---

[6]Our convention on vertex lattices slightly differ: the Hermtian form have integral values on our lattices Λ, which correspond to the dual lattices of those while in *loc. cit.*.

is called the type of Λ, which is an odd integer. For two vertex lattices Λ, Λ', we have $\mathcal{V}(\Lambda) \subseteq \mathcal{V}(\Lambda')$ if and only if $\Lambda \supseteq \Lambda'$; and $\mathcal{V}(\Lambda) \cap \mathcal{V}(\Lambda')$ is nonempty if and only if $\Lambda + \Lambda'$ is also a vertex lattice, in which case it is equal to $\mathcal{V}(\Lambda + \Lambda')$. We therefore have a *Bruhat–Tits stratification* by locally closed subvarieties,

$$\mathcal{N}_n^{\mathrm{red}} = \bigsqcup_{\Lambda \in \mathrm{Vert}(\mathbb{V})} \mathcal{V}(\Lambda)^\circ, \quad \mathcal{V}(\Lambda)^\circ := \mathcal{V}(\Lambda) \setminus \bigcup_{\Lambda \subsetneq \Lambda'} \mathcal{V}(\Lambda').$$

The irreducible components of $\mathcal{N}_n^{\mathrm{red}}$ are exactly the projective varieties $\mathcal{V}(\Lambda)$, where Λ runs over all vertex lattices of maximal type. The "dual graph" of the stratification is the Bruhat–Tits building of the unitary group $\mathrm{SU}(\mathbb{V})$.

For an O_F-lattice $\mathcal{E} \subseteq \mathbb{V}$ of rank $r \geq 1$, we denote $E = \mathcal{E} \otimes_{O_F} F$ and define a subset of $\mathrm{Vert}(\mathbb{V})$:

$$\mathrm{Vert}_{\mathcal{E}} := \{\Lambda \in \mathrm{Vert}(\mathbb{V}) \mid \mathcal{E} \subset \Lambda\}.$$

By [10, Prop. 4.1], the reduced subscheme $\mathcal{Z}(\mathcal{E})^{\mathrm{red}}$ of the KR cycle $\mathcal{Z}(\mathcal{E})$ is a union of Bruhat–Tits strata,

$$(5.15) \qquad \mathcal{Z}(\mathcal{E})^{\mathrm{red}} = \bigcup_{\Lambda \in \mathrm{Vert}_{\mathcal{E}}} \mathcal{V}(\Lambda) = \bigsqcup_{\Lambda \in \mathrm{Vert}_{\mathcal{E}}} \mathcal{V}(\Lambda)^\circ.$$

In fact, by [19, Thm. 10.1], we have for any $\Lambda \in \mathrm{Vert}(\mathbb{V})$

$$(5.16) \qquad\qquad\qquad \mathcal{Z}(\Lambda) \simeq \mathcal{V}(\Lambda)$$

as closed formal subschemes of $\mathcal{N}_n$, which provides a moduli theoretical interpretation of $\mathcal{V}(\Lambda)$.

Remark 5.7. In [24] (in the case $F_0 = \mathbb{Q}_p$) Vandenbergen proved that $\mathcal{Z}(\mathcal{E})^{\mathrm{red}}$ is connected when $E = \mathcal{E} \otimes_{O_F} F$ is a *non-degenerate* Hermitian subspace of $\mathbb{V}$. It is likely the method also proves the connectedness for any $\mathcal{E}$, particularly totally isotropic $\mathcal{E}$ for our interest in this paper.

Now we consider those $\Lambda \in \mathrm{Vert}_{\mathcal{E}}$ such that $\mathcal{E}$ is saturated in Λ:

$$\mathrm{Vert}_{\mathcal{E}}^\dagger := \{\Lambda \in \mathrm{Vert}_{\mathcal{E}} \mid \mathcal{E} = E \cap \Lambda\}.$$

For $\Lambda \in \mathrm{Vert}_{\mathcal{E}}$, we define

$$\mathcal{V}(\Lambda)_{\mathcal{E}}^\dagger = \mathcal{V}(\Lambda) \setminus \bigcup_{\Lambda \subsetneq \Lambda', \Lambda' \notin \mathrm{Vert}_{\mathcal{E}}^\dagger} \mathcal{V}(\Lambda').$$

It is easy to see that

$$\mathcal{V}(\Lambda)^\dagger_\mathcal{E} = \bigsqcup_{\Lambda \subsetneq \Lambda' \in \mathrm{Vert}^\dagger_\mathcal{E}} \mathcal{V}(\Lambda')^\circ.$$

Then by Prop. 5.4 and (5.15), we have a Bruhat–Tits stratification of $\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}}$.

$$(5.17) \qquad \mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} = \bigcup_{\Lambda \in \mathrm{Vert}^\dagger_\mathcal{E}} \mathcal{V}(\Lambda)^\dagger_\mathcal{E} = \bigsqcup_{\Lambda \in \mathrm{Vert}^\dagger_\mathcal{E}} \mathcal{V}(\Lambda)^\circ.$$

In fact, combining with (5.16) we have

$$(5.18) \qquad \mathcal{V}(\Lambda)^\dagger_\mathcal{E} = \mathcal{Z}(\Lambda) \cap \mathcal{Z}(\mathcal{E})^\dagger.$$

5.6. The morphism $\mathcal{N}^{\mathrm{red}}_{P,\mathcal{E}_\bullet} \simeq \mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} \to \mathcal{N}^{\mathrm{red}}_{n-2r}$

Recall by (5.7), (5.8) and Prop. 5.3 that we have a morphism

$$\mathcal{N}_{P,\mathcal{E}_\bullet} \simeq \mathcal{Z}(\mathcal{E})^\dagger \longrightarrow \mathcal{N}_{n-2r} \ .$$

This morphism appears in the formulation of the Arithmetic Fundamental Lemma conjecture in §7. Therefore, it is desirable to understand as much about it as possible. In this subsection, we make a first attempt to describe the induced morphism on their reduced schemes $\mathcal{N}^{\mathrm{red}}_{P,\mathcal{E}_\bullet} \simeq \mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}} \to \mathcal{N}^{\mathrm{red}}_{n-2r}$. The resulting morphism seems rather complicated and we can not prove any general result. Instead, we give some examples and raise some questions on the refined structure of the reduced scheme $\mathcal{N}^{\mathrm{red}}_{P,\mathcal{E}_\bullet} \simeq \mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}}$; see Remark 5.8 and 5.11 below. We point out that this subsection is not used in §7 and the reader who is only interested in the formulation of the Arithmetic Fundamental Lemma conjecture can safely skip it.

Denote by $\mathbb{V}^\flat = E^\perp/E$ with the induced Hermitian form. There is a map

$$(5.19) \qquad \mathrm{Vert}^\dagger_\mathcal{E} \longrightarrow \mathrm{Vert}(\mathbb{V}^\flat)$$

sending Λ to $\Lambda^\flat = (E^\perp \cap \Lambda)/E \cap \Lambda = (E^\perp \cap \Lambda)/\mathcal{E}$. By Lemma 3.1, $\Lambda^\flat$ is a vertex lattice and the type does not increase:

$$t(\Lambda^\flat) \le t(\Lambda).$$

Fix a $\Lambda \in \mathrm{Vert}_{\mathcal{E}}^{\dagger}$ and its image $\Lambda^{\flat} \in \mathrm{Vert}(\mathbb{V}^{\flat})$. Then, by (5.18), we obtain a morphism $\mathcal{V}(\Lambda)_{\mathcal{E}}^{\dagger} = \mathcal{Z}(\Lambda) \cap \mathcal{Z}(\mathcal{E})^{\dagger} \to \mathcal{N}_{m+1}$ and it is easy to see that the image lies in $\mathcal{Z}(\Lambda^{\flat}) = \mathcal{V}(\Lambda^{\flat})$. The resulting morphism is denoted by

$$(5.20) \qquad \pi_{\Lambda,\Lambda^{\flat}} : \mathcal{V}(\Lambda)_{\mathcal{E}}^{\dagger} \longrightarrow \mathcal{V}(\Lambda^{\flat}) \ .$$

One can describe $\pi_{\Lambda,\Lambda^{\flat}}$ in terms of the special lattice $A \in \mathbb{V}_{\breve{F}}$ (cf. the proof of Prop. 5.3), but it is not clear to the author what properties can be said about the morphism. The map $\pi_{\Lambda,\Lambda^{\flat}}$ behaves well with the change of vertex lattices $\Lambda \subset \Lambda'$ in $\mathrm{Vert}_{\mathcal{E}}$: there is a natural commutative diagram

$$\begin{array}{ccc}
\mathcal{V}(\Lambda')_{\mathcal{E}}^{\dagger} & \longrightarrow & \mathcal{V}(\Lambda'^{\flat}) \\
\downarrow & & \downarrow \\
\mathcal{V}(\Lambda)_{\mathcal{E}}^{\dagger} & \longrightarrow & \mathcal{V}(\Lambda^{\flat})
\end{array}$$

where the two vertical maps are closed immersions.

Remark 5.8. The reduced scheme $\mathcal{N}_n^{\mathrm{red}}$ is a certain affine Deligne–Lusztig variety (in mixed characteristic). It is not clear (to the author) whether the reduced scheme $\mathcal{Z}(\mathcal{E})^{\dagger,\mathrm{red}}$ admits a similar description, and whether one can use such description to study the map (5.20).

Example 5.9 (The case $r = 1$). To illustrate the map (5.20), we consider the rank one case $\mathcal{E} = \langle u \rangle$ for an isotropic vector u. Given $\Lambda \in \mathrm{Vert}_{\mathcal{E}}$, there are two cases

- $u \in \varpi \Lambda^{\vee}$,

- $u \notin \varpi \Lambda^{\vee}$.

In the former case, $(u, \Lambda) = \varpi O_F$, we have $t(\Lambda^{\flat}) = t(\Lambda) - 2$. Then

$$\mathcal{V}(\Lambda)_{\langle u \rangle}^{\dagger} = \mathcal{V}(\Lambda) \setminus \mathcal{V}(\Lambda + \langle u/\varpi \rangle).$$

The fiber of the map (5.20) seems rather complicated (for a general $t(\Lambda)$).

In the latter case, $(u, \Lambda) = O_F$, we have $t(\Lambda^{\flat}) = t(\Lambda)$. The map (5.20) is an isomorphism. (Note that, if Λ has maximal type n or $n-1$, then the latter case cannot happen.)

Example 5.10 (Disconnectedness of $\mathcal{Z}(u)^{\dagger} \subset \mathcal{N}_3$). Consider $n = 3$ and a rank one lattice $\mathcal{E} = \langle u \rangle$ spanned by a norm zero vector u. Then

$$\mathrm{Vert}_{\langle u \rangle}^{\dagger} = \{ \Lambda \in \mathrm{Vert}(\mathbb{V}) \mid u \in \Lambda \setminus \varpi \Lambda \}.$$

Let $\mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3} = \{\Lambda \in \mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3} \mid t(\Lambda) = 3\}$. Fixing a base point Λ_0 in this set, there is a natural bijection

$$\mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3} \xrightarrow{\ \sim\ } N(F_0)/N(O_{F_0})$$

where N is the unipotent radical of the parabolic subgroup of $U(\mathbb{V})$ corresponding to the flag $0 \subset \langle u \rangle_F \subset \langle u \rangle_F^\perp \subset \mathbb{V}$, and $N(O_{F_0}) := N(F_0) \cap \mathrm{GL}_{O_{F_0}}(\Lambda_0)$ is defined with respect to Λ_0.

Claim: $\mathcal{Z}(u)^{\dagger,\mathrm{red}}$ is disconnected of pure dimension one.

It follows that

$$(5.21) \qquad \mathcal{Z}(u)^{\dagger,\mathrm{red}} = \coprod_{\Lambda \in \mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3}} \mathcal{V}(\Lambda)_{\langle u \rangle}^\dagger.$$

To show the claim, first we note that there can not be zero dimensional connected components. In fact, for any type 1 lattice $\Lambda \in \mathrm{Vert}_{\langle u \rangle}^\dagger$, one can find a type 3 lattice Λ' such that $u \in \Lambda' \subset \Lambda$: if $u \notin \varpi \Lambda^\vee$, then we can take $\Lambda' = \langle u \rangle + \varpi \Lambda^\vee$; if $u \in \varpi \Lambda^\vee$ then we can take any one of the $q + 1$ lattices $\Lambda' \subset \Lambda$ of type 3.

It remains to show that the (punctured) DL curves $\mathcal{V}(\Lambda)^\dagger$ for $\Lambda \in \mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3}$ do not intersect. Suppose that $\Lambda_1, \Lambda_2 \in \mathrm{Vert}_{\langle u \rangle}^{\dagger, t=3}$ are adjacent, i.e., $\widetilde{\Lambda} = \Lambda_1 + \Lambda_2$ is a vertex of type 1. We need to show that $\frac{1}{\varpi} u \in \widetilde{\Lambda}$, which implies that the intersection point $\mathcal{V}(\Lambda_1) \cap \mathcal{V}(\Lambda_2) = \mathcal{V}(\widetilde{\Lambda})$ is "removed" in $\mathcal{Z}(u)^{\dagger,\mathrm{red}}$.

By $\Lambda_i^\vee = \varpi^{-1} \Lambda_i$ we have $\frac{1}{\varpi} u \in \Lambda_i^\vee$ and its image is nonzero in the quotient $\Lambda_i^\vee / \Lambda_i = \frac{1}{\varpi} \Lambda_i / \Lambda_i$. By

$$\Lambda_i \overset{1}{\hookrightarrow} \widetilde{\Lambda} \overset{1}{\hookrightarrow} \widetilde{\Lambda}^\vee \overset{1}{\hookrightarrow} \Lambda_i^\vee = \varpi^{-1} \Lambda_i.$$

and $\Lambda_1^\vee \cap \Lambda_2^\vee = \widetilde{\Lambda}^\vee$, we see that $\frac{1}{\varpi} u \in \widetilde{\Lambda}^\vee$. However, $(u, u) = 0$ hence $\frac{1}{\varpi} u$ mod Λ_1 is isotropic in the 2-dimensional k-space $\widetilde{\Lambda}^\vee / \Lambda_1$ inside $\Lambda_1^\vee / \Lambda_1$. Note that the 1-dimensional line $\Lambda^\vee / \Lambda_1$ inside $\widetilde{\Lambda}^\vee / \Lambda_1$ is characterized by $(x, x) = 0$. Hence $\frac{1}{\varpi} u \in \widetilde{\Lambda}$ as desired.

Remark 5.11. Similar to (5.21) there is a description of the set of connected components $\mathcal{Z}(u)^\dagger$ when $n = 4$. Is there a natural parameterization of the set of connected components of $\mathcal{Z}(u)^\dagger$ for an isotropic vector u (more generally, $\mathcal{Z}(\mathcal{E})^\dagger$ for a totally isotropic $\mathcal{E}$) when $n \geq 5$?

6. The FL conjecture of Liu

In this section, we recall the fundamental lemma (FL) conjecture of Liu in the Bessel case [13]. This also sets up the analytic side that will be used to formulate our AFL conjecture.

6.1. The Bessel subgroup

We recall [4, §12] for the Bessel subgroups (in the unitary case). Let F/F_0 be a separable quadratic extension of fields. Let V be an F/F_0-Hermitian space of dimension $n \geq 2$, with the sesquilinear pairing denoted by $(\cdot, \cdot)$. Following *loc. cit.* we will denote by $G(V)$ the unitary group $\mathrm{U}(V)$. Let E be a totally isotropic subspace of dimension r so that we have a flag $0 \subset E \subset E^\perp \subset V$. (Here $E^\perp$ denotes the orthogonal complement of E in V.) Let $W^\sharp = E^\perp/E$ be the quotient with the induced Hermitian structure:

$$(6.1) \qquad 0 \longrightarrow E \longrightarrow E^\perp \longrightarrow W^\sharp \longrightarrow 0.$$

Fix a non-isotropic line $L \subset W^\sharp$. Pulling-back the extension (6.1) along $L \hookrightarrow W^\sharp$ we get an extension, denoted by $E^\sharp$,

$$(6.2) \qquad 0 \longrightarrow E \longrightarrow E^\sharp \longrightarrow L \longrightarrow 0.$$

Let $W \subset W^\sharp$ be the orthogonal complement of L in $W^\sharp$. Then we have

$$W^\sharp = W \oplus L.$$

Now fix a complete flag of (necessarily isotropic) subspaces in E:

$$0 = E_0 \subset E_1 \subset \cdots \subset E_r = E.$$

It induces a (partial) flag of V

$$0 = V_0 \subset V_1 \subset \cdots \subset V_{2r+1} = V.$$

by setting

$$(6.3) \qquad V_i = \begin{cases} E_i, & 1 \leq i \leq r, \\ E_{2r+1-i}^\perp, & r+1 \leq i \leq 2r+1. \end{cases}$$

Then we define a subgroup H of $G(V)$ consisting of $g \in G(V)$ such that

$$\begin{cases} gE_i \subset E_i, g|_{E_i/E_{i-1}} = 1, & i = 1, \cdots, r, \\ gE^\sharp \subset E^\sharp, g|_{E^\sharp/E} = 1. \end{cases}$$

We call H a *Bessel subgroup* of $G(V)$ (note that we are abusing the terminology here as the name is more commonly used for H inside a product group, see the paragraph after (6.9)). Note that $g \in H$ also stabilizes $E_i^\perp$. By the isomorphism $E^\perp/E^\sharp \simeq W$, we have a quotient map

(6.4) $$\mathfrak{q} \colon H \longrightarrow\!\!\!\!\!\rightarrow G(W)$$

defined by $g \mapsto g|_{E^\perp/E^\sharp}$. (Note that all $g \in H$ automatically stabilizes $E^\perp$.) The unipotent radical of H, denoted by N, is then the kernel of the quotient map $H \to G(W)$.

Remark 6.1. Let $P = P_{1^r, m+1, 1^r}$ be the parabolic subgroup of $G(V)$ which stabilizes a complete flag of (necessarily isotropic) subspaces in E. The Levi subgroup M_P of P is naturally isomorphic to $\mathrm{Res}_{F/F_0}(\mathrm{GL}_1)^r \times G(W^\sharp)$. Then the group H fits into a Cartesian diagram

(6.5)
$$\begin{array}{ccc} H & \longrightarrow & P \\ \downarrow & & \downarrow \\ G(W) & \longrightarrow & \mathrm{Res}_{F/F_0}(\mathrm{GL}_1)^r \times G(W^\sharp). \end{array}$$

This provides a quick definition of H. However, it is then less obvious how to define the homomorphism (6.6) below.

Remark 6.2. In [4] the authors start with an embedding $W \subset V$ and an orthogonal decomposition $V = W \oplus W^\perp$ and $W^\perp = (E + E^\vee) \oplus L$. Then we have an isomorphism

$$H \simeq N \rtimes G(W),$$

where N is the unipotent radical of P in (6.5), and $G(W)$ (now as a subgroup of $G(V)$) acts on N by conjugation. It seems more natural to treat W as a subquotient rather than a subspace of V, as we will see in the moduli space of p-divisible group.

Let $N_{E^\sharp}$ denote the unipotent radical of the Borel subgroup $B_{E^\sharp}$ of the Weil restriction $\mathrm{Res}_{F/F_0}\,\mathrm{GL}(E^\sharp)$ stabilizing the complete flag $0 = E_0 \subset E_1 \subset \cdots \subset E_r = E \subset E_{r+1} = E^\sharp$. Then we have a quotient map

$$(6.6) \qquad\qquad\qquad u\colon H \longrightarrow\!\!\!\!\!\rightarrow N_{E^\sharp}$$

sending $g \in H$ to $g|_{E^\sharp}$. Finally, we fix a generic homomorphism (of algebraic groups)

$$(6.7) \qquad\qquad\qquad \lambda\colon N_{E^\sharp} \longrightarrow F.$$

Here by a "generic" homomorphism we mean that the stabilizer of the homomorphism under the natural action of the Levi torus $\mathrm{Res}_{F/F_0}(\mathrm{GL}_1)^{r+1}$ of $B_{E^\sharp}$ is as small as possible (in this case, the stabilizer is exactly the center of $\mathrm{Res}_{F/F_0}\,\mathrm{GL}(E^\sharp)$).

We now define an explicit homomorphism λ by choosing a basis $\{e_1, \cdots, e_r\}$ of E such that

$$E_i = \langle e_1, \cdots, e_i\rangle_F, \quad 1 \le i \le r,$$

and a basis $\{e\}$ of L. We let $e_{r+1} \in E^\perp$ be a lifting of e. For $u \in N_{E^\sharp}$, we write

$$u(e_{i+1}) \equiv e_{i+1} + a_{i+1,i}\, e_i \mod E_{i-1}, \quad a_{i+1,i} \in F_0, 1 \le i \le r.$$

Then we define λ by

$$(6.8) \qquad\qquad\qquad \lambda(u) = \sum_{i=1}^r a_{i+1,i}.$$

For later use we also fix a "dual basis" $\{e_1^\vee, \cdots, e_r^\vee\}$ (i.e. $\langle e_i, e_j^\vee\rangle = \delta_{ij}$) such that $e_{r+1} \perp e_i^\vee$ for all $i = 1, \cdots, r$. Denote $E^\vee = \langle e_1^\vee, \cdots, e_r^\vee\rangle_F$. We then lift $W^\sharp$ to the unique subspace $(E + E^\vee)^\perp$ of V and write $V = W^\sharp \oplus (E + E^\vee)$. In terms of the "dual basis" we have

$$(6.9) \qquad\qquad\qquad \lambda(u) = \sum_{i=1}^r (u(e_{i+1}), e_i^\vee).$$

We view H as a subgroup of $G = G(W) \times G(V)$ where the first factor is the natural projection (6.4). The subgroup H is called a *Bessel subgroup* of G (cf. [4, §12]). To rigidify the set up, in this paper let us simply assume that the discriminant of L is trivial. Then the pair $(H, \lambda \circ u)$, up to G-conjugacy,

depends only on the isomorphism class of V (or equivalently, only on the isomorphism class of W), which we assume to contain a totally isotropic subspace of dimension r.

We will be interested in the action of $H \times H$ on $G = G(W) \times G(V)$ by

$$(h_1, h_2) \cdot (h, g) := (h_1 h h_2^{-1}, h_1 g h_2^{-1}).$$

To simplify the action, we introduce the group

(6.10) $$\mathbf{H} := H \times_{G(W)} H = \{(h_1, h_2) \in H \times H \mid \mathfrak{q}(h_1) = \mathfrak{q}(h_2)\}.$$

Then (6.4) induces a surjective homomorphism denoted by the same symbol

(6.11) $$\mathfrak{q} : \mathbf{H} \longrightarrow G(W)$$

whose kernel is $N \times N$. It acts on $G(V)$ by

$$(h_1, h_2) \cdot g := h_1 g h_2^{-1}.$$

See [13, §4.2]. Then it is easy to see that the set of $(H \times H)(F_0)$-orbits in $G(F_0)$ is naturally bijective to the set of $\mathbf{H}(F_0)$-orbits in $G(V)(F_0)$. Henceforth, we will freely switch between the two orbit spaces.

6.2. Symmetric space

Consider the symmetric space

(6.12) $$S_n := \{\, \gamma \in \mathrm{Res}_{F/F_0} \mathrm{GL}_n \mid \gamma \overline{\gamma} = 1_n \,\}.$$

The group $\mathrm{Res}_{F/F_0} \mathrm{GL}_n$ acts on S_n by

$$g \cdot \gamma = g \gamma \overline{g}^{-1}.$$

As before we assume $n = m + 2r + 1$. Similar to the Bessel subgroup of the unitary group, we recall from [13, §2.1] the Bessel subgroup H' of $\mathrm{GL}_{n,F}$. Let

$$\{e_1, \cdots, e_n\}$$

be the standard basis of the n-dimensional vector space $V = F^n$. Let $E_i = \langle e_1, \cdots, e_i \rangle_F$ and $W = \langle e_{r+2}, \cdots, e_{n-r} \rangle_F \subset W^\sharp = \langle e_{r+1}, \cdots, e_{n-r} \rangle_F$. We

view $\mathrm{GL}(W)$ as a subgroup $\mathrm{GL}(W^\sharp)$ in the obvious way. Then the Bessel subgroup H' of $\mathrm{GL}_{n,F}$ consists of $g \in \mathrm{GL}_{n,F}$ such that

(6.13)
$$\begin{cases} gE_i \subset E_i, g|_{E_i/E_{i-1}} = 1, & i \in \{1, \cdots, r, n-r+1, \cdots, n\}, \\ g|_{E_{m+1+r}/E_r} \in \mathrm{GL}(W), \end{cases}$$

where we view $g|_{E_{m+1+r}/E_r}$ as an element in $\mathrm{GL}(W^\sharp)$ by the natural isomorphism $E_{m+1+r}/E_r \simeq W^\sharp$. The kernel of the projection $H' \to \mathrm{GL}(W) \simeq \mathrm{GL}_{m,F}$ (the isomorphism defined by the above basis) is the unipotent radical denoted by $U_{1^r,m+1,1^r}$. Then

$$H' = U_{1^r,m+1,1^r} \rtimes \mathrm{GL}_{m,F}.$$

We introduce the subgroup of $\mathrm{Res}_{F/F_0} H'$

$$\mathbf{H}' := \mathrm{Res}_{F/F_0} U_{1^r,m+1,1^r} \rtimes \mathrm{GL}_{m,F_0} \subset \mathrm{Res}_{F/F_0} H'.$$

with the induced action on S_n. In other words, $\mathbf{H}'$ is defined by (6.13) and further requiring that $g|_{E_{m+1+r}/E_r} \in \mathrm{GL}_{m,F_0}$. Denote the natural quotient map by

(6.14)
$$\mathfrak{q}' \colon \mathbf{H}' \longrightarrow\!\!\!\!\!\rightarrow \mathrm{GL}_{m,F_0} .$$

Let $\widetilde{E} = \langle e_1, \cdots, e_{r+1}, e_{n-r+1}, \cdots, e_n \rangle_F$, naturally identified with V/W. Let $U_{1^{2r+1}}$ denote the unipotent radical of the Borel subgroup of $\mathrm{GL}(\widetilde{E}) \simeq \mathrm{GL}_{2r+1}$ with respect to the complete flag

$$0 \subset \langle e_1 \rangle_F \subset \cdots \subset \langle e_1, \cdots, e_{r+1}, e_{n-r+1}, \cdots, e_n \rangle_F.$$

We have a surjective homomorphism

(6.15)
$$u' \colon \mathbf{H}' \longrightarrow\!\!\!\!\!\rightarrow U_{1^{2r+1}} ,$$

sending $g \in \mathbf{H}'$ to the composition

$$\widetilde{E} \longrightarrow V \xrightarrow{\ g\ } V \longrightarrow V/W \simeq \widetilde{E} .$$

(It is easy to verify that the composition indeed lies in the subgroup $U_{1^{2r+1}}$ of $\mathrm{GL}(\widetilde{E})$.) Similar to (6.8), we define a generic homomorphism,

(6.16)
$$\lambda' \colon U_{1^{2r+1}} \longrightarrow F$$

using the above basis $\{e_1, \cdots, e_{r+1}, e_{n-r+1}, \cdots, e_n\}$.

6.3. Regular orbit matching

We now consider the space of (regular) orbits for the two group actions introduced above:

(i) The action of $\mathbf{H}$ on $G(V)$.

(ii) The action of $\mathbf{H}'$ on S_n.

Recall from [13, Def. 4.5] that an element $\gamma \in S_n(F_0)$ (resp. $g \in G(V)(F_0)$) is called *pre-regular* if its stabilizer under the action of $\mathrm{Res}_{F/F_0} U_{1^r,m+1,1^r}$ (resp. $N \times N$) is trivial. By [13, Lem. 4.6], a $\mathrm{Res}_{F/F_0} U_{1^r,m+1,1^r}$-orbit of a pre-regular element contains a unique element of the form

(6.17)

$$
\gamma = \begin{bmatrix} & & & & & & t'_1 \\ & & & & & \cdots & \\ & & & & t'_r & & \\ & & & \gamma^\sharp & & & \\ & & (\bar{t}'_r)^{-1} & & & & \\ & \cdots & & & & & \\ (\bar{t}'_1)^{-1} & & & & & & \end{bmatrix}, \quad t'_i \in F^\times,\ \gamma^\sharp \in S_{m+1}(F_0).
$$

See [13, Rem. 4.8] for a more "intrinsic" definition of the invariants t'_i. Such a form is called the *normal form* (of any element in the orbit).

For the unitary groups, we take the fixed basis of E and $E^\vee$. By [13, Lem. 4.7], a $N \times N$-orbit of a pre-regular element contains a unique element of the form

(6.18) $$
g = \begin{bmatrix} & & & & & & t_1 \\ & & & & & \cdots & \\ & & & & t_r & & \\ & & & g^\sharp & & & \\ & & \bar{t}_r^{-1} & & & & \\ & \cdots & & & & & \\ \bar{t}_1^{-1} & & & & & & \end{bmatrix}, \quad t_i \in F^\times,\ g^\sharp \in G(W^\sharp)(F_0),
$$

which is called the *normal form* (of any element in the orbit).

Next we recall that the notion of *regular* (called regular semisimple in [31]) elements in S_{m+1} and $G(W^\sharp)$ (relative to the action of GL_m and $G(W)$ respectively). Using the basis of W and L we identify $\mathrm{End}_F(W \oplus L) \simeq \mathrm{Mat}_{m+1,F}$. Then an element $\xi \in \mathrm{Mat}_{m+1,F}$ is *regular* if $\{e, \xi e, \cdots, \xi^m e\}$ is an F-basis of

$W \oplus L$ and $\{e^\vee, e^\vee \xi, \cdots, e^\vee \xi^m\}$ is an F-basis of $W^\vee \oplus L^\vee$, where $e^\vee$ is a basis of $L^\vee = \mathrm{Hom}_F(L, F)$. Then we say that $g^\sharp \in G(W^\sharp)(F_0)$ (resp. $\gamma^\sharp \in S_{m+1}(F_0)$) is *regular* if it is regular as an element of $\mathrm{Mat}_{m+1, F}$. A regular $g^\sharp \in G(W^\sharp)(F_0)$ and $\gamma^\sharp \in S_{m+1}(F_0)$ are said to match if they are $\mathrm{GL}(W) \simeq \mathrm{GL}_m(F)$-conjugate when they are viewed as elements in $\mathrm{GL}_{m+1}(F)$.

Finally, we say that an element $\gamma \in S_n(F_0)$ (resp. $g \in G(V)$) is *regular* if it is pre-regular and $\gamma^\sharp \in S_{m+1}(F_0)$ (resp. $g^\sharp \in G(W^\sharp)(F_0)$) in its normal form is regular. Denote by $S_n(F_0)_{\mathrm{rs}}$ (resp. $G(V)(F_0)_{\mathrm{rs}}$) the subset of regular elements. Then $\mathbf{H}'(F_0)$ (resp. $\mathbf{H}(F_0)$) preserves $S_n(F_0)_{\mathrm{rs}}$ (resp. $G(V)(F_0)_{\mathrm{rs}}$). We will denote by $\left[S_n(F_0)\right]_{\mathrm{rs}}$ (resp. $\left[\mathrm{U}(V)(F_0)\right]_{\mathrm{rs}}$) the set of regular orbits.

A regular $g \in G(V)(F_0)$ and $\gamma \in S_n(F_0)$ are said to match if their normal forms (6.17) and (6.18) satisfy

1. $t_i = t_i'$ for all $i = 1, \cdots, r$, and

2. $g^\sharp \in G(W^\sharp)(F_0)$ and $\gamma^\sharp \in S_{m+1}(F_0)$ match.

By [13, Prop. 4.12] the matching relation defines a natural bijection of regular orbits,

$$(6.19) \qquad \coprod_V \left[\mathrm{U}(V)(F_0)\right]_{\mathrm{rs}} \xrightarrow{\;\sim\;} \left[S_n(F_0)\right]_{\mathrm{rs}},$$

where the disjoint union runs over the set of isometry classes of n-dimensional F/F_0-Hermitian spaces V containing a totally isotropic subspace of dimension r.

6.4. Orbital integral: smooth transfer

From now on we further assume that $m \geq 1$. We recall orbital integrals [13, §4]. Now let F/F_0 be a quadratic extension of p-adic fields. Then there are exactly two isometry classes of F/F_0-Hermitian spaces of dimension n, denoted by V_0 and V_1. When F/F_0 is unramified, we will assume that V_0 has a self-dual lattice. Then the orbit bijection (6.19) becomes

$$\left[(\mathrm{U}(V_0)(F_0)\right]_{\mathrm{rs}} \coprod \left[(\mathrm{U}(V_1)(F_0)\right]_{\mathrm{rs}} \xrightarrow{\;\sim\;} \left[S_n(F_0)\right]_{\mathrm{rs}}.$$

Fix a non-trivial continuous character $\psi_0 : F_0 \to \mathbb{C}^\times$, and set $\psi = \psi_0 \circ \mathrm{tr} : F \to \mathbb{C}^\times$. Abusing notation we also denote the character

$$(6.20) \qquad \psi \colon N_{E^\sharp}(F) \longrightarrow \mathbb{C}^\times,$$

sending $u \in N_{E^\sharp}(F)$ to $\psi \circ \lambda(u)$,

$$(6.21) \qquad\qquad\qquad \psi \colon H(F_0) \longrightarrow \mathbb{C}^\times$$

sending $h \in H(F_0)$ to $\psi \circ \lambda \circ u(h)$, and

$$(6.22) \qquad\qquad\qquad \psi \colon \mathbf{H}(F_0) \longrightarrow \mathbb{C}^\times$$

sending $(h_1, h_2) \in \mathbf{H}(F_0)$ to $\psi(h_1^{-1})\psi(h_2)$. Similarly, we denote

$$(6.23) \qquad\qquad\qquad \psi \colon \mathbf{H}'(F_0) \longrightarrow \mathbb{C}^\times$$

sending $\in \mathbf{H}'(F_0) \mapsto \psi \circ \lambda \circ u(h)$.

Let
$$\eta = \eta_{F/F_0} \colon F_0^\times \longrightarrow \{\pm 1\}$$

be the quadratic character associated to F/F_0 by local class field theory. Abusing notation we also denote the character

$$(6.24) \qquad\qquad\qquad \eta \colon \mathbf{H}'(F_0) \longrightarrow \{\pm 1\}$$

sending $h \in \mathbf{H}'(F_0)$ to $\eta \circ \det \circ \mathbf{q}'(h)$. Similarly, we let

$$(6.25) \qquad\qquad\qquad |\cdot| \colon \mathbf{H}'(F_0) \longrightarrow \mathbb{R}^\times$$

denote the character $h \mapsto |\det \circ \mathbf{q}'(h)|_{F_0}$.

For $V = V_0$ or V_1, $g \in G(V)(F_0)_{\mathrm{rs}}$ and $f \in C_c^\infty(G(V)(F_0))$, we define the orbital integral

$$\mathrm{Orb}(g, f) := \int_{\mathbf{H}(F_0)} f(h \cdot g)\psi(h)\, dh.$$

For $\gamma \in S_n(F_0)_{\mathrm{rs}}$, $f' \in C_c^\infty(S_n(F_0))$, and $s \in \mathbb{C}$, we define the orbital integral

$$(6.26) \qquad \mathrm{Orb}(\gamma, f', s) := \int_{\mathbf{H}'(F_0)} f'(h \cdot \gamma)\psi(h)\eta(h)|h|^s\, dh,$$

where $|\ |$ denotes the normalized absolute value on F_0. We define the special values

$$(6.27) \qquad \begin{aligned} \mathrm{Orb}(\gamma, f') &:= \omega(\gamma)\, \mathrm{Orb}(\gamma, f', 0) \quad \text{and} \\ \partial\mathrm{Orb}(\gamma, f') &:= \omega(\gamma)\, \frac{d}{ds}\Big|_{s=0} \mathrm{Orb}(\gamma, f', s), \end{aligned}$$

where $\omega(\gamma)$ is a transfer factor (defined in [13, §4.4, (4.17)] at least when F/F_0 is unramified and γ is a normal form).

Definition 6.3. A function $f' \in C_c^\infty(S_n(F_0))$ and a pair of functions $(f_0, f_1) \in C_c^\infty(\mathrm{U}(V_0)(F_0)) \times C_c^\infty(\mathrm{U}(V_1)(F_0))$ are (smooth) *transfers* of each other if for each $i \in \{0, 1\}$ and each $g \in \mathrm{U}(V_i)(F_0)_{\mathrm{rs}}$,

$$\mathrm{Orb}(g, f_i) = \mathrm{Orb}(\gamma, f')$$

where $\gamma \in S(F_0)_{\mathrm{rs}}$ matches g.

It is expected that for any given $(f_0, f_1) \in C_c^\infty(\mathrm{U}(V_0)(F_0)) \times C_c^\infty(\mathrm{U}(V_1)(F_0))$ or $f' \in C_c^\infty(S_n(F_0))$, a smooth transfer always exists, though this remains an open problem as of today.

Remark 6.4. We have assumed $m \geq 1$. The case $m = 0$ (so that $n = 2r+1$) requires little modification: in (6.19) there is only one Hermitian space V, i.e., the split one.

6.5. Liu's FL conjecture

We review the FL conjecture, cf. [13, §4.4]. From now on we assume that F/F_0 is an unramified quadratic extension of p-adic field for an *odd* prime p. We assume that the character $\psi_0 : F_0 \to \mathbb{C}^\times$ is of level zero (i.e., $\psi|_{O_{F_0}} = 1$ but $\psi|_{\varpi^{-1}O_{F_0}} \neq 1$). Fix a self-dual O_F-lattice $\Lambda_0 \subset V_0$ such that the intersection of the lattice $(\Lambda \cap E^\perp)/(\Lambda \cap E)$ in $W^\sharp$ with W remains a self-dual lattice in W. We denote its stabilizer by $K_0 \subset G(V_0)(F_0)$, a hyperspecial maximal compact open subgroup. Let $H(O_{F_0}) = K_0 \cap H(F_0)$ (note that H is a subgroup of $G(V_0)$) and $\mathbf{H}(O_{F_0}) = (H(O_{F_0}) \times H(O_{F_0})) \cap \mathbf{H}(F_0)$ (note that, by definition (6.10), $\mathbf{H}$ is a subgroup $H \times H$). Let $\mathbf{H}'(O_{F_0}) = \mathrm{GL}_n(O_F) \cap \mathbf{H}'(F_0)$ (note that $\mathbf{H}'$ is a subgroup of $\mathrm{Res}_{F/F_0} \mathrm{GL}_n$). We normalize the Haar measure on $\mathbf{H}(F_0)$ (resp. $\mathbf{H}'(F_0)$) such that $\mathrm{vol}(\mathbf{H}(O_{F_0})) = 1$ (resp. $\mathrm{vol}(\mathbf{H}'(O_{F_0})) = 1$). We have [13, §4.4]:

Conjecture 6.5 (Liu's Fundamental Lemma conjecture). *The characteristic function*

$$\mathbf{1}_{S_n(O_{F_0})} \in C_c^\infty(S_n(F_0))$$

transfers to the pair of functions

$$(\mathbf{1}_{K_0}, 0) \in C_c^\infty(G(V_0)(F_0)) \times C_c^\infty(G(V_1)(F_0)).$$

When $r = 0$ this is the Jacquet–Rallis Fundamental Lemma conjecture. When $r > 0$, the conjecture remains an open problem as of today.

Remark 6.6. Liu also allows the case $m = 0$ in [13, §4.4], which was already formulated by Jacquet and proved in some low dimensional cases in [8]. For the formulation of the AFL conjecture, we will always assume $m \geq 1$.

The easy part of the conjecture is known by [13, Prop. 4.16], i.e., when γ matches an element $g \in G(V_1)(F_0)$:

$$\mathrm{Orb}(\gamma, f') = 0.$$

In the next section, we will formulate a conjecture relating the first derivative $\partial \mathrm{Orb}(\gamma, f')$ to certain arithmetic intersection numbers.

6.6. FL: an interpretation as lattices counting

We continue with the set up from §6.5. Suppose now that V is the split Hermitian space and that the special vector $e \in W^\sharp$ has norm $(e, e) = 1$ (or a unit). In this subsection, we let $\mathcal{N}'_n$ denote the set $\mathrm{Vert}^0(V)$ of self-dual lattices $\Lambda \subset V$. As the notation suggests, $\mathcal{N}'_n$ is the set of connected components of the RZ space in the totally definite case, cf. Remark 5.2.

Recall we have fixed a basis $\{e_1, \cdots, e_r\}$ of E, a basis $\{e\}$ of L, and $e_{r+1} \in E^\sharp$ a lifting $e \in L$. We consider the O_F-lattices

$$\mathcal{E}_i^0 = \langle e_1, \cdots, e_i \rangle \subset E_i, \quad 1 \leq i \leq r + 1.$$

Denote

(6.28) $$\mathcal{E}^{0\flat} = \mathcal{E}_r^0 \subset E, \quad \mathcal{E}^0 = \mathcal{E}_{r+1}^0 \subset E^\sharp.$$

Let $\mathcal{N}'_P$ be the set of $0 = \Lambda_0 \subset \Lambda_1 \subset \cdots \subset \Lambda_{2r+1}$ such that $\Lambda_{2r+1} \in \mathcal{N}'_n$ and $\Lambda_i = V_i \cap \Lambda_{2r+1}$ for all $i = 1, \cdots, 2r + 1$. (Here recall V_i from (6.3).) Then the map $\Lambda_\bullet \mapsto \Lambda_{2r+1}$ defines a bijection:

$$\mathcal{N}'_P \overset{\sim}{\longrightarrow} \mathcal{N}'_n.$$

Therefore, we will freely switch between the two sets. Let $\mathcal{N}'_M$ denotes the set of tuples $(\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_m, \Lambda^\sharp)$ where $\mathcal{L}_i$ are rank one lattices in $V_i/V_{i-1} = E_i/E_{i-1}$ for $i = 1, \cdots, r$ and $\Lambda^\sharp$ is a self-dual lattice in $W^\sharp$. There is a natural map

$$\mathcal{N}'_P \longrightarrow \mathcal{N}'_M.$$

sending $\Lambda_\bullet$ to $\mathcal{L}_i = \Lambda_i/\Lambda_{i-1}$ for $i = 1, \cdots, r$ and $\Lambda^\sharp = \Lambda_{r+1}/\Lambda_r$ (this is self-dual by Lemma 3.1).

Let $\mathcal{N}'_m = \mathrm{Vert}^0(W)$ (resp. $\mathcal{N}'_{m+1} = \mathrm{Vert}^0(W^\sharp)$) denote the set of self-dual lattices in W (resp. in $W^\sharp$). We have an embedding

$$(6.29) \qquad\qquad \iota : \mathcal{N}'_m \longrightarrow \mathcal{N}'_{m+1}$$

with the image consisting of lattices $\Lambda^\sharp \in \mathrm{Vert}^0(W^\sharp)$ such that $e \in \Lambda^\sharp$. There is an embedding

$$(6.30) \qquad\qquad \iota_{\{e_1, \cdots, e_r\}} : \mathcal{N}'_{m+1} \longrightarrow \mathcal{N}'_M$$

with the image consisting of $(\mathcal{L}_1, \cdots, \mathcal{L}_r, \Lambda^\sharp)$ such that $\mathcal{L}_i = \langle e_i \rangle$. For the Bessel subgroup H, we define

$$(6.31) \qquad\qquad \begin{array}{ccc} \mathcal{N}'_H & \longrightarrow & \mathcal{N}'_P \\ \downarrow & & \downarrow \\ \mathcal{N}'_m & \longrightarrow & \mathcal{N}'_M \end{array}$$

where the bottom map is the composition of the two embeddings (6.29) and (6.30).

Unfolding the definition we see that $\mathcal{N}'_H$ is the set of lattices $\Lambda \in \mathcal{N}'_n = \mathrm{Vert}^0(V)$ such that, for $1 \leq i \leq r+1$, setting $\mathcal{E}_i := E_i \cap \Lambda$, then we have $\mathcal{E}_i/\mathcal{E}_{i-1} = \langle e_i \rangle$ (as lattices in E_i/E_{i-1}). (Here we warn the reader that, in (6.3), $V_i = E_i$ for $i \leq r$ but $V_{r+1} \neq E_{r+1} = E^\sharp$. This is the reason that here we write $\mathcal{E}_i$ rather than Λ_i.) We choose the first definition, which is more laborious, in order to be consistent with the arithmetic analog in the next section.

Let $\mathcal{N}'_{m,n} = \mathcal{N}'_m \times \mathcal{N}'_n$ with the natural action by the product group $(G(W) \times G(V))(F_0)$. We then have a map

$$\mathcal{N}'_H \longrightarrow \mathcal{N}'_{m,n}$$

sending $\Lambda \in \mathcal{N}'_H \mapsto (\Lambda^\flat, \Lambda)$. (Here via the embedding $\mathcal{N}'_H \hookrightarrow \mathcal{N}'_n$ we write an element in $\mathcal{N}'_H$ as $\Lambda \in \mathcal{N}'_n$.)

Let $\mathcal{FL}^\square_{E^\sharp}$ denote the set of (framed) complete flags $\mathcal{E}_\bullet$ of lattices in $E^\sharp$:

$$0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 \subset \cdots \subset \mathcal{E}_{r+1} = \mathcal{E}$$

such that $\mathcal{E}_i = \mathcal{E} \cap E_i$ is of rank i and $\mathcal{E}_i / \mathcal{E}_{i-1} = \langle e_i \rangle, 1 \leq i \leq r+1$. (The word "framed" refers to $\mathcal{E}_i / \mathcal{E}_{i-1} = \langle e_i \rangle$ for the fixed e_i.) Then there is a natural bijection

$$(6.32) \qquad \mathcal{FL}_{E^\sharp}^\square \simeq N_{E^\sharp}(F) / N_{E^\sharp}(O_F),$$

where the neutral point in the right hand side corresponds to the (framed) flag $\mathcal{E}_\bullet^0$ that we have fixed via (6.28). (Here $N_{E^\sharp}$ is defined in §6.1.) The character of $N_{E^\sharp}(F)$ defined by (6.20) induces a character

$$\psi : \mathcal{FL}_{E^\sharp}^\square \longrightarrow \mathbb{C}^\times.$$

There is a natural map

$$\mathcal{N}_H' \longrightarrow \mathcal{FL}_{E^\sharp}^\square$$

sending Λ to $\mathcal{E}_i = \Lambda \cap E_i, 1 \leq i \leq r$ and $\mathcal{E} = \Lambda \cap E^\sharp$. We also denote by

$$\psi : \mathcal{N}_H' \longrightarrow \mathbb{C}^\times$$

the pullback of the character of $\mathcal{FL}_{E^\sharp}^\square$.

Then we may interpret the orbital integral in §6.4 as a weighted lattice counting: for a regular $g \in G(V)$,

$$(6.33) \qquad \mathrm{Orb}(g, 1_{G(O_{F_0})}) = \sum_{\substack{(\Lambda, \Lambda') \in \mathcal{N}_H' \times \mathcal{N}_H' \\ \Lambda' = g\Lambda, \, \Lambda^\flat = \Lambda'^\flat}} \psi(\Lambda) \overline{\psi}(\Lambda').$$

As an analog of the intersection problem to be introduced in the next section, we show that the sum is finite.

Proposition 6.7. *Suppose that $g \in G(V)(F_0)$ is regular. Then the set*

$$(6.34) \qquad \mathcal{N}_H' \cap (1, g) \cdot \mathcal{N}_H' := \left\{ (\Lambda, \Lambda') \in \mathcal{N}_H' \times \mathcal{N}_H' \mid \Lambda' = g\Lambda, \, \Lambda^\flat = \Lambda'^\flat \right\}$$

is finite.

Proof. For $\Lambda \in \mathcal{N}_H'$ we denote

$$\mathcal{E} = \Lambda \cap E^\sharp, \quad \mathcal{E}^\flat = \Lambda \cap E$$

and denote by $\mathcal{E}'$ and $\mathcal{E}'^\flat$ the analogous lattices for $\Lambda' \in \mathcal{N}_H'$. Note that, by $\Lambda \in \mathcal{N}_H'$, $\mathcal{E}$ (resp. $\mathcal{E}^\flat$) induces a framed flag in $E^\sharp$ (resp. in E).

We may and do assume that g is a normal form (6.18). We may factorize the embedding $\mathcal{N}'_H \to \mathcal{N}_P$ into two steps by the two embeddings (6.29) and (6.30). Thus we define the fiber product

$$(6.35) \qquad\qquad \begin{array}{ccc} \mathcal{N}'_E & \longrightarrow & \mathcal{N}'_P \\ \downarrow & & \downarrow \\ \mathcal{N}'_{m+1} & \longrightarrow & \mathcal{N}'_M. \end{array}$$

We first consider the following set

$$(\mathcal{N}'_E \times \mathcal{N}'_E)_g := \left\{ (\Lambda, \Lambda') \in \mathcal{N}'_E \times \mathcal{N}'_E \mid \Lambda' = g\Lambda \right\}.$$

By definition we have $\mathcal{E}^\flat \subset \Lambda$ and $\mathcal{E}'^\flat \subset \Lambda'$. We thus obtain

$$\mathcal{E}'^\flat + g\mathcal{E}^\flat \subset \Lambda'.$$

In particular the lattice $\mathcal{E}'^\flat + g\mathcal{E}^\flat$ is integral, or equivalently $(\mathcal{E}'^\flat, g\mathcal{E}^\flat) \subset O_F$. Using the normal form (6.18) of g, it is easy to see that the integrality condition implies that both $\mathcal{E}^\flat$ and $\mathcal{E}'^\flat$ are contained in $\varpi^{-N} \mathcal{E}^{0\flat}$ for some large integer N (depending only on $t_1, \cdots, t_r$ in (6.18)). Here $\mathcal{E}^{0\flat}$ is the fixed lattice in (6.28). It follows that both $\mathcal{E}^\flat$ and $\mathcal{E}'^\flat$ contain $\varpi^N \mathcal{E}^{0\flat}$. Therefore, there are only finitely many possibilities for the lattices $\mathcal{E}^\flat, \mathcal{E}'^\flat$; moreover we have

$$(6.36) \qquad\qquad \varpi^N(\mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat}) \subset \Lambda'.$$

Now we return to prove the finiteness of $\mathcal{N}'_H \cap (1, g) \cdot \mathcal{N}'_H$. Now we have $e \in \Lambda^\sharp = \Lambda'^\sharp$. With the above result on $(\mathcal{N}'_E \times \mathcal{N}'_E)_g$, a similar integrality consideration shows that Λ contains a lifting of e that differs from e_{r+1} (the fixed lifting of e, which is orthogonal to $E + E^\vee$) by an element in $\varpi^{-M} \mathcal{E}^{0\flat}$ for some large integer M (depending only on g). It follows that $\varpi^M e_{r+1} \in \Lambda$ and the same holds for Λ'. Note that in the normal form, for any $i \geq 1$, the element $g^i e_{r+1}$ is a lifting of $(g^\sharp)^i e$, and it is orthogonal to $E + E^\vee$. Repeating the argument we may show by induction that

$$(6.37) \qquad\qquad \varpi^M g^i e_{r+1} \in \Lambda', \quad \forall i \geq 1.$$

By (6.36) and (6.37) we conclude that Λ' contains the lattice

$$(6.38) \qquad \Lambda(g) = \varpi^M \langle e_{r+1}, g e_{r+1}, \cdots g^m e_{r+1} \rangle + \varpi^N(\mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat}).$$

By the regularity of g, this last lattice $\Lambda(g)$ has full rank in V. By the self-duality of Λ', we conclude the desired finiteness of $\mathcal{N}'_H \cap (1, g) \cdot \mathcal{N}'_H$. □

In fact, the proof also reduces the "trivial" case of Liu's conjecture to the Jacquet–Rallis FL conjecture.

Proposition 6.8. *Let $g \in G(V)_{\mathrm{rs}}$ be a normal form (6.18). Assume that $\mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat}$ is a self-dual lattice in $E + gE$ (equivalently, in terms of (6.18), all t_i are units). Then Conjecture 6.5 for such g in the (m, n)-case reduces to Conjecture 6.5 for $g^\sharp$ in the $(m, m+1)$-case.*

Proof. Tracking the proof of Prop. 6.7, the assumption on g implies that the inclusion (6.36) becomes

$$(6.39) \qquad \mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat} \subset \Lambda'.$$

(Following the suggestion of the referee, here is a direct argument to show that, for $\Lambda, \Lambda' \in \mathcal{N}'_H$ such that $g\Lambda = \Lambda'$, the assumption on g (i.e., g is normal with all t_i units) implies that $\Lambda \cap E_{r+1} = \Lambda' \cap E_{r+1} = \mathcal{E}^0_{r+1}$. Let $1 \le i(\Lambda) \le r+1$ be the largest integer such that $\Lambda \cap E_i = \mathcal{E}^0_i$. Possibly after replacing g by its inverse, we may assume that $i(\Lambda) \le i(\Lambda')$. If $i := i(\Lambda) < r+1$, then $\Lambda \cap E_{i+1}$ contains an element of the form $e_{i+1} + a$ with $a \in E_i \setminus \mathcal{E}^0_i$. It follows that $g(e_{i+1} + a)$ cannot have integral pairing with $\mathcal{E}^0_i$. However, by $i(\Lambda) \le i(\Lambda')$, we have $\mathcal{E}^0_i \subset \Lambda'$ and hence $g(e_{i+1} + a) \in g\Lambda = \Lambda'$ cannot have integral pairing with Λ'. This contradicts to the integrality of the Hermitian form on Λ'.)

It follows from (6.39) that

$$\Lambda' = \Lambda'^\sharp \oplus (\mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat}), \quad \Lambda = \Lambda^\sharp \oplus (g^{-1}\mathcal{E}^{0\flat} + \mathcal{E}^{0\flat}),$$

where we have $\Lambda^\sharp, \Lambda'^\sharp \in \iota(\mathcal{N}'_m)$ (and ι is the embedding (6.29)). Hence the orbital integral (6.33) becomes

$$\mathrm{Orb}(g, 1_{G(O_{F_0})}) = \sum_{\substack{(\Lambda^\sharp, \Lambda'^\sharp) \in \iota(\mathcal{N}'_m) \times \iota(\mathcal{N}'_m) \\ \Lambda'^\sharp = g^\sharp \Lambda^\sharp, \, \Lambda^\flat = \Lambda'^\flat}} 1.$$

Similar argument reduces the orbital integral on S_n to that on S_{m+1}; we omit the details. □

Remark 6.9. The the Jacquet–Rallis FL conjecture (i.e., $(m, m+1)$-case of Conjecture 6.5) is proved by Yun–Gordan (for large p) [28, 6], and independently by Beuzart-Plessis (for $p > 2$) [1].

7. Intersection numbers of Bessel cycles

Let F/F_0 be an unramified quadratic extension of p-adic fields for an *odd* prime p. Our formulation of the AFL conjecture in the Bessel case will be modeled on the lattice counting interpretation of FL in §6.6.

7.1. The generalized diagonal cycle (or the Bessel cycle)

Let $m \in \mathbb{Z}_{\geq 1}, r \in \mathbb{Z}_{\geq 0}$ and $n = m + 1 + 2r$. Recall from §5.1 the unitary RZ space $\mathcal{N}_n$, and from §5.3 the filtered RZ space $\mathcal{N}_P$, where P stands for a parabolic subgroup $P_{1^r,m+1,1^r}$ of $G(V)$. We have $E_i = \mathrm{Hom}_{O_F}^\circ(\mathbb{E}, \mathbb{X}_i) \subset \mathbb{V}, i = 1, \cdots, r$, which are isotropic subspaces of $\mathbb{V} = \mathbb{V}_n$. We choose basis $\{e_1, \cdots, e_r\}$ of $E = E_r$ such that

$$E_i = \langle e_1, \cdots, e_i \rangle_F, \quad i = 1, \cdots, r.$$

Denote by $\mathbb{W}^\sharp = E^\perp/E$ and fix an isomorphism of $\mathbb{W}^\sharp$ with $\mathbb{V}_{m+1}$ defined by (5.2) with n replaced by $m + 1$. Fix a vector $e \in \mathbb{W}^\sharp$ such that (e, e) is a unit and let $\mathbb{W}^\sharp = \mathbb{W} \oplus \langle e \rangle_F$ be the orthogonal decomposition.

We recall the morphism (5.8)

$$(7.1) \qquad \mathcal{N}_P \longrightarrow \mathcal{N}_{M_P} := (\mathcal{M}_1)^r \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{N}_{m+1}.$$

Here $(\mathcal{M}_1)^r = \mathcal{M}_1 \times_{\mathrm{Spf}\, O_{\breve{F}}} \cdots \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{M}_1$ (with r copies of $\mathcal{M}_1$). Let $\mathcal{Z}_{m+1}(e)$ be the KR divisor associated to $e \in \mathbb{W}^\sharp$; then $\mathcal{Z}_{m+1}(e) \simeq \mathcal{N}_m$ and hence we have an induced embedding

$$\iota \colon \mathcal{N}_m \longrightarrow \mathcal{N}_{m+1}.$$

Similar to the definition of the Bessel subgroup H by (6.5), we define *the Bessel formal scheme $\mathcal{N}_H$* as the fiber product:

$$(7.2) \qquad \begin{array}{ccc} \mathcal{N}_H & \longrightarrow & \mathcal{N}_P \\ \downarrow & & \downarrow \\ \mathrm{Spf}\, O_{\breve{F}} \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{N}_m & \longrightarrow & (\mathcal{M}_1)^r \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{N}_{m+1} \end{array}$$

where on the bottom row the first factor $\mathrm{Spf}\, O_{\breve{F}} \to (\mathcal{M}_1)^r$ is the neutral component (with the chosen basis), cf. (5.9), and the second factor is ι above. Let

$$\mathcal{N}_{m,n} := \mathcal{N}_m \times_{\mathrm{Spf}\, O_{\breve{F}}} \mathcal{N}_n.$$

We have a morphism induced by (7.2) and $\mathcal{N}_P \to \mathcal{N}_n$,

$$\mathcal{N}_H \longrightarrow \mathcal{N}_{m,n}$$

whose image is called the Bessel cycle.

Remark 7.1. The Bessel cycle $\mathcal{N}_H$ may be viewed as a correspondence (in a loose sense) over the product $\mathcal{N}_m \times_{\mathrm{Spf}\,O_{\breve{F}}} \mathcal{N}_n$. In view of the arithmetic diagonal cycle in the AGGP conjecture [18], we may call it the *(local) "generalized (arithmetic) diagonal cycle"*.

We need to add a weight factor to the Bessel cycle $\mathcal{N}_H$ and therefore we need to find a partition of $\mathcal{N}_H$ into a union of open-and-closed formal subschemes, similar to (5.6). For this purpose we consider certain auxiliary formal schemes. Let $P' = P_{1^r,1,m,1^r}$ denote the subgroup $P \times_{G(W^\sharp)} G(W)$ of P. (However, note that P' is not a parabolic subgroup of $G(V)$.) Then we consider the fiber product, denoted by $\mathcal{N}_{P'}$,

(7.3)
$$\begin{array}{ccc} \mathcal{N}_{P'} & \longrightarrow & \mathcal{N}_P \\ \downarrow & & \downarrow \\ \mathcal{N}_m & \overset{\iota}{\longrightarrow} & \mathcal{N}_{m+1}. \end{array}$$

In terms of the moduli functor, for $S \in \mathtt{Nilp}$, $\mathcal{N}_{P'}(S)$ is the set of $(X, X_\bullet, \iota, \lambda, \rho)$ such that $X_{r+1}/X_r \in \mathcal{Z}_{m+1}(e)(S)$, i.e., it is a direct product (compatible with the additional structure)

(7.4)
$$X_{r+1}/X_r = \mathscr{E}_S \times X^\flat.$$

We have a pull-back of the bottom row

$$\begin{array}{ccccccccc} 0 & \longrightarrow & X_r & \longrightarrow & X'_{r+1} & \longrightarrow & \mathscr{E}_S & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle =} & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & X_r & \longrightarrow & X_{r+1} & \longrightarrow & X_{r+1}/X_r & \longrightarrow & 0. \end{array}$$

It follows that there is a morphism

(7.5)
$$\mathcal{N}_{P'} \longrightarrow \mathcal{N}_{1^{r+1}}$$

sending a tuple above to $(X'_{r+1}, 0 = X_0 \subset X_1 \subset \cdots \subset X_r \subset X'_{r+1}, \iota|_{X'_{r+1}}, \rho|_{X'_{r+1}})$. Recall that $\mathcal{N}_{1^{r+1}}$ is defined in §5.3 (note that $\mathcal{N}_{1^{r+1}}$ is associated to

a filtered RZ datum of EL type, and there is no polarization involved in its definition). Let $E^\sharp$ denote the preimage of the line $\langle e \rangle_F$ under the quotient map $E^\perp \to \mathbb{W}^\sharp$, and let $e_{r+1} \in E^\sharp$ be a lifting of e. Similar to (5.6), there is a natural bijection

$$(7.6) \qquad \mathcal{N}_{1^{r+1}} \simeq \coprod_{\mathcal{FL}_{E^\sharp}} \operatorname{Spf} O_{\breve{F}},$$

where $\mathcal{FL}_{E^\sharp}$ denotes the set of complete flags $\mathcal{E}_\bullet$ of lattices in $E^\sharp$:

$$0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 \subset \cdots \subset \mathcal{E}_{r+1} = \mathcal{E}$$

such that $\mathcal{E}_i = \mathcal{E} \cap E_i$ is of rank i. Let $\mathcal{N}_{E^\sharp}$ denote the subfunctor of $\mathcal{N}_{1^{r+1}}$ by imposing the additional condition that, for every $1 \leq i \leq r+1$, the quasi-homomorphism $e_i : \mathbb{E} \to \mathbb{X}_i/\mathbb{X}_{i-1}$ lifts (in the sense of §5.2) to a homomorphism $\mathscr{E}_S \to X_{i+1}/X_i$ but $\varpi^{-1} e_i$ does not (the source being X'_{r+1}/X_r when $i = r+1$). Then (7.6) induces a parameterization

$$(7.7) \qquad \mathcal{N}_{E^\sharp} \simeq \coprod_{\mathcal{FL}_{E^\sharp}^\square} \operatorname{Spf} O_{\breve{F}},$$

where $\mathcal{FL}_{E^\sharp}^\square$ is the subset of $\mathcal{FL}_{E^\sharp}$ consisting of flags $\mathcal{E}_\bullet$ satisfying

$$\mathcal{E}_i/\mathcal{E}_{i-1} = \langle e_i \rangle, \quad 1 \leq i \leq r+1.$$

There is a natural bijection

$$(7.8) \qquad \mathcal{FL}_{E^\sharp}^\square \simeq N_{E^\sharp}(F)/N_{E^\sharp}(O_F),$$

where $N_{E^\sharp}$ denotes the unipotent radical of the Borel subgroup (for the complete flag $E_\bullet$) of $\operatorname{GL}_F(E^\sharp) \simeq \operatorname{GL}_{n,F}$, with the basis $\{e_1, \cdots, e_{r+1}\}$ of $E^\sharp$.

Then the Bessel formal scheme $\mathcal{N}_H$ fits into a Cartesian diagram

$$(7.9) \qquad \begin{array}{ccc} \mathcal{N}_H & \longrightarrow & \mathcal{N}_{P'} \\ \downarrow & & \downarrow \\ \mathcal{N}_{E^\sharp} & \longrightarrow & \mathcal{N}_{1^{r+1}}. \end{array}$$

The morphism $\mathcal{N}_H \to \mathcal{N}_{E^\sharp}$ induces a partition as a disjoint union of open-and-closed formal subschemes

$$\mathcal{N}_H = \coprod_{\mathcal{E}_\bullet \in \mathcal{FL}_{E^\sharp}^\square} \mathcal{N}_{H,\mathcal{E}_\bullet}.$$

Proposition 7.2. *There is a natural isomorphism:*

$$\mathcal{N}_{H,\mathcal{E}_\bullet} \overset{\sim}{\longrightarrow} \mathcal{Z}(\mathcal{E})^\dagger.$$

Proof. The proof of Prop. 5.3 applies verbatim, noting that the norm of the special vector u is a unit. (Note that this is not a special case of Prop. 5.3, since $E^\sharp = \mathcal{E} \otimes_{O_F} F$ is not totally isotropic.) $\square$

Let $\mathcal{N}_{H,\mathcal{E}_\bullet}^\spadesuit$ be the (formal) scheme theoretic closure of $\mathcal{N}_{H,\mathcal{E}_\bullet}$ in $\mathcal{N}_{m,n}$, i.e., the smallest closed formal subscheme of $\mathcal{N}_{m,n}$ which the morphism $\mathcal{N}_{H,\mathcal{E}_\bullet} \to \mathcal{N}_{n,m}$ factors through. We also denote by $\mathcal{N}_H^\spadesuit$ the *disjoint union* of $\mathcal{N}_{H,\mathcal{E}_\bullet}^\spadesuit$ over all $\mathcal{E}_\bullet \in \mathcal{FL}_{E^\sharp}^\square$. Then there is a natural morphism $\mathcal{N}_H^\spadesuit \to \mathcal{N}_{m,n}$.

Remark 7.3. The formal scheme $\mathcal{N}_{H,\mathcal{E}_\bullet}^\spadesuit$ seems rather intractable, i.e., (to our knowledge) there is no moduli theoretical characterization.

Define a character

$$\psi : \mathcal{FL}_{E^\sharp}^\square \longrightarrow \mathbb{C}^\times$$

via the bijection (7.8) and the character of $N_{E^\sharp}(F)$ defined by (6.20).

The group $(G(\mathbb{W}) \times G(\mathbb{V}))(F_0)$ acts on $\mathcal{N}_{m,n}$. We are ready to define the intersection number, for a regular $g \in (G(\mathbb{W}) \times G(\mathbb{V}))(F_0)$

$$(7.10) \qquad \mathrm{Int}(g) := \sum_{(\mathcal{E}_\bullet,\mathcal{E}_\bullet') \in \mathcal{FL}_{E^\sharp}^\square \times \mathcal{FL}_{E^\sharp}^\square} \psi(\mathcal{E}_\bullet)\overline{\psi}(\mathcal{E}_\bullet') \, (\mathcal{N}_{H,\mathcal{E}_\bullet}^\spadesuit, g \cdot \mathcal{N}_{H,\mathcal{E}_\bullet'}^\spadesuit)_{\mathcal{N}_{m,n}}.$$

Here, for two closed formal subscheme $\mathcal{Z}_1, \mathcal{Z}_2$ of a regular formal scheme $\mathcal{X}$ over $\mathrm{Spf}\, O_{\breve{F}}$, we define their intersection number as

$$\begin{aligned}
(\mathcal{Z}_1, \mathcal{Z}_2)_{\mathcal{X}} &= \chi(\mathcal{O}_{\mathcal{Z}_1} \otimes_{\mathcal{O}_{\mathcal{X}}}^{\mathbb{L}} \mathcal{O}_{\mathcal{Z}_2}) \\
&= \sum_{i,j} (-1)^{i+j} \mathrm{length}_{O_{\breve{F}}} \mathrm{H}^j(\mathcal{X}, \mathrm{Tor}_i^{\mathcal{O}_{\mathcal{X}}}(\mathcal{O}_{\mathcal{Z}_1}, \mathcal{O}_{\mathcal{Z}_2})),
\end{aligned}$$

if the right hand side is finite (e.g., if $\mathcal{Z}_1 \cap \mathcal{Z}_2$ is a proper scheme over $\mathrm{Spf}\, O_{\breve{F}}$).

Remark 7.4. By Prop. 7.2, we could have simply taken the KR cycle to formulate the intersection problem, avoiding the filtered RZ spaces. However, the filtered RZ spaces seem more natural and amenable for generalization, as the example of Ginzburg–Rallis cycle in §4 shows.

7.2. The AFL conjectures in the Bessel case

Conjecture 7.5 (Arithmetic Fundamental Lemma conjecture for Bessel cycles). *(a) For a regular $g \in (G(\mathbb{W}) \times G(\mathbb{V}))(F_0)$, the disjoint union*

$$\coprod_{(\mathcal{E}_\bullet, \mathcal{E}'_\bullet) \in \mathcal{FL}^\square_{E^\sharp} \times \mathcal{FL}^\square_{E^\sharp}} \mathcal{N}^\spadesuit_{H, \mathcal{E}_\bullet} \cap g \cdot \mathcal{N}^\spadesuit_{H, \mathcal{E}'_\bullet}$$

is a proper scheme over $\operatorname{Spf} O_{\breve{F}}$.

(b) For a regular $g \in (G(\mathbb{W}) \times G(\mathbb{V}))(F_0)$ matching $\gamma \in S_n(F_0)_{\mathrm{rs}}$ where both g and γ are in normal forms, we have

$$\partial \mathrm{Orb}(\gamma, \mathbf{1}_{S_n(O_{F_0})}) = -\mathrm{Int}(g) \log q.$$

We have some indirect evidence towards part (a).

Proposition 7.6. *Let $g \in G(\mathbb{V})(F_0)_{\mathrm{rs}}$. Then the intersection $\mathcal{N}_H \cap (1, g) \cdot \mathcal{N}_H$ is a noetherian scheme.*

Proof. The proof of Prop. 6.7 applies to show that $\mathcal{N}_H \cap (1, g) \cdot \mathcal{N}_H$ is contained in $\mathcal{N}_m \times \mathcal{Z}(\Lambda(g))$ where $\Lambda(g)$ is the lattice defined by (6.38). Note that $\mathcal{Z}(\Lambda(g))$ is a noetherian proper scheme. Since $\mathcal{N}_H$ is a locally closed formal subscheme of $\mathcal{N}_n$ (via the natural morphism $\mathcal{N}_H \to \mathcal{N}_n$), the desired assertion follows. $\square$

In the formulation of the intersection problem in §7.1, we have fixed a basis $\{e_1, \cdots, e_r\}$ of E. Let $\mathcal{E}^{0\flat} \subset E$ be the lattice $\langle e_1, \cdots, e_r \rangle$. Note that for $g \in G(\mathbb{V})_{\mathrm{rs}}$, the sum $E + gE \subset \mathbb{V}$ is a non-degenerate Hermitian space with dimension $2r$.

Proposition 7.7. *Let $g \in G(\mathbb{V})(F_0)$ be a normal form (6.18). Assume that $\mathcal{E}^{0\flat} + g\mathcal{E}^{0\flat}$ is a self-dual lattice in $E + gE$. Then there is an isomorphism of formal schemes*

$$\mathcal{N}_H \times_{\mathcal{N}_{m,n}} (1, g) \cdot \mathcal{N}_H \xrightarrow{\sim} \Delta(\mathcal{N}_m) \times_{\mathcal{N}_{m,m+1}} (1, g^\sharp) \cdot \Delta(\mathcal{N}_m).$$

Here the right hand side is the intersection in the AFL conjecture for $U(m) \times U(m+1)$, i.e.

$$\Delta : \mathcal{N}_m \longrightarrow \mathcal{N}_{m,m+1} = \mathcal{N}_m \times_{\operatorname{Spf} O_{\breve{F}}} \mathcal{N}_{m+1} .$$

In particular, if we further assume that g is regular, then $\mathcal{N}_H \cap (1, g) \cdot \mathcal{N}_H$ is a proper scheme.

Proof. The proof of Prop. 6.8 applies. □

Remark 7.8. One may expect that, if the intersection $\mathcal{N}_H \cap (1, g) \cdot \mathcal{N}_H$ in Prop. 7.6 is already a proper scheme, then $\mathcal{N}_H^{\spadesuit} \times_{\mathcal{N}_{m,n}} (1, g) \cdot \mathcal{N}_H^{\spadesuit}$ is isomorphic to $\mathcal{N}_H \cap (1, g) \cdot \mathcal{N}_H$. But we do not know how to prove this.

There is another special case where we can at least show that the fiber product $\mathcal{N}_H^{\spadesuit} \times_{\mathcal{N}_{m,n}} (1, g) \cdot \mathcal{N}_H^{\spadesuit}$ is a proper scheme. This means that the intersection $\mathcal{N}_{H,\mathcal{E}_\bullet}^{\spadesuit} \cap (1, g) \cdot \mathcal{N}_{H,\mathcal{E}_\bullet'}^{\spadesuit}$ is a proper scheme for every pair $(\mathcal{E}_\bullet, \mathcal{E}_\bullet') \in \mathcal{FL}_{E^\sharp}^{\square} \times \mathcal{FL}_{E^\sharp}^{\square}$, and is empty for all but finitely many pairs. In particular, the intersection number in part (b) of the conjecture is well-defined.

Proposition 7.9. *Suppose $m = 1$ (so that $\mathcal{N}_{m,n} \simeq \mathcal{N}_n$). For $g \in G(\mathbb{V})(F_0)_{\mathrm{rs}}$, the formal scheme $\mathcal{N}_H^{\spadesuit} \times_{\mathcal{N}_{m,n}} (1, g) \cdot \mathcal{N}_H^{\spadesuit}$ is a proper scheme.*

Proof. By Prop. 7.2, the Zariski closure $\mathcal{N}_{H,\mathcal{E}_\bullet}^{\spadesuit}$ is a closed formal subscheme of the KR cycle $\mathcal{Z}(\mathcal{E})$ where $\mathcal{E} = \mathcal{E}_{r+1}$. Therefore, it suffices to show that

(a) $\mathcal{Z}(\mathcal{E}) \cap g\mathcal{Z}(\mathcal{E}')$ is empty except for finitely many pairs $(\mathcal{E}_\bullet, \mathcal{E}_\bullet') \in \mathcal{FL}_{E^\sharp}^{\square} \times \mathcal{FL}_{E^\sharp}^{\square}$, and

(b) $\mathcal{Z}(\mathcal{E}) \cap g\mathcal{Z}(\mathcal{E}')$ is a proper scheme.

The first assertion follows from the proof of Prop. 7.6 (hence Prop. 6.7). For the second one, the proof of Prop. 6.7 shows that $\mathcal{E} + g\mathcal{E}'$ is a lattice of full rank in $\mathbb{V}$. □

Acknowledgements

References

[1] R. BEUZART-PLESSIS, *A new proof of Jacquet-Rallis's fundamental lemma*, Duke Math. J., **170** (2021), no. 12, 2805–2814. MR4305382

[2] R. Beuzart-Plessis, P.-H. Chaudouard, M. Zydor, *The global Gan-Gross-Prasad conjecture for unitary groups: the endoscopic case*, Publ. Math. Inst. Hautes Études Sci., **135** (2022), 183–336. MR4426741

[3] R. Beuzart-Plessis, Y. Liu, W. Zhang, and X. Zhu, *Isolation of cuspidal spectrum, with application to the Gan–Gross–Prasad conjecture*, Ann. of Math(2), **194** (2021), no. 2, 519–584. MR4298750

[4] W. T. Gan, B. Gross, and D. Prasad, *Symplectic local root numbers, central critical L-values, and restriction problems in the representation theory of classical groups*, Astérisque, **346** (2012), 1–109. MR3202556

[5] D. Ginzburg, S. Rallis, *The exterior cube L-function for GL(6)*, Compos. Math., **123** (2000), no. 3, 243–272.

[6] J. Gordon, *Transfer to characteristic zero*, appendix to [28].

[7] B. Gross and D. Zagier, *Heegner points and derivatives of L-series*, Invent. Math., **84** (1986), no. 2, 225–320. MR0833192

[8] H. Jacquet, *Relative Kloosterman integrals for GL(3). II*, Canad. J. Math., **44** (1992), no. 6, 1220–1240.

[9] H. Jacquet, S. Rallis, *On the Gross-Prasad conjecture for unitary groups*, in On Certain L-Functions, Clay Math. Proc., vol. 13, Amer. Math. Soc., Providence, RI, 2011, pp. 205–264.

[10] S. Kudla and M. Rapoport, *Special cycles on unitary Shimura varieties I. Unramified local theory*, Invent. Math., **184** (2011), no. 3, 629–682. MR2800697

[11] Chao Li, Wei Zhang, *Kudla–Rapoport cycles and derivatives of local densities*, J. Amer. Math. Soc., **35** (2022), no. 3, 705–797.

[12] Chao Li, Yihang Zhu, *Remarks on the arithmetic fundamental lemma*, Algebra Number Theory, **11** (2017), no. 10, 2425–2445

[13] Y. Liu, *Relative trace formulae toward Bessel and Fourier-Jacobi periods of unitary groups*, Manuscripta Mathematica, **145** (2014), 1–69.

[14] Y. Liu, *Fourier-Jacobi cycles and arithmetic relative trace formula*, Cam. J. Math., **9** (2021), 1–147.

[15] E. Mantovan, *On non-basic Rapoport–Zink spaces*, Ann. Sci. Éc. Norm. Supér. (4), **41** (2008), no. 5, 671–716. MR2504431

[16] A. Mihatsch, *Relative unitary RZ-spaces and the arithmetic fundamental lemma*, Jussieu. J. Inst. Math. Jussieu, **21** (2022), no. 1, 241–301. MR4366338

[17] A. MIHATSCH, W. ZHANG, *On the Arithmetic Fundamental Lemma conjecture over a general p-adic field*, To appear JEMS, arxiv: 2104.02779.

[18] M. RAPOPORT, B. SMITHLING, and W. ZHANG, *Arithmetic diagonal cycles on unitary Shimura varieties*, Compos. Math., **156** (2020), no. 9, 1745–1824.

[19] M. RAPOPORT, U. TERSTIEGE, and W. ZHANG, *On the arithmetic fundamental lemma in the minuscule case*, Compos. Math., **149** (2013), no. 10, 1631–1666.

[20] M. RAPOPORT, E. VIEHMANN, *Towards a theory of local Shimura varieties*, Münster J. Math., **7** (2014), no. 1, 273–326.

[21] M. RAPOPORT and TH. ZINK, *Period spaces for p-divisible groups*, Annals of Mathematics Studies, vol. 141, Princeton University Press, Princeton, NJ, 1996. MR1393439

[22] Y. SAKELLARIDIS, *Spherical varieties and integral representations of L-functions*, Algebra Number Theory, **6** (2012), no. 4, 611–667.

[23] Y. SAKELLARIDIS, A. VENKATESH, *Periods and harmonic analysis on spherical varieties*, Astérisque, **396** (2017), viii+360 pp. MR3764130

[24] N. VANDENBERGEN, *On the global structure of special cycles on unitary Shimura varieties*, Canad. J. Math., **65** (2013), no. 5, 1125–1163. MR3095009

[25] I. VOLLAARD and T. WEDHORN, *The supersingular locus of the Shimura variety for* GU(1, n − 1) *II*, Invent. Math., **184** (2011), no. 3, 591–627. MR2800696

[26] C. WAN and L. ZHANG, *Periods of automorphic forms associated to strongly tempered spherical varieties*, arXiv:2102.03695.

[27] X. YUAN, S.-W. ZHANG, W. ZHANG, *The Gross–Zagier formula on Shimura curves*, Annals of Mathematics Studies, vol. 184, Princeton University Press, Princeton, NJ, 2013. MR3237437

[28] Z. YUN, *The fundamental lemma of Jacquet–Rallis in positive characteristics*, Duke Math. J., **156** (2011), no. 2, 167–228.

[29] S.-W. ZHANG, *Linear forms, algebraic cycles, and derivatives of L-series*, Sci. China Math., **62** (2019), no. 11, 2401–2408.

[30] W. ZHANG, *On arithmetic fundamental lemmas*, Invent. Math., **188** (2012), no. 1, 197–252.

[31] W. ZHANG, *Weil representation and Arithmetic Fundamental Lemma*, Ann. of Math., **193** (2021), no. 3, 863–978.

[32] ZHIYU ZHANG, *Maximal parahoric arithmetic transfers, resolutions and modularity.* arXiv:2112.11994.

Wei Zhang
Massachusetts Institute of Technology
Department of Mathematics
77 Massachusetts Avenue
Cambridge, MA 02139
USA
E-mail: weizhang@mit.edu