Co-rank 1 Arithmetic Siegel–Weil

by

Ryan C. Chen

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MATHEMATICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2025

Authored by: Ryan C. Chen
Department of Mathematics
April 28, 2025

Certified by: Wei Zhang
Professor of Mathematics, Thesis Supervisor

Accepted by: Zhiwei Yun
Professor of Mathematics
Graduate Co-Chair, Department of Mathematics

Co-rank 1 Arithmetic Siegel–Weil

by

Ryan C. Chen

Submitted to the Department of Mathematics
on April 28, 2025 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MATHEMATICS

ABSTRACT

We prove the arithmetic Siegel–Weil formula in co-rank 1, for Kudla–Rapoport special cycles on exotic smooth integral models of unitary Shimura varieties of arbitrarily large even arithmetic dimension. We also propose a construction for arithmetic special cycle classes associated to possibly singular matrices of arbitrary co-rank. Our arithmetic Siegel–Weil formula implies that degrees of Kudla–Rapoport arithmetic special 1-cycles are encoded in near-central first derivatives of unitary Eisenstein series Fourier coefficients.

The key input is a new limiting method at all places. On the analytic side, the limit relates local Whittaker functions on different groups. On the geometric side at nonsplit non-Archimedean places, the limit relates degrees of 0-cycles on Rapoport–Zink spaces and local contributions to heights of 1-cycles in mixed characteristic.

Thesis supervisor: Wei Zhang
Title: Professor of Mathematics

Acknowledgments

I thank my advisor Wei Zhang for suggesting this topic, for his dedicated support and constant enthusiasm, for insightful discussions throughout the entire course of this project, and for helpful comments on earlier drafts. I thank Tony Feng, Qiao He, Benjamin Howard, Ishan Levy, Chao Li, Keerthi Madapusi, Andreas Mihatsch, Bjorn Poonen, Siddarth Sankaran, Ananth Shankar, Yousheng Shi, Tonghai Yang, Shou-Wu Zhang, Zhiyu Zhang, and the anonymous referees for helpful comments, discussions, or suggestions.

I thank Bjorn Poonen, Zhiwei Yun, and Wei Zhang for serving on my thesis defense committee, and I thank George Lusztig, Davesh Maulik, and Wei Zhang for serving on my qualifying exam committee. I thank Tony Feng, Benjamin Howard, William Minicozzi, Michael Rapoport, Shou-Wu Zhang, and Wei Zhang for their support.

I thank everyone in the number theory group at MIT for providing a stimulating environment, and I thank the MIT math department staff for their work and help.

Lastly, I thank my parents and my brother Eric for their enduring support.

This work was partly supported by the National Science Foundation Graduate Research Fellowship under Grant Nos. DGE-1745302 and DGE-2141064. Parts of this work were completed at the Mathematical Sciences Research Institute (MSRI), now becoming the Simons Laufer Mathematical Sciences Institute (SLMath), and the Hausdorff Institute for Mathematics. I thank these institutes for their support and hospitality. The former is supported by the National Science Foundation (Grant No. DMS-1928930), and the latter is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/1 – 390685813.

I also thank the MIT math department and Presidential Fellowship for their support during graduate school.

Contents

List of Figures

Preliminary

1 Introduction

The landmark work of Gross and Zagier [GZ86] showed that Néron–Tate heights of Heegner points on elliptic curves over $\mathbb{Q}$ are encoded in the first central derivatives of associated Rankin–Selberg L-functions. After the work of Gross and Keating [GK93] on arithmetic intersection numbers for modular correspondences, Kudla proposed to recast such formulas in the language of special cycles on higher-dimensional Shimura varieties. This was originally formulated for integral models of orthogonal Shimura varieties in low dimensions by Kudla [Kud97a; Kud97b; Kud04] and the subsequent work of Kudla and Rapoport [KR99; KR00], where they pioneered the moduli definition of special cycles on integral models. Later, the attention was shifted to unitary Shimura varieties by Kudla–Rapoport in [KR11; KR14]. Along with other closely related predictions about special cycles (e.g. modularity of generating series), these ideas are now called *Kudla's program.* Kudla's program has played a role in a range of works, such as Gross–Zagier formulas on Shimura curves [YZZ13], the averaged Colmez conjecture [AGHMP18; YZ18], the arithmetic fundamental lemma [Zha21], results on the Beilinson–Bloch conjecture [LL21], and Picard rank jumps for K3 surfaces [SSTT22]. We refer to Li's excellent surveys [Li23; Li24] for more.

Our work is about *arithmetic Siegel–Weil formulas* in Kudla's program [Kud04, Problem 6], which (conjecturally) relate the first derivatives of Siegel Eisenstein series for unitary (resp. symplectic) groups with "arithmetic theta series" formed from special cycles on integral models of unitary (resp. orthogonal) Shimura varieties (see Eq. (1.3.3) below). These are closely parallel to the classical (resp. geometric) Siegel–Weil formulas, which state that special values of Eisenstein series encode representation numbers (resp. complex volumes) for lattices (resp. complex special cycles on Shimura varieties).

Our main results are stated in Section 1.4. Our proof is local in nature; we deduce our global arithmetic Siegel–Weil formula by formulating and proving key "local arithmetic Siegel–Weil formulas" at all places.

We now outline the rest of the introduction. Section 1.2 contains some background on classical and geometric Siegel–Weil. This is for comparison with arithmetic Siegel–Weil, and helps fix some notation. The material in Section 1.2 is mostly expository, but some of our formulations may be new, particularly in our normalizations for Eisenstein series. The same normalization choices play an amplified role in our main arithmetic Siegel–Weil results. We also mention other results (comparison of complex volume and degrees of complex zero cycles to Eisenstein series) which seem to be new or at least not explicit in the literature; see discussion following Eqs. (1.2.6) and (1.2.10).

Section 1.3 contains background on arithmetic Siegel–Weil formulas, and a brief overview of our main results. Section 1.4 contains more detailed statements of our main results, and Section 1.5 contains a comparison with existing literature on arithmetic Siegel–Weil. Sections 1.7 to 1.9 contain an overview of the new ideas in the formulation and proofs of our main theorems. Section 1.10 outlines the structure of the remainder of this paper.

1.1 Eisenstein series

In our work, we focus on the unitary/Hermitian case. For the introduction, fix an imaginary quadratic field $F/\mathbb{Q}$ with ring of integers $\mathcal{O}_F$ and odd discriminant Δ. Given $m \in \mathbb{Z}_{\geq 0}$ and an even integer $n \in \mathbb{Z}$, we consider the (normalized) *Siegel Eisenstein series*

$$E^*(z,s)^\circ_n := \Lambda_m(s)^\circ_n \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in P_1(\mathbb{Z}) \backslash SU(m,m)(\mathbb{Z})} \frac{\det(y)^{s-s_0}}{\det(cz+d)^n |\det(cz+d)|^{2(s-s_0)}} \tag{1.1.1}$$

for the group

$$H := U(m,m) := \left\{ h \in \operatorname{Res}_{\mathcal{O}_F/\mathbb{Z}} \mathrm{GL}_{2m} : {}^t\overline{h} \begin{pmatrix} 0 & 1_m \\ -1_m & 0 \end{pmatrix} h = \begin{pmatrix} 0 & 1_m \\ -1_m & 0 \end{pmatrix} \right\} \tag{1.1.2}$$

where $\Lambda_m(s)^\circ_n$ is the normalizing factor

$$\begin{aligned} \Lambda_m(s)^\circ_n := {} & \frac{(2\pi)^{m(m-1)/2}}{(-2\pi i)^{nm}} \pi^{m(-s+s_0)} |\Delta|^{m(m-1)/4 + \lfloor m/2 \rfloor (s+s_0)} \\ & \cdot \left(\prod_{j=0}^{m-1} \Gamma(s - s_0 + n - j) \cdot L(2s + m - j, \eta^{j+n}) \right). \end{aligned} \tag{1.1.3}$$

In (1.1.2), the notation 1_m stands for the $m \times m$ identity matrix, we wrote $SU(m,m) \subseteq U(m,m)$ for the determinant 1 subgroup, and we set $P_1 := P \cap SU(m,m)$ for the Siegel parabolic $P \subseteq U(m,m)$ (consisting of $m \times m$ block upper triangular matrices). The variable $s \in \mathbb{C}$ is a complex parameter, we set $s_0 = (n-m)/2$, and the element $z = x + iy$ lies in Hermitian upper-half space (i.e. $x \in \mathrm{Herm}_m(\mathbb{R})$ and $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$; the latter means that y is positive definite).[1] The symbol η denotes the quadratic character associated to $F/\mathbb{Q}$ (via class field theory). The sum in (1.1.1) is convergent for $\mathrm{Re}(s) > m/2$, and admits meromorphic continuation to all $s \in \mathbb{C}$. When $m = 1$, the expression in (1.1.1) is a classical Eisenstein series on the usual upper-half plane.

[1] Here, the notation Herm_m denotes a scheme over $\operatorname{Spec} \mathbb{Z}$, e.g. $\mathrm{Herm}_m(\mathbb{R})$ denotes $m \times m$ complex Hermitian matrices, and $\mathrm{Herm}_m(\mathbb{Q})$ denotes $m \times m$ Hermitian matrices with entries in F.

The normalized Eisenstein series has a symmetric functional equation

$$E^*(z,s)^\circ_n = (-1)^{m(m-1)(n-m-1)/2} E^*(z,-s)^\circ_n \tag{1.1.4}$$

as in Section 17.1. Our definition of the normalizing factor $\Lambda_m(s)^\circ_n$ is motivated by symmetry of global and local functional equations, along with certain local special value formulas; see Sections 13 to 17 for further discussion. The function $\Lambda_m(s)^\circ_n$ should be closely related with the L-function of an Artin–Tate motive attached to the group $U(m,m)$, in the sense of Gross [Gro97] (see [BH21, Remark 1.1.1]).

Given $T \in \mathrm{Herm}_m(\mathbb{Q})$, the Eisenstein series $E^*(z,s)^\circ_n$ has *T-th Fourier coefficient*

$$E^*_T(y,s)^\circ_n := 2^{m(m-1)/2}|\Delta|^{-m(m-1)/4} \int_{\mathrm{Herm}_m(\mathbb{Z})\backslash \mathrm{Herm}_m(\mathbb{R})} E^*(z,s)^\circ_n e^{-2\pi i \mathrm{tr}(Tz)} \, dx \tag{1.1.5}$$

for $z = x + iy$ in Hermitian upper-half space, where this integral is taken with respect to the Euclidean measure[2] on $\mathrm{Herm}_m(\mathbb{R})$. The integral is convergent for $\mathrm{Re}(s) > m/2$, and admits meromorphic continuation to all $s \in \mathbb{C}$. When $\det T \neq 0$, there is a factorization into normalized *local Whittaker functions*

$$E^*_T(y,s)^\circ_n = W^*_{T,\infty}(y,s)^\circ_n \prod_p W^*_{T,p}(s)^\circ_n \tag{1.1.6}$$

over all places (Part V).

For example, if $n = 2$ and $m = 1$, we have

$$\begin{aligned} W^*_{T,p}(s)^\circ_2 &= p^{v_p(T)(s+1/2)} \sigma_{-2s}(p^{v_p(T)}) \qquad \sigma_s(r) := \sum_{d \mid r} d^s \\ W^*_{T,\infty}(y,s)^\circ_2 &= \Gamma(s-1/2)^{-1} |4\pi T y|^{s-1/2} \int_a^\infty e^{-4\pi |T| y u} (u \pm 1)^{s+1/2} u^{s-3/2} \, du \end{aligned} \tag{1.1.7}$$

for any nonzero $T \in \mathbb{Z}$, where $a = 0$ and the sign $\pm$ is $+$ (resp. $a = 1$ and the sign $\pm$ is $-$) if $T > 0$ (resp. if $T < 0$). Here $W^*_{T,\infty}(y,s)^\circ_2$ (resp. $W^*_{T,p}(s)^\circ_2$) is a certain normalized Archimedean (resp. non-Archimedean) local Whittaker function.

1.2 Classical and geometric Siegel–Weil

Let V be an n-dimensional F-vector space, equipped with a non-degenerate Hermitian pairing $(-,-)$. Set $G = U(V)$ and assume $n > 0$. Fix a full-rank $\mathcal{O}_F$-lattice $L \subseteq V$. For simplicity,

[2] The factor $2^{m(m-1)/2}|\Delta|^{-m(m-1)/4}$ disappears in the (usual) equivalent adèlic formulation, upon taking a certain self-dual Haar measure. The adèlic formulations of (1.1.1) and (1.1.5) are used in Section 13.

we assume in the introduction that L is self-dual.[3] Write $K_{L,f} \subseteq G(\mathbb{A}_f)$ for the stabilizer of $L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$, where $\mathbb{A}_f$ denotes the finite adèle ring of $\mathbb{Q}$.

First consider the case where V is positive definite. Since we assumed L is self-dual, this forces $n \equiv 0 \pmod 4$ (by the global product formula for local invariants of Hermitian spaces). Given any positive definite Hermitian $\mathcal{O}_F$-lattice $\mathcal{L}$, we set

$$\mathcal{Z}_{T,\mathcal{L}} \coloneqq \{\underline{x} \in \mathcal{L}^m : (\underline{x}, \underline{x}) = T\}, \tag{1.2.1}$$

for Hermitian matrices T, where $(\underline{x}, \underline{x})$ denotes the Gram matrix[4] of $\underline{x}$. When $m \leq n$, we have

$$2\Lambda_n(0)^\circ_n = \sum_{\substack{\mathcal{O}_F\text{-lattices } \mathcal{L} \\ \text{self-dual, rank } n, \\ \text{positive definite}}} \frac{1}{|\operatorname{Aut}(\mathcal{L})|} = \#[G(\mathbb{Q})\backslash(G(\mathbb{A}_f)/K_{L,f})] \tag{1.2.2}$$

$$\frac{2\Lambda_n(0)^\circ_n}{\kappa\Lambda_m(s_0)^\circ_n} E^*_T(y, s_0)^\circ_n = \sum_{\substack{\mathcal{O}_F\text{-lattices } \mathcal{L} \\ \text{self-dual, rank } n, \\ \text{positive definite}}} \frac{|\mathcal{Z}_{T,\mathcal{L}}|}{|\operatorname{Aut}(\mathcal{L})|} \qquad \text{for any } T \in \operatorname{Herm}_m(\mathbb{Q}), \tag{1.2.3}$$

where $\kappa = 2$ (resp. $\kappa = 1$) if $m = n$ (resp. if $m < n$). The sums run over isomorphism classes of positive definite rank n self-dual Hermitian $\mathcal{O}_F$-lattices, the notation $\operatorname{Aut}(\mathcal{L})$ means the (unitary) automorphism group of $\mathcal{L}$. The symbols $\#[-]$ and $|-|$ mean groupoid and set cardinality, respectively. That is, we have

$$\frac{2\Lambda_n(0)^\circ_n}{\kappa\Lambda_m(s_0)^\circ_n} E^*(z, s_0)^\circ_n = \sum_{\substack{\mathcal{O}_F\text{-lattices } \mathcal{L} \\ \text{self-dual, rank } n, \\ \text{positive definite}}} \frac{1}{|\operatorname{Aut}(\mathcal{L})|} \Theta_{\mathcal{L}}(z) \tag{1.2.4}$$

which re-expresses the Eisenstein series at $s = s_0$ as a weighted sum of theta series for the lattices $\mathcal{L}$.

Equations (1.2.2) and (1.2.3) are special cases of (unitary analogues of) the classical *Siegel mass formula* and *Siegel–Weil formula* respectively. For Eq. (1.2.2), see Proposition 21.2.1. Equation (1.2.3) follows from [Ich04, Proposition 6.2], [Ich07, Theorem 1.1], and [Yam11, Theorem 2.2] (in combination with Eq. (1.2.2)).

Next, consider the case where V has arbitrary signature $(n-r, r)$. Since L was assumed self-dual, this forces $n \equiv 2^r \pmod 4$. There is an associated Hermitian symmetric domain $\mathcal{D}$

[3] We always mean self-dual for the bilinear *trace pairing* $\operatorname{tr}_{F/\mathbb{Q}}(v, w)$ unless otherwise specified; see conventions in Section 2.1.

[4] If $\underline{x}$ is the m-tuple $[x_1, \ldots, x_m]$, the notation $(\underline{x}, \underline{x})$ will mean the matrix with i, j-th entry (x_i, x_j). We often write e.g. $[x_1, \ldots, x_m]$ instead of $(x_1, \ldots, x_m)$ for tuples, to avoid confusion with Hermitian pairings $(-,-)$.

which parameterizes maximal negative definite subspaces of the complex Hermitian space $V_{\mathbb{R}}$. For sufficiently small open compact $K_f \subseteq K_{L,f}$ (so that we have manifolds instead of orbifolds, for simplicity), there is an associated complex Shimura variety

$$\mathrm{Sh}_{K_f,\mathbb{C}}(G) = [G(\mathbb{Q})\backslash(\mathcal{D} \times G(\mathbb{A}_f)/K_f)] \tag{1.2.5}$$

of dimension $(n-r)r$ (analytification suppressed from notation). In the signature $(n,0)$ and $(n-1,1)$ cases respectively, we have "geometric Siegel mass formulas"

$$2\Lambda_n(0)^\circ_n = \frac{\mathrm{vol}(\mathrm{Sh}_{K_f,\mathbb{C}}(G))}{[K_{L,f} : K_f]} \qquad -2\Lambda_n(0)^\circ_n = \frac{\mathrm{vol}(\mathrm{Sh}_{K_f,\mathbb{C}}(G))}{[K_{L,f} : K_f]} \tag{1.2.6}$$

where $\mathrm{vol}(\mathrm{Sh}_{K_{L,f},\mathbb{C}}(G))$ is the volume with respect to the Chern form of a certain dual tautological bundle. The case of signature $(n-1,1)$ may be extracted from [BH21, Theorem A], see Proposition 21.2.3.[5] The case of signature $(n,0)$ is an equivalent reformulation of the classical Siegel mass formula Eq. (1.2.2): if we allow the (stacky) level $K_f = K_{L,f}$, then there is a canonical equivalence of groupoids

$$\mathrm{Sh}_{K_L,\mathbb{C}}(G) \cong \left\{ \begin{array}{l} \text{Hermitian } \mathcal{O}_F\text{-lattices } \mathcal{L} \text{ which are} \\ \text{self-dual and signature } (n,0) \end{array} \right\} \tag{1.2.7}$$

in that case.

In *geometric Siegel–Weil formulas*, the sets $\mathcal{Z}_{T,\mathcal{L}}$ (from classical Siegel–Weil) are replaced by *special cycles* $\mathcal{Z}_{T,\mathbb{C}}$ over the Shimura variety, and the theta series $\Theta_{\mathcal{L}}(z)$ become *generating series of special cycles*. One can define $\mathcal{Z}_{T,\mathbb{C}}$ by the complex uniformization

$$\mathcal{Z}_{T,\mathbb{C}} \coloneqq \left[G(\mathbb{Q})\backslash \left(\coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_\infty) \times \mathcal{D}(\underline{x}_f) \right) \right] \tag{1.2.8}$$

where $\mathcal{D}(\underline{x}_\infty) \subseteq \mathcal{D}$ is the closed complex submanifold consisting of those complex lines perpendicular to all elements of the m-tuple $\underline{x}_\infty$, and

$$\mathcal{D}(\underline{x}_f) \coloneqq \{ g \in G(\mathbb{A}_f)/K_f : g^{-1}x_i \in L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}} \text{ for all } x_i \in \underline{x}_f \} \subseteq G(\mathbb{A}_f)/K_f, \tag{1.2.9}$$

with $\underline{x}_\infty$ and $\underline{x}_f$ denoting the images of $\underline{x}$ in $V(\mathbb{R})^m$ and $V(\mathbb{A}_f)^m$, respectively. The definition in Eq. (1.2.8) is (a reformulation of) a definition due to Kudla [Kud04] (there for GSpin), with unitary analogue as in [Liu11, §3]. We call $\mathcal{D}(\underline{x}_\infty)$ an Archimedean *local special cycle*

[5] In Proposition 21.2.3, note that we took vol(−) with respect to the tautological bundle. Here we are taking volume with respect to the dual bundle, which produces the minus sign in Eq. (1.2.6).

and $\mathcal{D}(\underline{x}_f)$ an "away-from-∞" local special cycle. There is a natural map $\mathcal{Z}_{T,\mathbb{C}} \to \mathrm{Sh}_{K_f,\mathbb{C}}$, which is a disjoint union of closed immersions of complex manifolds after possibly shrinking K_f.

A *geometric Siegel–Weil formula* for signature $(n-1,1)$ is an identity of the shape

$$-\frac{2\Lambda_n(0)_n^\circ}{\kappa\Lambda_m(s_0)_n^\circ} E_T^*(y,s_0)_n^\circ = \frac{\mathrm{vol}(\mathcal{Z}_{T,\mathbb{C}})}{[K_{L,f} : K_f]} \tag{1.2.10}$$

for $m \leq n-1$ (so $\kappa = 1$). In the case of signature $(n,0)$, the expression in Eq. (1.2.10) (without the minus sign on the left) is an equivalent reformulation of the classical Siegel–Weil formula Eq. (1.2.3): if we allow the (stacky) level $K_f = K_{L,f}$, there is a canonical equivalence of groupoids

$$\mathcal{Z}_{T,\mathbb{C}} \cong \left\{ \begin{array}{l} \text{pairs } (\mathcal{L}, \underline{x}), \text{ where } \mathcal{L} \text{ is a self-dual Hermitian } \mathcal{O}_F\text{-lattice} \\ \text{of signature } (n,0) \text{ and } \underline{x} \in \mathcal{L}^m \text{ is an } m\text{-tuple with } (\underline{x},\underline{x}) = T \end{array} \right\}. \tag{1.2.11}$$

Our presentation of the geometric Siegel–Weil formula in Eq. (1.2.10) may be nonstandard. Its appearance is intended to highlight the similarity with our formulation of arithmetic Siegel–Weil in Eq. (1.3.3).

Strictly speaking, geometric Siegel–Weil formulas in literature typically restrict to V satisfying Weil's convergence condition (meaning V anisotropic or $m < n-1$ in the signature $(n-1,1)$ Hermitian setup), see remarks following [Kud04, Theorem 4.1] and [Li24, Theorem 3.6.1]. It is also typical to phrase geometric Siegel–Weil formulas in terms of "coherent" Eisenstein series, while our $E^*(z,s)_n^\circ$ is described in terms of an incoherent adèlic Hermitian space (positive definite at ∞ and self-dual at all finite places), see Part V. Outside those cases available in the literature, geometric Siegel–Weil formulas may need additional care. For example, when $m = 1$ and $n = 2$ and $T = 0$ (which is essentially about "complex volume of modular curve"), the formula in Eq. (1.2.10) is only valid up to a non-holomorphic correction term $\frac{2h_F}{w_F} \cdot \frac{1}{8\pi y}$ on the left, where h_F (resp. w_F) is the class number of (resp. number of roots of unity in) $\mathcal{O}_F$. In this case, the right-hand side is $\frac{2h_F}{w_F} \cdot \frac{\zeta(-1)}{2} = \frac{-h_F}{12 w_F}$. Note that this non-holomorphic correction term does not seem to occur for the compact Shimura curves considered in [KRY06, (1.0.14)].

We will need the following geometric Siegel–Weil result which does not seem to be covered by the literature discussed in the previous paragraph: we prove Eq. (1.2.10) when T is nonsingular of rank $m = n-1$, see Proposition 21.1.1 (also complex uniformization in Section 12.3, as well as Eq. (22.1.2)); in that case, $\mathcal{Z}_{T,\mathbb{C}}$ is 0 dimensional. For example, when $n = 2$ and $\mathcal{O}_F^\times = \{\pm 1\}$, the special cycle $\mathcal{Z}_{T,\mathbb{C}}$ can be described in terms of Hecke translates of CM elliptic curves (Section 22.2), and Eq. (1.2.10) is then the (well-known) statement that the T-th Hecke correspondence (over the modular curve) has bidegree

$$-\frac{1}{h_F}\frac{2\Lambda_2(0)^\circ_2}{\Lambda_1(1/2)^\circ_2}E^*_T(y,1/2)^\circ_2 = \sigma_1(T) \tag{1.2.12}$$

for $T \in \mathbb{Z}_{>0}$. The extra factor of h_F accounts for multiple connected components in the Shimura variety, see Section 22.2.

We remark that our proof of Eq. (1.2.10) (for T nonsingular of rank $m = n-1$) is inspired by [LZ22a, Remark 4.6.2], and may be carried out using either complex or non-Archimedean (Rapoport–Zink) uniformization. We need that case of Eq. (1.2.10) as an ingredient for our main arithmetic Siegel–Weil results.

1.3 Arithmetic Siegel–Weil

Arithmetic Siegel–Weil formulas predict that the derivative of $E^*_T(y,s)^\circ_n$ at $s = s_0$ should encode *arithmetic degrees/volumes* of special cycles on integral models of Shimura varieties.

Since the work of Kudla–Rapoport [KR14] (also Rapoport–Smithling–Zhang [RSZ21]), it has been customary to consider special cycles $\mathcal{Z}(T) \to \mathcal{M}$ over (stacky) integral models $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ for Shimura varieties associated to $G' := \operatorname{Res}_{F/\mathbb{Q}} \mathbb{G}_m \times U(V)$, for signature $(n-1,1)$ non-degenerate $F/\mathbb{Q}$ Hermitian spaces V with pairing $(-,-)$. In the introduction, we assume V contains a full-rank self-dual[6] $\mathcal{O}_F$-lattice, and we take $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ to be the "exotic smooth" Rapoport–Smithling–Zhang (RSZ) integral model of relative dimension $n-1$ [RSZ21, §6].[7] When $n = 2$, the stack $\mathcal{M}$ is essentially a disjoint union of (stacky) modular curves (Example 3.2.2).

The stack $\mathcal{M}$ admits a moduli description: it parameterizes tuples $(A_0, \iota_0, \lambda_0, A, \iota, \lambda)$ where A_0 and A are abelian schemes (dimensions 1 and n respectively) with $\mathcal{O}_F$-actions ι_0 and ι, and with compatible quasi-polarizations λ_0 and λ. The datum $(A_0, \iota_0, \lambda_0, A, \iota, \lambda)$ satisfies a few additional conditions, which we postpone to Section 3.1.

The moduli stack $\mathcal{M}$ carries a natural family of Hermitian $\mathcal{O}_F$-lattices

$$\mathcal{L}at \to \mathcal{M} \qquad \mathcal{L}at := \underline{\operatorname{Hom}}_{\mathcal{O}_F}(A_0, A), \tag{1.3.1}$$

[6] We always mean *self-dual* for the bilinear *trace pairing* $\operatorname{tr}_{F/\mathbb{Q}}(v,w)$ unless otherwise specified; see conventions in Section 2.2.

[7] Our assumption on V forces $n \equiv 2 \pmod 4$. We allow a slightly more general setup in Section 3.1 for general V, at the cost of throwing out finitely many primes (particularly the ramified primes when n is odd). This does not affect the essential ideas of our method, which is local in nature.

e.g. as explained in Section 3.3 and Appendix C.1. Given any $T \in \mathrm{Herm}_m(\mathbb{Q})$, the associated *Kudla–Rapoport* special cycle $\mathcal{Z}(T) \to \mathcal{M}$ is defined as the substack

$$\mathcal{Z}(T) \coloneqq \{\underline{x} \in \mathcal{L}at^m : (\underline{x}, \underline{x}) = T\} \subseteq \mathcal{L}at^m \tag{1.3.2}$$

consisting of m-tuples $\underline{x}$ with Gram matrix T. More precisely, see Section 3.3. This is in close analogy with classical Siegel–Weil: there we considered $\mathcal{O}_F$-lattices varying in a given *genus*,[8] and here we are considering $\mathcal{O}_F$-lattices varying over the moduli stack $\mathcal{M}$. In the complex fiber, the special cycles $\mathcal{Z}(T)_{\mathbb{C}}$ recover the special cycles $\mathcal{Z}_{T,\mathbb{C}}$ appearing in Eq. (1.2.8), up to $\mathcal{M}_{K'_f,\mathbb{C}}$ being a finite cover of $\mathrm{Sh}_{K_f,\mathbb{C}}(G)$ (for suitable K'_f); see Section 12.3. The morphism $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F$ is smooth of relative dimension $n - 1 - \mathrm{rank}(T)$ in the generic fiber over $\operatorname{Spec} F$. If T is not positive semi-definite, then $\mathcal{Z}(T)$ is empty.

An *arithmetic Siegel–Weil formula* is an identity roughly of the shape

$$\frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=s_0} \frac{2\Lambda_n(s-s_0)^\circ_n}{\kappa \Lambda_m(s)^\circ_n} E^*_T(y,s)^\circ_n \stackrel{?}{=} \widehat{\mathrm{vol}}_{\widehat{\mathcal{E}}^\vee}([\widehat{\mathcal{Z}}(T)]). \tag{1.3.3}$$

Here we set $\kappa = 1$ (resp. $\kappa = 2$) if $m \neq n$ (resp. $m = n$). The right-hand side of Eq. (1.3.3) denotes an *arithmetic volume*, which is a real number "defined" by an arithmetic intersection product

$$\widehat{\mathrm{vol}}_{\widehat{\mathcal{E}}^\vee}([\widehat{\mathcal{Z}}(T)]) \text{“} \coloneqq \text{”} \widehat{\deg}([\widehat{\mathcal{Z}}(T)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{n-m}) \tag{1.3.4}$$

in an arithmetic Chow ring $\widehat{\mathrm{Ch}}^*(\mathcal{M})_{\mathbb{Q}}$ (roughly in the sense of Gillet–Soulé [GS87]) for a certain metrized tautological bundle $\widehat{\mathcal{E}}^\vee$ on $\mathcal{M}$ (the bundle $\widehat{\mathcal{E}}^\vee$ is discussed in Section 4.3). The notation $[\widehat{\mathcal{Z}}(T)]$ indicates a class in $\widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$, which is expected to involve $\mathcal{Z}(T)$ and some additional Archimedean data (e.g. from a Green current on the complex Shimura variety), as appearing in arithmetic intersection theory.

An expected application of arithmetic Siegel–Weil formulas is in the theory of *arithmetic theta lifting*. One expects to form automorphic *arithmetic theta series* as generating series

$$\widehat{\Theta} = \sum_T [\widehat{\mathcal{Z}}(T)] q^T \tag{1.3.5}$$

with "Fourier coefficients" $[\widehat{\mathcal{Z}}(T)]$ valued in the arithmetic Chow group $\widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$. These should be analogous to (weighted averages of) classical theta series, as in the classical Siegel–Weil formula Eq. (1.2.4). In analogy with classical theta lifting, one expects to use $\widehat{\Theta}$ as an integral kernel to lift $U(m,m)$ automorphic forms to elements of $\widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$. In

[8] In our previous setup, this meant the set of isomorphism classes of positive definite rank n self-dual $\mathcal{O}_F$-lattices.

analogy with the classical Rallis inner product formula, one expects to use the doubling method and arithmetic Siegel–Weil formulas to relate the derivative of an L-function with the arithmetic inner product of this arithmetic theta lift [Kud04, Part III]. We refer to [KRY06; BHKRY20II; LL21; LL22] for some cases where versions of this have been realized, with applications to Beilinson–Bloch. For modularity results on generating series of arithmetic divisors, see [KRY06; BBK07; BHKRY20; Qiu22].

We sketched the arithmetic Siegel–Weil formula as a rough expectation, because precise formulations remain open in the general case [Li24, Remark 4.4.2]. In general, it is necessary to renormalize or modify the Eisenstein series in a way which is not completely understood. In fact, our normalization on the left-hand side of Eq. (1.3.3) is already nonstandard (new). We are not certain about this normalization for arithmetic Siegel–Weil in general, but our Theorem A (i.e. (1.3.3) $m = n$ for T of co-rank 1, and $m = n - 1$ for T nonsingular; more discussion appears below) provides some evidence. The case of (1.3.3) with $m = n$ and T nonsingular also holds, as can be extracted from known theorems in the literature (see discussion following Eq. (1.4.9) below). Note also the similarity with our formulation of classical and geometric Siegel–Weil, as in Eqs. (1.2.3) and (1.2.10).

In general, posing a good (precise) definition of the arithmetic cycle class $[\widehat{\mathcal{Z}}(T)]$ is an open problem, especially for singular T (due to arithmetic-intersection-theoretic difficulties), and particularly in the unitary case or over general totally real fields (due to a certain class number phenomenon), see Section 4.4. Previous works used K-theoretic methods to define special cycle classes (e.g. [KR14] and [HM22]), and the works by Feng–Yun–Zhang (moduli of shtukas) [FYZ21; FYZ24] and Madapusi (Shimura varieties) [Mad23] have employed derived algebro-geometric methods to define special cycle classes. As of now, these constructions do not incorporate the Archimedean place, which would be needed for arithmetic intersection theory (e.g. there seems to be no "derived arithmetic intersection theory" at the moment). Garcia and Sankaran have defined (Archimedean) Green currents associated to singular T using the Mathai–Quillen theory of superconnections, but there has been no precise proposal to combine this with the above non-Archimedean theory.

We first propose a method to construct the arithmetic special cycle classes $[\widehat{\mathcal{Z}}(T)]$ for arbitrary T. Our proposed definition mixes the work of Garcia–Sankaran with K-theoretic methods for positive characteristic contributions. Our construction may need adjustment on compactifications of integral models, but we expect it to apply in already-compact situations (e.g. the Rapoport–Smithling–Zhang [RSZ21] setup for CM extensions of totally real fields $\neq \mathbb{Q}$).

The first part of our main theorem (Theorem A(1)) is a proof of (1) the arithmetic

Siegel–Weil formula when $m = n$ and $T \in \mathrm{Herm}_n(\mathbb{Q})$ is singular of co-rank 1. Most known (fully global) results concern special cycles $\mathcal{Z}(T) \to \mathcal{M}$ which are empty in the generic fiber. These previous results include the non-Archimedean Kudla–Rapoport conjectures (for $T \in \mathrm{Herm}_n(\mathbb{Q})$ nonsingular) proved by Li–Zhang [LZ22a] (and the ramified versions [HLSY23; LL22]), as well as the purely Archimedean results of Liu [Liu11] and Garcia–Sankaran [GS19]. Our theorem is the first (fully global) arithmetic Siegel–Weil result which involves mixed characteristic special cycles $\mathcal{Z}(T)$ on Shimura varieties of arbitrarily large dimension. We further discuss the comparison with previous literature in Section 1.5.

We also prove (2) the arithmetic Siegel–Weil formula when $m = n-1$ and T is nonsingular (Theorem A(2)). This is very closely related with our theorem for singular $T \in \mathrm{Herm}_n(\mathbb{Q})$ of co-rank 1, as we explain further in Section 1.4. This theorem implies that both the first derivative and the special value of a $U(n-1, n-1)$ Eisenstein series at the *non-central* point $s = s_0 = 1/2$ have geometric meaning; see discussion following Theorem A below.

As a byproduct of our methods, we prove (3) a version of the arithmetic Siegel–Weil formula (up to a volume constant which we did not calculate) for arbitrary m when T is nonsingular and not positive-definite (corresponding to a "purely Archimedean" arithmetic intersection number) (Theorem B). This purely Archimedean result is analogous to those in [GS19] (there in a situation with compact Shimura varieties, which need not apply in the setup above), but our method of proof is completely different and is insensitive to compactness.

More importantly, we propose and apply a new uniform strategy to prove (1), (2), and (3). This is the key conceptual novelty in our work. Our strategy is a certain "local limiting method" at all places, Archimedean and non-Archimedean. We further sketch this strategy in Section 1.7, and at a finer level of detail in Section 1.9.

1.4 Results

We describe our global results in more detail, retaining the notation from Section 1.3.

First, we propose a new candidate definition of arithmetic cycle classes

$$[\widehat{\mathcal{Z}}(T)] \coloneqq [\widehat{\mathcal{Z}}(T)_{\mathscr{H}}] + \sum_{p \text{ prime}} [{}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}] \in \widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}} \tag{1.4.1}$$

associated to arbitrary (possibly singular) T. Here, $[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}]$ is intended to describe "horizontal" contributions and ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ is intended to describe "vertical" contributions.

The vertical (positive characteristic) classes ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ will be constructed in Section 4.6. For each prime p, we give a (new) definition of an element ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}}$ ("vertical") lying in the dimension $n - m$ graded piece of the Grothendieck group (tensor

$\mathbb{Q}$) of coherent sheaves on $\mathcal{Z}(T)_{\mathbb{F}_p} := \mathcal{Z}(T) \times_{\operatorname{Spec}\mathbb{Z}} \operatorname{Spec}\mathbb{F}_p$. Our construction is based on a certain "p-local linear invariance", and is explained in Section 4.6.

As we explain in Section 4.5, the class $[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}]$ may be constructed using currents $g_{T,y}$ (associated to T and allowed to vary with a parameter $y \in \operatorname{Herm}_m(\mathbb{R})_{>0}$) satisfying a *modified current equation*, i.e. that

$$-\frac{1}{2\pi i}\partial\overline{\partial} g_{T,y} + \delta_{\mathcal{Z}(T)_{\mathbb{C}}} \wedge [c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{m-\operatorname{rank}(T)}] \tag{1.4.2}$$

is represented by a smooth (m,m)-form. For such currents, we apply the proposal of [GS19, §5.4] to the flat part[9] $\mathcal{Z}(T)_{\mathscr{H}}$ of $\mathcal{Z}(T)$.

For general T, there is no precise definition of $[\widehat{\mathcal{Z}}(T)]$ which has been proposed in the prior literature [Li24, Remark 4.4.2]. Our candidate definition may need modification on a compactification, but we expect it to apply in already-compact situations (e.g. the Rapoport–Smithling–Zhang [RSZ21] setup for CM extensions of totally real fields $\neq \mathbb{Q}$). In general, it may also be necessary to modify the Green currents differently than in [GS19, Definition 4.7]; see discussion below.

Currents satisfying Eq. (1.4.2) were constructed by Garcia and Sankaran [GS19], using the Mathai–Quillen theory of superconnections [GS19, (4.38)]. For their arithmetic Siegel–Weil results, however, they need a non "linearly invariant" modification of their current [GS19, Definition 4.7] (see discussion below).

We choose to instead use the star-product approach of Kudla [Kud97a] (as formulated by Liu for unitary groups [Liu11]) to define the currents $g_{T,y}$ for our arithmetic Siegel–Weil results. Traditionally, the star product approach was used for nonsingular T (or at least block diagonal T, with diagonal entries 0 or nonsingular). In Section 12.4, we give a (new) *linearly invariant* modification in the case of singular $T \in \operatorname{Herm}_n(\mathbb{Q})$ with rank $n-1$, which will appear in our arithmetic Siegel–Weil result for singular T.

As part of the expected automorphic behavior of $[\widehat{\mathcal{Z}}(T)]$, it is expected that these classes should satisfy a certain "linear invariance" property for the action[10] of $\operatorname{GL}_m(\mathcal{O}_F)$ on Hermitian matrices T. We verify this for the classes we define: for any $g_{T,y}$ satisfying

$$g_{T,y} = g_{{}^t\overline{\gamma}T\gamma, \gamma^{-1}y\,{}^t\overline{\gamma}^{-1}}, \tag{1.4.3}$$

[9]Given an algebraic stack $\mathcal{X}$ over a Dedekind domain R, its *flat part* or *horizontal part* of $\mathcal{X}_{\mathscr{H}}$ is the largest closed substack $\mathcal{X}_{\mathscr{H}} \subseteq \mathcal{X}$ which is flat over $\operatorname{Spec} R$. The stack $\mathcal{X}_{\mathscr{H}}$ is also the scheme-theoretic image of the generic fiber of $\mathcal{X}$. Given a formal algebraic stack $\mathcal{X}$ over $\operatorname{Spf} R$ for a complete discrete valuation ring R, its *flat part* or *horizontal part* $\mathcal{X}_{\mathscr{H}}$ is the largest closed substack $\mathcal{X}_{\mathscr{H}} \subseteq \mathcal{X}$ which is flat over $\operatorname{Spf} R$ (in the sense discussed in Section 11.7).

[10]For any $\gamma \in \operatorname{GL}_m(\mathcal{O}_F)$ and any Hermitian matrix $T \in \operatorname{Herm}_m(\mathbb{Q})$, we say e.g. that T and ${}^t\overline{\gamma}T\gamma$ are $\operatorname{GL}_m(\mathcal{O}_F)$-*equivalent*, and that they lie in the same $\operatorname{GL}_m(\mathcal{O}_F)$-*equivalence class*.

we show

$$[\widehat{\mathcal{Z}}(T)] = [\widehat{\mathcal{Z}}({}^t\overline{\gamma}T\gamma)] \tag{1.4.4}$$

where $[\widehat{\mathcal{Z}}(T)]$ is formed with respect to y and $[\widehat{\mathcal{Z}}({}^t\overline{\gamma}T\gamma)]$ is formed with respect to $\gamma^{-1}y^t\overline{\gamma}^{-1}$. In fact, we prove refined results: the vertical part at each prime p is linearly invariant on the level of Grothendieck groups Eq. (4.6.11), and the horizontal part is linearly invariant on its own (Section 4.5). The currents $g_{T,y}$ appearing in our main arithmetic Siegel–Weil results do satisfy the linear invariance property in (1.4.3); see Section 12.4. Note that the Garcia–Sankaran Green currents in [GS19, (4.38)] also satisfy the same linear invariance property (but the modified currents in [GS19, Definition 4.7] do not).

Due to non-properness of $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ in general, one should likely modify $[\widehat{\mathcal{Z}}(T)]$ on a suitable compactification of $\mathcal{M}$. If $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F$ is proper, however, we consider certain "arithmetic degrees without boundary contributions" (a real number)

$$\begin{aligned} \widehat{\deg}([\widehat{\mathcal{Z}}(T)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{n-m}) \coloneqq & \left(\int_{\mathcal{M}_{\mathbb{C}}} g_{T,y} \wedge c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{n-m} \right) \\ & + \widehat{\deg}((\widehat{\mathcal{E}}^\vee)^{n-\operatorname{rank}(T)}|_{\mathcal{Z}(T)_{\mathscr{H}}}) \\ & + \sum_{p \text{ prime}} \deg_{\mathbb{F}_p}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \cdot (\mathcal{E}^\vee)^{n-m}) \log p \end{aligned} \tag{1.4.5}$$

conditional on convergence of the integral, for a certain metrized tautological bundle $\widehat{\mathcal{E}}$ on $\mathcal{M}$ (Section 4.1) (we do check convergence of the integral in the settings of our arithmetic Siegel–Weil results). The middle term is mixed characteristic in nature: for $\operatorname{rank}(T) = n-1$, it is (essentially) a weighted sum of Faltings heights of abelian varieties (Remark 22.1.4). For proper $\mathcal{Z}(T) \to \mathcal{O}_F$, the quantity in (1.4.5) should coincide with the arithmetic degree (without boundary contributions) of a version of $[\widehat{\mathcal{Z}}(T)]$ on any reasonable compactification of $\mathcal{M}$.

Our main theorems concern the T-th Fourier coefficients $E^*_T(y,s)^\circ_n$ of $E^*(z,s)^\circ_n$. As above, we write h_F (resp. w_F) for the class number of (resp. number of roots of unity in) $\mathcal{O}_F$. The following is our main global theorem.

Theorem A (Co-rank 1 arithmetic Siegel–Weil)**.** *Assume the prime 2 splits in $\mathcal{O}_F$.*

(1) For any $T \in \operatorname{Herm}_n(\mathbb{Q})$ with $\operatorname{rank}(T) = n-1$ and any $y \in \operatorname{Herm}_n(\mathbb{R})_{>0}$, we have

$$\frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n = \widehat{\deg}([\widehat{\mathcal{Z}}(T)]). \tag{1.4.6}$$

(2) For any $T^\flat \in \operatorname{Herm}_{n-1}(\mathbb{Q})$ with $\det T^\flat \neq 0$ and any $y^\flat \in \operatorname{Herm}_{n-1}(\mathbb{R})_{>0}$, we have

$$2\frac{h_F}{w_F}\frac{d}{ds}\bigg|_{s=0} \left(\frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} E^*_{T^\flat}(y^\flat, s+1/2)^\circ_n \right) = \widehat{\deg}([\widehat{\mathcal{Z}}(T^\flat)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)). \tag{1.4.7}$$

This appears below as Theorem 22.1.1. Note that Theorem A(1) concerns the *central* derivative of a $U(n,n)$ Eisenstein series, while part Theorem A(2) concerns a *non-central* derivative of a $U(n-1,n-1)$ Eisenstein series. For $n \equiv 0 \pmod 4$, Theorem A(1) also holds in the sense that there is no self-dual $\mathcal{O}_F$-lattice of signature $(n-1,1)$ and the left-hand side is 0 (Remark 22.1.3).

Remark 1.4.1. In the situation of Theorem A(2), there is also a "geometric Siegel–Weil formula" as in (1.2.10) when we evaluate

$$2\frac{h_F}{w_F}\frac{\Lambda_n(s)_n^\circ}{\Lambda_{n-1}(s+1/2)_n^\circ}E^*_{T^\flat}(y^\flat, s+1/2)_n^\circ \tag{1.4.8}$$

at $s=0$; the resulting expression is exactly $-\deg_{\mathbb{C}} \mathcal{Z}(T^\flat)_{\mathbb{C}}$ (negative degree of complex fiber $\mathcal{Z}(T^\flat)_{\mathbb{C}}$, which is a proper and quasi-finite Deligne–Mumford stack over $\operatorname{Spec}\mathbb{C}$). In other words, both the special value and the first derivative at $s=1/2$ of the $U(n-1,n-1)$ Eisenstein series (normalized as in Eq. (1.4.8)) *simultaneously* have arithmetic-geometric meaning.

The above "geometric Siegel–Weil" formula is also needed as an ingredient in our proof of Theorem A, and will be treated in Section 21.1 via uniformization (Archimedean and non-Archimedean both work).

We highlight the simplicity of the analytic side in Theorem A(1). It is expected that arithmetic Siegel–Weil for integral models with bad reduction should be corrected on the analytic side, e.g. by special values of other Eisenstein series. See for example [HSY23; HLSY23] for bad reduction in the nonsingular case $\det T \neq 0$ for the central derivative at $s=0$ (i.e. T is $n \times n$), or [KRY06] for quaternionic Shimura curves. We do not know whether the analytic formulation [HSY23; HLSY23] is expected to hold for singular T.

We argue that arithmetic Siegel–Weil formulas should be simplest to formulate on integral models with everywhere good reduction, as in our case. We thus propose a precise formulation of the analytic side of the central derivative arithmetic Siegel–Weil formula in our setup.

Question (Arithmetic Siegel–Weil, central point)**.** *Let $T \in \operatorname{Herm}_n(\mathbb{Q})$ be arbitrary. For a suitable current $g_{T,y}$, a suitable compactification of $\mathcal{M}$, and a possibly modified class $[\widehat{\mathcal{Z}}(T)]$ on the compactification, do we have*

$$\frac{h_F}{w_F}\frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)_n^\circ \overset{?}{=} \widehat{\deg}([\widehat{\mathcal{Z}}(T)]). \tag{1.4.9}$$

Our theorem verifies this proposed arithmetic Siegel–Weil formula for all singular $T \in \operatorname{Herm}_n(\mathbb{Q})$ of rank $n-1$, in the sense of "arithmetic degrees without boundary contributions".

The formula also holds (in the same sense) for all nonsingular $T \in \mathrm{Herm}_n(\mathbb{Q})$. This latter case ("central derivative nonsingular arithmetic Siegel–Weil") is possibly considered known to experts up to a volume constant by collecting the local theorems in [Liu11; LZ22a; LL22]. This particular global statement does not appear in the literature, though other variants are available (e.g. for unramified CM fields F/F_0 with $F_0 \neq \mathbb{Q}$ [LZ22a] or on integral models with bad reduction and correction terms by special values of other Eisenstein series [HLSY23]). In one of our companion papers, we will compute the volume constant and explain how to extract the $\det T \neq 0$ case of (1.4.9) from the literature Remark 22.1.2.

A more optimistic version of Eq. (1.4.9) was given in Eq. (1.3.3) involving $T \in \mathrm{Herm}_m(\mathbb{Q})$ for arbitrary m, but we are less certain about the validity of that formulation in general.

In the general case of Eq. (1.4.9), we expect the current $g_{T,y}$ to be essentially the currents of [GS19, Definition 4.7], though $\mathfrak{g}(T, \mathbf{y}, \varphi_f)$ as defined in loc. cit. may need some modification (see above discussion on the non-"linearly invariant" modification of their current [GS19, Definition 4.7]). Since our main theorems take a different approach to define $g_{T,y}$, we do not pursue this issue further.

Part (2) of Theorem A is the special case of part (1) when $T = \mathrm{diag}(0, T^\flat)$ and $y = \mathrm{diag}(1, y^\flat)$. The geometric sides will agree essentially by definition Eq. (1.4.5). On the analytic side, the relation is provided by the formula

$$E^*_T(y, s)^\circ_n = \frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} E^*_{T^\flat}(y^\flat, s+1/2)^\circ_n - \frac{\Lambda_n(-s)^\circ_n}{\Lambda_{n-1}(-s+1/2)^\circ_n} E^*_{T^\flat}(y^\flat, s-1/2)^\circ_n \quad (1.4.10)$$

from Corollary 17.2.2, along with the functional equation $E^*_{T^\flat}(y^\flat, s)^\circ_n = E^*_{T^\flat}(y^\flat, -s)^\circ_n$. The general case of Theorem A is proved in a similar way as the special case $T = \mathrm{diag}(0, T^\flat)$, with an additional "local diagonalizability argument" (proof of Theorem 22.1.1) where the identity is proved modulo $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$ for any given p (and varying p removes the ambiguity).

It is also possible to formulate and prove Theorem A in terms of Faltings heights (i.e. replacing the middle term in (1.4.5) with the degree of the metrized Hodge bundle). The formulation in Theorem A seems more natural to us, but the version with Faltings heights is in Remark 22.1.4.

A simpler case of Theorem A is the case $n = 2$. When $\mathcal{O}_F^\times = \{\pm 1\}$, the Serre tensor construction gives an open and closed embedding $\mathscr{M}_0 \times_{\mathrm{Spec}\, \mathcal{O}_F} \mathscr{M}_{\mathrm{ell}} \to \mathcal{M}$, where $\mathscr{M}_0$ is the moduli stack of elliptic curves with signature $(1,0)$ action by $\mathcal{O}_F$ and $\mathscr{M}_{\mathrm{ell}}$ is the moduli stack of all elliptic curves, base-changed to $\mathcal{O}_F$ Section 22.2. In this case, the special cycle $\mathcal{Z}(j) \to \mathcal{M}$ for $j \in \mathbb{Z}_{>0}$ pulls back to the j-th Hecke correspondence. Then the proof of Theorem A gives the following corollary (appearing later as Corollary 22.2.2). One might think of this corollary as reformulating a result of Nakkajima–Taguchi [NT91] (they compute

Faltings heights of elliptic curves with CM by possibly non-maximal orders) by averaging over Hecke translates and expressing the result in terms of Eisenstein series Fourier coefficients.

Corollary 1.4.2. *Assume 2 is split in $\mathcal{O}_F$. Fix any elliptic curve E_0 over $\mathbb{C}$ with $\mathcal{O}_F$-action. For any integer $j > 0$, we have*

$$\sum_{E \xrightarrow{w} E_0} (h_{\mathrm{Fal}}(E) - h_{\mathrm{Fal}}(E_0)) = \frac{1}{2} \frac{d}{ds}\bigg|_{s=1/2} \left(j^{s+1/2} \sigma_{-2s}(j) \right) \tag{1.4.11}$$

where the sum runs over degree j isogenies $w \colon E \to E_0$ of elliptic curves.

The notation $h_{\mathrm{Fal}}(E)$ denotes the (stable) Faltings height of the elliptic curve E after descent to any number field, and similarly for E_0. The quantity $j^{s+1/2}\sigma_{-2s}(j)$ is the product of the normalized non-Archimedean local Whittaker functions in the j-th Fourier coefficient $E^*_j(z,s)^{\circ}_2$ (with $m = 1$), as in Eq. (1.1.7). The derivative of the Archimedean local Whittaker function $W^*_{j,\infty}(y,s)^{\circ}_2$ at $s = 1/2$ is also calculated explicitly and compared with its geometric counterpart (integral of Green function wedge Chern form on upper half-plane) in Section 19.2.

In fact, Corollary 1.4.2 holds if E_0 is replaced by any elliptic curve over $\overline{\mathbb{Q}}$. Indeed, it is known that the quantity on the left-hand side of Corollary 1.4.2 is constant as E_0 varies over all elliptic curves over $\overline{\mathbb{Q}}$. The case of isogenies of square-free degree appears in [SU99, Proposition 8.1], and the of cyclic isogenies appears in [Aut03, Corollaire 3.3]. Our Corollary 1.4.2 may be extracted from either of those cited results via a combinatorial calculation involving the usual Hecke algebra recurrence relations. The actual proof that we give (as a byproduct of the proof of our main theorem) is different, passing through local calculations on p-divisible groups and formal groups (e.g. as appearing in the calculation of Nakkajima–Taguchi).

Our purely Archimedean result (for arbitrary n and $m^\flat \geq 1$) is the following.

Theorem B (Archimedean arithmetic Siegel–Weil, nonsingular)**.** *Consider any integer $m^\flat$ with $1 \leq m^\flat \leq n$, and consider any $T^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{Q})$ which is nonsingular and not positive definite.*

(3) For any $y^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{R})_{>0}$, we have an equality of real numbers

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T^\flat)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{n-m^\flat}) := \int_{\mathcal{M}_{\mathbb{C}}} g_{T^\flat, y^\flat} \wedge c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{n-m^\flat} = (-1)^{n-m^\flat} C \cdot \frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=s^\flat_0} E^*_{T^\flat}(y^\flat, s)^{\circ}_n \tag{1.4.12}$$

where $s^\flat_0 := (n - m^\flat)/2$. Here $C \in \mathbb{Q}_{>0}$ is the volume constant from Lemma (1), for the Hermitian space V and $v_0 = \infty$ in the notation of loc. cit.. The constant C may depend on n and $m^\flat$ (and F), but does not otherwise depend on $T^\flat$.

This appears (in stronger form) as Theorem 22.1.6. That version applies for all n (even or not) and arbitrary level, as it is a statement about the complex Shimura variety. We gave the weaker version here to avoid more notation. Due to non-properness of $\mathcal{M}_{\mathbb{C}} \to \operatorname{Spec} \mathbb{C}$ for $n > 2$, the corresponding Archimedean Siegel–Weil result of [GS19] does not apply here if $n > 2$.

When $m^\flat = n$, our Theorem B follows from Liu's result [Liu11, Theorem 4.17]. We do not have a new proof of this case. Instead, we deduce our general result from his by a certain limiting argument. This is also our method at non-Archimedean places (replacing Liu's Archimedean results with the non-Archimedean results of Li–Zhang [LZ22a] and Li–Liu [LL22]). Our limiting method will be sketched further in Sections 1.7 and 1.9 below.

1.5 Previous work

We summarize what was previously known on arithmetic Siegel–Weil formulas. These were originally formulated for GSpin Shimura varieties (as opposed to the unitary Shimura varieties considered in Section 1.3); we call these the *orthogonal* and *unitary* cases respectively. In both cases, we write n for the arithmetic dimension of the Shimura varieties (i.e. complex dimension $n-1$).

The problem was initially studied in low-dimensional situations. For quaternionic Shimura curves, the full arithmetic Siegel–Weil formula has been proved in the influential work of Kudla–Rapoport–Yang [KRY04; KRY06]. For modular curves, the formula has been proved in the papers [Yan04; BF06; DY19; SSY23; Zhu23a; Zhu23b].

For Shimura varieties of complex dimension > 1, results on arithmetic Siegel–Weil formulas are currently incomplete. Most the available results concern the case $m = n$ and $\det T \neq 0$; we restrict to this case for the moment. Then $s_0 = 0$ is the central point and the special cycle $\mathcal{Z}(T) \to \mathcal{M}$ is empty in the generic fiber. The arithmetic cycle class $[\widehat{\mathcal{Z}}(T)]$ is thus "purely vertical", i.e. either purely in positive characteristic (non-Archimedean), or with $\mathcal{Z}(T)$ being empty with possibly nontrivial Green current (Archimedean).

The purely Archimedean case (with $\det T \neq 0$ and $s_0 = 0$) was proved by [Liu11; BY21] (unitary and orthogonal, respectively) using different methods. Garcia–Sankaran's Archimedean results apply here as well if the Shimura varieties are compact (more discussion below).

For unitary groups (with $\det T \neq 0$ and $s_0 = 0$), the purely non-Archimedean case for hyperspecial level was first proposed and studied by Kudla–Rapoport [KR11; KR14] at an odd inert prime, where they proved the formula when $\mathcal{Z}(T)$ has dimension 0 (reducing locally to the case $n = 2$). The case $n = 3$ at an odd inert prime was solved by Terstiege [Ter13]. The case of arbitrary n at odd inert primes was solved in the breakthrough work of Li and Zhang [LZ22a] by an inductive "uncertainty principle" strategy. This strategy was later adapted to solve the analogous problem at odd ramified primes [LL22; HLSY23]. We mention that the problem formulation itself needed to be resolved at ramified primes in the presence of bad reduction, and this was done in [HSY23] for the Krämer model. Split primes play a relatively trivial role when $\det T \neq 0$ and $s_0 = 0$. The timeline for non-Archimedean aspects of the GSpin arithmetic Siegel–Weil formula is similar, i.e. results for $\mathcal{Z}(T)$ of dimension 0 were obtained by Kudla–Rapport and Bruinier–Yang [KR99; KR00; BY21], the case $n = 3$ was resolved by Terstiege [Ter11], and the case of general n at hyperspecial level was resolved by Li and Zhang using (a modified version of) their "uncertainty principle" strategy [LZ22b].

We now drop the restrictions $\det T \neq 0$ and $s_0 = 0$. For the purpose of arithmetic theta lifting Eq. (1.3.5), it is desirable to also understand the special cycle classes $[\widehat{\mathcal{Z}}(T)]$ when $\det T = 0$, to fill out the complete arithmetic theta series. Much less is known about this case, which presents new difficulties on both the analytic and geometric sides. It also presents new opportunities: our arithmetic Siegel–Weil result for singular T relates Faltings heights and derivatives of Eisenstein series. Such formulas were observed by Kudla–Rapoport–Yang on Shimura curves [KRY04]; our result applies on unitary Shimura varieties of arbitrarily high dimension. These mixed characteristic phenomena are not visible from arithmetic Siegel–Weil for nonsingular T at the central point $s_0 = 0$ (which was "purely vertical").

We mention known partial results for singular T, besides the previously mentioned work on Shimura curves. There is concrete progress on the case $T = 0$, where the expected geometric side ("arithmetic volumes") has been computed for certain levels in the work of Hörmann and Bruinier–Howard [Hör14; BH21], with some partial results on the comparison with Eisenstein series. In the general case, an important advance was made by Garcia and Sankaran [GS19], who defined Green currents via superconnections and proved a purely Archimedean version of the arithmetic Siegel–Weil formula on compact Shimura varieties (e.g. when T is not positive semi-definite, giving an empty special cycle with possibly nontrivial Green current) via the classical Siegel–Weil formula.

Besides the partial results for $T = 0$, we are not aware of any previous arithmetic Siegel–Weil results which treat non-Archimedean (or combined Archimedean and non-Archimedean) aspects for singular T on Shimura varieties of complex dimension > 1. This is closely related to the following open problem: for Shimura varieties of complex dimension > 1, we are also unaware of any fully global arithmetic Siegel–Weil results (incorporating non-Archimedean places) at a non-central point $s_0 \neq 0$, besides the partial results in [BH21, Theorem C] (there for certain nonzero 1×1 matrices $T \in \mathbb{Z}$). As discussed at the end of Section 1.3, our main theorems make new contributions in both of these directions.

1.6 Non-Archimedean local main theorems

Our proof of Theorems A and B is local in nature. In Section 1.7, we outline the key (new) strategy for proving our main local theorems via our new limit argument. Further discussion at a finer level of detail appears in Section 1.9. In Section 1.8, we outline some of the (new) ideas involved in decomposing our main global theorems into our main local theorems (whose statements and proofs are both new) at every place.

Our non-Archimedean local main theorems relate Kudla–Rapoport local special cycles with *near-center* derivatives of local Whittaker functions.

For Section 1.6, we temporarily switch to local notation, e.g. let F be a finite étale $\mathbb{Q}_p$-algebra of degree 2. In Section 1.6 assume $p \neq 2$ unless $F/\mathbb{Q}_p$ is split (i.e. $F \cong \mathbb{Q}_p \times \mathbb{Q}_p$). If F is a field, we write $\breve{F}$ for the completion of the maximal unramified extension of F. If $F = \mathbb{Q}_p \times \mathbb{Q}_p$, we set $\breve{F} \coloneqq \breve{\mathbb{Q}}_p$ and make a choice of $F \to \breve{F}$.

Given any integer $n \geq 1$, there is an associated *Rapoport–Zink space* $\mathcal{N} \coloneqq \mathcal{N}(n-1,1)$, parameterizing certain polarized p-divisible groups X of height $2n$ and dimension n, with $\mathcal{O}_F$-action of signature $(n-1,1)$ and some additional conditions (Section 5.1). Here, we assume n is even if $F/\mathbb{Q}_p$ is ramified. This Rapoport–Zink space $\mathcal{N}$ is a formal scheme over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$, equidimensional of dimension n. Given framing p-divisible groups $\mathbf{X}_0$ and $\mathbf{X}$ over $\overline{\mathbb{F}}_p$, of dimensions 1 and n respectively, there is a space of *local special quasi-homomorphisms*

$$\mathbf{W} \coloneqq \operatorname{Hom}^0_F(\mathbf{X}_0, \mathbf{X}), \tag{1.6.1}$$

as in Section 5.2. This $\mathbf{W}$ is a non-degenerate $F/\mathbb{Q}_p$ Hermitian space of dimension n (resp. $n-1$) if $F/\mathbb{Q}_p$ is nonsplit (resp. split). Given any integer $m \geq 1$ and any m-tuple $\underline{\mathbf{x}} \in \mathbf{W}^m$, there is an associated *Kudla–Rapoport local special cycle* $\mathcal{Z}(\underline{\mathbf{x}}) \hookrightarrow \mathcal{N}$, which is the (closed) locus where every $\mathbf{x} \in \underline{\mathbf{x}}$ deforms to a homomorphism of p-divisible groups (Section 5.2).

The "expected codimension" of $\mathcal{Z}(\underline{\mathbf{x}}) \hookrightarrow \mathcal{N}$ is m, so we consider *derived local special cycle (classes)* ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}) \in \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{\mathbf{x}}))_{\mathbb{Q}} \coloneqq \mathrm{gr}_{n-m} K'_0(\mathcal{Z}(\underline{\mathbf{x}}))_{\mathbb{Q}}$ in codimension m (Section 5.5). If all elements of the tuple $\underline{\mathbf{x}} = [\mathbf{x}_1, \ldots, \mathbf{x}_m]$ are nonzero, then we have

$${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}) \coloneqq \mathcal{O}_{\mathcal{Z}(\mathbf{x}_1)} \otimes^{\mathbb{L}} \cdots \otimes^{\mathbb{L}} \mathcal{O}_{\mathcal{Z}(\mathbf{x}_m)}. \tag{1.6.2}$$

For any $x_i = 0$, it is usual to make a similar definition by replacing $\mathcal{O}_{Z(x_i)}$ with $[\mathcal{O}_{\mathcal{N}}] - [\mathcal{E}]$ for a certain *tautological bundle* $\mathcal{E}$ on $\mathcal{N}$ (Definition 5.1.9). For any proper closed subscheme $Z \subseteq \mathcal{N}$, there is a degree map $\deg_{\overline{\mathbb{F}}_p} \colon \mathrm{gr}^n_{\mathcal{N}} K'_0(Z)_{\mathbb{Q}} \to \mathbb{Q}$ given by the composition

$$\mathrm{gr}_0 K'_0(Z)_{\mathbb{Q}} \xrightarrow{\sim} \mathrm{gr}_0 K'_0(Z_{\overline{\mathbb{F}}_p})_{\mathbb{Q}} \to \mathrm{gr}_0 K'_0(\operatorname{Spec} \overline{\mathbb{F}}_p)_{\mathbb{Q}} = \mathbb{Q} \tag{1.6.3}$$

where the first arrow is induced by the dévissage pushforward isomorphism $K_0'(Z_{\overline{\mathbb{F}}_p}) \to K_0'(Z)$ and the second arrow is pushforward along $Z_{\overline{\mathbb{F}}_p} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ (e.g. induced by taking Euler characteristics of coherent sheaves on $Z_{\overline{\mathbb{F}}_p}$).

We are interested in "height contributions" for some local 1-cycle classes, namely ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)$ for $\underline{\mathbf{x}}^\flat \in \mathbf{W}^{n-1}$ with nonsingular Gram matrix. For this, write $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} \subseteq \mathcal{Z}(\underline{\mathbf{x}}^\flat)$ for the largest closed formal subscheme which is flat over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ ("horizontal") and write ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}} \in \operatorname{gr}^m_{\mathcal{N}} K_0'(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\overline{\mathbb{F}}_p})_{\mathbb{Q}}$ for the projection of ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)$ to the special fiber ("vertical") in the sense of Section 5.5.

The "vertical height contribution" will simply be the algebro-geometric intersection number

$$\deg_{\overline{\mathbb{F}}_p}({}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}} \cdot \mathcal{E}^\vee) \cdot \log p. \tag{1.6.4}$$

The "horizontal height contribution" is more delicate. The horizontal part $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$ is adic finite flat over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$, and we define a "horizontal height contribution" for each component $\mathcal{Z} \hookrightarrow \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$ of (the scheme associated to) $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$. When $n \geq 2$, the universal p-divisible group over each $\mathcal{Z}$ decomposes (Section 7.3) as $(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_F) \times \mathfrak{X}_0^{n-2}$ where $\mathfrak{X}_s$ (resp. $\mathfrak{X}_0$) is a height 2 dimension 1 *quasi-canonical lifting* of level $s \in \mathbb{Z}_{\geq 0}$ (resp. canonical lifting) in the sense of Gross [Gro86], and where $(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_F)$ denotes the *Serre tensor* construction (as recalled in Eq. (B.1.1)). The integer s may depend on $\mathcal{Z}$, so we write $s_{\mathcal{Z}}$ instead of s in Theorem C below.

The $\mathcal{O}_F$-module $\operatorname{Hom}(\mathfrak{X}_0, \mathfrak{X}_s)$ (by pre-composition) is free of rank one, with $\psi_s \in \operatorname{Hom}(\mathfrak{X}_0, \mathfrak{X}_s)$ being a generator if and only if ψ_s is an isogeny of degree p^s ("minimal degree"). We then define the "*local change of tautological height*"

$$\breve{\delta}_{\mathrm{tau}}(\psi_s) := \delta_{\mathrm{tau}}(s) \cdot \log p := \left(-\frac{1}{2} v_p(\deg \psi_s) + \frac{1}{[\breve{E}_s : \breve{\mathbb{Q}}_p]} \operatorname{length}_{\mathcal{O}_{\breve{E}_s}} (e^* \Omega^1_{\ker \psi_s / \operatorname{Spec} \mathcal{O}_{\breve{E}_s}}) \right) \log p \tag{1.6.5}$$

where $\breve{E}_s$ is the "field of definition" for $\mathfrak{X}_s$ (Section 7.2) and e^* denotes pullback along the identity section. We have $\delta_{\mathrm{tau}}(s) \in \mathbb{Q}$; this quantity only depends on $s \in \mathbb{Z}_{\geq 0}$, the prime p, and whether $F/\mathbb{Q}_p$ is inert/split/ramified. Then, we take the "horizontal height contribution" for $\mathcal{Z}$ to be

$$\deg(\mathcal{Z}) \cdot \delta_{\mathrm{tau}}(\mathcal{Z}) \cdot \log p \tag{1.6.6}$$

where $\deg(\mathcal{Z})$ is the degree of the adic finite flat morphism $\mathcal{Z} \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$.

The formula in (1.6.5) is reminiscent of the usual (global) formula for change of (stable) Faltings heights along an isogeny of abelian varieties. Unlike the global formula, which does not depend on the choice of isogeny (Faltings height is intrinsic), the quantity $\breve{\delta}_{\mathrm{tau}}(\psi_s)$ may change in general if ψ_s is not chosen to be an isogeny of minimal degree. For this reason, it is somewhat involved for us to verify that the "local change" $\breve{\delta}_{\mathrm{tau}}(\psi_s)$ for minimal degree ψ_s will indeed compute the "correct" local contribution to a global height, for our eventual global applications. This subtlety is treated in Part III. We also remark that a similar formulation for "local change of height" also appears for different special cycles in our forthcoming joint work [CLZ] on subleading terms of adjoint L-functions. There, such "local change of heights" appear in our formulation of new "near-center arithmetic fundamental lemmas" for near-center derivatives of orbital integrals.

On the analytic side, for integers $m \geq 0$, we consider a quasi-split unitary group $H \coloneqq U(m, m)$ for an $F/\mathbb{Q}_p$ Hermitian space of dimension $2m$ (Section 13). For standard sections $\Phi \in I(s, \chi)$ of a certain (local) degenerate principal series for the Siegel parabolic in H, we have *local Whittaker functions* $W_T(h, s, \Phi)$ associated to Hermitian matrices $T \in \mathrm{Herm}_m(\mathbb{Q}_p)$, with $h \in H(\mathbb{Q}_p)$ and $s \in \mathbb{C}$ (Section 15.3). These local Whittaker functions appear in Fourier coefficients of Eisenstein series. For the normalized spherical standard section Φ°, we then consider a *normalized spherical local Whittaker function*

$$W_T^*(s)_n^\circ \coloneqq \Lambda_T(s)_n^\circ W_T(1, s, \Phi^\circ) \tag{1.6.7}$$

for a *local normalizing factor* $\Lambda_T(s)_n^\circ$ (consisting of local L-factors, etc.) which produces a symmetric functional equation $s \leftrightarrow -s$ (Section 15.3) for nonsingular T.

With setup as above, our non-Archimedean "local arithmetic Siegel–Weil" theorem is the following.

Theorem C. *For any nonsingular $T^\flat \in \mathrm{Herm}_{n-1}(\mathbb{Q}_p)$ and any $\underline{\mathbf{x}}^\flat \in \mathbf{W}^{n-1}$ with Gram matrix $T^\flat$, we have*

$$-[\breve{F} : \breve{\mathbb{Q}}_p] \frac{d}{ds}\bigg|_{s=1/2} W_{T^\flat}^*(s)_n^\circ = \left(\frac{2}{[\breve{F} : \breve{\mathbb{Q}}_p]} \deg_{\overline{\mathbb{F}}_p}(\mathcal{E}^\vee \cdot {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}) + 2 \sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}} \deg(\mathcal{Z}) \cdot \delta_{\mathrm{tau}}(s_{\mathcal{Z}}) \right) \cdot \log p. \tag{1.6.8}$$

This appears below as Theorem 18.1.2, there stated in terms of local densities. When $n = 1$, both sides of (1.6.8) are zero.

We highlight the fact that $s = 1/2$ is *not* the central point of the functional equation for the local Whittaker function $W_{T^\flat}^*(s)_n^\circ$, so Theorem C concerns a *near-center* derivative. Moreover, the value of $W_{T^\flat}^*(s)_n^\circ$ at $s = 1/2$ (generally nonzero) *simultaneously* has geometric meaning: it equals $[\breve{F} : \breve{\mathbb{Q}}_p] \deg \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$, where $\deg \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$ denotes the degree of the adic finite flat morphism $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} \to \mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ (Lemma 18.1.3). This "geometric local Siegel–Weil" formula for the value at $s = 1/2$ was previously observed by Li–Zhang when $F/\mathbb{Q}_p$ is inert [LZ22a, Corollary 4.6.1]. Our Theorem C thus concerns a *subleading* term in the power series expansion at $s = 1/2$. These phenomena will reappear globally (Remark 1.4.1).

For a more concrete simple case, see Example 1.7.1 in the next section. We return to global notation below, e.g. F is an imaginary quadratic (global) field.

1.7 Strategy: overview

We now describe our key local strategy: "take a limit" (Figure 1).

In Figure 1 below, for a given place v of $\mathbb{Q}$, we consider $T^\flat \in \mathrm{Herm}_{n-1}(\mathbb{Q}_v)$ with $\det T^\flat \neq 0$, and $T = \mathrm{diag}(t, T^\flat)$ for suitable nonzero $t \in \mathbb{Q}_v$. On the left, the limit refers to $t \to 0$ in the v-adic topology (meaning the real topology if $v = \infty$). The upper horizontal arrow should be understood as a local version of Theorem A(2), and the lower horizontal arrow should be understood as the (known) local version of (1.4.9) when $\det T \neq 0$.

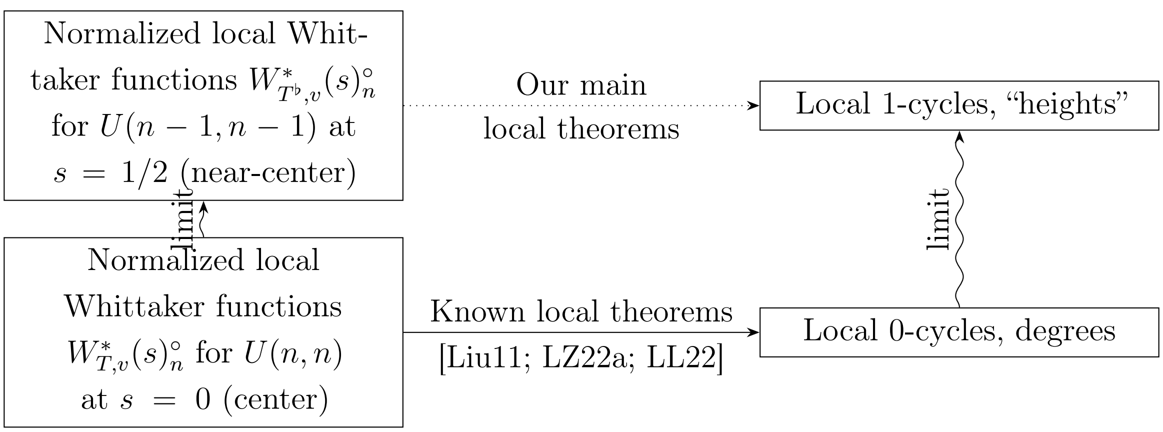

Figure 1: A local limiting method

This limiting method is the main conceptual novelty in our work, and is the key idea driving our main results. In Figure 1, the left vertical arrow and upper horizontal arrow are new in this work. In the right vertical arrow, the relation between limits and Faltings heights is also new in this work.

It is striking that the limiting method plays a similar role at all places, Archimedean and non-Archimedean. In the purely Archimedean case, i.e. when $v = \infty$ with $T^\flat$ nonsingular and not positive definite, we are able to run our limiting argument for special cycles (currents) in arbitrary dimension. This is why our purely Archimedean result (Theorem B) applies in arbitrary codimension.

For non-Archimedean places, it is interesting to ask whether the limiting method in Figure 1 can be adapted to the case of higher dimensional special cycles (corresponding to $T^\flat$ of smaller rank). Some key difficulties are mentioned in Remark 1.8.1.

We also mention a slight difference if $v = p$ is a prime split in $\mathcal{O}_F$. The known local theorems [Liu11; LZ22a; LL22] apply in the Archimedean, inert, and ramified cases respectively. In the split case, the lower left corner of Figure 1 will involve the special value $W^*_{T,v}(0)^\circ_n$ (while the derivative at $s = 0$ appears in the Archimedean, inert, and ramified cases). In the split case, the "known local theorem" in Figure 1 refers to a certain vanishing statement for a certain contribution to $W^*_{T,v}(0)^\circ_n$ ("vertical part" via an analogue of Cho–Yamauchi's formula; the vanishing is proved in Lemma 18.5.1) and emptiness of local 0-cycles.

In the next example, we sketch the strategy from Fig. 1 in a simple non-Archimedean case. The reader seeking a more detailed sketch of a more general setup (along with a comparison between the Archimedean and non-Archimedean strategies) may refer to Section 1.9.

Example 1.7.1. Suppose p is a prime which is inert in $\mathcal{O}_F$. Take $T^\flat$ and T as in Fig. 1, and assume t is such that T defines a non-split Hermitian space. We prove a limiting formula (inert case)

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,p}(s)^\circ_n = \lim_{t\to 0}\left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_n + (\log|t|_p - \log p)W^*_{T^\flat,p}(-1/2)^\circ_n\right), \qquad (1.7.1)$$

which appears in the text as Proposition 18.5.2 (there stated via local densities). This is the left vertical arrow in Fig. 1. Here $|-|_p$ is the usual p-adic norm.

The right vertical arrow in Fig. 1 asserts that the analytic limit formula in Eq. (1.7.1) has a geometric interpretation in terms of the local special cycles $\mathcal{Z}(\underline{\mathbf{x}})$.

To illustrate a relatively simple case, consider the case $n = 2$ (and $p \neq 2$). We then have $\mathcal{N} \cong \operatorname{Spf} \breve{\mathbb{Z}}_p[\![u]\!]$ (non-canonically). Consider

$$T^\flat = \begin{pmatrix} p \end{pmatrix} \qquad T = \operatorname{diag}(t, T^\flat) \qquad t = p^e \qquad (1.7.2)$$

for even integers $e \in \mathbb{Z}_{\geq 0}$, where $T^\flat$ is a 1×1 matrix. We have

$$W^*_{T^\flat,p}(s)^\circ_2 = p^{s+1/2} + p^{-s+1/2} \qquad W^*_{T,p}(s)^\circ_2 = p^{(e+1)s} - p^{-(e+1)s} + (1-p)p^{(e+1)s}\sum_{i=1}^{e}(-q^{-2s})^i.$$

Set $\mathcal{O}_{F_p} := \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$. With $\mathbf{V}_p$ denoting $\mathbf{W}$ from Section 1.6, let $\underline{\mathbf{x}} = [\mathbf{x}, \mathbf{x}^\flat] \in \mathbf{V}_p^n$ be any tuple with Gram matrix $(\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T$. Then $\mathcal{Z}(\mathbf{x}^\flat) \cong \operatorname{Spf} \mathcal{O}_{\breve{E}}$ for a certain degree $p+1$ extension $\breve{E}/\breve{\mathbb{Q}}_p$. Over $\mathcal{Z}(\mathbf{x}^\flat)$, the universal p-divisible group from $\mathcal{N}$ pulls back to $\mathfrak{X}_1 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$ (Serre tensor construction) where $\mathfrak{X}_1 \to \operatorname{Spf} \mathcal{O}_{\breve{E}}$ is a *quasi-canonical lifting* of level 1 in the sense of Gross [Gro86]. If $\mathfrak{X}_0$ denotes the canonical lifting, there is an isogeny $\psi_1 \colon \mathfrak{X}_0 \to \mathfrak{X}_1$ of minimal degree (unique up to $\mathcal{O}_{F_p}^\times$), with $\deg \psi_1 = p$. We have

$$\breve{\delta}_{\mathrm{Fal}}(\psi_1) \coloneqq \frac{1}{2}\log(\deg \psi_1) - \frac{1}{[\breve{E} : \breve{\mathbb{Q}}_p]}\mathrm{length}_{\mathcal{O}_{\breve{E}}}(e^* \Omega^1_{\ker \psi_1 / \mathcal{O}_{\breve{E}}}) \log p \tag{1.7.3}$$

$$= \left(\frac{1}{2} - \frac{1}{p+1}\right) \log p \tag{1.7.4}$$

where the second equality follows from a computation of Nakkajima–Taguchi [NT91] (with e^* denoting pullback along the identity section). We call the left-hand side a "local change of Faltings height"; its relation with (global) Faltings height is sketched in Section 1.8.

In this case, our geometric analogue of Eq. (1.7.1) is the limit formula

$$-\breve{\delta}_{\mathrm{Fal}}(\psi_1) \cdot \deg \mathcal{Z}(\mathbf{x}^\flat) = \lim_{\mathbf{x} \to 0} \left((\deg_{\overline{\mathbb{F}}_p}(\mathcal{O}_{\mathcal{Z}(\mathbf{x})} \otimes^{\mathbb{L}} \mathcal{O}_{\mathcal{Z}(\mathbf{x}^\flat)})) \cdot \log p + \frac{1}{2}(\log |t|_p - \log p) \cdot \deg \mathcal{Z}(\mathbf{x}^\flat) \right) \tag{1.7.5}$$

where the limit $t \to 0$ is p-adic. We have

$$\frac{1}{2}\frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_2 \overset{\text{[LZ22a]}}{=} (\deg_{\overline{\mathbb{F}}_p} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}})) \cdot \log p \qquad W^*_{T^\flat,p}(-1/2)^\circ_2 \overset{\text{[LZ22a]}}{=} \deg \mathcal{Z}(\mathbf{x}^\flat). \tag{1.7.6}$$

These may be thought of as (nonsingular, central point) "local arithmetic Siegel–Weil" and (nonsingular, near-central) "local geometric Siegel–Weil" formulas, respectively.

Applying our limiting formulas in Eqs. (1.7.1) and (1.7.5) (along with the local functional equation $W^*_{T^\flat,p}(s)^\circ_2 = W^*_{T^\flat,p}(-s)^\circ_2$) produces the formula

$$\frac{d}{ds}\bigg|_{s=1/2} W^*_{T^\flat,p}(s)^\circ_2 = 2\breve{\delta}_{\mathrm{Fal}}(\psi_1) \cdot \deg \mathcal{Z}(\mathbf{x}^\flat). \tag{1.7.7}$$

This is one form of our main local theorem in this simple special case (compare the global version, Theorem A(2), and the more general local formulation in Section 1.6).

In extremely impressionistic terms, the formula in Eq. (1.7.5) states that the "limit" of the special divisor $\lim_{\mathbf{x} \to 0} \mathcal{Z}(\mathbf{x})$ "converges" to some "local part" of a tautological bundle (essentially the Hodge bundle) when intersecting against $\mathcal{Z}(\mathbf{x}^\flat)$, after regularizing. We remark that the analogous "numerical limit" statement (without regularizing) is literally true if $\mathcal{Z}(\mathbf{x}^\flat)$ is replaced by any (proper) curve in the special fiber over $\mathcal{N}$. That case has a conceptual explanation: Grothendieck–Messing deformation theory. This discussion is continued in Eq. (1.9.13) and the text below in loc. cit..

1.8 Local-to-global

For the global-to-local reduction process, we use complex uniformization (Archimedean place) and Rapoport–Zink uniformization (non-Archimedean places). Unlike the previously known case $\det T \neq 0$ for $T \in \mathrm{Herm}_n(\mathbb{Q})$ (giving a purely vertical arithmetic special cycle class), we have a new mixed characteristic "horizontal" contribution in the Faltings height. While the Faltings height decomposes locally *after picking a section* of the metrized Hodge bundle, it does not admit an obvious *canonical* local decomposition. Such a canonical decomposition seems necessary for the comparison with local Whittaker functions (as appearing in Eisenstein series Fourier coefficients), which presumably does not retain information on which section was picked.

Instead, we do have a decomposition for the *difference* between (stable) Faltings heights of any two abelian varieties A_1, A_2 (over a number field E) in a fixed isogeny class. We may assume A_1 and A_2 have everywhere semi-abelian reduction after extending E. The difference of Faltings heights is then

$$h_{\mathrm{Fal}}(A_2) - h_{\mathrm{Fal}}(A_1) = \sum_{p \mid \deg \phi} a_p \log p = \frac{1}{[E : \mathbb{Q}]} \sum_p \sum_{\breve{w} \mid p} [\breve{E}_{\breve{w}} : \breve{\mathbb{Q}}_p] \breve{\delta}_{\mathrm{Fal}}(\phi_{\breve{w}}) \tag{1.8.1}$$

$$\breve{\delta}_{\mathrm{Fal}}(\phi_{\breve{w}}) := \frac{1}{2} \log(\deg \phi_{\breve{w}}) - \frac{1}{[\breve{E}_{\breve{w}} : \breve{\mathbb{Q}}_p]} \mathrm{length}_{\mathcal{O}_{\breve{E}_{\breve{w}}}} (e^* \Omega^1_{\ker \phi / \mathcal{O}_{\breve{E}_{\breve{w}}}}) \log p \tag{1.8.2}$$

for some $a_p \in \mathbb{Q}$ and any choice of isogeny $\phi \colon A_1 \to A_2$ (extend ϕ over $\mathcal{O}_E$). Here, the symbol e means the identity section, the inner sum in Eq. (1.8.1) runs over all prime ideals $\breve{w}$ of $E \otimes_{\mathbb{Q}} \breve{\mathbb{Q}}_p$ (with associated residue field $\breve{E}_{\breve{w}}$), and $\phi_{\breve{w}}$ means the associated isogeny of p-divisible groups $A_1[p^\infty] \to A_2[p^\infty]$ base-changed to $\mathrm{Spf}\, \mathcal{O}_{\breve{E}_{\breve{w}}}$.

The coefficients a_p do not depend on the choice of ϕ, by linear independence of $\log p$ for different p. This is also the reason why a difference of Faltings heights appears in Corollary 1.4.2. We then argue that these numbers $a_p \in \mathbb{Q}$ (averaged over the special cycle) can be calculated in a purely local way, in terms of local special cycles on Rapoport–Zink spaces. The argument we give is somewhat delicate, as we wish to avoid writing down explicit (global) isogenies $\phi \colon A_1 \to A_2$, so that we obtain a more local formulation. The quantities $\delta_{\mathrm{Fal}}(\phi_{\breve{w}})$ are amenable to (local) calculation on Rapoport–Zink spaces, but may depend on the chosen isogeny of p-divisible groups $\phi_{\breve{w}}$. We show that certain "minimal degree" isogenies $\phi_{\breve{w}}$ lift to global isogenies of p-divisible groups (Section 10.2) and can be used to calculate local contributions to (differences of) Faltings height.

The global-to-local reduction process from $h_{\mathrm{Fal}}(-)$ (Faltings heights) to $\delta_{\mathrm{Fal}}(\phi_{\breve{w}})$ (local quantities calculate-able on Rapoport–Zink spaces) for "minimal degree" $\phi_{\breve{w}}$ is the content of

Sections 9 and 10. The relation with global special cycles via uniformization is explained in Part IV. There is also the issue that the tautological bundle $\widehat{\mathcal{E}}^\vee$ is not the same as the metrized Hodge bundle (but is known to behave similarly, as first observed by Gross [Gro78] and studied further in [BHKRY20II]), so the (more natural) version with "tautological height" needs additional argument. The "tautological height" and Faltings height are treated in parallel in Sections 9 and 10.

While previous work for special points on Shimura curves [KRY04] also studied the change in Faltings heights along isogenies, our insistence on a purely local formulation is an important difference for our method. We only observe the limiting phenomena in Figure 1 (below) on a local level; this is what allows us to prove a theorem on Shimura varieties of arbitrarily large dimension.

Remark 1.8.1. After our work, the arithmetic Siegel–Weil formula in (1.4.9) remains open only for T with $\mathrm{rank}(T) \leq n-2$ (corresponding to special cycles $\mathcal{Z}(T)$ of dimension ≥ 1 in the generic fiber, if nonempty). We mention some of the difficulties for these higher dimensional special cycles, from our perspective.

Our strategy in Figure 1 strongly emphasizes *local* limits on both the analytic and geometric sides. On the geometric side at non-Archimedean places, this is possible in the rank $\geq n-1$ case because e.g. the special cycles are contained inside supersingular loci at all nonsplit primes. This allows us to describe global special cycles in terms of local special cycles on a single Rapoport–Zink space (at each prime).

Higher dimensional special cycles may pass through several strata, so a single Rapoport–Zink space should not be enough to capture all information. We are not sure whether it is sensible to "piece together" the intersection-theoretic information coming from several Rapoport–Zink spaces. Moreover, the corresponding local special cycles on Rapoport–Zink spaces may be non-proper (so it is unclear how to extract local intersection numbers in the more general situation).

For our method, it is also important to locally decompose Faltings heights in a canonical way. In the rank $n-1$ case, this was accomplished using the "change of Faltings height along an isogeny" formula. To generalize our limiting strategy to higher dimensional cycles, we may need a similar canonical local decomposition involving heights of higher dimensional cycles on the Shimura variety.

1.9 Strategy: local limit

For illustration purposes, we sketch the local limiting strategy described in Section 1.7 (particularly Fig. 1) at a finer level of detail. In Section 1.9, we let $F/\mathbb{Q}$ be an imaginary quadratic field, and allow n to be even or odd. We sketch the case where $v = \infty$ (Archimedean) and where $v = p$ is an odd prime inert in $\mathcal{O}_F$. In the main text, the inert/ramified/split cases are treated in parallel (Section 18). We hope that the similarities between the Archimedean and non-Archimedean cases are visible from the sketches below.

Case $v = \infty$. For purposes of exposition, we consider $T^\flat \in \mathrm{Herm}_n(\mathbb{R})$ and $t \in \mathbb{R}_{<0}$. Set $T \coloneqq \mathrm{diag}(t, T^\flat)$. Let V be any signature $(n-1,1)$ non-degenerate $F/\mathbb{Q}$ Hermitian space with pairing denoted $(-,-)$ (for any $n \geq 1$).

We prove the limiting identity (left vertical arrow in Figure 1)

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,\infty}(s)^\circ_n = \lim_{t\to 0^-}\left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n + (\log|t|_\infty + \log(4\pi e^\gamma))W^*_{T^\flat,\infty}(-1/2)^\circ_n\right) \tag{1.9.1}$$

where γ is the Euler–Mascheroni constant, and $|-|_\infty$ denotes the usual real norm. This formula appears below (more generally) as Proposition 19.1.2. The proof of this limiting formula is the bulk of the work at the Archimedean place. Note the similarity with the non-Archimedean version Eq. (1.9.8) (see also Section 15.6 for more comparisons).

On the geometric side of Figure 1, the local 0-cycles (resp. 1-cycles) should be interpreted as Green currents of top degree $(n-1,n-1)$ (resp. degree $(n-2,n-2)$) on the associated Hermitian symmetric domain $\mathcal{D}$ parameterizing maximal negative definite $\mathbb{C}$-linear subspaces of $V_\mathbb{R}$. Consider the signature $(n-1,1)$ complex Hermitian space $V_\mathbb{R}$, with Hermitian pairing $(-,-)$. Any tuple $\underline{x} \in V_\mathbb{R}$ with nonsingular Gram matrix has an associated Kudla Green current $[\xi(\underline{x})]$, studied by Liu [Liu11] in the unitary case. There is a certain local special cycle $\mathcal{D}(\underline{x}) \subseteq \mathcal{D}$ (complex submanifold which is the locus of $\mathbb{C}$-lines $z \in \mathcal{D}$ which are perpendicular to all elements of the tuple $\underline{x}$), arising in the complex uniformization of global special cycles.

Let $\underline{x}^\flat = [x^\flat_1, \ldots, x^\flat_{n-1}] \in V^{n-1}_\mathbb{R}$ be a tuple with Gram matrix $T^\flat$ and consider nonzero $x \in \mathrm{span}_\mathbb{C}(\underline{x}^\flat)^\perp \subseteq V_\mathbb{R}$ in the orthogonal complement. Set

$$\underline{x} = [x, x^\flat_1, \ldots, x^\flat_{n-1}] \qquad t \coloneqq (x,x) \qquad T \coloneqq \mathrm{diag}(t, T^\flat). \tag{1.9.2}$$

Liu's Archimedean local theorem [Liu11, Theorem 4.1.7] implies

$$\int_\mathcal{D} [\xi(\underline{x})] = \frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n. \tag{1.9.3}$$

We are using the star product construction of $[\xi(\underline{x})]$, which unfolds as

$$[\xi(\underline{x})] = [\xi(x)] * [\xi(\underline{x}^\flat)] = \omega(x) \wedge [\xi(\underline{x}^\flat)] + [\xi(x)] \wedge \delta_{\mathcal{D}(\underline{x}^\flat)} \tag{1.9.4}$$

where $\omega(x)$ is a $(1,1)$-form associated with x (Kudla–Millson form up to a normalization), $\delta_{\mathcal{D}(\underline{x}^\flat)}$ is a Dirac delta current, and $\xi(x)$ is a certain function on $\mathcal{D}$ with logarithmic singularity along $\mathcal{D}(x)$. The function $\xi(x)$ is expressed in terms of the exponential integral Ei. We have $\int_{\mathcal{D}} [\xi(x)] \wedge \delta_{\mathcal{D}(\underline{x}^\flat)} = -\operatorname{Ei}(4\pi t)$ and the limit formulas

$$\lim_{x \to 0} \omega(x) = c_1(\widehat{\mathcal{E}}^\vee) \qquad \lim_{u \to 0^-} (\operatorname{Ei}(u) - \log|u|) = \gamma \tag{1.9.5}$$

where $c_1(\widehat{\mathcal{E}}^\vee)$ denotes the Chern form of dual tautological bundle on $\mathcal{D}$ (as in Section 8). Under the assumption that $T^\flat$ is positive definite, we have $W^*_{T^\flat,\infty}(-1/2)^\circ_n = \deg \mathcal{D}(\underline{x}^\flat) = 1$ ("local geometric Siegel–Weil", i.e. $\mathcal{D}(\underline{x}^\flat)$ is a single point). We thus have

$$\int_{\mathcal{D}} c_1(\widehat{\mathcal{E}}^\vee) \wedge [\xi(\underline{x}^\flat)] = \lim_{x \to 0} \left(\left(\int_{\mathcal{D}} [\xi(\underline{x})] \right) + \log|t| + \log(4\pi e^\gamma) \right). \tag{1.9.6}$$

Using the functional equation $W^*_{T^\flat,\infty}(s)^\circ_n = W^*_{T^\flat,\infty}(-s)^\circ_n$, the limit formula in (1.9.1) now implies the following theorem.

Theorem (Archimedean local version of Theorem A)**.** *We have*

$$-\frac{d}{ds}\bigg|_{s=1/2} W^*_{T^\flat,\infty}(s)^\circ_n = \int_{\mathcal{D}} c_1(\widehat{\mathcal{E}}^\vee) \wedge [\xi(\underline{x}^\flat)]. \tag{1.9.7}$$

This is our main local Archimedean theorem for positive definite $T^\flat$ (i.e. the dotted arrow in Figure 1). This appears below as Theorem 19.1.1, which also includes a version for non-positive definite $T^\flat$. Limiting on the geometric side of Figure 1 was provided by a limiting property of the (normalized) Kudla–Millson form, i.e. $\omega(x) \to c_1(\widehat{\mathcal{E}}^\vee)$ as $x \to 0$. To compare with the limit of Whittaker function derivatives, we used the special value formula $W^*_{T^\flat,\infty}(-1/2)^\circ_n = 1$ ("local geometric Siegel–Weil") and the asymptotics of Ei.

Case $v = p$ is an odd prime inert in $\mathcal{O}_F$. For our non-Archimedean main local theorems, we run an argument similar (in spirit) to the Archimedean case where the star product of Green currents is replaced by a derived tensor product of complexes of coherent sheaves on Rapoport–Zink spaces.

Suppose $T^\flat \in \operatorname{Herm}_{n-1}(\mathbb{Q}_p)$ with $\det T^\flat \neq 0$, and consider $t \in \mathbb{Q}_p$ such that $T := \operatorname{diag}(t, T^\flat)$ has $\varepsilon(T) = -1$ (i.e. T defines a nonsplit Hermitian space; see Section 2.2). For the normalized local Whittaker functions defined in Section 15 below, we will prove the limit formula

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,p}(s)^\circ_n = \lim_{t \to 0} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_n + (\log|t|_p - \log p) W^*_{T^\flat,p}(-1/2)^\circ_n \right). \tag{1.9.8}$$

This appears as Proposition 18.5.2 below. Note the similarity with the Archimedean version Eq. (1.9.1) (see also Section 15.6 for comparisons with the ramified and split versions).

We set $\mathcal{N} := \mathcal{N}(n-1,1)$ for the Rapoport–Zink space $\mathcal{N}(n-1,1)$ from Section 5. We also use the notation on special cycles from loc. cit., e.g. $\mathbf{V}$ is the space of local special quasi-homomorphisms. (These notations were briefly sketched in Section 1.6.)

If $\underline{\mathbf{x}}$ is a basis for $\mathbf{V}$, then $\mathcal{Z}(\underline{\mathbf{x}})$ is a scheme with structure morphism $\mathcal{Z}(\underline{\mathbf{x}}) \to \operatorname{Spf} \breve{\mathbb{Z}}_p$ which is adic and proper [LZ22a, Lemma 2.10.1]. In this case, $\mathcal{Z}(\underline{\mathbf{x}})$ is thus a finite order thickening of its special fiber $\mathcal{Z}(\underline{\mathbf{x}})_{\overline{\mathbb{F}}_p}$, and there is a degree map $\deg_{\overline{\mathbb{F}}_p} \colon \operatorname{gr}_0 K_0'(\mathcal{Z}(\underline{\mathbf{x}}))_{\mathbb{Q}} \to \mathbb{Q}$ given by the composite

$$\operatorname{gr}_0 K_0'(\mathcal{Z}(\underline{\mathbf{x}}))_{\mathbb{Q}} \xrightarrow{\sim} \operatorname{gr}_0 K_0'(\mathcal{Z}(\underline{\mathbf{x}})_{\overline{\mathbb{F}}_p})_{\mathbb{Q}} \to \operatorname{gr}_0 K_0'(\operatorname{Spec} \overline{\mathbb{F}}_p)_{\mathbb{Q}} = \mathbb{Q} \tag{1.9.9}$$

where the first arrow is induced by the dévissage pushforward isomorphism $K_0'(\mathcal{Z}(\underline{\mathbf{x}})_{\overline{\mathbb{F}}_p}) \to K_0'(\mathcal{Z}(\underline{\mathbf{x}}))$ and the second arrow is pushforward along $\mathcal{Z}(\underline{\mathbf{x}})_{\overline{\mathbb{F}}_p} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ (e.g. induced by taking Euler characteristics of coherent sheaves on $\mathcal{Z}(\underline{\mathbf{x}})_{\overline{\mathbb{F}}_p}$).

Let $\underline{\mathbf{x}}^\flat = [\mathbf{x}_1^\flat, \ldots, \mathbf{x}_{n-1}^\flat] \in \mathbf{V}^{n-1}$ be a tuple with Gram matrix $T^\flat$ and consider nonzero $\mathbf{x} \in \operatorname{span}_{F_p}(\underline{\mathbf{x}}^\flat)^\perp \subseteq \mathbf{V}$ in the orthogonal complement. Set

$$\underline{\mathbf{x}} = [\mathbf{x}, \mathbf{x}_1^\flat, \ldots, \mathbf{x}_{n-1}^\flat] \qquad t := (\mathbf{x}, \mathbf{x}) \qquad T := \operatorname{diag}(t, T^\flat). \tag{1.9.10}$$

Li–Zhang's inert Kudla–Rapoport theorem [LZ22a, Theorem 1.2.1] implies

$$(\deg_{\overline{\mathbb{F}}_p} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}})) \cdot \log p = \frac{1}{2} \frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_n. \tag{1.9.11}$$

As an element of $\operatorname{gr}^n_{\mathcal{N}} K_0'(\mathcal{Z}(\underline{\mathbf{x}}))_{\mathbb{Q}}$, the derived tensor product unfolds as

$${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}) = [{}^{\mathbb{L}}\mathcal{Z}(\mathbf{x}) \otimes^{\mathbb{L}}_{\mathcal{O}_{\mathcal{N}}} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)] = [{}^{\mathbb{L}}\mathcal{Z}(\mathbf{x}) \otimes^{\mathbb{L}}_{\mathcal{O}_{\mathcal{N}}} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}] + [{}^{\mathbb{L}}\mathcal{Z}(\mathbf{x}) \otimes^{\mathbb{L}}_{\mathcal{O}_{\mathcal{N}}} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}]. \tag{1.9.12}$$

Here ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} = [\mathcal{O}_{\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}}] \in \operatorname{gr}^{n-1}_{\mathcal{N}} K_0'(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}})_{\mathbb{Q}}$ is the "horizontal part" of ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)$, with $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} \subseteq \mathcal{Z}(\underline{\mathbf{x}}^\flat)$ denoting the flat part, and ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}} \in \operatorname{gr}^{n-1}_{\mathcal{N}} K_0'(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\overline{\mathbb{F}}_p})_{\mathbb{Q}}$ is the "vertical part" of ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)$ (see [LZ22a, §5.2]; we are using the dévissage pushforward isomorphism $K_0'(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\overline{\mathbb{F}}_p}) \to K_0'(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}})$ where $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}$ is the vertical part from loc. cit.).

We then prove the limit formulas

$$\lim_{\mathbf{x} \to 0} (\deg_{\overline{\mathbb{F}}_p} [{}^{\mathbb{L}}\mathcal{Z}(\mathbf{x}) \otimes^{\mathbb{L}}_{\mathcal{O}_{\mathcal{N}}} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}]) = \deg_{\overline{\mathbb{F}}_p}(\mathcal{E}^\vee \cdot {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}) \tag{1.9.13}$$

$$\lim_{\mathbf{x} \to 0} \left((\deg_{\overline{\mathbb{F}}_p} [{}^{\mathbb{L}}\mathcal{Z}(\mathbf{x}) \otimes^{\mathbb{L}}_{\mathcal{O}_{\mathcal{N}}} {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}]) \cdot \log p + \frac{1}{2}(\log |t|_p - \log p) \cdot \deg(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}) \right) \tag{1.9.14}$$
$$= \sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}} \deg(\mathcal{Z}) \cdot \delta_{\mathrm{tau}}(\mathcal{Z}) \cdot \log p.$$

Here $\mathcal{E}^\vee$ is a certain dual tautological bundle on $\mathcal{N}$, the sum runs over components $\mathcal{Z}$ of (the finite scheme associated to) $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$, the notation $\deg(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}})$ means the degree of the adic finite flat morphism $\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} \to \operatorname{Spf} \breve{\mathbb{Z}}_p$ (and similarly for $\deg(\mathcal{Z})$), and $\delta_{\mathrm{tau}}(\mathcal{Z}) \in \mathbb{Q}$ is the appropriate "local change of tautological height". As in Section 7.3, each $\mathcal{Z}$ is associated with a quasi-canonical lifting of some level s, and our notation $\delta_{\mathrm{tau}}(\mathcal{Z})$ here is the $\delta_{\mathrm{tau}}(s)$ in Eq. (7.2.7).

The quantity $\delta_{\mathrm{tau}}(\mathcal{Z})$ (and the closely related "local change of Faltings height" δ_{Fal}) arise from our reduction process from mixed characteristic heights to local quantities. This local-to-global reduction is begun in Part III and completed in Part IV.

The "vertical" limit formula in (1.9.13) follows from a Grothendieck–Messing theory argument (such vertical limiting behavior was observed in the inert case by [LZ22a] via computation, and later in the ramified case by [LL22] via a linear-invariance argument), see Lemma 18.5.4. We prove the "horizontal" limit formula in (1.9.14) componentwise, i.e. we prove a refined limiting formula for each component $\mathcal{Z} \hookrightarrow \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}$ (Remark 18.5.5). Each $\mathcal{Z}$ embeds into a smaller Rapoport–Zink space of dimension 2, where we make a computation in terms of quasi-canonical liftings.

We have the formula $W^*_{T^\flat,p}(-1/2)^\circ_n = \deg(\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}} \to \operatorname{Spf} \breve{\mathbb{Z}}_p)$ ("local geometric Siegel–Weil"; right-hand side denotes degree of the indicated adic finite flat morphism) as in Lemma 18.1.3 (observed in the inert case by Li–Zhang [LZ22a, Corollary 4.6.1]). Using the functional equation $W^*_{T^\flat,p}(s)^\circ_n = W^*_{T^\flat,p}(-s)^\circ_n$, the limit formula in (1.9.8) now implies

$$-\frac{d}{ds}\bigg|_{s=1/2} W^*_{T^\flat,p}(s)^\circ_n = \left(2\deg_{\overline{\mathbb{F}}_p}(\mathcal{E}^\vee \cdot {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{V}}) + 2 \sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}(\underline{\mathbf{x}}^\flat)_{\mathscr{H}}} \deg(\mathcal{Z}) \cdot \delta_{\mathrm{tau}}(\mathcal{Z}) \right) \cdot \log p. \tag{1.9.15}$$

This is our main non-Archimedean local theorem for odd inert p (i.e. the dotted arrow in Figure 1), and appeared previously in Section 1.10. The analogous statement treating inert/split/ramified simultaneously is Theorem 18.1.2 in this paper (there stated in terms of local densities). Limiting on the geometric side of Figure 1 was provided by the formulas in (1.9.13) and (1.9.14).

1.10 Outline

We briefly summarize the remaining contents of the paper. Further explanations may be found at the beginning of some parts and sections.

This work is divided into Parts I through VII and appendices. We hope that each part may be read mostly independently for the reader willing to assume a few results from other parts.

In Part I, Section 3, we set up the global moduli stacks (RSZ) and special cycles (KR) appearing in our main global theorems. In Section 4, we define the associated arithmetic special cycle classes and discuss arithmetic degrees.

In Part II, we set up the analogous local special cycles on Rapoport–Zink spaces (inert/ramified/split) and Hermitian symmetric domains. The case of split primes is less well-studied in the literature than the inert/ramified cases (we need uniformization in a non supersingular situation at split primes). Section 6 contains some new results on decomposing local special cycles into quasi-canonical lifting cycles at split primes, which we need later. These are analogous to known results at inert and ramified primes (Section 7.3), though our method of proof is different.

In Part III, we begin the reduction process from global heights in mixed characteristic to quantities computable in terms of local special cycles. We study "local change of heights" along isogenies, in a way suitable for formulation of our main local theorems.

In Part IV, we discuss complex and Rapoport–Zink uniformization of special cycles in our setup, and finish the reduction process from global heights/intersections to local quantities. Strictly speaking, the Rapoport–Zink uniformization we need at split places does not seem covered by the literature (not supersingular locus). We treat inert/ramified/split in parallel. Most of the time, we disallow $p = 2$ only in the ramified case. We explain a modified Green current for singular T (of rank $n - 1$ and size $n \times n$) in Section 12.4.

Part V discusses $U(m, m)$ Siegel–Weil Eisenstein series. To formulate and prove our main results, it is extremely important that we normalize the Eisenstein series and local Whittaker functions (e.g. by certain L-factors). We pin down explicit precise normalizations, guided by special value formulas and symmetric functional equations. We also study (normalized) Fourier coefficients for singular T (focusing on rank $m - 1$ and size $m \times m$), and give formulas needed for our main results. Section 15.6 collects several limiting formulas for local Whittaker functions (the left vertical arrow in Figure 1), whose proofs appear later.

Part VI contains the heart of this work. Here, we prove our main local identities at inert/ramified/split and Archimedean places via the local limiting method sketched in Sections 1.7 and 1.9.

In Part VII, we first give some special value formulas (local and geometric Siegel–Weil, Sections 20 and 21) which are needed to prove our arithmetic Siegel–Weil theorems. The finale occurs in Section 22.1, where we collect our local main theorems to prove our (global) arithmetic Siegel–Weil theorems. This proof relies on results from almost all preceding sections. Section 22.2 contains a reformulation of our arithmetic Siegel–Weil results in the special case $n = 2$, via an exceptional comparison with Hecke translates of CM elliptic curves.

The appendices may be technically useful. Appendix A explains the setup we use for K_0 groups of Deligne–Mumford stacks. Appendix B concerns p-divisible groups, where we fix some notation and record some (presumably standard) facts. Appendix C contains some notation on abelian schemes, and records a proof for quasi-compactness of special cycles (which does not seem explicitly available in the literature).

Our algebro-geometric conventions follow the Stacks project [SProject] unless stated otherwise.

2 Conventions on Hermitian spaces and lattices

2.1 Hermitian, alternating, symmetric

Consider a Dedekind domain $\mathcal{O}_{F_0}$ with fraction field F_0. Let F be a finite étale F_0-algebra of degree 2, i.e. either F is a degree 2 separable field extension of F_0, or $F \cong F_0 \times F_0$. Let $\mathcal{O}_F \subseteq F$ be the integral closure of $\mathcal{O}_{F_0}$ in F. Write $a \mapsto a^\sigma$ for the nontrivial involution of F over F_0, and $\mathrm{tr}\colon F \to F_0$ for the trace map $a \mapsto a + a^\sigma$.

Assume that the different ideal $\mathfrak{d}$ may be generated by one element u such that $u^\sigma = -u$, and fix such a u. This is always possible if $\mathcal{O}_F$ is a free $\mathcal{O}_{F_0}$-module: in this case, if $\{\alpha, \beta\}$ is any $\mathcal{O}_{F_0}$-basis for $\mathcal{O}_F$, we may take $u = \alpha\beta^\sigma - a^\sigma\beta$.

Let L be a finite locally free $\mathcal{O}_F$-module of constant rank. If F_0 has characteristic $\neq 2$, the following data are equivalent.

(1) A *Hermitian pairing* on L, i.e. an $\mathcal{O}_{F_0}$-bilinear map $(-,-)\colon L \times L \to F$ satisfying

$$(x, ay) = a(x, y) \qquad (y, x) = (x, y)^\sigma \tag{2.1.1}$$

for all $a \in \mathcal{O}_F$ and $x, y \in L$.

(2) An $\mathcal{O}_F$*-compatible alternating pairing* on L, i.e. an $\mathcal{O}_{F_0}$-bilinear map $\langle -,- \rangle\colon L \times L \to F_0$ satisfying

$$\langle ax, y \rangle = \langle x, a^\sigma y \rangle \qquad \langle y, x \rangle = -\langle x, y \rangle \tag{2.1.2}$$

for all $a \in \mathcal{O}_F$ and $x, y \in L$.

(3) An $\mathcal{O}_F$*-compatible symmetric pairing* on L, i.e. an $\mathcal{O}_{F_0}$-bilinear map $(-,-)\colon L \times L \to F_0$ satisfying

$$(ax, y) = (x, a^\sigma y) \qquad (y, x) = (x, y) \tag{2.1.3}$$

for all $a \in \mathcal{O}_F$ and $x, y \in L$.

If L is equipped with any of the equivalent data above, we say that L is a *Hermitian $\mathcal{O}_F$-lattice* (or *Hermitian $\mathcal{O}_F$-module*). Note that our Hermitian pairings $(-,-)$ are conjugate linear in the first argument. We pass between these pairings using the formulas (depending on the choice of u)

$$\begin{aligned} 2(x,y) &= (x,y) - u^{-1}(ux,y) & \langle x,y\rangle &= (u^{-1}x,y) & (x,y) &= \mathrm{tr}((x,y)) \\ 2(x,y) &= \langle ux,y\rangle - u\langle x,y\rangle & \langle x,y\rangle &= -\mathrm{tr}((x,y)u^{-1}) & (x,y) &= \langle ux,y\rangle \end{aligned}$$

and this will be freely used in the paper. The choice of u plays a limited role for us, so we generally suppress it.

We say that $(-,-)$ is the associated *trace pairing*, and otherwise avoid the notation $(-,-)$ outside of Section 2.1.

Given any tuple $\underline{x} = [x_1, \ldots, x_m] \in L^m$, its *Gram matrix* is the matrix $(\underline{x}, \underline{x}) = T$ with i,j-th entry $T_{i,j} = (x_i, x_j)$. We write $L_F \coloneqq L \otimes_{\mathcal{O}_F} F$ and say that a Hermitian $\mathcal{O}_F$-module L is *non-degenerate* if the Gram matrix for any F-basis of L_F has nonzero determinant. Given non-degenerate Hermitian F-modules V and V' with Hermitian pairings $(-,-)$ and $(-,-)'$, there is a canonical σ-linear involution of F-modules

$$\mathrm{Hom}_F(V, V') \xrightarrow{f \mapsto f^\dagger} \mathrm{Hom}_F(V', V) \qquad \begin{array}{l} \text{such that } (fx, y')' = (x, f^\dagger y') \\ \text{for all } x \in V \text{ and } y' \in V'. \end{array} \tag{2.1.4}$$

The notation $\mathrm{Hom}_F(V, V')$ and $\mathrm{Hom}_F(V', V)$ does not include any requirement on preserving Hermitian pairings.

Given a non-degenerate Hermitian $\mathcal{O}_F$-lattice L, we always form its *dual lattice* L^* with respect to the trace pairing $(-,-)$, i.e.

$$L^* \coloneqq \{x \in L_F : \mathrm{tr}(x,y) \in \mathcal{O}_{F_0} \text{ for all } y \in L\}. \tag{2.1.5}$$

The dual lattice $L^\vee$ with respect to $(-,-)$ is the same as the dual lattice for $\langle -,-\rangle$. We have $L^\vee = uL^*$ (as sublattices of L_F). If the dual $L^\vee$ with respect to $(-,-)$ or $\langle -,-\rangle$ is intended, we will state this explicitly. We say that L is *self-dual* if $L = L^*$.

As a typical example of passing between $(-,-)$ and $\langle -,-\rangle$, suppose $\mathcal{O}_{F_0} = \mathbb{Z}$ and suppose $\mathcal{O}_F$ is the ring of integers in an imaginary quadratic field $F/\mathbb{Q}$. Let (A, ι, λ) be a *Hermitian abelian variety* (Definition 3.1.1) over an algebraically closed field k of characteristic $\neq p$, i.e. A is an abelian variety over k with an action $\iota \colon \mathcal{O}_F \to \mathrm{End}(A)$, and λ is an $\mathcal{O}_F$-compatible quasi-polarization on A. After picking a trivialization $\mathbb{Z}_p(1) \cong \mathbb{Z}_p$ of p-th power roots of unity over k, the polarization λ induces an $(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)$-compatible alternating pairing on the Tate module $T_p(A)$, so we automatically view $T_p(A)$ as a Hermitian $(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)$-lattice without further mention. If (A', ι', λ') is another such Hermitian p-divisible group, note that the induced Hermitian pairing on $\mathrm{Hom}_{\mathcal{O}_F}(T_p(A), T_p(A'))$ does not depend on the choice of trivialization $\mathbb{Z}_p(1) \cong \mathbb{Z}_p$ or the choice of u.

The notation $\mathrm{Herm}_n(\mathcal{O}_{F_0})$ means the set of $n \times n$ Hermitian matrices with coefficients in $\mathcal{O}_F$ (i.e. $T \in M_{n,n}(\mathcal{O}_F)$ satisfying $T = {}^t\overline{T}$ where ${}^t\overline{T}$ means conjugate transpose). Here we are considering the subfunctor $\mathrm{Herm}_n \subseteq \mathrm{Res}_{\mathcal{O}_F/\mathcal{O}_{F_0}} M_{n,n}$ of the Weil restriction (of $n \times n$ matrices $M_{n,n}$). We adhere strictly to this notation (when $\mathcal{O}_F$ is understood), e.g. $\mathrm{Herm}_n(\mathbb{R})$ will typically mean $n \times n$ complex Hermitian matrices when $\mathcal{O}_F/\mathcal{O}_{F_0} = \mathbb{C}/\mathbb{R}$ is understood.

2.2 Lattices for local fields

Continuing in the setup of Section 2.1, suppose F_0 is a local field. Let $\eta \colon F_0^\times \to \{\pm 1\}$ be the character associated to F/F_0 by local class field theory. Given a non-degenerate Hermitian F-module V of rank n, define its *local invariant*

$$\varepsilon(V) \coloneqq \eta((-1)^{n(n-1)/2} \det T) \in \{\pm 1\} \tag{2.2.1}$$

where T is the Gram matrix of any basis for V. This is normalized so that $\varepsilon(V) = 1$ for the Hermitian F-module V given by the antidiagonal unit Gram matrix. Rank n non-degenerate Hermitian F-modules V and V' are isomorphic if and only if $\varepsilon(V) = \varepsilon(V')$. If $T \in \mathrm{Herm}_n(F_0)$ is a Hermitian matrix (with entries in F) satisfying $\det T \neq 0$, we set $\varepsilon(T) \coloneqq \eta((-1)^{n(n-1)/2} \det T)$.

Next, assume F_0 is non-Archimedean and that $\mathcal{O}_{F_0} \subseteq F_0$ is its ring of integers. Write q for the residue cardinality of $\mathcal{O}_{F_0}$. If q is even, we require F/F_0 to be unramified. Let $\varpi_0 \in \mathcal{O}_{F_0}$ and $\varpi \in \mathcal{O}_F$ be uniformizers (meaning $\varpi \in \varpi_0 \mathcal{O}_F^\times$ in the unramified cases) satisfying $\varpi^\sigma = -\varpi$. If a non-degenerate Hermitian F-module V contains a full rank self-dual $\mathcal{O}_F$-lattice, then $\varepsilon(V) = 1$.

The "norm" $\|-\|$ on a Hermitian F-module V with pairing $(-,-)$ is given by

$$\|x\| \coloneqq q^{-v_{\varpi_0}((x,x))/2} \tag{2.2.2}$$

where v_{ϖ_0} is the ϖ_0-adic valuation, normalized so that $v_{\varpi_0}(\varpi_0) = 1$.

Given a non-degenerate Hermitian $\mathcal{O}_F$-lattice L of rank n, we set $\varepsilon(L) \coloneqq \varepsilon(L_F)$. By a *lattice* or *sublattice* $L' \subseteq L$, we mean any $\mathcal{O}_F$-submodule which is finite free of constant rank (similarly for lattices or sublattices in L_F). If L' has rank n, we say that L' is *full rank* in L_F. A sublattice $L' \subseteq L$ is *saturated* if $ax \in L'$ with $a \in F^\times$ and $x \in L$ implies $x \in L'$ (equivalently, L' is a direct summand of L).

We say that L is *integral* if $L \subseteq L^*$. If F/F_0 is nonsplit, we say that L is *almost self-dual* if $L \subseteq L^*$ and $\mathrm{length}_{\mathcal{O}_F}(L^*/L) = 1$. We say that a non-degenerate integral lattice L is *maximal integral* if any integral lattice $L' \subseteq L_F$ with $L \subseteq L'$ satisfies $L = L'$.

If L is a non-degenerate Hermitian $\mathcal{O}_F$-lattice, we define the *valuation* $\mathrm{val}(L) \in \frac{1}{2}\mathbb{Z}$ such that

$$q^{-\mathrm{val}(L)} = \mathrm{vol}(L) \tag{2.2.3}$$

where $\mathrm{vol}(L)$ is the volume of L for the self-dual Haar measure on L_F with respect to the pairing $x, y \mapsto \psi(\mathrm{tr}(x,y))$ for any unramified (unitary) additive character $\psi \colon F_0 \to \mathbb{C}^\times$. If L is integral, we have $q^{2\mathrm{val}(L)} = |L^*/L|$. If F/F_0 is unramified, we have $\mathrm{val}(L) \in \mathbb{Z}$. Given $x \in L$, we write $\langle x \rangle \subseteq L$ for the rank one $\mathcal{O}_F$-submodule generated by x. If $(x,x) \neq 0$, we set $\mathrm{val}(x) \coloneqq \mathrm{val}(\langle x \rangle)$ (and otherwise set $\mathrm{val}(x) = \infty$).

Continuing to assume L is non-degenerate and integral, we define its sequence of *fundamental invariants* to be the unique sequence of integers $(a_1, \ldots, a_n)$ with $0 \leq a_1 \leq \cdots \leq a_n$ such that $L^*/L \cong \oplus_{i=1}^n \mathcal{O}_F/\varpi^{a_i}$ (where n is the rank of L). Two non-degenerate integral Hermitian $\mathcal{O}_F$-lattices of the same rank are isomorphic if and only if they have the same sequence of fundamental invariants (in the unramified case, this follows from diagonalizability of Hermitian lattices; in the ramified case, this follows from [Jac62, Proposition 4.3, Proposition 8.1] (see also [LL22, Lemma 2.12])). We set

$$t(L) \coloneqq |\{a_i \in \{a_1, \ldots, a_n\} : a_i \neq 0\}| \in \mathbb{Z} \qquad a_{\max}(L) \coloneqq a_n \tag{2.2.4}$$

and refer to $t(L)$ as the *type* of L. If F/F_0 is ramified, recall that $t(L)$, $2\mathrm{val}(L)$, and n all have the same parity (follows from [Jac62, Proposition 4.3, Proposition 8.1]).

Given a finite length $\mathcal{O}_F$-module M, we define $\ell(M) \in \mathbb{Z}$ such that

$$q^{\ell(M)} = |M| \tag{2.2.5}$$

where $|M|$ denotes the cardinality of M.

The above terminology is adapted from e.g. [LZ22a] (inert), [FYZ21] (inert and split), [LL22] (ramified). We made slight modifications to give a uniform description (e.g. our $\mathrm{val}(L)$ is half of the $\mathrm{val}(L)$ appearing in [LL22], and our $\ell(M)$ differs by a factor of 2 from some of the references).

Part I

Global special cycles

3 Moduli stacks of abelian varieties

We discuss Kudla–Rapoport (KR) global special cycles on Rapoport–Smithling–Zhang (RSZ) smooth integral models of unitary Shimura varieties (which may be stacks). Fix an imaginary quadratic field extension $F/\mathbb{Q}$ with ring of integers $\mathcal{O}_F$ and write $a \mapsto a^\sigma$ for the nontrivial automorphism σ of F. We write $\Delta \in \mathbb{Z}_{<0}$ and $\sqrt{\Delta} \in \mathcal{O}_F$ (pick a square root) for (generators of the) discriminant and different, respectively.

3.1 Integral models

Definition 3.1.1. Let S be a scheme over $\operatorname{Spec} \mathcal{O}_F$. By a *Hermitian abelian scheme* over S, we mean a tuple (A, ι, λ) where

A	is an abelian scheme over S of constant relative dimension n
$\iota \colon \mathcal{O}_F \to \operatorname{End}(A)$	is a ring homomorphism
$\lambda \colon A \to A^\vee$	is a quasi-polarization satisfying:

> (Action compatibility) The Rosati involution $\dagger$ on $\operatorname{End}^0(A)$ satisfies $\iota(a)^\dagger = \iota(a^\sigma)$ for all $a \in \mathcal{O}_F$.

An *isomorphism* of Hermitian abelian schemes is an isomorphism of abelian schemes which respects the $\mathcal{O}_F$-actions and polarizations (exactly). For fixed $n \geq 1$, the *moduli stack of Hermitian abelian schemes* $\mathscr{M}$ is the stack[11] in groupoids over $\operatorname{Spec} \mathcal{O}_F$ with

$$\mathscr{M}(S) \coloneqq \{\text{groupoid of relative } n\text{-dimensional Hermitian abelian schemes over } S\} \quad (3.1.1)$$

for $\mathcal{O}_F$-schemes S.

For an integer r with $0 \leq r \leq n$, we next consider

> (Kottwitz $(n-r, r)$ signature condition) For all $a \in \mathcal{O}_F$, the characteristic polynomial of $\iota(a)$ acting on $\operatorname{Lie} A$ is $(x-a)^{n-r}(x-a^\sigma)^r \in \mathcal{O}_S[x]$

for pairs (A, ι), where $A \to S$ is a relative n-dimensional abelian scheme with $\mathcal{O}_F$-action ι, and S is an $\mathcal{O}_F$-scheme. Here we view $\mathcal{O}_S$ as an $\mathcal{O}_F$-algebra via the structure map $S \to \operatorname{Spec} \mathcal{O}_F$. This defines a substack[12]

$$\mathscr{M}(n-r, r) \subseteq \mathscr{M} \quad (3.1.2)$$

[11] By a *stack in groupoids* over some base scheme S, we always mean a (not necessarily algebraic) stack in groupoids as in [SProject, Definition 02ZI] over the fppf site $(Sch/S)_{fppf}$.

[12] A *substack* will always mean a strictly full substack.

consisting of Hermitian abelian schemes of signature $(n-r,r)$. The inclusion $\mathscr{M}(n-r,r) \to \mathscr{M}$ is representable by schemes (in the sense of [SProject, Section 04ST]) and is a closed immersion. There is an isomorphism[13] $\mathscr{M}(n-r,r) \to \mathscr{M}(r,n-r)$ given by $(A,\iota,\lambda) \mapsto (A,\iota\circ\sigma,\lambda)$.

For any integer $d \geq 1$, there is a substack $\mathscr{M}^{(d)} \subseteq \mathscr{M}$ consisting of Hermitian abelian schemes (A,ι,λ) where λ is polarization of constant degree $\deg\lambda \coloneqq \deg\ker\lambda = d$. If $\mathcal{A}_{n,d}$ (over $\operatorname{Spec}\mathcal{O}_F$) denotes the moduli stack of (relative) n-dimensional abelian schemes equipped with a polarization of degree d, the forgetful map $\mathscr{M}^{(d)} \to \mathcal{A}_{n,d}$ is representable by schemes, finite, and unramified (e.g. via Lemma C.2.3). Hence $\mathscr{M}^{(d)}$ is a Noetherian Deligne–Mumford stack which is separated and finite type over $\operatorname{Spec}\mathcal{O}_F$ (because this is true of $\mathcal{A}_{n,d}$ as proved with level structure in the classical [MFK94, §7.2 Theorem 7.9]; one can deduce the stacky version upon inverting primes dividing the level, taking stack quotients, and patching over $\operatorname{Spec}\mathcal{O}_F$).

We set

$$\mathscr{M}(n-r,r)^{(d)} \coloneqq \mathscr{M}(n-r,r) \cap \mathscr{M}^{(d)} \tag{3.1.3}$$

where the right-hand side is an intersection of substacks of $\mathscr{M}$. There is an open and closed disjoint union decomposition[14]

$$\mathscr{M}^{(d)}[1/\Delta] = \coprod_{(n-r,r)} \mathscr{M}(n-r,r)^{(d)}[1/\Delta] \tag{3.1.4}$$

over $\operatorname{Spec}\mathcal{O}_F[1/\Delta]$, where the disjoint union runs over all possible signatures $(n-r,r)$.

The structure morphism $\mathscr{M}(n-r,r)^{(d)}[1/(d\Delta)] \to \operatorname{Spec}\mathcal{O}_F[1/\Delta]$ is smooth of relative dimension $(n-r)r$ (e.g. by Remark 3.5.6 below; recall that being smooth of some relative dimension may be checked fppf locally on the target for morphisms of algebraic stacks). We set $\mathscr{M}_0 \coloneqq \mathscr{M}(1,0)^{(1)}$. The structure morphism $\mathscr{M}_0 \to \operatorname{Spec}\mathcal{O}_F$ is proper, quasi-finite,[15] and étale by [How12, Proposition 3.1.2] or [How15, Proposition 2.1.2].

Given any non-degenerate Hermitian $\mathcal{O}_F$-lattice L of rank n and signature $(n-r,r)$, we define an associated substack

$$\mathcal{M} \subseteq \mathscr{M}_0 \times_{\operatorname{Spec}\mathcal{O}_F} \mathscr{M}(n-r,r) \tag{3.1.5}$$

[13]As in the Stacks project (e.g. [SProject, Section 04XA]), we often abuse terminology and say "isomorphism" of stacks instead of "equivalence".

[14]Here, the notation $\mathscr{M}^{(d)}[1/\Delta]$ means $\mathscr{M}^{(d)} \times_{\operatorname{Spec}\mathcal{O}_F} \operatorname{Spec}\mathcal{O}_F[1/\Delta]$. We often use such shorthand, along with subscripts for base change, e.g. $\mathscr{M}^{(d)}_S \coloneqq \mathscr{M}^{(d)} \times_{\operatorname{Spec}\mathcal{O}_F} S$ over an understood base.

[15]Following the Stacks project [SProject, Definition 0CHU], we require that finite morphisms of algebraic stacks are by definition (relatively) representable by schemes. The morphism $\mathscr{M}_0 \to \operatorname{Spec}\mathcal{O}_F$ is not finite in this sense, because $\mathscr{M}_0$ is not a scheme. Nevertheless, we continue to use terminology like "representable by schemes and finite" for morphisms of stacks which are not necessarily algebraic.

as follows (cf. [KR14, Proposition 2.12], there in a principally polarized situation). Write $(-,-)$ for the pairing on L. Let b_L be the smallest positive integer such that $b_L \cdot (-,-)$ is $\mathcal{O}_F$-valued. Let L' be the Hermitian $\mathcal{O}_F$ lattice which is the $\mathcal{O}_F$-module L but with Hermitian pairing $b_L \cdot (-,-)$. Form the dual lattice $L'^{\vee}$ of L' with respect to the Hermitian pairing, and set $d'_L \coloneqq |L'^{\vee}/L'|$.

If L is self-dual of signature $(n-1,1)$ and $2 \nmid \Delta$, we set $d_L \coloneqq 1$ (the *exotic smooth* setup for even n, see Section 3.2). Otherwise, let $d_L \in \mathbb{Z}_{>0}$ be the product of ramified primes and the primes p for which $L \otimes_{\mathbb{Z}} \mathbb{Z}_p$ is not self-dual.

Definition 3.1.2. Let $\mathcal{M} \subseteq \mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(n-r,r)[1/(d_L\Delta)]$ be the substack

$$\mathcal{M}(S) \coloneqq \left\{ (A_0, \iota_0, \lambda_0, A, \iota, \lambda) : \begin{array}{l} \operatorname{Hom}_{\mathcal{O}_F \otimes \hat{\mathbb{Z}}^p}(T^p(A_{0,\overline{s}}), T^p(A_{\overline{s}})) \cong L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \\ \text{for every geometric point } \overline{s} \text{ of } S \text{, with } p = \operatorname{char}(\overline{s}), \\ \text{and } b_L \cdot \lambda \text{ is a polarization of degree } d'_L \end{array} \right\} \tag{3.1.6}$$

for schemes S over $\operatorname{Spec} \mathcal{O}_F[1/(d_L\Delta)]$, where

$$(A_0, \iota_0, \lambda_0) \in \mathscr{M}_0(S) \qquad (A, \iota, \lambda) \in \mathscr{M}(n-r,r)(S). \tag{3.1.7}$$

Warning 3.1.3. Whenever L satisfies the even rank *exotic smooth* setup (Section 3.2), we will extend $\mathcal{M}$ to a smooth Deligne–Mumford stack which surjects onto $\operatorname{Spec} \mathcal{O}_F$ (see loc. cit.). In that case, we will override the notation here: after Section 3.2, the notation $\mathcal{M}$ will always denote the exotic smooth moduli stack for such L. The restriction of the exotic smooth moduli stack over $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$ will recover the stack in Definition 3.1.2.

In the definition of $\mathcal{M}$, the notation $\operatorname{Hom}_{\mathcal{O}_F \otimes \hat{\mathbb{Z}}^p}(T^p(A_{0,\overline{s}}), T^p(A_{\overline{s}})) \cong L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p$ asserts the existence of isomorphisms of Hermitian lattices, and the elements of $\operatorname{Hom}_{\mathcal{O}_F \otimes \hat{\mathbb{Z}}^p}(T^p(A_{0,\overline{s}}), T^p(A_{\overline{s}}))$ are not required to respect Hermitian pairings. As usual, $T^p(-)$ is the away-from-p adèlic Tate module (if $p = 0$, this is over the full finite adèles) and $\hat{\mathbb{Z}}^p = \prod_{\ell \neq p} \mathbb{Z}_\ell$. Note that $\mathcal{M}$ depends only on the *adèlic isomorphism class*[16] of L. The stack $\mathcal{M}$ (also the extension in Section 3.2) and its special cycles will be the global moduli stacks of main interest in this work. We generally suppress L from notation, but sometimes write $\mathcal{M}^L$ instead of $\mathcal{M}$ to emphasize L dependence.

We claim that $\mathcal{M}$ is a Noetherian Deligne–Mumford stack which is separated and smooth of relative dimension $(n-r)r$ over $\operatorname{Spec} \mathcal{O}_F[1/(d_L\Delta)]$. Indeed, there is an open and closed

[16] We say that non-degenerate Hermitian $\mathcal{O}_F$ lattices L and L' are *adèlically isomorphic* (or are in the same *adèlic isomorphism class*) if there exist isomorphisms of $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$-Hermitian lattices $L \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong L' \otimes_{\mathbb{Z}} \mathbb{Z}_p$ for every prime p, as well as isomorphisms of $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{R}$-Hermitian spaces $L \otimes_{\mathbb{Z}} \mathbb{R} \cong L' \otimes_{\mathbb{Z}} \mathbb{R}$ (classical terminology: *genus*).

disjoint union decomposition

$$\mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(n-r,r)^{(d)}[1/(d\Delta)] = \coprod_{L''} \mathcal{M}^{L''} \tag{3.1.8}$$

running over representatives L'', one for each adèlic isomorphism class of non-degenerate Hermitian $\mathcal{O}_F$-lattices of signature $(n-r,r)$ satisfying $L'' \subseteq L''^{\vee}$ and $|L''^{\vee}/L''| = d$. We have used flatness of $\mathscr{M}(n-r,r)^{(d)}[1/(d\Delta)] \to \operatorname{Spec} \mathcal{O}_F[1/(d\Delta)]$ in the open and closed decomposition (to lift to characteristic 0; cf. [KR14, Proposition 2.12] [RSZ18, Remark 4.2]). With notation as above, the map

$$\begin{array}{ccc} \mathcal{M}^L & \longrightarrow & \mathcal{M}^{L'} \\ (A_0, \iota_0, \lambda_0, A, \iota, \lambda) & \longmapsto & (A_0, \iota_0, \lambda_0, A, \iota, b_L \lambda) \end{array} \tag{3.1.9}$$

is an isomorphism for any L, after restricting to $\operatorname{Spec} \mathcal{O}_F[1/(d_L \Delta)]$.

Remark 3.1.4. If L has rank $n = 1$, we can construct $\mathcal{M}$ without discarding any primes. Then $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ is smooth, by smoothness of $\mathscr{M}(1,0)^{(d)} \to \operatorname{Spec} \mathcal{O}_F$ for any $d \in \mathbb{Z}_{>0}$.

In the next lemma, A_0^{σ} is the abelian scheme A_0 but with $\mathcal{O}_F$-action $\iota_0 \circ \sigma$.

Lemma 3.1.5. *Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n and signature $(n-r,r)$, with associated moduli stack $\mathcal{M}$. There exists a finite degree field extension E/F and $(A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(\mathcal{O}_E[1/(d_L \Delta)])$ such that A is $\mathcal{O}_F$-linearly isogenous to $A_0^{n-r} \times (A_0^{\sigma})^r$. In particular, $\mathcal{M}$ is nonempty.*

Proof. First consider $\kappa = \mathbb{C}$ (equipped with a morphism $\mathcal{O}_F \to \mathbb{C}$). Fix the trivializations of roots of unity $\mathbb{Z}/d\mathbb{Z} \xrightarrow{\sim} \boldsymbol{\mu}_d(\mathbb{C})$ sending $1 \mapsto e^{-2\pi i/d}$.

Choose $\sqrt{\Delta}$ to be the square-root whose image under $F \to \mathbb{C}$ has positive imaginary part. We pass between Hermitian and alternating forms using the generator $\sqrt{\Delta}$ of the different ideal (Section 2.1). Express L as a triple (L, ι, λ) where $\iota \colon \mathcal{O}_F \to \operatorname{End}_{\mathbb{Z}}(L)$ is an action and λ is a $\mathcal{O}_F$-compatible alternating pairing on L.

Take $(A_0, \iota_0, \lambda_0)$ to be the complex elliptic curve $\mathbb{C}/\mathcal{O}_F$. If $L_0 \coloneqq \mathcal{O}_F$ is the rank one Hermitian $\mathcal{O}_F$-lattice with Hermitian pairing $(x,y) = x^{\sigma} y$, we have $H_1(A_0, \mathbb{Z}) \cong L_0$ as Hermitian lattices.

Take any orthogonal decomposition $L_F = W \oplus W^{\perp}$ where W is positive definite of rank $n-r$ and $W^{\perp}$ is negative definite of rank r. Define the $\mathbb{C} = \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{R}$-action on $L \otimes_{\mathbb{Z}} \mathbb{R}$ to agree with ι on $W \otimes_F \mathbb{C}$ and to agree with $\iota \circ \sigma$ on $W^{\perp} \otimes_F \mathbb{C}$. This complex structure gives a tuple (A, ι, λ), where $A \coloneqq (L \otimes_{\mathbb{Z}} \mathbb{R})/L$ is an abelian variety with $\mathcal{O}_F$-action ι and action compatible quasi-polarization λ. We have $H_1(A,\mathbb{Z}) \cong L$ as Hermitian lattices. By the usual comparison of $H_1(-,\mathbb{Z})$ with p-adic Tate modules [Mum85, §24 Theorem 1], we conclude $(A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(\mathbb{C})$.

We claim that A is $\mathcal{O}_F$-linearly isogenous to $A_0^{n-r} \times (A_0^{\sigma})^r$. Indeed, any $\mathcal{O}_F$-linear inclusion $\mathcal{O}_F^{n-r} \hookrightarrow L \cap W$ and any σ-linear inclusion $\mathcal{O}_F^r \hookrightarrow L \cap W^{\perp}$ will define an $\mathcal{O}_F$-linear isogeny $A_0^{n-r} \times (A_0^{\sigma})^r \to A$.

Since A_0 is defined over some number field $\overline{\mathbb{Q}}$, it follows that A and any isogeny $A_0^{n-r} \times (A_0^{\sigma})^r \to A$ may also be defined over $\overline{\mathbb{Q}}$ (here using characteristic zero, so the kernel of the isogeny is étale). Descend these objects to some number field E.

Over a number field, it is a classical fact that any elliptic curve with $\mathcal{O}_F$-action has everywhere potentially good reduction [Deu41]. After extending E if necessary, we thus obtain $(A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(\mathcal{O}_E[1/(d_L \Delta)])$ (the $\mathcal{O}_F$-actions extend by the Néron mapping property, and the polarizations extend to polarizations as in the proof of [FC90, Theorem 1.9]). The Néron mapping property extends the isogeny $A_0^{n-r} \times (A_0^{\sigma})^r \to A$ over $\operatorname{Spec} \mathcal{O}_E[1/(d_L \Delta)]$. $\square$

The preceding lemma will provide a base point for non-Archimedean uniformization (Section 11.3). For arbitrary L and specializing to any geometric point of characteristic $p > 0$, the abelian variety A of Lemma 3.1.5 is supersingular (resp. ordinary) if p is nonsplit (resp. split) by classical results of Deuring [Deu41] on endomorphism rings of elliptic curves.

Notation 3.1.6. Given a commutative ring R with an automorphism $\sigma \colon R \to R$ (e.g. $R = \mathcal{O}_F$), given a presheaf of modules $\mathcal{F}$ on a scheme S over $\operatorname{Spec} R$, and given an action $\iota \colon R \to \operatorname{End}(\mathcal{F})$ (with $\mathcal{F}$ viewed as a presheaf of abelian groups), we say the R action via ι is *R-linear* (resp. *σ-linear*) if $\iota(a) = a$ (resp. $\iota(a) = a^{\sigma}$) for all $a \in R$. Here we view $\mathcal{O}_S$ as an R-algebra via the structure morphism $S \to \operatorname{Spec} R$.

Given any $(A, \iota, \lambda) \in \mathscr{M}(n-r, r)[1/\Delta]$ for a base scheme S, there is a canonical eigenspace decomposition

$$\operatorname{Lie} A = (\operatorname{Lie} A)^+ \oplus (\operatorname{Lie} A)^- \tag{3.1.10}$$

characterized by $(\operatorname{Lie} A)^+$ (resp. $(\operatorname{Lie} A)^-$) being rank $n-r$ (resp. rank r) and the $\mathcal{O}_F$ action via ι on $(\operatorname{Lie} A)^+$ (resp. $(\operatorname{Lie} A)^-$) being $\mathcal{O}_F$-linear (resp. σ-linear).

Definition 3.1.7. By the *tautological bundle* on $\mathscr{M}(n-r, r)[1/\Delta]$, we mean the rank r locally free sheaf $\mathscr{E}$ (for the fppf topology) whose dual is given by the assignment $\mathscr{E}^{\vee} := (\operatorname{Lie} A)^-$ for $(A, \iota, \lambda) \in \mathscr{M}(n-r, r)[1/\Delta](S)$ for $\mathcal{O}_F$-schemes S.

3.2 Exotic smoothness

Our main results at ramified primes eventually restrict to even n and residue characteristic $\neq 2$. For this reason, we require n even and $2 \nmid \Delta$ in Section 3.2.

Notation 3.2.1 (Exotic smooth setup, even rank)**.** A non-degenerate Hermitian $\mathcal{O}_F$-lattice L of even rank n satisfies the *even rank exotic smooth* setup if $2 \nmid \Delta$, the signature of L is $(n-1,1)$, and L is self-dual (for the trace pairing).

In the exotic smooth setup, we recall how $\mathcal{M}$ can be extended to a certain smooth integral model over $\operatorname{Spec} \mathcal{O}_F$. We consider arbitrary signature $(n-r,r)$. For Hermitian abelian schemes (A, ι, λ), we consider

> (Polarization condition $\circ$) The quasi-polarization $|\Delta| \cdot \lambda$ is a polarization and we have $\ker(|\Delta| \cdot \lambda) = A[\sqrt{\Delta}]$.

We write $\mathscr{M}(n-r,r)^{\mathrm{Kot},\circ}$ for the substack of $\mathscr{M}(n-r,r)$ consisting of Hermitian abelian schemes (A, ι, λ) where λ satisfies polarization condition $\circ$ from above. Here "Kot" indicates that we have "only" imposed the Kottwitz signature condition. For $d = |\Delta|^n$, the map

$$\begin{aligned} \mathscr{M}(n-r,r)^{\mathrm{Kot},\circ} &\longrightarrow \mathscr{M}(n-r,r)^{(d)} \\ (A, \iota, \lambda) &\longmapsto (A, \iota, |\Delta| \cdot \lambda) \end{aligned} \tag{3.2.1}$$

is representable by closed immersions of schemes. In particular, $\mathscr{M}(n-r,r)^{\mathrm{Kot},\circ}$ is also a separated Deligne–Mumford stack which is finite type over $\operatorname{Spec} \mathcal{O}_F$. The restriction $\mathscr{M}(n-r,r)^{\mathrm{Kot},\circ}[1/\Delta] \to \mathscr{M}(n-r,r)^{(d)}[1/\Delta]$ is an open immersion. If κ is any algebraically closed field of characteristic 0, an object $(A, \iota, \lambda) \in \mathscr{M}(n-r,r)$ lies in $\mathscr{M}(n-r,r)^{\mathrm{Kot},\circ}$ if and only if the Hermitian $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$-lattice $T_p(A)$ is self-dual (for the trace pairing) for all p. In particular, we have

$$\mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(n-r,r)^{\mathrm{Kot},\circ}[1/\Delta] = \mathcal{M}^L \tag{3.2.2}$$

where L is a representative for the unique adèlic isomorphism class of self-dual signature $(n-r,r)$ non-degenerate Hermitian $\mathcal{O}_F$-lattices (if it exists). Such L exists if and only if $n \equiv 2^r \pmod 4$ due to the global product formula for local invariants of Hermitian spaces.

Now we restrict to signature $(n-1,1)$ and $n \geq 2$. Let $\mathscr{M}(n-1,1)^{\circ}$ be the flat part of $\mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}$, i.e. the scheme-theoretic image of the generic fiber. Equivalently, this is the largest closed substack which is flat over $\operatorname{Spec} \mathcal{O}_F$.

Example 3.2.2. Suppose $\mathcal{O}_F^{\times} = \{\pm 1\}$ (i.e. further exclude $F = \mathbb{Q}[\sqrt{-3}]$). If $\mathscr{M}_{\mathrm{ell}} \to \operatorname{Spec} \mathcal{O}_F$ denotes the moduli stack of elliptic curves base-changed to $\operatorname{Spec} \mathcal{O}_F$, the *Serre tensor* *construction* $E \mapsto E \otimes_{\mathbb{Z}} \mathcal{O}_F$ defines an open and closed immersion $i_{\mathrm{Serre}} \colon \mathscr{M}_{\mathrm{ell}} \to \mathscr{M}(1,1)^{\circ}$ (22.2.2). If we replace $\mathcal{O}_F$ by (representatives of) fractional ideal classes for $\mathcal{O}_F$, we obtain an isomorphism $\coprod_{\mathrm{Cl}(\mathcal{O}_F)} \mathscr{M}_{\mathrm{ell}} \to \mathscr{M}(1,1)^{\circ}$, where $\mathrm{Cl}(\mathcal{O}_F)$ is the class group. This is [KR14, Proposition 14.4]. The local analogue (e.g. Lemma 5.6.2) will play an important role in this work. In Section 22.2, we revisit this description of $\mathscr{M}(1,1)^{\circ}$ to restate our main theorem in the simplest case.

Rapoport–Smithling–Zhang have given a moduli description [RSZ21, §6] for $\mathscr{M}(n-1,1)^{\circ}$. They define[17] a closed substack $\mathscr{M}(n-1,1)^{\mathrm{RSZ}} \subseteq \mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}$ with the same generic fiber, with $\mathscr{M}(n-1,1)^{\mathrm{RSZ}} \to \operatorname{Spec} \mathcal{O}_F$ smooth of relative dimension $n-1$. This moduli description is as follows: given any scheme S over $\operatorname{Spec} \mathcal{O}_F$, the groupoid $\mathscr{M}(n-1,1)^{\mathrm{RSZ}}(S) \subseteq \mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}(S)$ is the full subcategory consisting of tuples (A, ι, λ) such that the action $\iota \colon \mathcal{O}_F \to \mathrm{End}(A)$ satisfies:

(1) (Pappas wedge condition) For all $a \in \mathcal{O}_F$, the action of $\iota(a)$ on $\operatorname{Lie} A$ satisfies

$$\bigwedge^2 (\iota(a) - a) = 0 \quad \text{and} \quad \bigwedge^n (\iota(a) - a^{\sigma}) = 0.$$

(2) (PRRSZ spin condition) For every geometric point $\overline{s}$ of S, the action of $(\iota(a) - a)$ on $\operatorname{Lie} A_{\overline{s}}$ is nonzero for some $a \in \mathcal{O}_F$.

The signature condition implies that the equation involving $\bigwedge^n$ in the wedge condition is automatic, and that the wedge condition is empty if $n = 2$. The wedge and spin conditions are automatic (given the signature condition) over $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$, i.e. $\mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}[1/\Delta] = \mathscr{M}(n-1,1)^{\mathrm{RSZ}}[1/\Delta]$. For closedness of the spin condition, we refer to the closedness assertion in [RSZ21, Theorem 5.4]. The acronym PRRSZ stands for Pappas, Rapoport, Richarz, Smithling, and Zhang. We have $\mathscr{M}(n-1,1)^{\circ} = \mathscr{M}(n-1,1)^{\mathrm{RSZ}}$ by agreement in the generic fiber, flatness, and closedness.

We define the *exotic smooth moduli stack*

$$\mathcal{M}^{\circ} := \mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(n-1,1)^{\circ} \tag{3.2.3}$$

associated to any self-dual lattice L of signature $(n-1,1)$. The structure morphism $\mathcal{M}^{\circ} \to \operatorname{Spec} \mathcal{O}_F$ is smooth, by the discussion above.

[17] Strictly speaking, Rapoport–Smithling–Zhang normalize their polarization differently (i.e. our λ is their $|\Delta|^{-1}\lambda$). Their convention is more common elsewhere in the literature, and is of course equivalent to our formulation. We prefer our normalization, which seems more natural for our main results on the comparison with Eisenstein series Fourier coefficients. A related remark is [LL22, Footnote 9].

Notation 3.2.3. From here on, we always write $\mathcal{M}$ instead of $\mathcal{M}^\circ$ if L satisfies the even rank exotic smooth setup (we are overriding previous notation, see Warning 3.1.3; i.e. $\mathcal{M}[1/\Delta]$ recovers the moduli stack in Definition 3.1.2). Recall that we have set $d_L := 1$ for L satisfying the even rank exotic smooth setup.

Remark 3.2.4. Suppose L satisfies the even rank exotic smooth setup, and form the associated moduli stack $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$. Then Lemma 3.1.5 holds for every geometric point $\operatorname{Spec} \kappa \to \operatorname{Spec} \mathcal{O}_F$ by the same proof verbatim (replacing $d_L \Delta$ in loc. cit. with the number 1).

We have the following analogue of (3.1.10): set

$$(\operatorname{Lie} A)^+ := \bigcap_{a \in \mathcal{O}_F} \ker(\iota(a) - a)|_{\operatorname{Lie} A} \tag{3.2.4}$$

for objects $(A, \iota, \lambda) \in \mathscr{M}(n-1,1)^\circ(S)$ over $\mathcal{O}_F$-schemes S.

Lemma 3.2.5. *For objects $(A, \iota, \lambda) \in \mathscr{M}(n-1,1)^\circ(S)$ over $\mathcal{O}_F$-schemes S, the subsheaf $(\operatorname{Lie} A)^+ \subseteq \operatorname{Lie} A$ is a local direct summand of rank $n-1$ whose formation commutes with arbitrary base change. The $\mathcal{O}_F$ action via ι on $(\operatorname{Lie} A)^+$ (resp. the line bundle $(\operatorname{Lie} A)/(\operatorname{Lie} A)^+$) is $\mathcal{O}_F$-linear (resp. σ-linear).*

Proof. This lemma (and its proof) is a global analogue of [LL22, Lemma 2.36] (the latter is an analogous statement on a Rapoport–Zink space).

Fix $a \in \mathcal{O}_F$ such that $\{1, a\}$ forms a $\mathbb{Z}$-basis of $\mathcal{O}_F$. We have exact sequences

$$\begin{gathered} 0 \longrightarrow (\operatorname{Lie} A)^+ \longrightarrow \operatorname{Lie} A \xrightarrow{\iota(a)-a} \operatorname{im}(\iota(a) - a) \longrightarrow 0 \\ 0 \longrightarrow \operatorname{im}(\iota(a) - a) \longrightarrow \operatorname{Lie} A \longrightarrow \operatorname{coker}(\iota(a) - a) \longrightarrow 0 \end{gathered} \tag{3.2.5}$$

of quasi-coherent sheaves on S. The wedge and spin conditions imply that $(\operatorname{Lie} A)^+$ has rank $n-1$ if $S = \operatorname{Spec} k$ for a field k. If S is an arbitrary reduced scheme, the rank constancy of $\operatorname{coker}(\iota(a) - a)$ on geometric points implies that $\operatorname{coker}(\iota(a) - a)$ is finite locally free of rank $n-1$ (e.g. by [SProject, Lemma 0FWG]). Hence, when S is reduced, every sheaf appearing in (3.2.5) is finite locally free, with $(\operatorname{Lie} A)^+$, $\operatorname{im}(\iota(a)-a)$, and $\operatorname{coker}(\iota(a)-a)$ having ranks $n-1$, 1, and $n-1$ respectively. Thus the exact sequences of (3.2.5) remain exact after pullback along any morphism of schemes $S' \to S$ (where S is reduced but S' is not necessarily reduced). For arbitrary S (not necessarily reduced), the morphism $S \to \mathscr{M}(n-1,1)^\circ$ corresponding to (A, ι, λ) factors through a regular (hence reduced) locally Noetherian scheme fppf locally on S (since the moduli stack $\mathscr{M}(n-1,1)^\circ$ is smooth over $\operatorname{Spec} \mathcal{O}_F$). These considerations show that $(\operatorname{Lie} A)^+ \subseteq \operatorname{Lie} A$ is a local direct summand of rank $n-1$ whose formation commutes with arbitrary base change.

It is clear that the $\mathcal{O}_F$ action via ι on $(\operatorname{Lie} A)^+$ is $\mathcal{O}_F$-linear. To show that the action on $(\operatorname{Lie} A)/(\operatorname{Lie} A)^+$ is σ-linear, it is enough to check the case where S is an integral scheme (argue fppf locally as above). When S is integral, the σ-linearity follows from the $(n-1,1)$ signature condition on $\operatorname{Lie} A$. □

Definition 3.2.6. By the *tautological bundle* on $\mathscr{M}(n-1,1)^\circ$, we mean the rank one locally free sheaf $\mathscr{E}$ whose dual is $\mathscr{E}^\vee := (\operatorname{Lie} A)/(\operatorname{Lie} A)^+$ for $(A, \iota, \lambda) \in \mathscr{M}(n-1,1)^\circ(S)$ for $\mathcal{O}_F$-schemes S.

The restriction of $\mathscr{E}$ to $\mathscr{M}(n-1,1)^\circ[1/\Delta]$ coincides with the restriction of the tautological bundle defined in Definition 3.1.7.

Remark 3.2.7. We mention how $\mathscr{M}(n-1,1)^{\mathrm{RSZ}}$ relates to other moduli stacks. We caution that the terms "Pappas model" and "Krämer model" in the literature may refer to variants, e.g. using principal polarizations.

Let $\mathscr{M}(n-1,1)^{\mathrm{Pap}} \subseteq \mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}$ be the closed substack ("Pappas model", named for the work [Pap00]) where we impose the wedge condition (for both n even and odd) but not the spin condition.

Let $\mathscr{M}(n-1,1)^{\mathrm{Krä}}$ be the stack in groupoids over $\operatorname{Spec} \mathcal{O}_F$ ("Krämer model", named for the work [Krä03]) consisting of tuples $(A, \iota, \lambda, \mathcal{F})$ for $(A, \iota, \lambda) \in \mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}(S)$ and $\mathcal{F} \subseteq \operatorname{Lie} A$ a ι-stable local direct summand of rank $n-1$, such that the $\mathcal{O}_F$ action via ι on $\mathcal{F}$ (resp. $(\operatorname{Lie} A)/\mathcal{F}$) is $\mathcal{O}_F$-linear (resp. σ-linear).

We have a diagram

$$\begin{tikzcd}
\mathscr{M}(n-1,1)^{\mathrm{RSZ}} \times_{\mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}} \mathscr{M}(n-1,1)^{\mathrm{Krä}} \arrow[rr, hook] \arrow[d] \arrow[drr, phantom, "\lrcorner"] & & \mathscr{M}(n-1,1)^{\mathrm{Krä}} \arrow[d] \\
\mathscr{M}(n-1,1)^{\mathrm{RSZ}} \arrow[r, hook] & \mathscr{M}(n-1,1)^{\mathrm{Pap}} \arrow[r, hook] & \mathscr{M}(n-1,1)^{\mathrm{Kot},\circ}
\end{tikzcd} \tag{3.2.6}$$

where the horizontal arrows are closed immersions, the vertical arrows are forgetful, and the outer square is 2-Cartesian. The left vertical arrow is an isomorphism (by Lemma 3.2.5, i.e. $\mathcal{F} = (\operatorname{Lie} A)^+$) and the inclusion $\mathscr{M}(n-1,1)^{\mathrm{RSZ}} \hookrightarrow \mathscr{M}(n-1,1)^{\mathrm{Pap}}$ is also an open immersion. All arrows are isomorphisms after base-change to $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$.

3.3 Special cycles

Fix a non-degenerate Hermitian $\mathcal{O}_F$-lattice L of rank n, with associated moduli stack $\mathcal{M}$. The following definition of special cycles is due to Kudla–Rapoport [KR14, Definition 2.8] (there in a principally polarized situation). The notation $(\underline{x}, \underline{x})$ was explained in Section C.1.

Definition 3.3.1 (Kudla–Rapoport special cycles)**.** Given an integer $m \geq 0$, let $T \in \mathrm{Herm}_m(\mathbb{Q})$ be an $m \times m$ Hermitian matrix (with coefficients in F). The *Kudla–Rapoport (KR) special cycle* $\mathcal{Z}(T)$ is the stack in groupoids over $\operatorname{Spec} \mathcal{O}_F$ defined as follows: for schemes S over $\operatorname{Spec} \mathcal{O}_F$, we take $\mathcal{Z}(T)(S)$ to be the groupoid

$$\mathcal{Z}(T)(S) := \left\{ (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) : \begin{array}{c} (A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(S) \\ \underline{x} = [x_1, \ldots, x_m] \in \mathrm{Hom}_{\mathcal{O}_F}(A_0, A)^m \\ (\underline{x}, \underline{x}) = T \end{array} \right\}. \tag{3.3.1}$$

We sometimes refer to elements $x \in \mathrm{Hom}_{\mathcal{O}_F}(A_0, A)$ as *special homomorphisms.*

Example 3.3.2. Suppose $2 \nmid \Delta$, and consider L which is self-dual of signature $(1,1)$. Let $j \in \mathbb{Z}_{>0}$ be any positive integer. If $\mathcal{O}_F^\times = \{\pm 1\}$, consider the inclusion

$$\mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}_{\mathrm{ell}} \xrightarrow{1 \times i_{\mathrm{Serre}}} \mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(1,1)^\circ = \mathcal{M} \tag{3.3.2}$$

with i_{Serre} as in Example 3.2.2. Then $\mathcal{Z}(j) \to \mathcal{M}$ pulls back to the j-th Hecke correspondence over the left-hand side, parameterizing triples (E_0, E, w) where E_0 and E are elliptic curves, E_0 has $\mathcal{O}_F$ action of signature $(1,0)$, and $w \colon E \to E_0$ is an isogeny of degree j. This is [KR14, Proposition 14.5]. We revisit this example in Section 22.2, where we restate our main theorem in the simplest case via this description.

In the situation of Definition 3.3.1, recall $\mathrm{End}_{\mathcal{O}_F}(A_0) = \mathcal{O}_F$ (if the right-hand side is abuse of notation for global sections of the constant sheaf $\mathcal{O}_F$ on S). If the Hermitian pairing on L is $\mathcal{O}_F$-valued, we thus have $\mathcal{Z}(T) = \emptyset$ unless T has coefficients in $\mathcal{O}_F$. If L is self-dual and $2 \nmid \Delta$, we have $\mathcal{Z}(T) = \emptyset$ unless $\sqrt{\Delta} \cdot T$ has coefficients in $\mathcal{O}_F$. Positivity of the Rosati involution also implies that the special cycle $\mathcal{Z}(T)$ is empty unless T is positive semi-definite of rank $\leq n$.

By Lemma C.2.3, the forgetful map $\mathcal{Z}(T) \to \mathcal{M}$ is representable by schemes, finite, and unramified (and of finite presentation). Hence $\mathcal{Z}(T)$ is a separated Deligne–Mumford stack of finite type over $\operatorname{Spec} \mathcal{O}_F$. We will soon verify that $\mathcal{Z}(T)$ is smooth (after base change) over an explicit nonempty open subscheme of $\operatorname{Spec} \mathcal{O}_F$ (Lemma 3.5.5).

We refer to $\mathcal{Z}(T) \to \mathcal{M}$ as being a cycle over $\mathcal{M}$, although it is not literally a cycle (where the precise version of *cycle* means a formal linear combination of integral closed substacks). We also do not wish to take pushforward, which may lose information. However, since $\mathcal{Z}(T) \to \mathcal{M}$ is finite and unramified, this morphism is étale locally on the target a disjoint union of closed immersions [SProject, Lemma 04HJ]. For a more explicit version with level structure, see Lemma 3.4.5 below.

We record a few miscellaneous facts which will be used later. If $T_i \in \mathrm{Herm}_{m_i}(\mathbb{Q})$ for $i = 1, \ldots, j$ with $m \coloneqq m_1 + \cdots + m_j$ and all $m_i > 0$, then there is an identification

$$\mathcal{Z}(T_1) \times_{\mathcal{M}} \cdots \times_{\mathcal{M}} \mathcal{Z}(T_j) \cong \coprod_{\substack{T \in \mathrm{Herm}_m(\mathbb{Q}) \\ \text{satisfying (3.3.4)}}} \mathcal{Z}(T) \tag{3.3.3}$$

where the disjoint union runs over T of the form

$$T = \begin{pmatrix} T_1 & & * \\ & \ddots & \\ * & & T_j \end{pmatrix} \tag{3.3.4}$$

(i.e. whose block diagonal entries are given by the T_i).

We write $\mathcal{Z}(T)_{\mathscr{H}} \subseteq \mathcal{Z}(T)$ for the largest closed substack flat over $\operatorname{Spec} \mathcal{O}_F$, and call $\mathcal{Z}(T)_{\mathscr{H}}$ a *horizontal special cycle* or the *flat part*. There is a decomposition of $\mathcal{Z}(T)$ as a scheme-theoretic union of closed substacks[18]

$$\mathcal{Z}(T) = \mathcal{Z}(T)_{\mathscr{H}} \cup \bigcup_p \mathcal{Z}(T)_{\mathscr{V},p} \tag{3.3.5}$$

where $\mathcal{Z}(T)_{\mathscr{V},p} \coloneqq \mathcal{Z}(T) \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Z}/p^{e_p}\mathbb{Z}$ for a choice of $e_p \gg 0$ (notation e_p is temporary). This follows from quasi-compactness of $\mathcal{Z}(T)$ (which also ensures that we may take $e_p = 0$ for all but finitely many p). We think of (3.3.5) as decomposition into a "horizontal part" and "vertical parts". A similar horizontal/vertical decomposition for local special cycles on Rapoport–Zink spaces is [LZ22a, §2.9] (inert case).

While the horizontal part $\mathcal{Z}(T)_{\mathscr{H}}$ is defined canonically, the vertical parts $\mathcal{Z}(T)_{\mathscr{V},p}$ depend on e_p as above. We will mostly work with the "derived vertical special cycle classes" from Section 4.6, which do not depend on such a choice of e_p.

[18]By the *scheme-theoretic union* of finitely many closed substacks $\mathcal{Z}_i$ of a Deligne–Mumford stack $\mathcal{Z}$, we mean the closed substack whose ideal sheaf is given by intersecting the ideal sheaves of $\mathcal{Z}_i$ on the small étale site of $\mathcal{Z}$.

Given any $T \in \mathrm{Herm}_m(\mathbb{Q})$ and $\gamma \in M_{n,n}(\mathcal{O}_F)$, there is a commutative diagram

$$\begin{tikzcd}[column sep=small] (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) \arrow[rr, mapsto] & & (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x} \cdot \gamma) \\ \mathcal{Z}(T) \arrow[rr] \arrow[dr] & & \mathcal{Z}({}^t\overline{\gamma} T \gamma) \arrow[dl] \\ & \mathcal{M} & \end{tikzcd} \tag{3.3.6}$$

induced by γ. Below, we set $\mathcal{O}_{F,(p)} \coloneqq \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$.

Lemma 3.3.3. *Fix any $T \in \mathrm{Herm}_m(\mathbb{Q})$ and $\gamma \in M_{n,n}(\mathcal{O}_F)$. Consider the induced map*

$$\mathcal{Z}(T) \to \mathcal{Z}({}^t\overline{\gamma} T \gamma). \tag{3.3.7}$$

(1) This map is a finite morphism of algebraic stacks. If moreover $\gamma \in \mathrm{GL}_m(F)$ (resp. $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$) then the map is a closed immersion (resp. isomorphism).

(2) Assume $\gamma \in \mathrm{GL}_m(F)$, and let $N \in \mathbb{Z}$ be the product of primes p such that $\gamma \notin \mathrm{GL}_m(\mathcal{O}_{F,(p)})$. Then the restriction $\mathcal{Z}(T)[1/N] \to \mathcal{Z}({}^t\overline{\gamma} T \gamma)[1/N]$ is an open immersion.

Proof. (1) The map $\mathcal{Z}(T) \to \mathcal{Z}(T)({}^t\overline{\gamma} T \gamma)$ is finite because the projections to $\mathcal{M}$ are finite (finiteness for morphisms of algebraic stacks may be checked fppf locally on the target, so we reduce to the case of schemes). If $\gamma \in \mathrm{GL}_m(F)$, then $\mathcal{Z}(T) \to \mathcal{Z}(T)({}^t\overline{\gamma} T \gamma)$ is a monomorphism of algebraic stacks (check via the moduli description), and any proper monomorphism of algebraic stacks is a closed immersion. If $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$, there is an inverse $\mathcal{Z}({}^t\overline{\gamma} T \gamma) \to \mathcal{Z}(T)$ sending $\underline{x} \mapsto \underline{x} \cdot \gamma^{-1}$.

(2) Consider the substack $\mathcal{Z} \subseteq \mathcal{Z}({}^t\overline{\gamma} T \gamma)$ consisting of tuples $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{w})$ such that $\underline{w} \cdot \gamma^{-1} \in \mathrm{Hom}_{\mathcal{O}_F}(A_0, A)^m$ (i.e. that the tuple of quasi-homomorphisms $\underline{w} \cdot \gamma^{-1}$ is a tuple of homomorphisms). The closed immersion $\mathcal{Z}(T) \to \mathcal{Z}({}^t\overline{\gamma} T \gamma)$ maps isomorphically onto $\mathcal{Z}$.

A quasi-homomorphism $f \colon B \to B'$ of abelian schemes (over any base scheme S) is a homomorphism if and only the induced quasi-homomorphisms of p-divisible groups $f[p^\infty] \colon B[p^\infty] \to B'[p^\infty]$ are homomorphisms for all primes p. This is a closed condition on S for each prime p (e.g. the quasi-homomorphism version of [RZ96, Proposition 2.9]). This is also an open condition for any prime p which is invertible on S (by étaleness of the p-divisible groups). If p is a prime such that $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)})$, then the tuple $\underline{w} \cdot \gamma^{-1}$ induces a tuple of quasi-homomorphisms $A_0[p^\infty] \to A[p^\infty]$ consisting of homomorphisms. $\square$

3.4 Level structure

Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n, and form the associated moduli stack $\mathcal{M}$. We discuss level structure for $\mathcal{M}$.

Set $V = L \otimes_{\mathcal{O}_F} F$, and form the unitary group $U(V)$ (over $\operatorname{Spec}\mathbb{Q}$). Let $L_0 = \mathcal{O}_F$ any self-dual Hermitian $\mathcal{O}_F$-lattice of rank 1. We set

$$K_{L_0,p} \coloneqq \operatorname{Stab}_{U(V_0)(\mathbb{Q}_p)}(L_0 \otimes \mathbb{Z}_p) \qquad K_{L,p} \coloneqq \operatorname{Stab}_{U(V)(\mathbb{Q}_p)}(L \otimes \mathbb{Z}_p)$$
$$K_{L_0,f} = \prod_p K_{L_0,p} \qquad K_{L,f} = \prod_p K_{L,p}$$

for all p, where $\operatorname{Stab}_{U(V)(\mathbb{Q}_p)}(L \otimes \mathbb{Z}_p)$ denotes the stabilizer of $L \otimes \mathbb{Z}_p$ in $U(V)(\mathbb{Q}_p)$, etc.. We say that $K_{L,f} \subseteq U(V)(\mathbb{A}_f)$ is the *adèlic stabilizer* of L. Set $K'_{L,f} = K_{L_0,f} \times K_{L,f}$. Note that there is no dependence (up to functorial isomorphism) on the choice of L_0, or the choice of L within its adèlic isomorphism class. We use the usual notation where $K^p_{L,f}$ means to omit the p-th factor in the product, etc..

For integers $N \geq 1$, we define the "principal congruence subgroups"

$$K_p(N) \coloneqq \ker(K_{L,p} \to \mathrm{GL}(L \otimes \mathbb{Z}_p/N\mathbb{Z}_p)) \qquad K_f(N) = \prod_p K_{L,p}(N)$$

(suppressing L dependence from notation) and similarly define $K_{0,p}(N_0)$ and $K_f(N_0)$ for $N_0 \geq 1$. Given a pair $N' = (N_0, N)$ of integers $N_0, N \geq 1$, we set $K'_f(N') \coloneqq K_{0,f}(N_0) \times K_f(N)$ and $K'_p(N') \coloneqq K_{0,p}(N_0) \times K_p(N)$, etc.. Given $N' = (N_0, N)$, we sometimes abuse notation, e.g. $N' \geq a$ means $N_0, N \geq a$, and the notation $\mathcal{X}[1/N']$ for an algebraic stack $\mathcal{X}$ will mean inverting all primes dividing N'. If $N_0 = N$, we write $K'_f(N) \coloneqq K'_f(N')$.

Let $K'_f = K_{0,f} \times K_f \subseteq K'_{L,f}$ be any open compact subgroup which admits product factorizations $K_{0,f} = \prod_p K_{0,p}$ and $K_f = \prod_p K_p$. We set $K'_p \coloneqq K_{0,p} \times K_p$, etc.. Let $N_{K'_f}$ be the product of primes p for which $K_{0,p} \neq K_{L_0,p}$ or $K_p \neq K_{L,p}$. We say that K'_f is *standard* at p if $p \nmid N_{K'_f}$.

Notation 3.4.1. For $K'_f \subseteq K'_{L,f}$ as above, we reserve the term *small* or *small level* to mean that $K'_f \subseteq K'_f(N')$ for some $N' \geq 3$.

Consider $\alpha = (A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(S)$ for some scheme S. Suppose p is a prime which is invertible on S. For any integer $e \geq 0$, we consider the fppf sheaf

$$\underline{\mathrm{Level}}(p^e) \subseteq \underline{\mathrm{Isom}}(A_0[p^e], L_0 \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z}) \times \underline{\mathrm{Isom}}(A[p^e], (L_0 \otimes_{\mathcal{O}_F} L) \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z}) \tag{3.4.1}$$

on S, where $\underline{\mathrm{Level}}(p^e)$ is the open and closed subfunctor corresponding to isomorphisms $A_0[p^e] \to L_0/p^e L_0$ and $A[p^e] \to L/p^e L$ which lift to $\mathcal{O}_F$-linear isomorphisms

$$T_p(A_{0,\overline{s}}) \xrightarrow{\sim} L_0 \qquad \text{unitary up to scalar}$$
$$\operatorname{Hom}_{\mathcal{O}_F}(T_p(A_{0,\overline{s}}), T_p(A_{\overline{s}})) \xrightarrow{\sim} L \qquad \text{unitary}$$

over every geometric point $\overline{s}$ of S. Since L_0 is rank one, the "unitary up to scalar" condition is automatic.

Consider any open compact subgroup $K'_p \subseteq K'_{L,p}$. For each $e \geq 0$, we write K'_p mod p^e (temporary notation) for the image of K'_p in $\mathrm{GL}_1(L_0 \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z}) \times \mathrm{GL}_n(L \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z})$. There is a canonical action of K'_p mod p^e on $\underline{\mathrm{Level}}(p^e)$.

Definition 3.4.2 (Level structure)**.** Let $K'_f \subseteq K'_{L,p}$ be any factorizable open compact subgroup, as above. Consider an object $\alpha = (A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(S)$ for some scheme $S \to \operatorname{Spec} \mathcal{O}_F[1/N_{K'_f}]$.

If p is prime which is invertible on S, the *sheaf of level K'_p structures for α* is the quotient (coequalizer)[19]

$$\underline{\mathrm{Level}}_{K'_p} := (K'_p \bmod p^e)_{e \geq 0} \backslash (\underline{\mathrm{Level}}(p^e))_{e \geq 0} \tag{3.4.2}$$

in the category of pro-objects for the category of fppf sheaves on S. If p is not invertible on S, we let $\underline{\mathrm{Level}}_{K'_p}$ be the constant sheaf valued in a singleton set.

The *sheaf of level K'_f structures for α* is the product

$$\underline{\mathrm{Level}}_{K'_f} := \prod_p \underline{\mathrm{Level}}_{K'_p} \tag{3.4.3}$$

over all primes p. This is an fppf sheaf on S, and is locally constant in the étale topology.

A K'_p *level structure* (rep. K'_f *level structure*) for α is a global section of $\underline{\mathrm{Level}}_{K'_p}$ (resp. $\underline{\mathrm{Level}}_{K'_f}$).

Remark 3.4.3. Let K'_f, α, and S be as in Definition 3.4.2. Fix a geometric point $\overline{s}$ of S, with $\mathrm{char}(\overline{s}) = p \geq 0$. Assume moreover that S is connected. In this case, giving a K'_f-level structure for α is (canonically) same as giving a pair $(\tilde{\eta}_0, \tilde{\eta})$ where $\tilde{\eta}_0$ (resp. $\tilde{\eta}$) is a $\pi_{1,\text{ét}}(S, \overline{s})$-stable $K^p_{0,f}$-orbit (resp. K^p_f-orbit) of isomorphisms

$$\begin{aligned} \eta_0 \colon T^p(A_{0,\overline{s}}) &\xrightarrow{\sim} L_0 \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p & &\text{unitary up to scalar} \\ \eta \colon \operatorname{Hom}_{\mathcal{O}_F \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}_p}(T^p(A_{0,\overline{s}}), T^p(A_{\overline{s}})) &\xrightarrow{\sim} L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p & &\text{unitary.} \end{aligned}$$

In the notation of Definition 3.4.2, note that $K'_p = K'_p(p^e)$ implies $\underline{\mathrm{Level}}_{K'_p} = \underline{\mathrm{Level}}(p^e)$. In Remark 3.4.3, note that the "unitary up to scalar" condition is automatic because L_0 has rank 1. Even when S is not connected, we abuse notation and write $(\tilde{\eta}_0, \tilde{\eta})$ for level K'_f structure in the sense of Section 3.4

[19]Note that this quotient $\underline{\mathrm{Level}}_{K'_p}$ is (isomorphic to) a sheaf (not just a pro-sheaf), so it is sensible to refer to $\underline{\mathrm{Level}}_{K'_p}$ an fppf sheaf on S (rather than a pro-object).

Given an open compact K'_f as in Definition 3.4.2, we now define a stack in groupoids $\mathcal{M}_{K'_f}$ over $\operatorname{Spec} \mathcal{O}_F[1/(d_L N_{K'_f})]$ with

$$\mathcal{M}_{K'_f}(S) \coloneqq \{(\alpha, \tilde{\eta}_0, \tilde{\eta}) : \alpha \in \mathcal{M}(s) \text{ and } (\tilde{\eta}_0, \tilde{\eta}) \text{ a level } K'_f \text{ structure for } \alpha\} \tag{3.4.4}$$

for schemes S over $\operatorname{Spec} \mathcal{O}_F[1/(d_L N_{K'_f})]$. Given $T \in \mathrm{Herm}_m(\mathbb{Q})$, we write $\mathcal{Z}(T)_{K'_f} \coloneqq \mathcal{Z}(T) \times_{\mathcal{M}} \mathcal{M}_{K'_f}$ ("level K'_f special cycle").

Write $\mathcal{A}_{n,d,N}$ for the moduli stack over $\operatorname{Spec} \mathcal{O}_F[1/N]$ of (relative) n-dimensional abelian schemes A with degree d polarization and a chosen isomorphism $A[N] \to L \otimes \mathbb{Z}/N\mathbb{Z}$ of group schemes (not necessarily compatible with symplectic pairings). We similarly form $\mathcal{A}_{1,1,N_0}$ using the lattice L_0 (and pick a basis of L_0 for convenience). Recall that $\mathcal{A}_{n,d,N}$ is representable by a separated Deligne–Mumford stack of finite type over $\operatorname{Spec} \mathcal{O}_F[1/N]$, and that $\mathcal{A}_{n,d,N}$ is a scheme quasi-projective over $\operatorname{Spec} \mathcal{O}_F[1/N]$ if $N \geq 3$ (see [MFK94, §7.2 Theorem 7.9]).

Let $b_L, d_L \in \mathbb{Z}_{>0}$ be associated to L, as discussed before Definition 3.1.2. If $K'_f(N')$ is the principal congruence subgroup of some level $N' = (N_0, N)$, consider the forgetful morphism

$$\mathcal{M}_{K'_f(N')} \longrightarrow \mathcal{A}_{1,1,N_0} \times \mathcal{A}_{n,d_L,N} \tag{3.4.5}$$

which forgets the $\mathcal{O}_F$-actions and sends $\lambda \mapsto b_L \lambda$. For level structure, see the description in Eq. (3.4.1). The induced map

$$\mathcal{M}_{K'_f(N')} \to \mathcal{M}[1/N'] \times_{(\mathcal{A}_{1,1} \times \mathcal{A}_{n,d_L})} (\mathcal{A}_{1,1,N_0} \times \mathcal{A}_{n,d_L,N})[1/d_L] \tag{3.4.6}$$

is representable by schemes and is an open and closed immersion. Hence $\mathcal{M}_{K'_f(N')} \to \mathcal{A}_{1,1,N_0} \times \mathcal{A}_{n,d_L,N}[1/d_L]$ is finite (and representable by schemes).

Lemma 3.4.4.

(1) *For any open compact subgroup K'_f as in Definition 3.4.2, the stack $\mathcal{M}_{K'_f}$ is a separated Deligne–Mumford stack of finite type over* $\operatorname{Spec} \mathcal{O}_F$. *If K'_f is small, then $\mathcal{M}_{K'_f}$ is a quasi-projective scheme over* $\operatorname{Spec} \mathcal{O}_F$.

(2) *For any inclusion $K'_f \subseteq K''_f$, the forgetful morphism $\mathcal{M}_{K'_f} \to \mathcal{M}_{K''_f}[1/N_{K'_f}]$ (i.e. expand a K'_f-orbit to a K''_f orbit) is finite étale of degree $|K''_f/K'_f|$. If $K'_f \subseteq K''_f$ is a normal subgroup, then $\mathcal{M}_{K'_f} \to \mathcal{M}_{K''_f}[1/N_{K'_f}]$ is a torsor for the finite discrete group K''_f/K'_f.*

Proof. The second sentence in part (2) is clear from construction (and makes sense before we know these stacks are algebraic). When $K'_f = K'_f(N')$ for some N', the claims in part (1) follow from (3.4.6).

For general K'_f, select $N' = (N_0, N)$ such that $\mathcal{M}_{K'_f(N')}$ is a scheme and $K'_f(N') \subseteq K'_f$. Then $\mathcal{M}_{K'_f(N')} \to \mathcal{M}_{K'_f}[1/N']$ is a torsor for the finite discrete group $K'_f/K'_f(N')$ (in particular, finite étale), and hence admits the stack quotient presentation $\mathcal{M}_{K'_f}[1/N'] \cong [\mathcal{M}_{K'_f(N')}/(K'_f/K'_f(N'))]$, which shows that $\mathcal{M}_{K'_f}[1/N']$ is Deligne–Mumford. Picking another $M' = (M_0, M)$ such that $\gcd(N_0 N, M_0 M)$ is only divisible by primes dividing $N_{K'_f}$, we find that $\mathcal{M}_{K'_f}[1/M']$ is Deligne–Mumford as well. These two charts show that $\mathcal{M}_{K'_f}$ is Deligne–Mumford, as well as separated and finite type over $\operatorname{Spec} \mathcal{O}_F$.

If $K'_f \subseteq K''_f$, then for any scheme S with a morphism $S \to \mathcal{M}_{K''_f}$, the 2-fiber product $\mathcal{M}_{K'_f} \times_{\mathcal{M}_{K''_f}} S$ is fibered in setoids, hence equivalent to a sheaf (of sets). But since $\mathcal{M}_{K'_f(N')} \to \mathcal{M}_{K'_f}[1/N']$ is a $K'_f/K'_f(N')$-torsor and affine morphisms satisfy fpqc descent [SProject, Section 0244], we conclude that $\mathcal{M}_{K'_f}[1/N'] \times_{\mathcal{M}_{K''_f}} S$ is represented by a scheme. As above, we may pick some other M' to patch and show that the morphism $\mathcal{M}_{K'_f} \to \mathcal{M}_{K''_f}[1/N_{K'_f}]$ is representable by schemes. Since $\mathcal{M}_{K'_f(N')} \to \mathcal{M}_{K'_f}[1/N']$ and $\mathcal{M}_{K'_f(N')} \to \mathcal{M}_{K''_f}[1/N']$ are both finite étale surjections, we conclude that $\mathcal{M}_{K'_f} \to \mathcal{M}_{K''_f}[1/N_{K'_f}]$ is finite étale by varying N' again (using standard facts like [SProject, Lemma 02K6, Lemma 01KV, Lemma 0AH6, Lemma 02LS]). The remaining claims follow from this. $\square$

Lemma 3.4.5. *Fix any prime p and a matrix $T \in \operatorname{Herm}_m(\mathbb{Q})$ with $m \geq 0$. The morphism $\mathcal{Z}(T)_{K'_f(p^e)} \to \mathcal{M}_{K'_f(p^e)}$ is a disjoint union of closed immersions for all $e \gg 0$.*

Proof. For $e \in \mathbb{Z}_{\geq 0}$, we define a stack (used only in this proof) $\mathcal{M}(p^e)$ over $\operatorname{Spec} \mathcal{O}_F[1/(d_L p)]$ as follows. For schemes S over $\operatorname{Spec} \mathcal{O}_F[1/(d_L p)]$, we take $\mathcal{M}(p^e)(S)$ to be the groupoid

$$\mathcal{M}(p^e)(S) := \left\{ (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) : \begin{array}{c} (A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{M}(S) \\ \underline{x} = [x_1, \ldots, x_m] \in \operatorname{Hom}_{\mathcal{O}_F}(A_0[p^e], A[p^e])^m \end{array} \right\}. \tag{3.4.7}$$

We have a commutative diagram

$$\begin{tikzcd} & \mathcal{M}(p^e) \arrow[d] \\ \mathcal{Z}(T)[1/p] \arrow[ur] \arrow[r] & \mathcal{M} . \end{tikzcd} \tag{3.4.8}$$

The forgetful morphism $\mathcal{M}(p^e) \to \mathcal{M}[1/p]$ is representable by schemes and a finite étale surjection. Thus, $\mathcal{M}(p^e)$ is representable by a separated Deligne–Mumford stack of finite type over $\operatorname{Spec} \mathcal{O}_F$.

We claim that $\mathcal{Z}(T) \to \mathcal{M}(p^e)$ is a closed immersion for $e \gg 0$. This may be checked fppf locally on the target. Suppose $S \to \mathcal{M}[1/p]$ is an fppf cover by a Noetherian scheme

S (possible since $\mathcal{M}$ is locally Noetherian and quasi-compact). It is enough to check that $\mathcal{Z}(T) \times_{\mathcal{M}} S \to \mathcal{M}(p^e) \times_{\mathcal{M}} S$ is a closed immersion of schemes. Since the morphism $\mathcal{Z}(T) \to \mathcal{M}$ (resp. $\mathcal{M}(p^e) \to \mathcal{M}[1/p]$) is finite and unramified (resp. finite), we conclude that $\mathcal{Z}(T)[1/p] \to \mathcal{M}(p^e)$ is also finite and unramified.

To prove the claim, it remains only to check that the morphism of schemes $\mathcal{Z}(T) \times_{\mathcal{M}} S \to \mathcal{M}(p^e) \times_{\mathcal{M}} S$ is universally injective for $e \gg 0$ (for morphisms of schemes, being a closed immersion is the same as being proper, unramified, and universally injective [SProject, Lemma 04XV]).

We first show that universal injectivity holds fiber-wise over every point $s \in S$ for e sufficiently large (with e possibly depending on s). For any point s on S with residue field $k(s)$, we know that $\mathcal{Z}(T)_{k(s)} \to \mathcal{M}(p^e)_{k(s)}$ is universally injective for $e \gg 0$ (possibly depending on s) because the map $\mathrm{Hom}(A_1, A_2) \to \mathrm{Hom}(T_p(A_1), T_p(A_2))$ is injective for abelian varieties A_1, A_2 over any field of characteristic $\neq p$ (apply this over a geometric point mapping to s and use finiteness of $\mathcal{Z}(T)$).

Being universally injective may be checked fiber-wise over S, so we need to show that there is a value of e which works for all points $s \in S$ simultaneously. We can select $e \gg 0$ so that $\mathcal{Z}(T) \times_{\mathcal{M}} S \to \mathcal{M}(p^e) \times_{\mathcal{M}} S$ is universally injective (hence a closed immersion) over the generic point of each irreducible component of S. For such e, a limiting argument ("spreading out") implies that $\mathcal{Z}(T) \times_{\mathcal{M}} S \to \mathcal{M}(p^e) \times_{\mathcal{M}} S$ is a closed immersion over an open dense subset of S. Applying Noetherian induction on S proves the claim.

To finish the proof of the lemma, we observe that $\mathcal{M}(p^e) \times_{\mathcal{M}} \mathcal{M}_{K'_f(p^e)} \to \mathcal{M}_{K'_f(p^e)}$ is a finite disjoint union of isomorphisms, corresponding to the constant sheaf valued in $\mathrm{Hom}_{\mathcal{O}_F}(L_0 \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z}, L \otimes_{\mathbb{Z}} \mathbb{Z}/p^e\mathbb{Z})^m$. Hence $\mathcal{Z}(T)_{K'_f(p^e)} \to \mathcal{M}_{K'_f(p^e)}$ is a disjoint union of closed immersions. $\square$

3.5 Generic smoothness

We explain a generic smoothness result for special cycles (Lemma 3.5.5). The other lemmas are auxiliary. The proof proceeds by reducing to p-divisible groups over a base where p is locally nilpotent, and then checking formal smoothness using Serre–Tate and Grothendieck–Messing deformation theory.

We first consider p-divisible groups (see Appendix B.1 for a review of terminology). For primes p, set $\mathcal{O}_{F_p} \coloneqq \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Suppose $p \nmid \Delta$ and consider schemes S over $\mathrm{Spf}\, \mathcal{O}_{F_p}$, i.e. S is a scheme over $\mathrm{Spec}\, \mathcal{O}_F$ with p locally nilpotent on S. We consider tuples (Y, ι, λ) over S

where

$$
\begin{array}{lll}
Y & \text{is a } p\text{-divisible group over } S \text{ of height } 2n \text{ and dimension } n & \\
\iota \colon \mathcal{O}_{F_p} \to \operatorname{End}(Y) & \text{is an action satisfying the } (n-r,r) \text{ Kottwitz signature condition, i.e. for all } a \in \mathcal{O}_{F_p}, \text{ the characteristic polynomial of } \iota(a) \text{ acting on } \operatorname{Lie} Y \text{ is } (x-a)^{n-r}(x-a^{\sigma})^r \in \mathcal{O}_S[x] & (3.5.1) \\
\lambda \colon Y \xrightarrow{\sim} Y^{\vee} & \text{is a principal polarization whose Rosati involution } \dagger \text{ on } \operatorname{End}(Y) \text{ satisfies } \iota(a)^{\dagger} = \iota(a^{\sigma}) \text{ for all } a \in \mathcal{O}_{F_p}. &
\end{array}
$$

In the signature condition described above, we view $\mathcal{O}_S$ as an $\mathcal{O}_{F_p}$-algebra via the structure map $S \to \operatorname{Spf} \mathcal{O}_{F_p}$.

Parts of the next formal smoothness result (Lemma 3.5.1) may exist in some form in the literature, see e.g. discussion about formal smoothness for "unramified Rapoport–Zink data" in [RZ96, 3.82] and the reference to [Kot92, §5] given there.

Following [SProject, Section 04EW], we use the term *thickening* to refer to a closed immersion which is a homeomorphism on underlying topological spaces, and the term *first order thickening* for a thickening defined by a square zero ideal.

Let S be a scheme over $\operatorname{Spf} \mathcal{O}_{F_p}$, and suppose (Y, ι, λ) is a tuple over S as in (3.5.1). There is an associated deformation functor $\operatorname{Def}_{(Y,\iota,\lambda)}$ (possibly non-standard usage, and it will not appear after Lemma 3.5.1) which sends a thickening $S \to S'$ to the set of (isomorphism classes of) lifts of (Y, ι, λ) to S'. Write $S[\epsilon]$ and $S[\epsilon, \epsilon']$ as shorthand for $S \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Z}[\epsilon]/(\epsilon^2)$ and $S \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Z}[\epsilon, \epsilon']/(\epsilon^2, \epsilon\epsilon', \epsilon'^2)$, respectively. In the proof of Lemma 3.5.1, we will see that the canonical map

$$
\operatorname{Def}_{(Y,\iota,\lambda)}(S[\epsilon, \epsilon']) \to \operatorname{Def}_{(Y,\iota,\lambda)}(S[\epsilon]) \times_{\operatorname{Def}_{(Y,\iota,\lambda)}(S)} \operatorname{Def}_{(Y,\iota,\lambda)}(S[\epsilon']) \tag{3.5.2}
$$

is an isomorphism. More generally, if M is a finite rank free $\mathcal{O}_S$-module and $\mathcal{O}_S \oplus M$ denotes the quasi-coherent $\mathcal{O}_S$-algebra with M an ideal of square zero, we will see that the functor $M \mapsto \operatorname{Def}_{(Y,\iota,\lambda)}(\underline{\operatorname{Spec}}_S(\mathcal{O}_S \oplus M))$ preserves fiber products over the base $M = 0$ (note that this holds when $\operatorname{Def}_{(Y,\iota,\lambda)}$ is replaced by any scheme, and this is essentially the method of proof). Here $\underline{\operatorname{Spec}}_S$ denotes relative Spec.

For any scheme S over $\operatorname{Spf} \mathcal{O}_{F_p}$, the above considerations imply that the set $\operatorname{Def}_{(Y,\iota,\lambda)}(S[\epsilon])$ has the natural structure of a $\Gamma(S, \mathcal{O}_S)$-module in the standard way (as a "tangent space") as in [SGA3II, Proposition 3.6] or [SProject, Section 06I2].

Lemma 3.5.1. *Let p be a prime which is unramified in $\mathcal{O}_F$. The deformation problem for triples as in* (3.5.1) *is formally smooth of relative dimension $(n-r)r$ in the following sense. Let S be any scheme over $\operatorname{Spf} \mathcal{O}_{F_p}$, and let (Y, ι, λ) be a triple over S as in* (3.5.1)*.*

(1) The triple (Y, ι, λ) lifts along any first order thickening of affine schemes $S \to S'$, i.e. the map $\operatorname{Def}_{(Y,\iota,\lambda)}(S') \to \operatorname{Def}_{(Y,\iota,\lambda)}(S)$ is surjective.

(2) When $S = \operatorname{Spec} \kappa$ for a field κ, the κ vector space $\operatorname{Def}_{(Y,\iota,\lambda)}(\kappa[\epsilon])$ has dimension $(n-r)r$.

If $(n-r)r = 0$, then (Y, ι, λ) lifts uniquely along any first order thickening of schemes $S \to S'$.

Proof. We study this lifting problem for p-divisible groups in terms of Grothendieck–Messing deformation theory. Let $S \to S'$ be a first order thickening of schemes (not necessarily affine). View $S \hookrightarrow S'$ as a PD thickening, with trivial PD structure on the square zero ideal of the thickening.

Write $\mathbb{D}(Y)$ for the covariant Dieudonné crystal of Y, and write $\mathbb{D}(Y)(S)$ and $\mathbb{D}(Y)(S')$ for the evaluation of this crystal on the PD thickenings $\mathrm{id}\colon S \to S$ and $S \hookrightarrow S'$ respectively. We have a short exact sequence of $\mathcal{O}_S$-modules given by the Hodge filtration

$$0 \to \Omega_{Y^\vee} \to \mathbb{D}(Y)(S) \to \operatorname{Lie}_Y \to 0 \tag{3.5.3}$$

with $\Omega_{Y^\vee} = (\operatorname{Lie}_{Y^\vee})^\vee$ and each $\mathcal{O}_S$-module above being finite locally free.

We may decompose the Hodge filtration into eigenspaces with respect to the action $\iota\colon \mathcal{O}_{F_p} \to \operatorname{End}(Y)$ (and the structure morphism $S \to \operatorname{Spec} \mathcal{O}_{F_p}$). We use superscripts $(-)^+$ and $(-)^-$ to denote these eigenspaces, where $\mathcal{O}_{F_p}$ acts linearly (resp. σ-linearly) on $(-)^+$ (resp. $(-)^-$) via ι. Then we have short exact sequences

$$0 \to \Omega^+_{Y^\vee} \to \mathbb{D}(Y)(S)^+ \to \operatorname{Lie}^+_Y \to 0 \tag{3.5.4}$$
$$0 \to \Omega^-_{Y^\vee} \to \mathbb{D}(Y)(S)^- \to \operatorname{Lie}^-_Y \to 0 \tag{3.5.5}$$

where each $\mathcal{O}_S$-module above is finite locally free and, for example, we have $\mathbb{D}(Y) = \mathbb{D}(Y)^+ \oplus \mathbb{D}(Y)^-$. From left to right, the modules in (3.5.4) have ranks r, n, and $n-r$, and the modules in (3.5.5) have ranks $n-r$, n, and r respectively.

Using the polarization λ, we may identify (3.5.4) with the dual of (3.5.5). There is a choice of sign in this identification, which plays essentially no role in this proof.

We have $\mathbb{D}(Y)(S')|_S \cong \mathbb{D}(Y)(S)$ canonically (as $\mathbb{D}(Y)$ is a crystal), and Grothendieck–Messing theory implies that lifting (Y, ι, λ) to S' is the same as lifting the Hodge filtration (3.5.3) compatibly with the action ι and the polarization λ. Compatibility with the ι action means that we should lift the eigenspace decomposition in (3.5.4) and (3.5.5), and compatibility with the polarization λ means that the resulting exact sequences should again be dual to each other (as determined by λ). It is equivalent to lift either one of the exact sequences of (3.5.4) and (3.5.5) (one determines the other) to a filtration of $\mathbb{D}(Y)(S')^+$ or $\mathbb{D}(Y)(S')^-$ respectively (with no additional restrictions).

Consider the lifting problem for, say, the + eigenspace of the Hodge filtration as in (3.5.4). Zariski locally on S', this lifting problem may be identified with the problem of lifting an S point to an S' point on the Grassmannian parametrizing rank r subbundles of the rank n trivial bundle. This Grassmannian is smooth of relative dimension $(n-r)r$, which proves the claims in the lemma statement. $\square$

The next three lemmas are used to prove Lemma 3.5.4. This latter lemma is in turn used in the proof of generic smoothness in Lemma 3.5.5, to reduce to bases where p is locally nilpotent for some unramified prime p. This will allow us to reduce to formal smoothness for deformations of p-divisible groups (with certain additional structure) as proved in Lemma 3.5.1.

Lemma 3.5.2. *Let A be an adic Noetherian ring, and let X be a locally Noetherian scheme over* $\operatorname{Spec} A$*. If $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ is flat, then $X \to \operatorname{Spec} A$ is flat at every point of X which lies over* $\operatorname{Spf} A$*. If $X \to \operatorname{Spec} A$ is locally of finite type, then the same holds with "flat" replaced by "smooth".*

Proof. Here, flatness (resp. smoothness) of $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ is equivalent to the requirement that, for every scheme T with a map $T \to \operatorname{Spf} A$, the base changed map $X_T \to T$ is flat (resp. smooth).

We first check the flatness assertion. Passing to an affine open of X, we may reduce to the case where $X = \operatorname{Spec} B$ for a Noetherian ring B. Let $I \subseteq A$ be an ideal of definition. Then $X_{\operatorname{Spf} A}$ is described by a completed tensor product, and we have $X_{\operatorname{Spf} A} = \hat{B}$ where $\hat{B}$ is the I-adic completion of B. Since B is a Noetherian ring, the canonical map $B \to \hat{B}$ is flat. Since $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ is flat, we know that $A \to \hat{B}$ is a flat ring map. We conclude that B is flat over A at every prime in the image of $\operatorname{Spec} \hat{B} \to \operatorname{Spec} B$. These are precisely the points of X lying over $\operatorname{Spf} A$.

Next, assume $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ is smooth. By Noetherianity of A, the map $f \colon X \to \operatorname{Spec} A$ is locally of finite presentation. We have just shown that $X \to \operatorname{Spec} A$ is flat at every point $x \in X$ which lies over $\operatorname{Spf} A$. Thus, for such $x \in X$, the map $f \colon X \to \operatorname{Spec} A$ is smooth at x if and only if $X_{f(x)} \to \operatorname{Spec} k(f(x))$ is smooth at x, where $k(f(x))$ denotes the residue field of $f(x)$. But since $\operatorname{Spec} k(f(x)) \to \operatorname{Spec} A$ factors through $\operatorname{Spf} A \to \operatorname{Spec} A$, we conclude that $X_{f(x)} \to \operatorname{Spec} k(f(x))$ is indeed smooth. $\square$

Lemma 3.5.3. *Let $f\colon X \to Y$ be a locally of finite type (resp. locally of finite presentation) morphism of schemes, and assume that Y is a Jacobson scheme. Then f is smooth (resp. flat) if and only if f is smooth (resp. flat) at every point of X which lies over a closed point of Y.*

Proof. Since f is locally of finite type (resp. locally of finite presentation), we know that X is a Jacobson scheme (i.e. closed points are dense in every closed subset). Since f is smooth (resp. flat) on an open subset of X, it is enough to check that f is smooth (resp. flat) at every closed point of X. As f is locally of finite type and Y is Jacobson, we know that f maps closed points to closed points [SProject, Lemma 01TB] which gives the lemma claim. □

Lemma 3.5.4. *Let $\mathcal{X}$ be an algebraic stack, let Y be a Jacobson locally Noetherian scheme, and let $f\colon \mathcal{X} \to Y$ be a morphism which is locally of finite type. For points $y \in Y$, write $\widehat{\mathcal{O}}_{Y,y}$ for the completion of the local ring at y. Then $\mathcal{X} \to Y$ is smooth (resp. flat) if and only if $\mathcal{X}_{\operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}} \to \operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}$ is smooth (resp. flat) for every closed point $y \in Y$.*

Proof. Select any scheme U with a surjective smooth morphism $U \to \mathcal{X}$. Then $U \to Y$ is a locally of finite type morphism of Jacobson locally Noetherian schemes, and $\mathcal{X} \to Y$ is smooth (resp. flat) if and only if $U \to Y$ is smooth (resp. flat). By Lemma 3.5.3, we may check smoothness (resp. flatness) of $U \to Y$ at points of U lying over closed points of Y. If $x \in U$ and $y = f(x)$, then $U \to Y$ is smooth (resp. flat) at x if and only if $U_{\operatorname{Spec} \widehat{\mathcal{O}}_{Y,y}} \to \operatorname{Spec} \widehat{\mathcal{O}}_{Y,y}$ is smooth at x (first checking flatness, then checking smoothness in the fiber over the closed point). For any $y \in Y$, Lemma 3.5.2 implies that $U \to Y$ is smooth (resp. flat) at all points $x \in U$ lying over y if and only if $U_{\operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}} \to \operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}$ is smooth (resp. flat). By Lemma 3.5.3, we then see that $U \to Y$ is smooth (resp. flat) if and only if $U_{\operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}} \to \operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}$ is smooth (resp. flat) for every closed point $y \in Y$. This is equivalent to the condition that $\mathcal{X}_{\operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}} \to \operatorname{Spf} \widehat{\mathcal{O}}_{Y,y}$ is smooth (resp. flat) for all closed points $y \in Y$, since $U \to \mathcal{X}$ is a smooth surjection. □

Lemma 3.5.5. *Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice, with associated moduli stack $\mathcal{M}$. Fix $T \in \operatorname{Herm}_m(F)$.*

Then there exists $N \in \mathbb{Z}$ such that $\mathcal{Z}(T)[1/(Nd_L\Delta)]$ is either empty[20] *or smooth of relative dimension $(n - r - \operatorname{rank}(T))r$ over $\operatorname{Spec} \mathcal{O}_F[1/(Nd_L\Delta)]$.*

We may take N such that for $p \nmid Nd_L\Delta$, there exists $g \in \mathrm{GL}_m(\mathcal{O}_{F_p})$ with

$$ {}^t\overline{g}Tg = \begin{pmatrix} \mathrm{Id}_{\operatorname{rank}(T)} & 0 \\ 0 & 0 \end{pmatrix} \tag{3.5.6}$$

[20]Following the Stacks project [SProject, Definition 0055], our convention is that $\dim \emptyset = -\infty$.

where ${}^t\overline{g}$ denotes the conjugate transpose of g.

Proof. Fix a prime $p \nmid d_L\Delta$ such that there exists $g \in \mathrm{GL}_m(\mathcal{O}_{F_p})$ satisfying (3.5.6). By Lemma 3.5.4, it is enough to check that the base change $\mathcal{Z}(T)_{\mathrm{Spf}\,\mathcal{O}_{F_p}}$ is either empty or smooth of relative dimension $(n-r-\mathrm{rank}(T))r$ over $\mathrm{Spf}\,\mathcal{O}_{F_p}$.

The morphism $\mathcal{Z}(T)_{\mathrm{Spf}\,\mathcal{O}_{F_p}} \to \mathrm{Spf}\,\mathcal{O}_{F_p}$ is representable by algebraic stacks and locally of finite presentation. Thus $\mathcal{Z}(T)_{\mathrm{Spf}\,\mathcal{O}_{F_p}} \to \mathrm{Spf}\,\mathcal{O}_{F_p}$ is smooth if and only if it is formally smooth [SProject, Lemma 0DP0].

Let $S \to S'$ be a first order thickening of affine schemes, and assume S' is equipped with a morphism to $\mathrm{Spf}\,\mathcal{O}_{F_p}$. To check formal smoothness of $\mathcal{Z}(T)_{\mathrm{Spf}\,\mathcal{O}_{F_p}} \to \mathrm{Spf}\,\mathcal{O}_{F_p}$, we need to show that every object $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) \in \mathcal{Z}(T)(S)$ admits a lift to S'.

Form $(X_0, \iota_0, \lambda_0)$, where $X_0 = A_0[p^\infty]$ is the p-divisible group of A_0 with induced action $\iota_0 \colon \mathcal{O}_{F_p} \to \mathrm{End}(X_0)$ and principal polarization $\lambda_0 \colon X_0 \to X_0^\vee$. Similarly associate (X, ι, λ) to (A, ι, λ), where $X = A[p^\infty]$ is the p-divisible group of A. Note that the polarization $\lambda \colon X \to X^\vee$ is principal because $p \nmid d_L\Delta$. Write also $\underline{x} = [x_1, \ldots, x_m]$ for the corresponding m-tuple of morphisms $x_i \colon X_0 \to X$. By Serre–Tate, lifting $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x})$ from S to S' is the same as lifting $(X_0, \iota_0, \lambda_0, X, \iota, \lambda, \underline{x})$ from S to S'.

Using an element $g \in \mathrm{GL}_m(\mathcal{O}_{F_p})$ satisfying (3.5.6) as a "change of basis" for X_0^m, we obtain an $\mathcal{O}_F$-linear "orthogonal splitting" $X \cong X_0^{\mathrm{rank}(T)} \times Y$. That is, Y is a p-divisible group with action $\iota_Y \colon \mathcal{O}_{F_p} \to \mathrm{End}(Y)$, and a principal polarization $\lambda_Y \colon Y \to Y^\vee$ whose Rosati involution $\dagger$ satisfies $\iota_Y(a)^\dagger = \iota_Y(a^\sigma)$ for all $a \in \mathcal{O}_F$. Under the described identification $X \cong X_0^{\mathrm{rank}(T)} \times Y$, the polarization λ on X is given by $(\lambda_0)^{\mathrm{rank}(T)} \times \lambda_Y$. The map $\underline{x} \colon X_0^m \to X$ may be identified with the projection $X_0^m \to X_0^{\mathrm{rank}(T)}$ onto the first $\mathrm{rank}(T)$ factors, followed by the canonical inclusion $X_0^{\mathrm{rank}(T)} \to X_0^{\mathrm{rank}(T)} \times Y$.

Note that the actions of $\mathcal{O}_{F_p}$ on X_0, X, and Y have signatures $(1,0)$, $(n-r,r)$, and $(n-r-\mathrm{rank}(T), r)$ respectively (in the sense of (3.5.1)). These considerations also show that $\mathrm{rank}(T) \leq n-r$ if $\mathcal{Z}(T)_{\mathrm{Spf}\,\mathcal{O}_{F_p}}$ is nonempty.

The triple $(X_0, \iota_0, \lambda_0)$ admits a unique lift to S' as in Lemma 3.5.1. The projection map $\underline{x} \colon X_0^m \to X_0^{\mathrm{rank}(T)}$ clearly lifts to S' as well. So it remains only to lift (Y, ι_Y, λ_Y) from S to S'. Such a lift exists by formal smoothness of the corresponding deformation problem described in Lemma 3.5.1. We apply the same lemma to compute tangent spaces (e.g. after passing to an étale cover by a scheme), which shows that the relative dimension is $(n-r-\mathrm{rank}(T))r$. $\square$

Remark 3.5.6. Taking $T = \emptyset$ (or $T = 0$) in Lemma 3.5.5 and varying over non-degenerate Hermitian $\mathcal{O}_F$-lattices L satisfying $L \subseteq L^\vee$ and $|L^\vee/L| = d$, we see that $\mathscr{M}(n-r,r)^{(d)} \to \mathrm{Spec}\,\mathcal{O}_F[1/(d\Delta)]$ is smooth of relative dimension $(n-r)r$ for every $d \in \mathbb{Z}_{\geq 0}$. If $\mathcal{M}$ is associated with any non-degenerate Hermitian $\mathcal{O}_F$-lattice L (not necessarily with $L \subseteq L^\vee$), this then implies that $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F[1/(d_L \Delta)]$ is smooth of relative dimension $(n-r)r$.

4 Arithmetic special cycle classes

4.1 Hermitian vector bundles

Given a smooth algebraic stack $\mathcal{X}$ over $\operatorname{Spec} \mathbb{C}$, a *Hermitian vector bundle* $\widehat{\mathcal{E}}$ on $\mathcal{X}$ is the following functorial assignment: for every morphism $S \to \mathcal{X}$ with S a scheme smooth over $\operatorname{Spec} \mathbb{C}$, the assignment gives a vector bundle on S with a smooth Hermitian metric on the analytification. We sometimes write $\widehat{\mathcal{E}} = (\mathcal{E}, \|-\|)$ where $\mathcal{E}$ is the underlying vector bundle on $\mathcal{X}$ and $\|-\|$ is the norm associated with the smooth Hermitian metric (on the analytification), defined functorially.

Let (R, Σ, c_∞) be an arithmetic ring in the sense of Gillet–Soulé [GS90, §3.1], i.e. R is an excellent regular Noetherian integral domain (e.g. Dedekind domains with fraction field of characteristic 0 or fields), Σ is a finite nonempty set of injective homomorphisms $\tau \colon R \to \mathbb{C}$, and $c_\infty \colon \mathbb{C}^\Sigma \to \mathbb{C}^\Sigma$ is a conjugate-linear involution of $\mathbb{C} \otimes_{\mathbb{Z}} R$-algebras. Write K for the fraction field of R.

Suppose $\mathcal{X}$ is an algebraic stack which is flat and finite type over $\operatorname{Spec} R$. Write $\mathcal{X}_\tau \coloneqq \mathcal{X} \times_{\operatorname{Spec} K, \tau} \operatorname{Spec} \mathbb{C}$. Assume that the generic fiber $\mathcal{X}_K$ is smooth over $\operatorname{Spec} K$. A *Hermitian vector bundle* on $\mathcal{X}$ is a vector bundle $\mathcal{E}$ on $\mathcal{X}$ equipped with a smooth Hermitian metric on $\mathcal{E}|_{\mathcal{X}_\tau}$ for each $\tau \in \Sigma$, such that this collection of metrics is conjugation invariant (meaning c_∞-invariant). We write $\widehat{\operatorname{Pic}}(\mathcal{X})$ for the group of (isomorphism classes of) Hermitian line bundles, with group structure given by the tensor product. We use the notation $\widehat{\mathcal{L}} = (\mathcal{L}, \|-\|)$ for a Hermitian line bundle with underlying line bundle $\mathcal{L}$ and $\|-\|$ its norm (meaning a collection of norms $\{\|-\|_\tau\}_{\tau \in \Sigma}$). We also write $\|-\|_\infty = \prod_{\tau \in \Sigma} \|-\|_\tau$.

Next, we discuss stacky degrees of Hermitian line bundles. We fix the arithmetic ring (R, Σ, c_∞) associated with $R = \mathcal{O}_K[1/N]$ for a number field K and an integer $N \in \mathbb{Z}$, i.e. Σ is the set of all embeddings $\tau \colon K \to \mathbb{C}$, and c_∞ is induced by complex conjugation. For the rest of Section 4.1, we assume that

$\mathcal{X}$ is a reduced 1-dimensional Deligne–Mumford stack which is proper and flat over $\operatorname{Spec} R$. (4.1.1)

Here, dimension is used in an absolute sense rather than a relative one (e.g. we could have $\mathcal{X} = \operatorname{Spec} R$).

Let $\widehat{\mathcal{L}} = (\mathcal{L}, \|-\|)$ be a Hermitian line bundle on $\mathcal{X}$. For each complex place v of $\mathcal{O}_K$, we set $\|-\|_v := \|-\|_{\tau_1} \|-\|_{\tau_2}$ where $\tau_1, \tau_2 \colon K \to \mathbb{C}$ are the two embeddings corresponding to v. The Arakelov *arithmetic degree* $\widehat{\deg}(\widehat{\mathcal{L}})$ of $\widehat{\mathcal{L}}$ on $\mathcal{X}$ is valued in $\mathbb{R}_N := \mathbb{R}/(\sum_{p|N} \mathbb{Q} \cdot \log p)$, and may be described as follows. If $\mathcal{X} = \operatorname{Spec} \mathcal{O}_E[1/N]$ for a number field E, we have the standard definition

$$\widehat{\deg}(\widehat{\mathcal{L}}) := -\sum_{v \nmid N} \log \|s\|_v \quad \text{with} \quad \|s\|_v := q_v^{-\operatorname{ord}_v(s)} \text{ if } v < \infty \tag{4.1.2}$$

where the sum runs over all places v of E (Archimedean included) with $v \nmid N$, the quantity q_v is the cardinality of the residue field at v, and s is any rational section of $\mathcal{L}$.

If $\mathcal{X}$ is integral (equivalently, reduced and irreducible by quasi-separatedness), select any number field E with a finite surjection $\operatorname{Spec} \mathcal{O}_E[1/N] \to \mathcal{X}$ and set

$$\widehat{\deg}(\widehat{\mathcal{L}}) := \frac{1}{\deg(E/\mathcal{X}_K)} \widehat{\deg}(\widehat{\mathcal{L}}|_{\operatorname{Spec} \mathcal{O}_E[1/N]}) \tag{4.1.3}$$

where $\deg(E/\mathcal{X}_K)$ denotes the degree of the finite étale morphism $\operatorname{Spec} E \to \mathcal{X}_K$. This definition does not depend on the choice of E or the morphism $\operatorname{Spec} \mathcal{O}_E[1/N] \to \mathcal{X}$, since any two such choices may be covered by finite surjections from a third such choice $\operatorname{Spec} \mathcal{O}_{E'}[1/N]$ (and these finite surjections can be made compatible with the maps to $\mathcal{X}$).

Remark 4.1.1. Existence of E as in the preceding paragraph follows from some general theory. Indeed, a general fact about Noetherian Deligne–Mumford stacks with separated diagonal [LMB00, Théorème 16.6] implies that there exists a scheme Z with a morphism $Z \to \mathcal{X}$ which is finite, surjective, and generically étale (in the sense that $Z_{\mathcal{U}} \to \mathcal{U}$ is étale for a dense open substack $\mathcal{U} \subseteq \mathcal{X}$). Selecting an irreducible component of Z which surjects onto $\mathcal{X}$, we may assume that Z is also integral. Thus Z is a 1-dimensional integral scheme which is proper and flat over $\operatorname{Spec} R$. Such a map $Z \to \operatorname{Spec} R$ must be quasi-finite, hence also finite. If E denotes the fraction field of Z, the normalization of Z must be $\tilde{Z} = \operatorname{Spec} \mathcal{O}_E[1/N]$.

One can check that the definition of stacky arithmetic degree in (4.1.3) recovers the definition of [KRY04, (4.4)] and [KRY06, §2.1] which counts geometric points weighted by orders of automorphism groups.[21]

In the general case when $\mathcal{X}$ is not necessarily irreducible, we take

$$\widehat{\deg}(\widehat{\mathcal{L}}) := \sum_{\xi} \widehat{\deg}(\widehat{\mathcal{L}}|_{\mathcal{X}_\xi})$$

[21] In loc. cit., the functorial definition of Hermitian line bundle should also include the additional assumption of complex conjugation invariance as above.

where the sum runs over generic points ξ of irreducible components $\mathcal{X}_\xi$ of $\mathcal{X}$.

The preceding discussion showed that $\mathcal{X}$ admits a finite surjection from a scheme which is finite over $\operatorname{Spec} R$, hence $\mathcal{X} \to \operatorname{Spec} R$ is also quasi-finite (in the sense of [SProject, Definition 0G2M]).

Suppose $\mathcal{X}'$ and $\mathcal{X}$ are Deligne–Mumford stacks which both satisfy the hypotheses from (4.1.1), and consider a morphism $f \colon \mathcal{X}' \to \mathcal{X}$ over $\operatorname{Spec} R$. Let $\widehat{\mathcal{L}}$ be a Hermitian line bundle on $\mathcal{X}$. First consider the case where $\mathcal{X}$ is irreducible. We say that the morphism $\mathcal{X}'_K \to \mathcal{X}_K$ has *degree* $\deg(\mathcal{X}'_K/\mathcal{X}_K) \coloneqq \deg(\mathcal{X}'_K/K)/\deg(\mathcal{X}_K/K)$ (with stacky degrees of 0-cycles over fields as in (A.1.10)). We have $\widehat{\deg}(f^*\widehat{\mathcal{L}}) = \deg(\mathcal{X}'_K/\mathcal{X}_K)\widehat{\deg}(\widehat{\mathcal{L}})$. Next, consider the general case where $\mathcal{X}$ is not necessarily irreducible. We say that $\mathcal{X}'_K \to \mathcal{X}_K$ has *constant degree* d if $\mathcal{X}'_K \times_{\mathcal{X}_K} \mathcal{X}_{\xi,K} \to \mathcal{X}_{\xi,K}$ has degree d for every irreducible component $\mathcal{X}_\xi$ of $\mathcal{X}$. In this case, we have

$$\widehat{\deg}(f^*\widehat{\mathcal{L}}) = d \cdot \widehat{\deg}(\widehat{\mathcal{L}}). \tag{4.1.4}$$

4.2 Arithmetic Chow rings

We fix definitions for arithmetic Chow rings with rational coefficients.

Let (R, Σ, c_∞) be an arithmetic ring (Section 4.1) and write K for the fraction field of R. Suppose X is a scheme which is separated, flat, and finite type over $\operatorname{Spec} R$ with smooth and quasi-projective generic fiber X_K. There are associated Gillet–Soulé *arithmetic Chow groups* $\widehat{\operatorname{Ch}}^m(X)$ in codimensions $m \geq 0$. If X is moreover regular, these groups form an *arithmetic Chow ring* $\widehat{\operatorname{Ch}}^*(X)_{\mathbb{Q}}$ (with $\mathbb{Q}$-coefficients) [GS90, Theorem 4.2.3].

Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n, with associated moduli stack $\mathcal{M}$. Consider the arithmetic ring (R, Σ, c_∞) associated with $R = \operatorname{Spec} \mathcal{O}_F[1/d_L]$. We define *arithmetic Chow groups* for $\mathcal{M}$ by limiting over level structure: for any nonzero integer N, set

$$\widehat{\operatorname{Ch}}^*(\mathcal{M}[1/N])_{\mathbb{Q}} \coloneqq \lim_{K'_f} \widehat{\operatorname{Ch}}^*(\mathcal{M}_{K'_f}[1/N])_{\mathbb{Q}} \tag{4.2.1}$$

where K'_f varies over all small levels as in Section 3.4 (so that each $\mathcal{M}_{K'_f}$ is a scheme). Similar limiting procedures appeared in [BBK07] and [BH21, §4.4]; see also [Gil09] for more on arithmetic Chow rings of Deligne–Mumford stacks.

Since $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is smooth, we know that $\mathcal{M}$ is regular. Hence we obtain an *arithmetic Chow ring* $\widehat{\operatorname{Ch}}^*(\mathcal{M}[1/N])_{\mathbb{Q}}$ via the intersection product for each $\widehat{\operatorname{Ch}}^*(\mathcal{M}_{K'_f}[1/N])_{\mathbb{Q}}$.

Suppose $\mathcal{Z} \to \mathcal{M}$ is a finite morphism of algebraic stacks with $\mathcal{Z} \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ proper and $\mathcal{Z}$ equidimensional of dimension d. Then we define the *height* of $\mathcal{Z}$ with respect to any

Hermitian line bundle $\widehat{\mathcal{L}}$ on $\mathcal{M}$ as follows: if $\mathcal{Z}_{K'_f} := \mathcal{Z} \times_{\mathcal{M}} \mathcal{M}_{K'_f}$, the quantity

$$\widehat{\deg}(\widehat{\mathcal{L}}^d|_{\mathcal{Z}}) := \frac{1}{[K'_f(1) : K'_f]} \widehat{\deg}(\widehat{\mathcal{L}}^d|_{\mathcal{Z}_{K'_f}}) \in \mathbb{R}_{d_L N_{K'_f}} = \mathbb{R}/(\textstyle\sum_{p | d_L N_{K'_f}} \mathbb{Q} \cdot \log p) \qquad (4.2.2)$$

does not depend on the choice of small level K'_f, where $\widehat{\deg}(\widehat{\mathcal{L}}^d|_{\mathcal{Z}_{K'_f}})$ is the arithmetic height as in [BGS94, Proposition 2.3.1, Remarks(ii)] (see also [Zha95]) calculated by replacing $\mathcal{Z}_{K'_f}$ with its pushforward cycle on $\mathcal{M}_{K'_f}$. Varying K'_f, we obtain the *height* $\widehat{\deg}(\widehat{\mathcal{L}}^d|_{\mathcal{Z}}) \in \mathbb{R}_{d_L}$. We will be primarily interested in the case where $d = 1$ with $\mathcal{Z}$ reduced and flat over $\operatorname{Spec} \mathcal{O}_F[1/d_L]$. In this case, $\widehat{\deg}(\widehat{\mathcal{L}}|_{\mathcal{Z}})$ is the (stacky) arithmetic degree of $\widehat{\mathcal{L}}$ restricted to $\mathcal{Z}$, as discussed in Section 4.1.

4.3 Hodge bundles

Given an abelian scheme $\pi \colon A \to S$ of relative dimension n, we consider the *Hodge bundles* $\Omega_A := \pi_* \Omega^1_{A/S}$ and $\omega_A := \pi_* \bigwedge^n \Omega_A$. Here Ω_A and ω_A are locally free of ranks n and 1 respectively. If $e \colon S \to A$ denotes the identity section, there are canonical isomorphisms $\Omega_A \cong e^* \Omega_{A/S} \cong (\operatorname{Lie} A)^\vee$. These Hodge bundles are (contravariantly) functorial in A, and their formation commutes with arbitrary base change [BBM82, Proposition 2.5.2].

When S is a scheme which is smooth over $\operatorname{Spec} \mathbb{C}$, the analytification of ω_A may be equipped with a Hermitian metric (*Faltings* or *Hodge* metric), normalized as follows. The fiber of ω_A over any $s \in S(\mathbb{C})$ is canonically identified with $H^0(A_s, \omega_{A_s})$. Viewing $H^0(A_s, \omega_{A_s})$ as the 1-dimensional $\mathbb{C}$-vector space of holomorphic n-forms on $A_s(\mathbb{C})$, we take the metric on ω_A at s to be

$$(\alpha, \beta) = \left(\frac{i}{2\pi} \right)^n \int_{A_s(\mathbb{C})} \beta \wedge \bar{\alpha} \qquad (4.3.1)$$

for $\alpha, \beta \in H^0(A_s, \omega_{A_s})$. We call the resulting Hermitian line bundle $\widehat{\omega}_A := (\omega_A, \|{-}\|)$ a *metrized Hodge bundle*. This metric on ω_A is functorial for isomorphisms between abelian schemes A over S.

Next, suppose $S = \operatorname{Spec} \mathbb{C}$ and suppose $\lambda \colon A \to A^\vee$ is a quasi-polarization. There is an associated Hermitian metric on $\Omega_A^\vee \cong \operatorname{Lie}(A)$ which we normalize as follows: if λ is a polarization and $\lambda(a) = (t_a^* \mathcal{L}) \otimes \mathcal{L}^{-1}$ for an ample line bundle $\mathcal{L}$ on A (where t_a is translation by a), then the Chern class $c_1(\mathcal{L}) \in H^2(A, \mathbb{Z})$ defines a $\mathbb{Z}$-valued alternating form ψ on $H_1(A, \mathbb{Z})$ (sign conventions etc. as in [Mum85, §I], except our Hermitian pairings are conjugate linear in the first variable) and a positive definite Hermitian pairing

$$(x, y) := \pi \sqrt{|\Delta|} \big(\psi(ix, y) - i\psi(x, y) \big) \qquad (4.3.2)$$

on $\operatorname{Lie} A$. If instead λ is a quasi-polarization, the associated Hermitian pairing is m^{-1} times the Hermitian pairing of $m\lambda$ for any $m \in \mathbb{Z}_{>0}$ such that $m\lambda$ is a polarization. If $\dim A = 1$ and if λ is the unique principal polarization of A, the induced dual metric on $\Omega_A \cong (\operatorname{Lie} A)^\vee$ is $\sqrt{|\Delta|}^{-1}$ times the Faltings metric (cf. the proof of [BHKRY20II, Lemma 5.1.4]).

Over an arbitrary smooth scheme $S \to \operatorname{Spec} \mathbb{C}$, any quasi-polarization $\lambda \colon A \to A^\vee$ defines a smooth Hermitian metric on $\operatorname{Lie} A$, given fiberwise by the construction above. We also call the resulting Hermitian vector bundle $\widehat{\Omega}^\vee_A$ a *metrized Hodge bundle*.

Next, let (R, Σ, c_∞) be an arithmetic ring, and suppose $\mathcal{X}$ is an algebraic stack which is flat and finite type over $\operatorname{Spec} R$, and whose generic fiber is smooth. Suppose we are given a relative abelian scheme $\mathcal{A}$ over $\mathcal{X}$ (equivalently, a functorial assignment of abelian schemes $A \to S$ to objects $x \in \mathcal{X}(S)$, for R-schemes S). Formation of the metrized Hodge bundle $\widehat{\omega}_A$ is functorial, hence defines a *metrized Hodge bundle* $\widehat{\omega}$ on $\mathcal{X}$. If $\mathcal{A}$ is equipped with a quasi-polarization, then there is similarly a *metrized Hodge bundle* $\widehat{\Omega}^\vee$ on $\mathcal{X}$.

Let L be a non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n, with associated moduli stack $\mathcal{M}$. If L is self-dual and signature $(n-1,1)$, we write $\mathscr{E}$ for the pullback of the tautological bundle (Definition 3.2.6) along $\mathcal{M} \to \mathscr{M}(n-1,1)^\circ$. Otherwise, we write $\mathscr{E}$ for the pullback of the tautological bundle (Definition 3.1.7) along $\mathcal{M} \to \mathscr{M}(n-r,r)[1/\Delta]$. We similarly write Ω for the pullback to $\mathcal{M}$ of the Hodge bundle from $\mathscr{M}(n-r,r)[1/\Delta]$ in the non exotic smooth case (resp. $\mathscr{M}(n-1,1)^\circ$ in the exotic smooth case) to $\mathcal{M}$ (where Ω is the rank n Hodge bundle). After base change to $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$, the (dual) Hodge bundle $\Omega^\vee$ on $\mathcal{M}$ admits a (canonical) decomposition $\Omega^\vee[1/\Delta] = \Omega^\vee[1/\Delta]^+ \oplus \Omega^\vee[1/\Delta]^-$ (with $\Omega^\vee[1/\Delta]^- = \mathscr{E}^\vee[1/\Delta]$) where the $\mathcal{O}_F$-action (via ι) on $\Omega^\vee[1/\Delta]^+$ (resp. $\Omega^\vee[1/\Delta]^-$) is $\mathcal{O}_F$-linear (resp. σ-linear).

Equip $\Omega^\vee[1/\Delta]$ with the Hermitian metric which is

$$(x, y) \coloneqq 4\pi^2 e^\gamma \sqrt{|\Delta|}\big(\psi(ix, y) - i\psi(x, y)\big) \tag{4.3.3}$$

in complex fibers (i.e. multiply the Hermitian metric from (4.3.2) by $4\pi e^\gamma$) where γ is the Euler–Mascheroni constant. We remark that the normalization constant $4\pi e^\gamma$ has appeared previously in similar contexts, e.g. [KRY04, (0.4)] [KRY06, §7] [BHKRY20, §7.2]. We refer to loc. cit. for possible conceptual explanations of this constant.

Then $\mathscr{E}^\vee[1/\Delta] \subseteq \Omega^\vee[1/\Delta]$ (via decomposition in previous paragraph) inherits a Hermitian metric by restriction. This makes $\mathscr{E}^\vee$ into a Hermitian vector bundle $\widehat{\mathscr{E}}^\vee$.

Write Ω_0 for the Hodge bundle on $\mathscr{M}_0$. Equip the dual $\Omega_0^\vee$ with the metric described fiberwise by (4.3.2), giving a Hermitian line bundle $\widehat{\Omega}_0^\vee$. By the *metrized dual tautological bundle* on $\mathcal{M}$, we mean

$$\widehat{\mathcal{E}}^\vee \coloneqq \widehat{\Omega}_0^\vee \otimes \widehat{\mathscr{E}}^\vee \tag{4.3.4}$$

where we have suppressed pullbacks from notation. The metric on $\widehat{\mathcal{E}}^\vee$ is the tensor product of the metrics described above. This definition of $\widehat{\mathcal{E}}^\vee$ is similar to [BHKRY20, §2.4, §7.2], though in a different setup (we are considering not-necessarily principal polarizations). Taking a dual gives the *metrized tautological bundle* $\widehat{\mathcal{E}}$. For more discussion on these metric normalization choices, see Section 12.2. We write $\mathcal{E}^\vee$ for underlying line bundle of $\widehat{\mathcal{E}}^\vee$.

For readers interested in Faltings height, we also consider the metrized Hodge (determinant) bundle $\widehat{\omega}$ on $\mathcal{M}$, which is pulled back from the Hodge determinant bundle ω on $\mathscr{M}(n-r,r)$ and with metric normalized as in (4.3.1).

Suppose $A_0 \to \operatorname{Spec} \mathcal{O}_E$ is any (relative) elliptic curve with $\mathcal{O}_F$-action, where E is a number field. If $\widehat{\omega}_{A_0}$ denotes the associated metrized Hodge bundle (normalized as in (4.3.1)), we recall that the *Faltings height* of A_0 is

$$h^{\mathrm{CM}}_{\mathrm{Fal}} := \frac{1}{[E:\mathbb{Q}]}\widehat{\deg}(\widehat{\omega}_{A_0}) = \frac{1}{2}\frac{L'(1,\eta)}{L(1,\eta)} + \frac{1}{2}\frac{\Gamma'(1)}{\Gamma(1)} + \frac{1}{4}\log|\Delta| - \frac{1}{2}\log(2\pi) \tag{4.3.5}$$

where η is the quadratic character associated to $F/\mathbb{Q}$, and Γ is the usual gamma function. This comes from the classical Chowla–Selberg formula (the statement above is as in [KRY04, Proposition 10.10]). It will be convenient to define the height constants

$$h^{\mathrm{CM}}_{\mathrm{tau}} := -h^{\mathrm{CM}}_{\mathrm{Fal}} + \frac{1}{4}\log|\Delta| - \frac{1}{2}\log(4\pi e^{\gamma}) \qquad h^{\mathrm{CM}}_{\mathcal{E}^\vee} := -h^{\mathrm{CM}}_{\mathrm{Fal}} - \frac{1}{4}\log|\Delta| + h^{\mathrm{CM}}_{\mathrm{tau}}. \tag{4.3.6}$$

These will re-appear in Sections 11.9 and 22.1.

4.4 Arithmetic special cycle classes

Let L be a non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n, with associated moduli stack $\mathcal{M}$. For the rest of Section 4, we assume L has signature $(n-1,1)$. Consider an $m \times m$ Hermitian matrix $T \in \mathrm{Herm}_m(\mathbb{Q})$, assume $m \leq n$, and form the associated special cycle $\mathcal{Z}(T) \to \mathcal{M}$. One expects to be able to construct an associated *arithmetic special cycle class* $[\widehat{\mathcal{Z}}(T)] \in \widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$.

For arbitrary singular T, there is no proposed definition of $[\widehat{\mathcal{Z}}(T)]$ in the literature. In general, $\mathcal{Z}(T)_{\mathscr{H}}$ has larger-than-expected dimension. The stack $\mathcal{Z}(T)$ could also have components with larger-than-expected dimension in positive characteristic (occurs already for nonsingular T). Available methods in the literature for treating the non-Archimedean theory (K-theoretic and derived algebro-geometric) do not incorporate the Archimedean place in general, as needed for arithmetic intersection theory (see introduction).

The analogue of the "linear invariance" approach of [KRY04, §6.4] (there for Shimura curves) is to first define $[\widehat{\mathcal{Z}}(T^\flat)]$ for nonsingular $T^\flat$, to consider ${}^t\overline{\gamma}T\gamma = \operatorname{diag}(0, T^\flat)$ for some

$\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$ with $T^\flat$ nonsingular, and to define $[\widehat{\mathcal{Z}}(T)]$ by intersecting $[\widehat{\mathcal{Z}}(T^\flat)]$ with a power of some metrized tautological bundle (possibly with additional Archimedean adjustment). This is not literally possible in the unitary setting, where $\mathcal{O}_F$ may have class number $\neq 1$ (in particular, γ as above may not exist). One also needs to verify independence of the choice of γ.

For arbitrary $T \in \mathrm{Herm}_m(\mathbb{Q})$, we propose to define $[\widehat{\mathcal{Z}}(T)]$ as a sum

$$[\widehat{\mathcal{Z}}(T)] := [\widehat{\mathcal{Z}}(T)_{\mathscr{H}}] + \sum_{\substack{p \text{ prime} \\ p \nmid d_L}} [{}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}] \in \widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}. \tag{4.4.1}$$

We construct $[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}]$ using the horizontal part $\mathcal{Z}(T)_{\mathscr{H}}$ and an appropriate Green current $g_{T,y}$ (4.5.5) with an additional parameter $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$. The element $[{}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}]$ arises from a class ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}}$ corresponding to the "vertical part" of $\mathcal{Z}(T)$ at p (4.6.10). The classes ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ will be zero for all but finitely many primes p. We define ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ using a "p-local" variant of the linear invariance strategy above.

We will show that $[\widehat{\mathcal{Z}}(T)]$ satisfies the "linear invariance" property

$$[\widehat{\mathcal{Z}}(T)] = [\widehat{\mathcal{Z}}({}^t\overline{\gamma}T\gamma)] \tag{4.4.2}$$

for all $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$, where $[\widehat{\mathcal{Z}}(T)]$ is formed with respect to $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$ and $[\widehat{\mathcal{Z}}(T)]$ is formed with respect to $\gamma^{-1}y{}^t\overline{\gamma}^{-1}$.

In fact, we prove refined statements. We show

$$[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}] = [\widehat{\mathcal{Z}}({}^t\overline{\gamma}T\gamma)_{\mathscr{H}}] \tag{4.4.3}$$

for the Green currents $g_{T,y}$ defined in Section 12.4 (where the current $g_{{}^t\overline{\gamma}T\gamma,\gamma^{-1}y{}^t\overline{\gamma}^{-1}}$ is used on the right-hand side). Moreover, we show $g_{T,y} = g_{{}^t\overline{\gamma}T\gamma,\gamma^{-1}y{}^t\overline{\gamma}^{-1}}$ (Section 12.4); this property is also satisfied for the Garcia–Sankaran currents in [GS19, (4.38)] (which we do not use for our arithmetic Siegel–Weil results).

For any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$, we show that the pullback

$$\mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \leftarrow \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}({}^t\overline{\gamma}T\gamma)_{\mathbb{F}_p})_{\mathbb{Q}} \tag{4.4.4}$$

along $\mathcal{Z}(T)_{\mathbb{F}_p} \to \mathcal{Z}({}^t\overline{\gamma}T\gamma)_{\mathbb{F}_p}$ (defined in (3.3.6)) sends ${}^{\mathbb{L}}\mathcal{Z}({}^t\overline{\gamma}T\gamma)_{\mathscr{V},p}$ to ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ (see (4.6.11)).

4.5 Horizontal arithmetic special cycle classes

Consider any $m \times m$ Hermitian matrix $T \in \mathrm{Herm}_m(\mathbb{Q})$. The horizontal arithmetic special cycle class $[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}]$ should involve some extra Archimedean data, e.g. from a Green current $g_{T,y}$ (which we allow to depend on a parameter $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$, as is typical in the literature).

Given an equidimensional complex manifold X, recall that a (p,q)-*current* on X is a continuous linear map $\Omega_c^{\dim X-p,\dim X-q}(X) \to \mathbb{C}$ on compactly supported smooth forms of degree $(\dim X - p, \dim X - q)$, where $\Omega_c^{\dim X-p,\dim X-q}(X)$ has the usual colimit topology. A (p,p)-current is *real* if it is induced by a continuous real-valued linear map on real (p,p)-forms. Given a top degree current g on X (i.e. a distribution), we say that g is *integrable* or that $\int_X g$ *converges* (possibly non-standard usage) if g extends (necessarily uniquely) to a continuous map $C_{b_1}^\infty(X) \to \mathbb{C}$, where

$$C_{b_1}^\infty(X) := \{f \in C^\infty(X) : |f(x)| \leq 1 \text{ for all } x \in X\} \tag{4.5.1}$$

with topology given by sup-norms ranging over all compact subsets $K \subseteq X$. In this case, we write $\int_X g$ for the value of g on $1 \in C_{b_1}^\infty(X)$. Suppose α is a (measurable) locally L^1 form of top degree on X. If α is integrable, then the associated distribution $[\alpha]$ on X is integrable, and we have $\int_X [\alpha] = \int_X \alpha$. We use the orientation and sign conventions of [GS90].

Returning to the moduli stack $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ from above, choose any embedding $F \to \mathbb{C}$ and form the base changes $\mathcal{M}_{\mathbb{C}} := \mathcal{M} \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$ and $\mathcal{Z}(T)_{\mathbb{C}} := \mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$, etc.. By a (p,q)-*current* on $\mathcal{M}_{\mathbb{C}}$, we mean a system of currents $g = (g_{K'_f})_{K'_f} = (\Omega_c^{n-1-p,n-1-q}(\mathcal{M}_{K'_f,\mathbb{C}}) \to \mathbb{C})_{K'_f}$ compatible with pullback of currents as we vary K'_f among all small levels. We say a (p,q)-current on $\mathcal{M}_{\mathbb{C}}$ is *real* if the associated current at each small level K'_f is real. If g is a current of top degree on $\mathcal{M}$, its *integral* is defined as

$$\int_{\mathcal{M}_{\mathbb{C}}} g := \frac{1}{[K'_{L,f} : K'_f]} \int_{\mathcal{M}_{K'_f,\mathbb{C}}} g_{K'_f} \tag{4.5.2}$$

for any sufficiently small level K'_f (conditional on convergence). This definition does not depend on the choice of small level.

Suppose $g_{T,y}$ is any real $(m-1, m-1)$ current on $\mathcal{M}_{\mathbb{C}}$ which satisfying a *modified current equation*, i.e. such that

$$-\frac{1}{2\pi i}\partial\overline{\partial} g_{T,y} + \delta_{\mathcal{Z}(T)_{\mathbb{C}}} \wedge [c_1(\widehat{\mathcal{E}}_{\mathbb{C}}^\vee)^{m-\operatorname{rank}(T)}] \tag{4.5.3}$$

is (represented by) a smooth (m,m)-form (for all small levels K'_f), where $c_1(\widehat{\mathcal{E}}_{\mathbb{C}}^\vee)$ is the Chern form of the Hermitian line bundle $\widehat{\mathcal{E}}_{\mathbb{C}}^\vee$. We call $g_{T,y}$ a *Green current*. We typically write $g_{T,y}$ instead of g_{T,y,K'_f} to lighten notation, for understood level K'_f.

For each small level K'_f, pick a representative $(\mathcal{Z}_0, g_0)$ for the self-intersection arithmetic cycle class $\widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{m-\operatorname{rank}(T)} \in \widehat{\operatorname{Ch}}^{m-\operatorname{rank}(T)}(\mathcal{M}_{K'_f})_{\mathbb{Q}}$. We can assume that $\mathcal{Z}_0$ intersects $\mathcal{Z}(T)_{K'_f}$ properly in the generic fiber (moving lemma) and that g_0 is a Green form of logarithmic type for $\mathcal{Z}_0$ (in the sense of [GS90, §4]).

The intersection pairing for Chow groups with supports (as in [GS90, §4]) gives a class $\mathcal{Z}(T)_{K'_f,\mathscr{H}} \cdot \mathcal{Z}_0 \in \mathrm{Ch}^m_{\mathcal{Z}(T)_{K'_f,\mathscr{H}} \cap \mathcal{Z}_0}(\mathcal{M}_{K'_f})_{\mathbb{Q}}$. We set

$$[\widehat{\mathcal{Z}}(T)_{K'_f,\mathscr{H}}] \coloneqq [(\mathcal{Z}(T)_{K'_f,\mathscr{H}} \cdot \mathcal{Z}_0, g_{T,y} + g_0 \wedge \delta_{\mathcal{Z}(T)_{K'_f,\mathbb{C}}})] \in \widehat{\mathrm{Ch}}^m(\mathcal{M}_{K'_f}[1/N])_{\mathbb{Q}} \tag{4.5.4}$$

(where we have suppressed the $1/N$ notation from the left). As in [GS19, (5.158)], a short computation (using well-definedness of arithmetic intersection products) shows that this class does not depend on the choice of $(\mathcal{Z}_0, g_0)$. One can verify that $g_{T,y} + g_0 \wedge \delta_{\mathcal{Z}(T)_{K'_f,\mathbb{C}}}$ satisfies a Green current equation for $(\mathcal{Z}(T)_{K'_f} \cap \mathcal{Z}_0)_{\mathbb{C}}$ by combining the Green current equation for g_0 with the modified current equation of $g_{T,y}$ (see also [GS19, §5.4]).

These classes $[\widehat{\mathcal{Z}}(T)_{K'_f,\mathscr{H}}]$ thus form a compatible system as K'_f varies, and hence give an element

$$[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}] \coloneqq ([\widehat{\mathcal{Z}}(T)_{K'_f,\mathscr{H}}])_{K'_f} \in \widehat{\mathrm{Ch}}^m(\mathcal{M}[1/N])_{\mathbb{Q}}. \tag{4.5.5}$$

This construction of $[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}]$ is essentially that of [GS19, §5.4]. If $g_{T,y} = g_{{}^t\overline{\gamma}T\gamma,\gamma^{-1}y{}^t\overline{\gamma}^{-1}}$, note that we automatically have the "linear invariance" equality

$$[\widehat{\mathcal{Z}}(T)_{\mathscr{H}}] = [\widehat{\mathcal{Z}}({}^t\overline{\gamma}T\gamma)_{\mathscr{H}}]. \tag{4.5.6}$$

Currents $g_{T,y}$ satisfying (4.5.3) were studied by Garcia–Sankaran [GS19, Theorem 1.1 and §4]. We choose to use the star-product approach of Kudla [Kud97b] to define currents $g_{T,y}$ for our main results (for $\operatorname{rank} T \geq n-1$ or $\det T \neq 0$ with T not positive definite), and postpone the explicit description of $g_{T,y}$ to Section 12.4 (12.4.11). Our definition of $g_{T,y}$ is that of [Liu11, Theorem 4.20] in the nonsingular cases. When $T \in \mathrm{Herm}_n(\mathbb{Q})$ is singular with $\operatorname{rank}(T) = n-1$, our definition is new (still based on star products). These Green currents satisfy $g_{T,y} = g_{{}^t\overline{\gamma}T\gamma,\gamma^{-1}y{}^t\overline{\gamma}^{-1}}$ (Section 12.4), so linear invariance (4.5.6) is satisfied.

4.6 Vertical special cycle classes

We define vertical special cycle classes via K_0 groups. We remind the reader that $d_L \in \mathbb{Z}$ is an integer associated to the lattice L as discussed before Definition 3.1.2. Recall that $d_L = 1$ if L is self-dual of signature $(n-1,1)$ with $2 \nmid \Delta$.

For our notation and definitions regarding K_0-groups for Deligne–Mumford stacks, we refer to Appendix A. Note that the stacky K_0 groups we use are different from those used in [HM22]. It is also possible to avoid stacky K_0 groups entirely by working with compatible systems of classes in towers of level structure.

Fix any prime $p \nmid d_L$ and set

$$\mathcal{M}_{(p)} \coloneqq \mathcal{M} \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Z}_{(p)} \qquad \mathcal{Z}(T)_{(p)} \coloneqq \mathcal{Z}(T) \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Z}_{(p)} \tag{4.6.1}$$

for any $T \in \mathrm{Herm}_m(\mathbb{Q})$. Since $\mathcal{M}_{(p)}$ admits a finite étale cover by a scheme (add away-from-p level structure), we may consider filtrations for K_0' groups as in Definition A.1.2.

Let $T \in \mathrm{Herm}_m(\mathbb{Q})$ be any $m \times m$ Hermitian matrix (with entries in F). We first describe a "p-local" derived special cycle class ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p)} \in F^m_{\mathcal{M}_{(p)}} K_0'(\mathcal{Z}(T)_{(p)})_{\mathbb{Q}}$ (with $F^m_{\mathcal{M}_{(p)}} \coloneqq F_{n-m}$ denoting the m-th step of the codimension filtration) before extracting a "vertical" piece.

For any $t \in \mathbb{Q}$, we define ${}^{\mathbb{L}}\mathcal{Z}(t)_{(p)} \in F^1_{\mathcal{M}_{(p)}} K_0'(\mathcal{Z}(t)_{(p)})_{\mathbb{Q}}$ to be the element

$$
{}^{\mathbb{L}}\mathcal{Z}(t)_{(p)} \coloneqq \begin{cases} [\mathcal{O}_{\mathcal{Z}(t)_{(p)}}] & \text{if } t \neq 0 \\ [\mathcal{O}_{\mathcal{M}_{(p)}}] - [\mathcal{E}] & \text{if } t = 0. \end{cases}
$$

Write $t_1, \ldots, t_m$ for the diagonal entries of T. Using the intersection pairing of Lemma A.2.1 as well as compatibility with dimension filtrations from Lemma A.2.2 via Lemma 3.4.5, we form the intersection ${}^{\mathbb{L}}\mathcal{Z}(t_1)_{(p)} \cdots {}^{\mathbb{L}}\mathcal{Z}(t_m)_{(p)}$ and define ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p)}$ by the restriction

$$
\begin{array}{ccc} F^m_{\mathcal{M}} K_0'(\mathcal{Z}(t_1)_{(p)} \times_{\mathcal{M}_{(p)}} \cdots \times_{\mathcal{M}_{(p)}} \mathcal{Z}(t_m)_{(p)})_{\mathbb{Q}} & \longrightarrow & F^m_{\mathcal{M}_{(p)}} K_0'(\mathcal{Z}(T)_{(p)})_{\mathbb{Q}} \\ {}^{\mathbb{L}}\mathcal{Z}(t_1)_{(p)} \cdots {}^{\mathbb{L}}\mathcal{Z}(t_m)_{(p)} & \longmapsto & {}^{\mathbb{L}}\mathcal{Z}(T)_{(p)}. \end{array} \tag{4.6.2}
$$

This displayed restriction map comes from the disjoint union decomposition in (3.3.3). We call ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p)}$ the *p-local derived special cycle class*[22] associated with T.

The following lemma is a "p-local" version of linear invariance, and is proved using a variant on ideas from [How19; HM22]. The map in (4.6.3) was defined in (3.3.6).

Lemma 4.6.1. *Given any $T \in \mathrm{Herm}_m(\mathbb{Q})$ and any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$, the pullback along*

$$
\mathcal{Z}(T)_{(p)} \to \mathcal{Z}({}^t\overline{\gamma} T \gamma)_{(p)} \tag{4.6.3}
$$

sends ${}^{\mathbb{L}}\mathcal{Z}({}^t\overline{\gamma} T\gamma)_{(p)}$ to ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p)}$.

Proof. By Lemma 3.3.3, we know that (4.6.3) is an open and closed immersion.

The ring $\mathcal{O}_{F,(p)}$ is a Euclidean domain, with Euclidean function $\phi(a) \coloneqq \sum_{\mathfrak{p}_i} v_{\mathfrak{p}_i}(a) \cdot f_i$ for nonzero $a \in \mathcal{O}_{F,(p)}$ (summing over primes $\mathfrak{p}_i$ in $\mathcal{O}_F$ lying over p, with residue cardinality p^{f_i}). Row reducing via the Euclidean algorithm shows that $\mathrm{GL}_m(\mathcal{O}_{F,(p)})$ is generated by elementary matrices.

[22]This is the construction of [HM22, Definition 5.1.3] (there for orthogonal Shimura varieties). This construction also underlies the intersection numbers considered in [KR14] for non-degenerate T. We differ slightly from those references by localizing at p, since we will only be interested in the "vertical" part of ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p)}$. The "horizontal part" is accounted for by Section 4.5.

Any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)})$ may thus be expressed as $\gamma = \gamma_1\gamma_2^{-1}$ where each γ_1 and γ_2 are products of elementary matrices lying in $\mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$. If moreover $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$, the commutative diagram

$$\begin{tikzcd} & \mathcal{Z}({}^t\overline{\gamma}_1 T \gamma_1)_{(p)} & \\ \mathcal{Z}(T)_{(p)} \arrow[ur] \arrow[rr] & & \mathcal{Z}({}^t\overline{\gamma} T \gamma)_{(p)} \arrow[ul] \end{tikzcd} \tag{4.6.4}$$

shows that it is enough to prove the lemma when $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$ is an elementary matrix.

If γ is a permutation matrix, the lemma is clear. Next, consider $a \in \mathcal{O}_{F,(p)}^\times \cap \mathcal{O}_F$. For any $t \in \mathbb{Q}$, note that $\mathcal{Z}(t)_{(p)} \to \mathcal{Z}(\overline{a}ta)_{(p)}$ is an open and closed immersion (by Lemma 3.3.3 again). This fact implies the present lemma for the case where $\gamma = \mathrm{diag}(a, 1, \ldots, 1)$.

It remains to check the case where γ is an elementary unipotent matrix. This case follows as in the analogous result [HM22, Proposition 5.4.1] (there for GSpin).[23] The latter is proved using methods from [How19] (the analogous local linear invariance result on Rapoport–Zink spaces). We are also using global analogues of [LL22, Lemma 2.36, Lemma 2.37, Lemma 2.41] (there about a tautological bundle on an exotic smooth Rapoport–Zink space) which may be proved similarly, e.g. our Lemma 3.2.5 replaces [LL22, Lemma 2.36] in the global setup. Alternatively, linear invariance for $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$ should also follow from the derived algebro-geometric methods in [Mad23]. $\square$

Next, we define a *derived vertical special cycle class*

$$^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \mathrm{gr}^m_{\mathcal{M}} K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \tag{4.6.5}$$

at p, where $\mathcal{Z}(T)_{\mathbb{F}_p} \coloneqq \mathcal{Z}(T) \times_{\mathrm{Spec}\,\mathbb{Z}} \mathrm{Spec}\,\mathbb{F}_p$.

First consider the case where $\det T \neq 0$. Using Lemmas 3.5.5 and A.1.5 as well as (3.3.5), we decompose

$$\mathrm{gr}^m_{\mathcal{M}_{(p)}} K_0'(\mathcal{Z}(T)_{(p)})_{\mathbb{Q}} = \mathrm{gr}^m_{\mathcal{M}_{(p)}} K_0'(\mathcal{Z}(T)_{(p),\mathscr{H}})_{\mathbb{Q}} \oplus \mathrm{gr}^m_{\mathcal{M}} K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \tag{4.6.6}$$

into a "horizontal" part and a "vertical" part. This uses nonsingularity of T (via Lemma 3.5.5), so that $\mathcal{Z}(T)_{(p),\mathscr{H}} \cap \mathcal{Z}(T)_{\mathbb{F}_p}$ is of dimension $< n - m$. We are also using the dévissage pushforward identification $K_0'(\mathcal{Z}(T)_{\mathbb{F}_p}) \xrightarrow{\sim} K_0'(\mathcal{Z}(T)_{\mathscr{V},p})$, with $\mathcal{Z}(T)_{\mathscr{V},p}$ as in (3.3.5). The

[23] Strictly speaking, our setup for stacky K_0' groups is slightly different from that of [HM22], see Appendix A. This makes no difference in the proof of the cited result. Alternatively, one can replace $\mathcal{M}_{(p)}$ by a finite étale cover by a scheme to reduce to the case of schemes, where our setup agrees with [HM22, §A.2].

above decomposition of $\mathrm{gr}^m_{\mathcal{M}_{(p)}} K'_0(\mathcal{Z}(T)_{(p)})_{\mathbb{Q}}$ is independent of the choice of e_p in (3.3.5). We define ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ to be given by the projection

$$\begin{aligned} \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{(p)})_{\mathbb{Q}} &\longrightarrow \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \\ {}^{\mathbb{L}}\mathcal{Z}(T)_{(p)} &\longmapsto {}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}. \end{aligned} \tag{4.6.7}$$

If $T = \mathrm{diag}(0, T^\flat)$ (with $\det T^\flat \neq 0$), we set

$${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \coloneqq (\mathcal{E}^\vee)^{m-\mathrm{rank}(T)} \cdot {}^{\mathbb{L}}\mathcal{Z}(T^\flat)_{\mathscr{V},p} \in \mathrm{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \tag{4.6.8}$$

where $\mathcal{E}^\vee$ stands for the class $[\mathcal{O}_{\mathcal{M}_{(p)}}] - [\mathcal{E}] \in F^1_{\mathcal{M}_{(p)}}(\mathcal{M}_{(p)})_{\mathbb{Q}}$. Given arbitrary T (not necessarily block diagonal), select any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$ such that

$${}^t\overline{\gamma} T \gamma = \mathrm{diag}(0, T'^\flat) \tag{4.6.9}$$

where $\det T'^\flat \neq 0$. Set $T' \coloneqq \mathrm{diag}(0, T'^\flat)$. We define ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ to be the pullback class

$$\begin{aligned} \mathrm{gr}^m_{\mathcal{M}}(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} &\longleftarrow \mathrm{gr}^m_{\mathcal{M}}(\mathcal{Z}(T')_{\mathbb{F}_p})_{\mathbb{Q}} \\ {}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} &\longleftarrow\!\shortmid {}^{\mathbb{L}}\mathcal{Z}(T')_{\mathscr{V},p} \end{aligned} \tag{4.6.10}$$

along the open and closed immersion $\mathcal{Z}(T)_{\mathbb{F}_p} \to \mathcal{Z}(T')_{\mathbb{F}_p}$ induced by γ.

By Lemma 4.6.1 (applied to $T'^\flat$, in the notation above), the preceding definition of ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$ does not depend on the choice of γ. Moreover, the class $\mathcal{Z}(T)_{\mathscr{V},p}$ is *linearly invariant* in the following sense: given any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F,(p)}) \cap M_{m,m}(\mathcal{O}_F)$ (no additional assumptions on ${}^t\overline{\gamma} T \gamma$), the pullback along

$$\mathcal{Z}(T)_{\mathbb{F}_p} \to \mathcal{Z}({}^t\overline{\gamma} T \gamma)_{\mathbb{F}_p} \tag{4.6.11}$$

sends ${}^{\mathbb{L}}\mathcal{Z}({}^t\overline{\gamma} T \gamma)_{\mathscr{V},p}$ to ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$. This follows from the construction of ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$.

We see that ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} = 0$ for all but finitely many primes p. Taking pushforward, we obtain associated cycle classes $[{}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}] \in \widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$. The preceding constructions may be repeated (essentially verbatim) with K'_f level structure away from p. The resulting classes ${}^{\mathbb{L}}\mathcal{Z}(T)_{(p),K'_f}$ and ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p,K'_f}$ will be compatible with pullback for varying level.

4.7 Degrees of arithmetic special cycles

The moduli stack $\mathcal{M} \to \mathrm{Spec}\, \mathcal{O}_F[1/d_L]$ considered in Section 4.4 may not be proper. For a robust arithmetic degree theory via arithmetic Chow groups for arbitrary $T \in \mathrm{Herm}_m(\mathbb{Q})$, one might instead consider arithmetic special cycle classes on a suitably compactified moduli space.

If the special cycle $\mathcal{Z}(T)$ is already proper over $\operatorname{Spec} \mathcal{O}_F[1/d_L]$, we can directly define the *arithmetic degree without boundary contributions* which should result from a compactification: set

$$\begin{aligned}\widehat{\deg}([\widehat{\mathcal{Z}}(T)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{n-m}) \coloneqq &\left(\int_{\mathcal{M}_{\mathbb{C}}} g_{T,y} \wedge c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{n-m} \right) \\ &+ \widehat{\deg}((\widehat{\mathcal{E}}^\vee)^{n-\operatorname{rank}(T)}|_{\mathcal{Z}(T)_{\mathscr{H}}}) \\ &+ \sum_{\substack{p \text{ prime} \\ p \nmid d_L}} \deg_{\mathbb{F}_p}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \cdot (\mathcal{E}^\vee)^{n-m}) \log p\end{aligned} \tag{4.7.1}$$

conditional on convergence of the integral. Since compactification of $\mathcal{M}$ plays no other role in this work, we take this approach. As in Section 4.5, the notation $\mathcal{M}_{\mathbb{C}}$ mean $\mathcal{M} \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$ for a choice of $F \to \mathbb{C}$ (the choice does not matter).

The quantity in (4.7.1) is an element of $\mathbb{R}_{d_L} = \mathbb{R}/(\sum_{p|d_L} \mathbb{Q}\cdot\log p)$. Here $\widehat{\deg}((\widehat{\mathcal{E}}^\vee)^{n-m}|_{\mathcal{Z}(T)_{\mathscr{H}}})$ is the height of $\mathcal{Z}(T)_{\mathscr{H}}$ with respect to the metrized tautological bundle $\widehat{\mathcal{E}}^\vee$ (Sections 4.2 and 4.3). The symbol ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \cdot (\mathcal{E}^\vee)^{n-m}$ is shorthand for the intersection product

$${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \cdot ([\mathcal{O}_{\mathcal{M}}] - [\mathcal{E}])^{n-m} \in \operatorname{gr}^n_{\mathcal{M}} K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} = \operatorname{gr}_0 K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}}, \tag{4.7.2}$$

defined in Lemma A.2.1. With $\mathcal{Z}(T)_{\mathbb{F}_p}$ viewed as a proper scheme over $\mathbb{F}_p$, the notation $\deg_{\mathbb{F}_p}$ refers to the degree map $\deg \colon \operatorname{gr}_0 K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} \to \mathbb{Q}$ as defined in (A.1.12).

Certainly $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is proper if $\mathcal{Z}(T)$ is empty (e.g. if T is not positive semidefinite). In this case, the right-hand side of (4.7.1) consists only of the integral $\int_{\mathcal{M}_{\mathbb{C}}} g_{T,y} \wedge c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{n-m}$.

We show below that $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is also proper if $\operatorname{rank}(T) \geq n-1$, so (4.7.1) applies in this case as well.

Lemma 4.7.1. *Fix a Hermitian matrix $T \in \operatorname{Herm}_m(\mathbb{Q})$ with $\operatorname{rank}(T) \geq n-1$ and $m \geq 0$. Let κ be a field, and consider $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) \in \mathcal{Z}(T)(\kappa)$. There exists an $\mathcal{O}_F$-linear isogeny $(A_0)^{n-1} \times A^- \to A$ where A^- is an elliptic curve with $\mathcal{O}_F$-action. After replacing κ by a finite extension, we may take $A^- = A_0^\sigma$ where $A_0^\sigma = A_0$ but with $\mathcal{O}_F$-action $\iota_0 \circ \sigma$.*

Proof. Write $\underline{x} = [x_1, \ldots, x_m]$, where the x_i are $\mathcal{O}_F$-linear homomorphisms $x_i \colon A_0 \to A$. Since T has rank $\geq n-1$, we may assume (rearranging the elements x_i if necessary) that $\underline{x}^\flat = [x_1, \ldots, x_{n-1}]$ has nonsingular Gram matrix $(\underline{x}^\flat, \underline{x}^\flat)$. Then the map

$$f \colon A \xrightarrow{\sqrt{\Delta} \circ (x_1^\dagger \times \cdots \times x_{n-1}^\dagger)} A_0^{n-1} \tag{4.7.3}$$

is a homomorphism and a surjection of fppf sheaves. We form the "isogeny complement" in the standard way, i.e. we let A^- be the reduced connected component of $\ker f$. If $j \colon A^- \to A$ is the natural inclusion, then the map $(A_0)^{n-1} \times A^- \xrightarrow{x_1 \times \cdots \times x_{n-1} \times j} A$ is an $\mathcal{O}_F$-linear isogeny.

Note that A^- is $\mathcal{O}_F$-linearly isogenous to an elliptic curve with $\mathcal{O}_F$ action of signature $(0,1)$: if $\mathrm{char}(k) = p > 0$ with p nonsplit in $\mathcal{O}_F$, then A^- is supersingular, so apply Skolem–Noether to $\mathrm{End}(A^-) \otimes \mathbb{Q}$; otherwise, A^- automatically has signature $(0,1)$ because A has signature $(n-1,1)$.

If κ is algebraically closed, any two elliptic curves over κ with $\mathcal{O}_F$-action of the same signature are $\mathcal{O}_F$-linearly isogenous. This is classical: lift to characteristic zero to reduce to $\kappa = \mathbb{C}$ (the moduli stack $\mathscr{M}_0 \to \operatorname{Spec} \mathcal{O}_F$ is étale; more classically, see Deuring [Deu41]); recall that elliptic curves over $\mathbb{C}$ with $\mathcal{O}_F$-action are defined and isogenous over $\overline{\mathbb{Q}}$. By a standard limiting argument, we conclude that A^- and A_0^σ are $\mathcal{O}_F$-linearly isogenous over a finite extension of the (not necessarily algebraically closed) original field κ. □

Remark 4.7.2. If p is a prime which splits in $\mathcal{O}_F$ and if $\mathrm{rank}(T) \geq n$, then $\mathcal{Z}(T)_{\mathbb{F}_p} = \emptyset$. This is a standard argument (e.g. [KR14, Lemma 2.21]): if κ is a field of characteristic p and $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) \in \mathcal{Z}(T)(\kappa)$, arguing as in Lemma 4.7.1 shows that A is $\mathcal{O}_F$-linearly isogenous to A_0^n. This contradicts Lemma 4.7.1, because there is no nonzero $\mathcal{O}_F$-linear map $A_0 \to A_0^\sigma$ as A_0 and A_0^σ have $\mathcal{O}_F$-action of opposite signature (e.g. there are no nonzero maps of the underlying ordinary p-divisible groups).

We say a characteristic $p > 0$ geometric point $(A_0, \iota_0, \lambda_0, A, \iota, \lambda)$ of $\mathcal{M}$ lies in the *supersingular locus* if A_0 and A are supersingular abelian varieties (i.e. the associated p-divisible groups are supersingular). The following corollary also holds in arbitrary signature $(n-r, r)$ (i.e. all but the last sentence of Lemma 4.7.1 is valid for arbitrary signature $(n-r,r)$).

Corollary 4.7.3. *Let p be a prime which is nonsplit in $\mathcal{O}_F$. Fix $T \in \mathrm{Herm}_m(\mathbb{Q})$ with $\mathrm{rank}(T) \geq n-1$ and $m \geq 0$. The morphism $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p} \to \mathcal{M}_{\overline{\mathbb{F}}_p}$ factors (set-theoretically) through the supersingular locus on $\mathcal{M}_{\overline{\mathbb{F}}_p}$.*

Proof. Follows from Lemma 4.7.1 and Deuring's classical results on endomorphisms of elliptic curves in positive characteristic [Deu41] (i.e. over a field of characteristic $p > 0$, the p-divisible group of an elliptic curve with $\mathcal{O}_F$-action is supersingular (resp. ordinary) if p is nonsplit (resp. split) in $\mathcal{O}_F$). Here we used the notation $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p} \coloneqq \mathcal{Z}(T) \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \overline{\mathbb{F}}_p$ and similarly for $\mathcal{M}$. □

Lemma 4.7.4. *Fix $T \in \mathrm{Herm}_m(\mathbb{Q})$ with $\mathrm{rank}(T) \geq n-1$ and $m \geq 0$. Then the horizontal special cycle $\mathcal{Z}(T)_{\mathscr{H}}$ is proper and quasi-finite over $\operatorname{Spec} \mathcal{O}_F[1/d_L]$.*

Proof. By Lemma 3.5.5, we know the generic fiber $\mathcal{Z}(T)_{\mathscr{H},F} \to \operatorname{Spec} F$ is smooth of relative dimension 0. Hence each generic point of $\mathcal{Z}(T)_{\mathscr{H}}$ is the image of a map $\operatorname{Spec} E \to \mathcal{Z}(T)$ for some number field E, corresponding to an object $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \underline{x}) \in \mathcal{Z}(T)(E)$. By Lemma 4.7.1, we know that A is isogenous to a product of elliptic curves with complex multiplication by $\mathcal{O}_F$. It is a classical result of Deuring that such elliptic curves have everywhere potentially good reduction, so A_0 and A have everywhere potentially good reduction [Deu41; ST68]. Enlarging E if necessary, we can thus extend $\operatorname{Spec} E \to \mathcal{Z}(T)$ to a morphism $\operatorname{Spec} \mathcal{O}_E[1/d_L] \to \mathcal{Z}(T)$ (the Néron mapping property ensures that the datum $(\iota_0, \lambda_0, \iota, \lambda, \underline{x})$ extends as well; the polarizations must extend to polarizations as in the proof of [FC90, Theorem 1.9]).

Hence each irreducible component of $\mathcal{Z}(T)_{\mathscr{H}}$ may be covered by a morphism $\operatorname{Spec} \mathcal{O}_E[1/d_L] \to \mathcal{Z}(T)$ for some number field E. Since $\mathcal{Z}(T)$ is quasi-compact and separated, this implies that $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is proper and quasi-finite. $\square$

Lemma 4.7.5. *For $m \geq 0$, suppose $T \in \mathrm{Herm}_m(\mathbb{Q})$ has $\mathrm{rank}(T) \geq n-1$. Then the structure map $\mathcal{Z}(T) \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is proper.*

Proof. We already know that the horizontal part $\mathcal{Z}(T)_{\mathscr{H}}$ is proper over $\operatorname{Spec} \mathcal{O}_F[1/d_L]$, so it suffices to check that every irreducible component of $\mathcal{Z}(T)$ in characteristic $p \nmid d_L$ is proper over $\operatorname{Spec} \mathbb{F}_p$. It is enough to check that $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ is proper by fpqc descent (e.g. use away-from-p level structure to replace $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p}$ with a finite cover by a scheme, then use fpqc descent for morphisms of schemes). It is enough to check properness of the map $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p,\mathrm{red}} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ on reduced substacks (e.g. by local Noetherianity of these algebraic stacks, or because $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ is locally of finite type).

The supersingular locus on $\mathcal{M}_{\overline{\mathbb{F}}_p}$ is proper (follows from the proof of [Oor74, Theorem 1.1a]; and finiteness of the forgetful map $\mathscr{M}^{(d)} \to \mathcal{A}_{n,d}$ in Section 3.1). Properness of $\mathcal{Z}(T)_{\overline{\mathbb{F}}_p} \to \operatorname{Spec} \overline{\mathbb{F}}_p$ now follows from Corollary 4.7.3. $\square$

Part II

Local special cycles

5 Moduli spaces of p-divisible groups

We review some unitary Rapoport–Zink spaces [RZ96] and their special cycles, which will be used in p-adic uniformization of the moduli stacks from Section 3. Some notation on p-divisible groups is collected in Appendix B.1.

Fix a prime p and let $F/\mathbb{Q}_p$ be a degree 2 étale algebra, i.e. $F/\mathbb{Q}_p$ is an inert quadratic extension, a ramified quadratic extension, or $F = \mathbb{Q}_p \times \mathbb{Q}_p$. If $F/\mathbb{Q}_p$ is ramified, we assume $p \neq 2$. Write $a \mapsto a^\sigma$ for the nontrivial automorphism σ of F over $\mathbb{Q}_p$. We write $\mathcal{O}_F$ for the integral closure of $\mathbb{Z}_p$ in F (hence $\mathcal{O}_F = \mathbb{Z}_p \times \mathbb{Z}_p$ in the split case). If $F/\mathbb{Q}_p$ is ramified, we write ϖ for a uniformizer of $\mathcal{O}_F$ satisfying $\varpi^\sigma = -\varpi$.

We use the usual notation $\breve{\mathbb{Q}}_p$ for the completion of the maximal unramified extension of $\mathbb{Q}_p$, with ring of integers $\breve{\mathbb{Z}}_p$.

For $F/\mathbb{Q}_p$ nonsplit, let $\breve{F}$ be the completion of the maximal unramified extension of F. If $F/\mathbb{Q}_p$ is split, set $\breve{F} = \breve{\mathbb{Q}}_p$, and view $\breve{F}$ as an F-algebra by choosing one of the two morphisms of $\mathbb{Q}_p$-algebras $F \to \breve{F}$. We also equip $\breve{F}$ with the structure of a $\breve{\mathbb{Q}}_p$-algebra (taking the identity map if $F/\mathbb{Q}_p$ is split).

In all cases, let $\mathcal{O}_{\breve{F}} \subseteq \breve{F}$ be the ring of integers and let $\overline{k}$ be the residue field of $\breve{F}$. There is a canonical map $\mathcal{O}_F \to \mathcal{O}_{\breve{F}}$ (using the above choice of $F \to \breve{F}$ when $F/\mathbb{Q}_p$ is split).

We write $\Delta \subseteq \mathbb{Z}_p$ (resp. $\mathfrak{d} \subseteq \mathcal{O}_F$) for the discriminant ideal (resp. different ideal), which is $\Delta = \mathbb{Z}_p$ and $\mathfrak{d} = \mathcal{O}_F$ in the split case. We also abuse notation and write $\mathfrak{d}$ for a chosen generator of the different ideal satisfying $\mathfrak{d}^\sigma = -\mathfrak{d}$, taking $\mathfrak{d} = \varpi$ in the ramified case.

In the split case, let e^+ (resp. e^-) be the nontrivial idempotent in $\mathcal{O}_F$ which maps to $1 \in \mathcal{O}_{\breve{F}}$ (resp. $0 \in \mathcal{O}_{\breve{F}}$). Given an $\mathcal{O}_F$-module M, we write $M = M^+ \oplus M^-$ where e^+ projects to M^+ and e^- projects to M^-. We use similar notation $f = f^+ \oplus f^-$ for morphisms $f \colon M \to M'$ of $\mathcal{O}_F$-modules. We often use this for p-divisible groups X with $\mathcal{O}_F$ action, e.g. $X = X^+ \times X^-$ (and similarly for $\mathcal{O}_F$-linear quasi-homomorphisms).

5.1 Rapoport–Zink spaces

Definition 5.1.1. Let S be a formal scheme over $\operatorname{Spf} \mathbb{Z}_p$ and let $n \geq 1$ be an integer. By a *Hermitian p-divisible group* over S, we mean a tuple (X, ι, λ) where

X	is a p-divisible group over S of height $2n$ and dimension n
$\iota \colon \mathcal{O}_F \to \operatorname{End}(X)$	is a ring homomorphism
$\lambda \colon X \to X^\vee$	is a quasi-polarization satisfying:

(1) (Action compatibility) The Rosati involution $\dagger$ on $\operatorname{End}^0(X)$ satisfies $\iota(a)^\dagger = \iota(a^\sigma)$ for all $a \in \mathcal{O}_F$.

An *isomorphism* of Hermitian p-divisible groups is an isomorphism of underlying p-divisible groups which respects the $\mathcal{O}_F$-actions and polarizations.

In Part II, we only consider Hermitian p-divisible groups over formal schemes S equipped with a morphism $S \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$, and we assume that X is supersingular (resp. ordinary) if $F/\mathbb{Q}_p$ is nonsplit (resp. split).

We primarily discuss Hermitian p-divisible groups satisfying either of the following two conditions.

(2) (Principal polarization) The quasi-polarization λ is a principal polarization.

(2°) (Polarization condition $\circ$) Assume n is even if $F/\mathbb{Q}_p$ is ramified. The quasi-polarization $\Delta\lambda$ is a polarization, and $\ker(\Delta\lambda) = X[\iota(\mathfrak{d})]$.

In these cases, we say that (X, ι, λ) is *principally polarized* or *$\circ$-polarized* respectively.[24] If $F/\mathbb{Q}_p$ is ramified, we also use the alternative terminology *ϖ^{-1}-modular*.

Given an integer r with $0 \leq r \leq n$, we next consider

(1) (Kottwitz $(n-r, r)$ signature condition) For all $a \in \mathcal{O}_F$, the characteristic polynomial of $\iota(a)$ acting on $\operatorname{Lie} X$ is $(T-a)^{n-r}(T-a^\sigma)^r \in \mathcal{O}_S[T]$.

for pairs (X, ι), i.e. n-dimensional p-divisible groups X over a formal scheme S with action $\iota \colon \mathcal{O}_F \to \operatorname{End}(X)$. Here we view $\mathcal{O}_S$ as an $\mathcal{O}_F$-algebra via $S \to \operatorname{Spf} \mathcal{O}_{\breve{F}} \to \operatorname{Spec} \mathcal{O}_F$.

If (X, ι, λ) is a Hermitian p-divisible group of signature $(n-r, r)$, then $(X^\sigma, \iota^\sigma, \lambda^\sigma)$ with

$$X^\sigma = X \qquad \iota^\sigma = \iota \circ \sigma \qquad \lambda^\sigma = \lambda \tag{5.1.1}$$

[24] The local analogue of Footnote 17 applies as well.

is a Hermitian p-divisible group of signature $(r, n-r)$. We use similar notation $(X, \iota) \leftrightarrow (X^\sigma, \iota^\sigma)$ without the presence of a polarization. This allows us to switch between signatures $(n-r, r)$ and signature $(r, n-r)$ (e.g. for comparison with the literature).

From here on, we always implicitly restrict to signature $(n-1, 1)$ (and even n) when discussing $\circ$-polarized Hermitian p-divisible groups for ramified $F/\mathbb{Q}_p$. In this case, we also impose

(2) (Pappas wedge condition) For all $a \in \mathcal{O}_F$, the action of $\iota(a)$ on $\operatorname{Lie} X$ satisfies

$$\bigwedge^2 (\iota(a) - a) = 0 \quad \text{and} \quad \bigwedge^n (\iota(a) - a^\sigma) = 0.$$

(3) (PRRSZ spin condition) For every geometric point $\overline{s}$ of S, the action of $(\iota(a) - a)$ on $\operatorname{Lie} X_{\overline{s}}$ is nonzero for some $a \in \mathcal{O}_F$

The signature $(n-1, 1)$ condition implies that the equation involving $\bigwedge^n$ in the wedge condition is automatic, and that the wedge condition is empty if $n = 2$.

We temporarily allow $p = 2$ even if $F/\mathbb{Q}_p$ is ramified. Recall that there exists a supersingular (resp. ordinary) p-divisible group $\mathbf{X}_0$ of height 2 and dimension 1 over $\overline{k}$, and that $\mathbf{X}_0$ is unique up to isomorphism (this also holds for any algebraically closed field κ over $\overline{k}$). In the supersingular case $\mathbf{X}_0$ is given by a Lubin–Tate formal group law, and in the ordinary case we have $\mathbf{X}_0 \cong \boldsymbol{\mu}_{p^\infty} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}$. We have $\operatorname{End}(\mathbf{X}_0) \cong \mathcal{O}_D$ (resp. $\operatorname{End}(\mathbf{X}_0) \cong \mathbb{Z}_p \times \mathbb{Z}_p$) in the supersingular case (resp. ordinary case) where $\mathcal{O}_D$ is the unique maximal order in the quaternion division algebra D over $\mathbb{Q}_p$ (e.g. [Gro86] or [Wew07, §1]).

Quasi-polarizations on $\mathbf{X}_0$ exist and are unique up to $\mathbb{Q}_p^\times$ scalar, and there exists a principal polarization $\lambda_{\mathbf{X}_0}$ on $\mathbf{X}_0$ (unique up to $\mathbb{Z}_p^\times$ scalar). The induced Rosati involution on $\operatorname{End}(\mathbf{X}_0)$ is the standard involution (this can be verified on the Dieudonné module, see e.g. [RSZ17, Page 2205]), hence induces the nontrivial Galois involution on F for any embedding $F \hookrightarrow \operatorname{End}^0(\mathbf{X}_0)$ (if such an embedding exists).

From now on, we assume $\mathbf{X}_0$ is supersingular (resp. ordinary) if $F/\mathbb{Q}_p$ is nonsplit (resp. split). Then there exists an embedding $j \colon \mathcal{O}_F \hookrightarrow \operatorname{End}(\mathbf{X}_0)$. Given such a j, form $(\mathbf{X}_0, j)$ and $(\mathbf{X}_0^\sigma, j^\sigma)$ as above. There is an $\mathcal{O}_F$-linear isogeny of degree $|\Delta|_p^{-1}$

$$\begin{aligned} \mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F &\longrightarrow \mathbf{X}_0 \times \mathbf{X}_0^\sigma \\ x \otimes a &\longmapsto (j(a)x, j(a^\sigma)x) \end{aligned} \tag{5.1.2}$$

where $X \otimes_{\mathbb{Z}_p} \mathcal{O}_F$ is the Serre tensor p-divisible group (B.1.1), with its Serre tensor $\mathcal{O}_F$-action. See also [KR11, Lemma 6.2] (there in the inert case for $p \neq 2$, but the version in (5.1.2) allows for $p = 2$).

Suppose $\lambda_{\mathbf{X}_0}$ is a principal polarization of $\mathbf{X}_0$. Under the map in (5.1.2), the ($\mathcal{O}_F$-action compatible) product polarization $\lambda_{\mathbf{X}_0} \times \lambda_{\mathbf{X}_0}^\sigma$ on $\mathbf{X}_0 \times \mathbf{X}_0^\sigma$ pulls back to the polarization

$$\lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}} \colon \mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F \to (\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F^*) \cong \mathbf{X}_0^\vee \otimes_{\mathbb{Z}_p} \mathcal{O}_F^* \tag{5.1.3}$$

where $\mathcal{O}_F^* \coloneqq \mathrm{Hom}_{\mathbb{Z}_p}(\mathcal{O}_F, \mathbb{Z}_p)$ and $\lambda_{\mathrm{tr}} \colon \mathcal{O}_F \to \mathcal{O}_F^*$ is induced by the symmetric bilinear pairing $\mathrm{tr}_{F/\mathbb{Q}_p}(a^\sigma b)$ on $\mathcal{O}_F$. Indeed, after picking a $\mathbb{Z}_p$-basis $\{1, \alpha\}$ for $\mathcal{O}_F$ to identify $\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F \cong \mathbf{X}_0^2$, the map in (5.1.2) is given by the matrix

$$\phi = \begin{pmatrix} 1 & \alpha \\ 1 & \alpha^\sigma \end{pmatrix} \in M_{2,2}(\mathcal{O}_F) \tag{5.1.4}$$

and $\det \phi$ generates the different ideal of $\mathcal{O}_F/\mathbb{Z}_p$ (so Smith normal form shows $\deg \phi = |\Delta|_p^{-1}$). Identifying $\mathbf{X}_0^\vee \otimes_{\mathbb{Z}_p} \mathcal{O}_F^* \cong \mathbf{X}_0^{\vee 2}$ using the basis of $\mathcal{O}_F^*$ dual to $\{1, \alpha\}$, the preceding claim about pullback polarizations follows because $({}^t\phi^\sigma)\phi$ (where ${}^t\phi^\sigma$ means conjugate transpose) is the Gram matrix for the basis $\{1, \alpha\}$ and the trace pairing on $\mathcal{O}_F$.

If $p \neq 2$, the polarization $\lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}}$ coincides with the polarization on $\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F$ described in [KR11, (6.2)] (inert case, with modification as in [RSZ17, Footnote 4]) and [RSZ17, (3.4)] (ramified case, though we normalize differently).

Suppose $\iota_{\mathbf{X}_0} \colon \mathcal{O}_F \hookrightarrow \mathrm{End}(\mathbf{X}_0)$ is an action of signature $(1, 0)$. The pair $(\mathbf{X}_0, \iota_{\mathbf{X}_0})$ exists and is unique up to isomorphism. In the split case, the element $\iota_{\mathbf{X}_0}(e^+)$ is projection to $\boldsymbol{\mu}_{p^\infty}$ and $\iota_{\mathbf{X}_0}(e^-)$ is projection to $\underline{\mathbb{Q}_p/\mathbb{Z}_p}$. In the ramified case, note that $(\mathbf{X}_0, \iota_{\mathbf{X}_0})$ is simultaneously of signature $(1, 0)$ and $(0, 1)$.

Fix $(\mathbf{X}_0, \iota_{\mathbf{X}_0})$ as above, and form $(\mathbf{X}_0^\sigma, \iota_{\mathbf{X}_0}^\sigma)$. We have

$$\mathrm{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}_0^\sigma) \cong \begin{cases} \mathcal{O}_F & \text{if } F/\mathbb{Q}_p \text{ is nonsplit} \\ 0 & \text{if } F/\mathbb{Q}_p \text{ is split} \end{cases} \tag{5.1.5}$$

as $\mathcal{O}_F$-modules by precomposition. Using $\mathrm{End}(\mathbf{X}_0) \cong \mathcal{O}_D$ in the nonsplit cases, we find that the $\mathcal{O}_F$-module $\mathrm{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}_0^\sigma)$ is generated by any isogeny of degree p if $F/\mathbb{Q}_p$ is inert, and is generated by an isomorphism if $F/\mathbb{Q}_p$ is ramified (namely any element $a \in \mathcal{O}_D^\times$ such that conjugation by a induces the nontrivial Galois involution on F). We have $\mathrm{End}_{\mathcal{O}_F}(\mathbf{X}_0) = \mathcal{O}_F$ in all cases.

Suppose $\lambda_{\mathbf{X}_0}$ is a principal polarization of $\mathbf{X}_0$. If $F/\mathbb{Q}_p$ is unramified, the triple $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0})$ is unique (up to isomorphism): given another polarization $\lambda'_{\mathbf{X}_0}$, we have $\lambda_{\mathbf{X}_0}^{-1} \circ \lambda'_{\mathbf{X}_0} \in \mathbb{Z}_p^\times$ and know that the norm map $N_{F/\mathbb{Q}_p} \colon \mathcal{O}_F^\times \to \mathbb{Z}_p^\times$ is surjective. If $F/\mathbb{Q}_p$ is ramified, the same reasoning shows that there are two choices of $\lambda_{\mathbf{X}_0}$ (differing by a $\mathbb{Z}_p^\times$ scalar) because $N_{F/\mathbb{Q}_p}(\mathcal{O}_F^\times) \subseteq \mathbb{Z}_p^\times$ has index 2. Fix a choice of $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0})$.

We now re-impose our running assumption that $p \neq 2$ if $F/\mathbb{Q}_p$ is ramified. Fix any $\circ$-polarized Hermitian p-divisible group $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ of signature $(n-r, r)$ over $\overline{k}$. Such triples exist and are unique up to F-linear quasi-isogenies preserving polarizations exactly. This uniqueness may be proved via Dieudonné theory: see [Vol10, §1] (inert case, but we allow $p = 2$ by the same proof) and [RSZ17, Proposition 3.1] [RSZ18, §6] (ramified case). In the split case, we have a stronger uniqueness statement.

Lemma 5.1.2. *For $F/\mathbb{Q}_p$ split, the triple $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ is unique up to isomorphism. This also holds over any algebraically closed field extension κ of $\overline{k}$.*

Proof. Decompose $\mathbf{X} = \mathbf{X}^+ \times \mathbf{X}^-$ using the idempotents in the $\mathcal{O}_F = \mathbb{Z}_p \times \mathbb{Z}_p$ action given by $\iota_{\mathbf{X}}$. Then $\mathbf{X}^+$ and $\mathbf{X}^-$ are the unique ordinary p-divisible groups over κ of height n and the correct dimension ($n-r$ and r, respectively). Uniqueness of $\lambda_{\mathbf{X}}$ (up to isomorphism) corresponds to the following fact: there is a unique self-dual Hermitian $\mathcal{O}_F$-lattice (up to isomorphism) of any given rank. $\square$

For existence, we may construct $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ as follows. For $F/\mathbb{Q}_p$ unramified, we can take $\mathbf{X} = (\mathbf{X}_0)^{n-r} \times (\mathbf{X}_0^{\sigma})^r$ (with the product $\mathcal{O}_F$-action and polarization).

For $F/\mathbb{Q}_p$ ramified, we can take $\mathbf{X} = (\mathbf{X}_0)^{n-2} \times (\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F)$ using the Serre tensor construction (B.1.1). The $\mathcal{O}_F$-action $\iota_{\mathbf{X}}$ is diagonal on $(\mathbf{X}_0)^{n-2}$ and given by the Serre tensor $\mathcal{O}_F$-action on $\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F$. We can take the product quasi-polarization $\lambda_{\mathbf{X}}$ of $\mathbf{X}$ given by

$$\begin{gathered} -\iota_{\mathbf{X}_0}(\varpi)^{-2} \circ (\lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}}) \quad \text{on } \mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F, \text{ and} \\ \underbrace{\begin{pmatrix} 0 & \lambda_{\mathbf{X}_0} \circ -\iota_{\mathbf{X}_0}(\varpi)^{-1} \\ \lambda_{\mathbf{X}_0} \circ \iota_{\mathbf{X}_0}(\varpi)^{-1} & 0 \end{pmatrix} \times \cdots \times \begin{pmatrix} 0 & \lambda_{\mathbf{X}_0} \circ -\iota_{\mathbf{X}_0}(\varpi)^{-1} \\ \lambda_{\mathbf{X}_0} \circ \iota_{\mathbf{X}_0}(\varpi)^{-1} & 0 \end{pmatrix}}_{(n-2)/2 \text{ times}} \quad \text{on } \mathbf{X}_0^{n-2}. \end{gathered} \tag{5.1.6}$$

This is the construction of [RSZ17, §3.3] (but rescaled).

Given a principally polarized triple $(X_0, \iota_0, \lambda_0)$ of signature $(1, 0)$ over some base scheme S, a *framing similitude quasi-isogeny* ρ_0 is an F-linear quasi-isogeny $X_{0,\overline{S}} \to \mathbf{X}_{0,\overline{S}}$ such that

$$\rho_0^*(\lambda_{\mathbf{X}_0,\overline{S}}) = b\lambda_{0,\overline{S}} \quad \text{for some } b \in \mathbb{Q}_p^{\times} \tag{5.1.7}$$

where the subscript indicates base-change to $\overline{S} \coloneqq S \times_{\operatorname{Spec} \mathcal{O}_{\breve{F}}} \operatorname{Spec} \overline{k}$ (and where $b \in \mathbb{Q}_p^{\times}$ really means a section of the constant sheaf). We call $(X_0, \iota_0, \lambda_0, \rho)$ a *framed similitude* tuple. An *isomorphism* of framed similitude tuples $f \colon (X_0, \iota_0, \lambda_0, \rho_0) \to (X_0', \iota_0', \lambda_0', \rho_0')$ is an $\mathcal{O}_F$-linear isomorphism of p-divisible groups $f \colon X_0 \to X_0'$ such that $f^*(\lambda_0')$ and λ_0 agree up to $\mathbb{Z}_p^{\times}$-scalar and also $\rho_0' \circ f_{\overline{S}} = \rho_0$.

Given a $\circ$-polarized triple (X, ι, λ) of signature $(n-r, r)$ over some base scheme S, we define a *similitude framing quasi-isogeny* $\rho \colon X_{\overline{S}} \to \mathbf{X}_{\overline{S}}$ in the same way. A *framing quasi-isogeny* $\rho \colon X_{\overline{S}} \to \mathbf{X}_{\overline{S}}$ is given by the stricter requirement $b \in \mathbb{Z}_p^\times$. In these two cases, we call the datum $(X, \iota, \lambda, \rho)$ a *framed similitude* tuple and a *framed tuple*, respectively. In both cases, isomorphisms of two such tuples are defined as before: isomorphisms of p-divisible groups which are $\mathcal{O}_F$-linear, preserve polarizations up to $\mathbb{Z}_p^\times$, and commute with framings.

Definition 5.1.3. We consider three *Rapoport–Zink spaces* over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$, given by the (set-valued) functors

$$\begin{aligned}
\mathcal{N}(1,0)'(S) &\coloneqq \{\text{isomorphism classes of framed similitude tuples } (X_0, \iota_0, \lambda_0, \rho_0) \text{ over } S\} \\
\mathcal{N}(n-r,r)'(S) &\coloneqq \{\text{isomorphism classes of framed similitude tuples } (X, \iota, \lambda, \rho) \text{ over } S\} \\
\mathcal{N}(n-r,r)(S) &\coloneqq \{\text{isomorphism classes of framed tuples } (X, \iota, \lambda, \rho) \text{ over } S\}
\end{aligned}$$

for schemes S over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$. Here, signature $(1,0)$ and principal polarizations are understood for $\mathcal{N}(1,0)'$. Signature $(n-r,r)$ and $\circ$-polarizations are understood for $\mathcal{N}(n-r,r)'$ and $\mathcal{N}(n-r,r)$.

These Rapoport–Zink spaces do not depend on the choices of framing objects (up to functorial isomorphism). The functor $\mathcal{N}(n-r,r)$ is canonically isomorphic to its variant where instead we require framing quasi-isogenies and isomorphisms of framed tuples to preserve polarizations exactly (not just up to $\mathbb{Z}_p^\times$ scalar). If S is a formal scheme, we also write e.g. $\mathcal{N}(n-r,r)(S) \coloneqq \operatorname{Hom}(S, \mathcal{N}(n-r,r))$.

Lemma 5.1.4. *Each of $\mathcal{N}(1,0)'$, $\mathcal{N}(n-r,r)'$, and $\mathcal{N}(n-r,r)$ is represented by a locally Noetherian formal scheme which is formally locally of finite type and separated over* $\operatorname{Spf} \mathcal{O}_{\breve{F}}$. *Each irreducible component of the reduced subschemes is projective over $\overline{k}$.*

Proof. Representability, local Noetherianity, and formally locally of finite type-ness follow via [RZ96, Theorem 2.16]; various closedness statements can be checked via [RZ96, Proposition 2.9], which holds verbatim with "isogeny" replaced by "homomorphism". Projectivity of the reduced irreducible components follows from [RZ96, Proposition 2.32], also using [RSZ17, Proposition 3.8] in the ramified case. Separatedness now follows because this can be checked on underlying reduced subschemes (then apply the valuative criterion). $\square$

Lemma 5.1.5. *The formal scheme $\mathcal{N}(n-r,r)$ is regular and the structure morphism $\mathcal{N}(n-r,r) \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$ is formally smooth of relative dimension $(n-r,r)$.*

Proof. We know the structure map $\mathcal{N}(n-r,r) \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$ is formally smooth of relative dimension $(n-r,r)$ via [Mih22, Proposition 1.3] in the unramified case (also Section 3.5, where we allow $p = 2$) and [RSZ17, Proposition 3.8] for the ramified case. We conclude that $\mathcal{N}(n-r,r)$ is regular because the map to $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ is formally smooth and formally locally of finite type. $\square$

For $F/\mathbb{Q}_p$ ramified, the Rapoport–Zink space $\mathcal{N}(n-1,1)$ is often called *exotic smooth* in the literature, following the terminology of [RSZ17].

Lemma 5.1.6. *There is an isomorphism*

$$\begin{aligned} F^\times/\mathcal{O}_F^\times &\xrightarrow{\quad\sim\quad} \mathcal{N}(1,0)' \\ a &\longmapsto (\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0}, a \cdot \mathrm{id}_{\mathbf{X}_0}) \end{aligned} \tag{5.1.8}$$

where the left-hand side is viewed as a constant formal scheme over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$.

Proof. In the nonsplit case, [How19, Proposition 2.1] states that $\mathcal{N}(1,0)'$ is a disjoint union of copies of $\operatorname{Spf} \mathcal{O}_{\breve{F}}$, so it is enough to check the claim on $\overline{k}$-points. The claim on $\overline{k}$-points follows from uniqueness of the triple $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0})$ (up to isomorphism preserving polarizations up to $\mathbb{Z}_p^\times$ scalar) and the equality $\operatorname{End}_F^0(\mathbf{X}_0) = F$.

The split case holds via the following similar argument. The map $\mathcal{N}(1,0)'(\kappa) \to \mathcal{N}(1,0)'(\kappa')$ is bijective for any extension of algebraically closed fields $\kappa \subseteq \kappa'$ (essentially by uniqueness of the triple $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0})$, which holds over any algebraically closed field of characteristic p). So the reduced subscheme $\mathcal{N}(1,0)'_{\mathrm{red}}$ is isomorphic to a (discrete) disjoint union of copies of $\operatorname{Spec} \overline{k}$. We also see that the map $F^\times/\mathcal{O}_F^\times \to \mathcal{N}(1,0)'$ is bijective on $\overline{k}$-points (this follows as in the nonsplit case). To finish, note that $\mathcal{N}(1,0)'$ is formally étale over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ (e.g. by Grothendieck–Messing theory as in Section 3.5). $\square$

Definition 5.1.7. By the *canonical lifting* of $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0})$, we mean the tuple $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \lambda_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ corresponding (via Lemma 5.1.6) to the unique section $\operatorname{Spf} \mathcal{O}_{\breve{F}} \to \mathcal{N}(1,0)'$ associated to the element $1 \in F^\times/\mathcal{O}_F^\times$.

Definition 5.1.8. We define the open and closed subfunctor $\mathcal{N}' \subseteq \mathcal{N}(1,0)' \times \mathcal{N}(n-r,r)'$ as

$$\mathcal{N}'(S) := \left\{ (X_0, \iota_0, \lambda_0, \rho_0, X, \iota, \lambda, \rho) : \begin{array}{l} \rho_0^*(\lambda_{\mathbf{X}_0, \overline{S}}) = b_0 \lambda_{0,\overline{S}} \quad \rho^*(\lambda_{\mathbf{X},\overline{S}}) = b \lambda_{\overline{S}} \\ \text{with } b_0 = b \text{ in } \mathbb{Q}_p^\times/\mathbb{Z}_p^\times \end{array} \right\}$$

for schemes S over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$.

With b as above and $a \in F^\times$ any element with $N_{F/\mathbb{Q}_p}(a) = b$ in $\mathbb{Q}_p^\times/\mathbb{Z}_p^\times$, there is an isomorphism

$$\begin{aligned} \mathcal{N}' &\xrightarrow{\quad\sim\quad} \mathcal{N}(1,0)' \times \mathcal{N}(n-r,r) \\ (X_0, \iota_0, \lambda_0, \rho_0, X, \iota, \lambda, \rho) &\longmapsto (X_0, \iota_0, \lambda_0, \rho_0, X, \iota, \lambda, a^{-1}\rho). \end{aligned} \tag{5.1.9}$$

Whenever we write $(X_0, \iota_0, \lambda_0, \rho_0, X, \iota, \lambda, \rho)$ for a (functorial) point of $\mathcal{N}'$, we mean an object as on the left of (5.1.9) (i.e. ρ preserves polarizations up to $\mathbb{Q}_p^\times$ scalar).

The functorial assignment $(X, \iota, \lambda, \rho) \mapsto \operatorname{Lie} X$ defines a locally free sheaf Lie on $\mathcal{N}(n-r, r)$. In the case of signature $(n-1, 1)$, there is a unique maximal local direct summand $\mathcal{F} \subseteq \operatorname{Lie}$ of rank $n-1$ such that the ι action on $\mathcal{F}$ (resp. $\operatorname{Lie}/\mathcal{F}$) is $\mathcal{O}_F$-linear (resp. σ-linear). The ramified case is proved in [LL22, Lemma 2.36] (and in the unramified case, we have a canonical eigenspace decomposition $\operatorname{Lie} = \mathcal{F} \oplus (\operatorname{Lie}/\mathcal{F})$ for the $\mathcal{O}_F$-action).

Consider the canonical lifting $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \lambda_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ (Definition 5.1.7). Given any $\operatorname{Spf} \mathcal{O}_{\breve{F}}$-scheme S, we write $\mathbb{D}(\mathfrak{X}_{0,S})(S)$ for evaluation of the (covariant) Dieudonné crystal $\mathbb{D}(\mathfrak{X}_{0,S})$ at $\mathrm{id}\colon S \to S$, with associated Hodge filtration step $F^0\mathbb{D}(\mathfrak{X}_{0,S})(S)$. The assignment $S \mapsto F^0\mathbb{D}(\mathfrak{X}_{0,S})(S)$ defines a (trivial) line bundle on $\mathcal{N}$, which we denote $\operatorname{Lie}_0^\vee$. The principal polarization $\lambda_{\mathfrak{X}_0}$ induces an identification $(\operatorname{Lie} \mathfrak{X}_{0,S})^\vee \cong (\operatorname{Lie} \mathfrak{X}_{0,S}^\vee)^\vee$ and the latter is $F^0\mathbb{D}(\mathfrak{X}_{0,S})(S)$.

Definition 5.1.9. The *tautological bundle* $\mathcal{E}$ on $\mathcal{N}(n-1,1)$ is the line bundle whose dual is $\mathcal{E}^\vee \coloneqq \underline{\operatorname{Hom}}(\operatorname{Lie}_0^\vee, \operatorname{Lie}/\mathcal{F})$.

The definition of $\mathcal{E}$ is taken from [How19, Definition 3.4] (at least in the inert case). The line bundle $\mathcal{E}$ on $\mathcal{N}(n-1,1)$ is a local analogue of the global tautological bundle (Section 4.3, also Definitions 3.1.7 and 3.2.6). We are recycling the notation $\mathcal{E}$ (but the global tautological bundle pulls back to $\mathcal{E}$ under Rapoport–Zink uniformization, see e.g. Section 11.8).

5.2 Local special cycles

We define certain local special cycles on Rapoport–Zink spaces, following [KR11, Definition 3.2] (there in the inert case). Retain notation from Section 5.1.

The space of *local special quasi-homomorphisms* means the F-module

$$\mathbf{W} \coloneqq \operatorname{Hom}_F^0(\mathbf{X}_0, \mathbf{X}). \tag{5.2.1}$$

If $F/\mathbb{Q}_p$ is nonsplit, then $\mathbf{W}$ is free of rank n (see also [RSZ17, Lemma 3.5] in the ramified case). If $F/\mathbb{Q}_p$ is split, then $\mathbf{W}$ is a free F-module of rank $n-r$ (because $\operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}_0^\sigma) = 0$ in the split case, in contrast with $\operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}_0^\sigma) \cong \mathcal{O}_F$ in the nonsplit cases). In the split case only, set $\mathbf{W}^\perp \coloneqq \operatorname{Hom}_F^0(\mathbf{X}_0^\sigma, \mathbf{X})$. In the nonsplit cases, set $\mathbf{W}^\perp = 0$.

Set

$$\mathbf{V} = \mathbf{W} \oplus \mathbf{W}^\perp \qquad \mathbf{V}_0 = \operatorname{Hom}_F^0(\mathbf{X}_0, \mathbf{X}_0). \tag{5.2.2}$$

In all cases, these are free F-modules of rank n and 1, respectively.

We equip $\mathbf{W}$, $\mathbf{W}^\perp$, and $\mathbf{V}_0$ with the (non-degenerate) Hermitian pairings $(x, y) = x^\dagger y \in \mathrm{End}^0_F(\mathbf{X}_0) = F$. We give $\mathbf{V}$ the Hermitian form making $\mathbf{W}$ and $\mathbf{W}^\perp$ orthogonal. We have $\varepsilon(\mathbf{V}) = (-1)^r$ if $F/\mathbb{Q}_p$ is nonsplit (resp. $\varepsilon(\mathbf{V}) = 1$ if $F/\mathbb{Q}_p$ is split). This follows upon inspecting the explicit framing tuples constructed in Section 5.1 (see [RSZ17, Lemma 3.5] for the ramified case).

Definition 5.2.1 (Kudla–Rapoport local special cycles)**.** Given any set $L \subseteq \mathbf{W}$, there is an associated *local special cycle*

$$\mathcal{Z}(L)' \subseteq \mathcal{N}' \quad (\text{resp. } \tilde{\mathcal{Z}}(L) \subseteq \mathcal{N}(n-r, r)) \tag{5.2.3}$$

which is the subfunctor consisting of tuples $(X_0, \iota_0, \lambda_0, X, \iota, \lambda, \rho)$ (resp. $(X, \iota, \lambda, \rho)$) over schemes S over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ such that, for all $x \in L$, the quasi-homomorphism

$$\rho^{-1} \circ x_{\overline{S}} \circ \rho_0 \colon X_{0,\overline{S}} \to X_{\overline{S}} \quad (\text{resp. } \rho^{-1} \circ x_{\overline{S}} \circ \rho_{\mathfrak{X}_0,\overline{S}} \colon \mathfrak{X}_{0,\overline{S}} \to X_{\overline{S}}) \tag{5.2.4}$$

lifts to a homomorphism $X_0 \to X$ (resp. $\mathbf{X}_{0,S} \to X$). Here $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \lambda_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ is the canonical lifting.

In the preceding definition, such lifts are unique (if they exist) by Drinfeld rigidity for quasi-homomorphisms of p-divisible groups. We know that $\mathcal{Z}(L)' \subseteq \mathcal{N}'$ and $\tilde{\mathcal{Z}}(L) \subseteq \mathcal{N}(n-r, r)$ are closed subfunctors (hence locally Noetherian formal schemes) by [RZ96, Proposition 2.9] for quasi-homomorphisms. From the definition, it is clear that $\mathcal{Z}(L)'$ and $\tilde{\mathcal{Z}}(L)$ depend only on the $\mathcal{O}_F$-span of L.

The isomorphism $\mathcal{N}' \xrightarrow{\sim} \mathcal{N}(1,0)' \times \mathcal{N}(n-r, r)$ of (5.1.9) induces an isomorphism

$$\mathcal{Z}(L)' \xrightarrow{\sim} \mathcal{N}(1,0)' \times \tilde{\mathcal{Z}}(L). \tag{5.2.5}$$

Lemma 5.2.2. *Let $L \subseteq \mathbf{W}$ be any subset. If $\mathcal{Z}(L)' \neq \emptyset$, then $(x, y) \in \mathfrak{d}^{-1}$ for all $x, y \in L$.*

Proof. If $\mathcal{Z}(L)' \neq \emptyset$, then $\mathcal{Z}(L)(\overline{k})' \neq \emptyset$ because $\mathcal{Z}(L)' \to \mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ is formally locally of finite type. If $\mathcal{Z}(L)(\overline{k})' \neq \emptyset$ and $x, y \in L$, we find $\mathfrak{d} x^\dagger y \in \mathrm{End}_{\mathcal{O}_F}(\mathbf{X}_0) = \mathcal{O}_F$ by the $\circ$-polarization condition defining $\mathcal{N}(n-r, r)$, where $\mathfrak{d}$ is the different ideal. $\square$

If $F/\mathbb{Q}_p$ is nonsplit, set $\mathcal{Z}(L) \coloneqq \tilde{\mathcal{Z}}(L)$ for any subset $L \subseteq \mathbf{W}$. If $F/\mathbb{Q}_p$ is split, we will instead define $\mathcal{Z}(L)$ as a certain open and closed subfunctor (see (5.4.4)) for later notational uniformity.

In all cases, we write $\mathcal{Z}(L)_{\mathscr{H}} \subseteq \mathcal{Z}(L)$ (*horizontal special cycle*) for the flat part of $\mathcal{Z}(L)$, i.e. the largest closed formal subscheme which is flat over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$.

5.3 Actions on Rapoport–Zink spaces

Consider the groups

$$I_0 \coloneqq \{\gamma_0 \in \mathrm{End}^0_F(\mathbf{X}_0) : \gamma_0^\dagger \gamma_0 \in \mathbb{Q}_p^\times\} \qquad I \coloneqq \{\gamma \in \mathrm{End}^0_F(\mathbf{X}) : \gamma^\dagger \gamma \in \mathbb{Q}_p^\times\} \tag{5.3.1}$$

$$I' \coloneqq \{(\gamma_0, \gamma) \in I_0 \times I : \gamma_0^\dagger \gamma_0 = \gamma^\dagger \gamma\} \qquad I_1 \coloneqq \{\gamma \in \mathrm{End}^0_F(\mathbf{X}) : \gamma^\dagger \gamma = 1\}. \tag{5.3.2}$$

We have $I_0 = F^\times$ (canonically). Using this identification, there is an isomorphism $I' \to I_0 \times I_1$ given by $(\gamma_0, \gamma) \mapsto (\gamma_0, \gamma_0^{-1}\gamma)$. We have actions

$$\begin{aligned} I &\circlearrowright \mathcal{N}(n-r,r)' & \qquad I_1 &\circlearrowright \mathcal{N}(n-r,r) \\ (X,\iota,\lambda,\rho) &\mapsto (X,\iota,\lambda,\gamma\circ\rho) & \qquad (X,\iota,\lambda,\rho) &\mapsto (X,\iota,\lambda,\gamma\circ\rho) \end{aligned} \tag{5.3.3}$$

$$\begin{aligned} &I' \circlearrowright \mathcal{N}' \\ &(X_0,\iota_0,\lambda_0,\rho_0,X,\iota,\lambda,\rho) \mapsto (X_0,\iota_0,\lambda_0,\gamma_0\circ\rho_0,X,\iota,\lambda,\gamma\circ\rho). \end{aligned} \tag{5.3.4}$$

These actions are compatible with the isomorphisms $I' \cong I_0 \times I_1$ and $\mathcal{N}' \cong \mathcal{N}(1,0)' \times \mathcal{N}(n-r,r)$. We have isomorphisms

$$I_0 \xrightarrow{\sim} GU(\mathbf{V}_0) \qquad I_1 \xrightarrow{\sim} U(\mathbf{W}) \times U(\mathbf{W}^\perp) \tag{5.3.5}$$

where $\gamma \in I_1$ acts on $\mathbf{V}$ as $x \mapsto \gamma \circ x$, and similarly for $\mathbf{V}_0$.

For any subset $L \subseteq \mathbf{W}$ with associated local special cycles $\mathcal{Z}'(L) \subseteq \mathcal{N}'$ and $\tilde{\mathcal{Z}}(L) \subseteq \mathcal{N}(n-r,r)$, the actions of I' and I_1 described above satisfy

$$(\gamma_0,\gamma)(\mathcal{Z}(L)') = \mathcal{Z}(\gamma L \gamma_0^{-1})' \quad \gamma(\tilde{\mathcal{Z}}(L)) = \tilde{\mathcal{Z}}(\gamma L). \tag{5.3.6}$$

We will also have $\gamma(\mathcal{Z}(L)) = \mathcal{Z}(\gamma(L))$ (already checked in the nonsplit cases; in the split case, this will be clear from the definition, see (5.4.4)).

5.4 Discrete reduced subschemes

In the nonsplit cases (at least if $p \neq 2$), the reduced subscheme $\mathcal{N}(n-1,1)_{\mathrm{red}}$ of $\mathcal{N}(n-1,1)$ admits a stratification by Deligne–Lusztig varieties, described by a certain Bruhat–Tits building [VW11; Wu16]. Later, we will use these results implicitly via citation to [LZ22a; LL22].

In this section, we further discuss some cases where the reduced subscheme is discrete (continuing to allow $p = 2$ if $F/\mathbb{Q}_p$ is unramified).

In the split case, set

$$\mathbf{L} \coloneqq \operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}) \qquad \mathbf{L}^{\perp} \coloneqq \operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0^{\sigma}, \mathbf{X}). \tag{5.4.1}$$

In the nonsplit case, define $\mathbf{L}$ in the same way but take $\mathbf{L}^{\perp} \coloneqq 0$. Let $K_{1,\mathbf{L}} \subseteq U(\mathbf{W})$ and $K_{1,\mathbf{L}^{\perp}} \subseteq U(\mathbf{W}^{\perp})$ be the respective stabilizers.

Lemma 5.4.1. *Consider signature $(n-r,r) = (1,1)$ if $F/\mathbb{Q}_p$ is nonsplit (resp. any signature $(n-r,r)$ if $F/\mathbb{Q}_p$ is split).*

(1) The framing object $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ is unique up to isomorphism. This also holds over any algebraically closed field κ over $\overline{k}$, at least if $F/\mathbb{Q}_p$ is unramified.

(2) The reduced scheme $\mathcal{N}(n-r,r)_{\mathrm{red}}$ is discrete (i.e. a disjoint union of copies of $\operatorname{Spec} \overline{k}$). If $F/\mathbb{Q}_p$ is inert (resp. ramified), then $\mathcal{N}(1,1)_{\mathrm{red}}$ is one point (resp. two points).

(3) The lattices $\mathbf{L} \subseteq \mathbf{W}$ and $\mathbf{L}^{\perp} \subseteq \mathbf{W}^{\perp}$ are maximal integral lattices. In the nonsplit cases, $\mathbf{L} \subseteq \mathbf{W} = \mathbf{V}$ is the unique maximal integral lattice.

(4) The group $I_1 \cong U(\mathbf{W}) \times U(\mathbf{W}^{\perp})$ acts transitively on $\mathcal{N}(n-r,r)(\overline{k})$. Consider the resulting surjection

$$\begin{aligned} \mathcal{N}(n-r,r)(\overline{k}) &\longrightarrow U(\mathbf{W})/K_{1,\mathbf{L}} \times U(\mathbf{W}^{\perp})/K_{1,\mathbf{L}^{\perp}} \\ (\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}}, (\gamma, \gamma^{\perp})) &\longmapsto (\gamma, \gamma^{\perp}). \end{aligned} \tag{5.4.2}$$

If $F/\mathbb{Q}_p$ is unramified, this map is a bijection. If $F/\mathbb{Q}_p$ is ramified, this map is 2-to-1. If $F/\mathbb{Q}_p$ is nonsplit, the set $U(\mathbf{W})/K_{1,\mathbf{L}} \times U(\mathbf{W}^{\perp})/K_{1,\mathbf{L}^{\perp}}$ has size 1.

(5) Consider the bijective identification

$$\begin{aligned} U(\mathbf{W})/K_1(\mathbf{L}) \times U(\mathbf{W}^{\perp})/K_1(\mathbf{L}^{\perp}) &\longleftrightarrow \left\{ \begin{array}{l} \text{maximal full-rank integral } \mathcal{O}_F\text{-lattices } N \subseteq \mathbf{V} \\ \text{where } N = M \oplus M^{\perp} \text{ with} \\ M \subseteq \mathbf{W} \text{ and } M^{\perp} \subseteq \mathbf{W}^{\perp} \end{array} \right\} \\ (\gamma, \gamma^{\perp}) &\longmapsto \gamma(\mathbf{L}) \oplus \gamma^{\perp}(\mathbf{L}^{\perp}). \end{aligned}$$

Given any subset $L \subseteq \mathbf{W}$, the subset $\tilde{\mathcal{Z}}(L)(\overline{k}) \subseteq \mathcal{N}(n-r,r)(\overline{k})$ is identified (via (5.4.2)) with the pre-image of the set of lattices $\{N : L \subseteq N\}$.

Proof.

(1) In the inert case, this follows from Dieudonné theory as in [Vol10, Proposition 1.10] (but we allow $p = 2$ by the same method), diagonalizability of Hermitian $\mathcal{O}_F$-lattices, and the following fact: consider the rank 2 Hermitian $\mathcal{O}_F$-lattice Λ with pairing $(-,-)$ specified by the Gram matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}, \tag{5.4.3}$$

and also write $(-,-)$ for the induced pairing on $\Lambda \otimes_{\mathcal{O}_F} W(\kappa)[1/p]$ (which is "sesquilinear" for the Frobenius on $W(\kappa)[1/p]$). If $x \in \Lambda \otimes_{\mathcal{O}_F} W(\kappa)[1/p]$ is any element with $(x, x) \in W(k)$, then $x \in \Lambda \otimes_{\mathcal{O}_F} W(\kappa)$ (so $\Lambda \otimes_{\mathcal{O}_F} W(\kappa)$ satisfies a certain "unique maximal integral lattice" property). This computation shows that $\mathcal{N}(1,1)(\overline{k})$ is a single point if $F/\mathbb{Q}_p$ is inert.

The ramified case follows from [RSZ17, Lemma 6.1]. The split case was already verified in Lemma 5.1.2.

(2) In the unramified case, part (1) implies $\mathcal{N}(n-r,r)(\overline{k}) \to \mathcal{N}(n-r,r)(\kappa)$ is bijective for every algebraically closed field extension κ over $\overline{k}$. Discreteness then follows because $\mathcal{N}(n-r,r)_{\mathrm{red}} \to \operatorname{Spec} \overline{k}$ is locally of finite type. If $F/\mathbb{Q}_p$ is inert, $\mathcal{N}(1,1)(\overline{k})$ being a single point was already discussed above. The ramified case is [RSZ17, Lemma 6.1].

(3) By part (1), we may assume $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ is the explicit tuple constructed in Section 5.1. The claim can then be verified explicitly, using (5.1.5) and the surrounding discussion. In the split case, $\mathbf{L}$ and $\mathbf{L}^{\perp}$ will be self-dual. In the inert case, $\mathbf{L}$ admits a Gram matrix with basis $\operatorname{diag}(1,p)$. In the ramified case, $\mathbf{L}$ admits a Gram matrix with basis $\operatorname{diag}(1,-a)$ for some $a \in \mathbb{Z}_p^{\times}$ which is a non-norm, i.e. $a \notin N_{F/\mathbb{Q}_p}(F^{\times})$. This claim in the ramified case follows from the observation that $\mathbf{L}$ is integral and that $\mathbf{L} \subseteq \operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0, \mathbf{X}_0 \times \mathbf{X}_0^{\sigma}) \subseteq \varpi^{-1}\mathbf{L}$ (as the isogeny in (5.1.2) has kernel contained in the ϖ-torsion subgroup).

(4) Transitivity of the I_1 action is immediate from part (1). Note also $\operatorname{End}(\mathbf{X}) \cap (U(\mathbf{W}) \times U(\mathbf{W}^{\perp})) \subseteq K_{1,\mathbf{L}} \times K_{1,\mathbf{L}^{\perp}}$, so the displayed map is well-defined. In the nonsplit cases, the assertions follow from parts (2) and (3). Bijectivity in the split case follows because we then have $\operatorname{End}(\mathbf{X}) \cap (U(\mathbf{W}) \times U(\mathbf{W}^{\perp})) = K_{1,\mathbf{L}} \times K_{1,\mathbf{L}^{\perp}}$.

(5) Follows from the previous parts, i.e. $\tilde{\mathcal{Z}}(L)(\overline{k})$ corresponds to $(\gamma, \gamma^{\perp})$ such that $L \subseteq \gamma(\mathbf{L})$. □

Suppose $F/\mathbb{Q}_p$ is split and $L \subseteq \mathbf{W}$ is any subset (with arbitrary signature $(n-r,r)$). We take $\mathcal{Z}(L) \subseteq \tilde{\mathcal{Z}}(L)$ to be the open and closed subfunctor corresponding (via Lemma 5.4.1(4)) to the locus where $\gamma^{\perp}(\mathbf{L}^{\perp}) = \mathbf{L}^{\perp}$. By the previous discussion, there is an isomorphism of formal schemes

$$\begin{aligned} \tilde{\mathcal{Z}}(L) &\longrightarrow \mathcal{Z}(L) \times U(\mathbf{V}^{\perp})/K_{1,\mathbf{L}^{\perp}} \\ (\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}}, (\gamma, \gamma^{\perp})) &\longmapsto ((\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}}, (\gamma, 1)), \gamma^{\perp}). \end{aligned} \tag{5.4.4}$$

In this case, we have a canonical bijection

$$\mathcal{Z}(L)(\overline{k}) = \{M \subseteq \mathbf{W} : \text{full rank self-dual lattice with } L \subseteq M\} \tag{5.4.5}$$

via Lemma 5.4.1.

Lemma 5.4.2. *Suppose $F/\mathbb{Q}_p$ is split. If $L \subseteq \mathbf{W}$ is an $\mathcal{O}_F$-lattice of full rank (i.e. rank $n - r$), then $\mathcal{Z}(L)(\overline{k})$ is a finite set.*

Proof. Our task is to show that the right-hand side of (5.4.5) is finite. For such M, we must have $L \subseteq M \subseteq M^{\vee} \subseteq L^{\vee}$ where $L^{\vee}$ and $M^{\vee}$ denote the dual lattices. If $L \not\subseteq L^{\vee}$ then $\mathcal{Z}(L)$ is empty. Otherwise, $L^{\vee}/L$ is an $\mathcal{O}_F$-module of finite length, so there are only finitely many possibilities for M. □

5.5 Horizontal and vertical decomposition

For a locally Noetherian formal scheme $\mathcal{X}$, viewed as a ringed space with structure sheaf $\mathcal{O}_{\mathcal{X}}$, we write

$$K_0'(\mathcal{X}) := K_0(Coh(\mathcal{O}_{\mathcal{X}})) \qquad F_d K_0'(\mathcal{X}) \subseteq K_0'(\mathcal{X}) \tag{5.5.1}$$

for the K_0 group of coherent $\mathcal{O}_{\mathcal{X}}$-modules and the subgroup generated by coherent sheaves supported in (formal scheme-theoretic) dimension $\leq d$, respectively. If $\mathcal{X}$ is moreover formally locally of finite type over $\operatorname{Spf} R$ for a complete discrete valuation ring R, we say that $\mathcal{X}$ is *equidimensional of dimension* n if every open formal subscheme of $\mathcal{X}$ has dimension n. In this case, if $\mathcal{Z} \to \mathcal{X}$ is an adic finite morphism of locally Noetherian formal schemes, we write

$$F^m_{\mathcal{X}} K_0'(\mathcal{Z}) := F_{n-m} K_0'(\mathcal{Z}) \qquad \mathrm{gr}^m_{\mathcal{X}} K_0'(\mathcal{Z}) := F^m_{\mathcal{X}} K_0'(\mathcal{Z}) / F^{m+1}_{\mathcal{X}} K_0'(\mathcal{Z}). \tag{5.5.2}$$

We often work with these groups tensor $\mathbb{Q}$, written as $K_0'(\mathcal{X})_{\mathbb{Q}}$, etc.

The discussion in Section 5.1 implies that the Rapoport–Zink space $\mathcal{N}(n-r, r)$ is equidimensional of dimension $(n-r)r + 1$. For the rest of Section 5.5 we fix signature $(n-1, 1)$ and use the shorthand $\mathcal{N} := \mathcal{N}(n-1, 1)$. The material below is a local analogue of Section 4.6.

Assume $F/\mathbb{Q}_p$ is nonsplit for the moment. For any nonzero $x \in \mathbf{W}$, the local special cycle $\mathcal{Z}(x)$ is a Cartier divisor on $\mathcal{N}$ for any nonzero $x \in \mathbf{W}$ ([KR11, Proposition 3.5] (inert) [How19, Proposition 4.3] (inert allowing $p = 2$), and also [LL22, Lemma 2.40] (ramified exotic smooth)). For any $x \in \mathbf{W}$, set

$$
{}^{\mathbb{L}}\mathcal{Z}(x) := \begin{cases} \mathcal{O}_{\mathcal{Z}(x)} & \text{if } x \neq 0 \\ (\cdots 0 \to \mathcal{E} \xrightarrow{0} \mathcal{O}_{\mathcal{N}} \to 0 \cdots) & \text{if } x = 0 \end{cases} \tag{5.5.3}
$$

in $D^b_{Coh(\mathcal{O}_{\mathcal{Z}(x)})}(\mathcal{O}_{\mathcal{N}})$ (bounded derived category of $\mathcal{O}_{\mathcal{N}}$-modules with cohomology sheaves coherent and supported along $\mathcal{O}_{\mathcal{Z}(x)}$), where the $\mathcal{O}_{\mathcal{N}}$ term is in degree 0. For any tuple $\underline{x} \in \mathbf{W}^m$, we then consider the *derived local special cycle*

$$
{}^{\mathbb{L}}\mathcal{Z}(\underline{x}) := {}^{\mathbb{L}}\mathcal{Z}(x_1) \otimes^{\mathbb{L}} \cdots \otimes^{\mathbb{L}} {}^{\mathbb{L}}\mathcal{Z}(x_m) \in D^b_{Coh(\mathcal{O}_{\mathcal{Z}(\underline{x})})}(\mathcal{O}_{\mathcal{N}}) \tag{5.5.4}
$$

Its image ${}^{\mathbb{L}}\mathcal{Z}(\underline{x}) \in K'_0(\mathcal{Z}(\underline{x}))_{\mathbb{Q}}$ lies in $F^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x}))_{\mathbb{Q}}$ by multiplicativity of the codimension filtration[25] and depends only on $\mathrm{span}_{\mathcal{O}_F}(\underline{x})$ ("linear invariance") by [How19, Theorem B] (inert) and [LL22, Proposition 2.33] (ramified exotic smooth).

Continuing to assume $F/\mathbb{Q}_p$ is nonsplit, assume $\underline{x} \in \mathbf{W}^m$ spans a non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank $m^\flat$. We define certain *derived vertical local special cycles* ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} \in \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}$ as follows.

For integers $e \gg 0$, we have a scheme-theoretic union decomposition (Lemma 11.7.5)

$$
\mathcal{Z}(\underline{x}) = \mathcal{Z}(\underline{x})_{\mathscr{H}} \cup \mathcal{Z}(\underline{x})_{\mathscr{V}} \tag{5.5.5}
$$

where $\mathcal{Z}(\underline{x})_{\mathscr{H}}$ is the flat part of $\mathcal{Z}(\underline{x})$, i.e. the largest closed formal subscheme which is flat over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$, and $\mathcal{Z}(\underline{x})_{\mathscr{V}} := \mathcal{Z}(\underline{x})_{\mathrm{Spf}\, \mathcal{O}_{\breve{F}}/p^e}$ for $e \gg 0$. Since $\mathcal{Z}(\underline{x})_{\mathscr{H}}$ is equidimensional of dimension $n - m^\flat$ (Lemma 11.7.4), and since $\mathcal{Z}(\underline{x})_{\mathscr{H}} \cap \mathcal{Z}(\underline{x})_{\mathscr{V}}$ has dimension $\leq n - m^\flat - 1$, there is an induced decomposition

$$
\mathrm{gr}^{m^\flat}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})) = \mathrm{gr}^{m^\flat}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\mathscr{H}}) \oplus \mathrm{gr}^{m^\flat}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}}) \tag{5.5.6}
$$

independent of e (cf. [Zha21, Lemma B.1][26]). Here we have used the pushforward dévissage isomorphism $K'_0(\mathcal{Z}(\underline{x})_{\overline{k}}) \xrightarrow{\sim} K'_0(\mathcal{Z}(\underline{x})_{\mathscr{V}})$ for K'_0 groups.

[25] There is a technicality here, as $\mathcal{N}$ is a formal scheme rather than a scheme. So we instead prove filtration multiplicativity via uniformization (Corollary 11.7.8) by reducing to the analogous filtration multiplicativity statement for global special cycles. We will make a few other forward references to Section 11.7 where we verify some properties of local special cycles via uniformization.

[26] Strictly speaking, our setup for K'_0 groups may be different from Zhang's in non quasi-compact settings. The proof of the cited lemma is the same in our setup.

If $m = m^\flat$, we define ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$ to be given by the projection

$$\begin{aligned} \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x}))_{\mathbb{Q}} &\longrightarrow \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}} \\ {}^{\mathbb{L}}\mathcal{Z}(\underline{x}) &\longmapsto {}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}. \end{aligned} \tag{5.5.7}$$

By the linear invariance property for ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})$ discussed above, the class ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$ depends only on $\mathrm{span}_{\mathcal{O}_F}(\underline{x})$.

For possibly $m \neq m^\flat$, we say that $\underline{x} = [x_1, \ldots, x_m]$ is in *minimal form* if $\underline{x}^\flat := [x_{m-m^\flat+1}, \ldots, x_m]$ satisfies $\mathrm{span}_{\mathcal{O}_F}(\underline{x}^\flat) = \mathrm{span}_{\mathcal{O}_F}(\underline{x})$. In this case, set $\underline{x}^\# := [x_1, \ldots, x_{m-m^\flat}]$ and define[27]

$${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} := {}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\#) \cdot {}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\flat)_{\mathscr{V}} \in \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}. \tag{5.5.8}$$

For $\underline{x}$ possibly not in minimal form, select any $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$ such that $\underline{x} \cdot \gamma$ is in minimal form, and set ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} := {}^{\mathbb{L}}\mathcal{Z}(\underline{x} \cdot \gamma)_{\mathscr{V}}$ (note $\mathcal{Z}(\underline{x}) = \mathcal{Z}(\underline{x} \cdot \gamma)$).

We claim that ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$ depends only on m and $\mathrm{span}_{\mathcal{O}_F}(\underline{x})$, and not on the choice of $\underline{x}$ or a minimal form ("linear invariance"). For $\underline{x}$ in minimal form and with notation as above, we already explained that ${}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\flat)_{\mathscr{V}}$ depends only on $\mathrm{span}_{\mathcal{O}_F}(\underline{x})$. Recall $\mathcal{Z}(\underline{x}) = \mathcal{Z}(\underline{x}^\flat)$. Consider any element x_i of the tuple $\underline{x}^\#$. Then $x_i \in \mathrm{span}_{\mathcal{O}_F}(\underline{x}^\flat) = \mathrm{span}_{\mathcal{O}_F}(\underline{x})$. In particular, $\mathcal{Z}(\underline{x}^\flat) \subseteq \mathcal{Z}(x_i)$. But Grothendieck–Messing theory provides a canonical isomorphism

$$\mathcal{E}|_{\mathcal{Z}(x_i)} \xrightarrow{\sim} \mathcal{I}(x_i)/\mathcal{I}(x_i)^2 \tag{5.5.9}$$

if $x_i \neq 0$, where $\mathcal{I}(x_i) \subseteq \mathcal{O}_{\mathcal{N}}$ is the ideal sheaf of the Cartier divisor $\mathcal{Z}(x_i) \subseteq \mathcal{N}$ (follows from [How19, Definition 4.2] (inert) and [LL22, Lemma 2.39]). Hence we have

$${}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\#)|_{\mathcal{Z}(\underline{x}^\flat)} \cong {}^{\mathbb{L}}\mathcal{Z}(\underline{0}_{m-m^\flat})|_{\mathcal{Z}(\underline{x}^\flat)} \tag{5.5.10}$$

as elements of $D^b_{Coh}(\mathcal{O}_{\mathcal{Z}(\underline{x}^\flat)})$, where $\underline{0}_{m-m^\flat} \in \mathbf{W}^{m-m^\flat}$ is the tuple with all entries equal to 0. Then we have

$${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} = {}^{\mathbb{L}}\mathcal{Z}(\underline{0}_{m-m^\flat}) \cdot {}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\flat)_{\mathscr{V}} \in \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}. \tag{5.5.11}$$

We have explained that the right-hand side does not depend on any auxiliary choices.

---

[27] One needs to show that the map $\alpha \mapsto {}^{\mathbb{L}}\mathcal{Z}(\underline{x}^\#) \cdot \alpha$ sends $F^{m^\flat+1}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}} \to F^{m+1}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}$. This is clear if $m^\flat \geq n-1$, but we do not know a proof of this in general as $\mathcal{N}$ is a formal scheme and not a scheme. Since we are mostly interested in the case $m^\flat \geq n-1$, we do not pursue this point further. Even when $m^\flat = n-1$ and $m = n$, one still needs to check that ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$ lies in $F^n_{\mathcal{N}}(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}$ (rather than $F^{n-1}_{\mathcal{N}}(\mathcal{Z}(\underline{x})_{\overline{k}})_{\mathbb{Q}}$). This follows e.g. because $\mathcal{Z}(\underline{x})_{\overline{k}}$ is a Noetherian scheme (Lemma 11.7.3) whose reduced irreducible components are projective over $\overline{k}$. The definition of ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$ should thus be treated as conditional unless $m = m^\flat$ or $m^\flat \geq n-1$.

Next, suppose $F/\mathbb{Q}_p$ is split, and assume $\underline{x} \in \mathbf{W}^m$ has $\mathcal{O}_F$-span which is a lattice of rank $n-1$ (full rank). We have $\mathrm{gr}^{n-1}_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}}) = 0$ for dimension reasons (the reduced subscheme of $\mathcal{N}$ is dimension 0, see Section 5.4 and Lemma 11.7.3). Constructing ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} \in \mathrm{gr}^m_{\mathcal{N}} K'_0(\mathcal{Z}(\underline{x})_{\overline{k}})$ as above gives the *derived vertical local special cycle* ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} = 0$.

Next, consider $\underline{x} \in \mathbf{W}^m$ which is a basis for its $\mathcal{O}_F$-span $L^\flat \coloneqq \mathrm{span}_{\mathcal{O}_F}(\underline{x})$. If $F/\mathbb{Q}_p$ is split, we also assume $m = n-1$. In this situation, we set ${}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \coloneqq {}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}}$, since the latter depends only on $L^\flat$. If $n = 2$ and $m = 1$, we have ${}^{\mathbb{L}}\mathcal{Z}(\underline{x})_{\mathscr{V}} = 0$ since the reduced subscheme $\mathcal{N}_{\mathrm{red}}$ has dimension 0 (Section 5.4) and since $\mathcal{N}$ has dimension 2 in this case.

5.6 Serre tensor and signature $(1,1)$

The case of signature $(1,1)$ plays an important role for describing local special cycles via the Serre tensor construction.

As above, let $\mathbf{X}_0$ be the unique supersingular (resp. ordinary) p-divisible group over $\overline{k}$ of height 2 and dimension 1 if $F/\mathbb{Q}_p$ is nonsplit (resp. split). For schemes S over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$, we consider pairs (X, ρ) where X is a p-divisible group over S and $\rho \colon X_{\overline{S}} \to \mathbf{X}_{0,\overline{S}}$ is any quasi-isogeny.

We form the Rapoport–Zink space $\widetilde{\mathcal{N}}_{2,1}$ over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$, given by

$$\widetilde{\mathcal{N}}_{2,1}(S) \coloneqq \{\text{isomorphism classes of framed tuples } (X, \rho) \text{ over } S\}. \tag{5.6.1}$$

This is a locally Noetherian formal scheme which is formally locally of finite type over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ (via the by-now standard representability result [RZ96, Theorem 2.16]). There is an isomorphism of formal schemes $\mathrm{Isog}^0(\mathbf{X}_0) \to \widetilde{\mathcal{N}}_{2,1}$ given by $\rho \mapsto (\mathbf{X}_0, \rho)$, where $\mathrm{Isog}^0(\mathbf{X}_0)$ is viewed as a constant formal scheme. Indeed, this follows as in the proof of Lemma 5.4.1 by uniqueness of $\mathbf{X}_0$ over any algebraically closed field κ (any quasi-endomorphism of $\mathbf{X}_0$ descends to $\overline{k}$ as well, as may be checked on isocrystals).

We let $\mathcal{N}_{2,1} \subseteq \widetilde{\mathcal{N}}_{2,1}$ be the open and closed locus where the framing ρ is fiberwise an isomorphism. Then $\mathcal{N}_{2,1}$ is representable by a formal scheme, and there is a (non-canonical) isomorphism $\mathcal{N}_{2,1} \cong \mathrm{Spf}\, \mathcal{O}_{\breve{F}}[\![t]\!]$ (e.g. by Grothendieck–Messing theory).

For arbitrary signature $(n-r, r)$ in the split case, we can form $\widetilde{\mathcal{N}}_{n,r}$ and $\mathcal{N}_{n,r}$ as above, where we replace $\mathbf{X}_0$ with the unique ordinary p-divisible group of height n and dimension r. The previous assertions for $\widetilde{\mathcal{N}}_{2,1}$ and $\mathcal{N}_{2,1}$ hold in this case as well, except we now have $\mathcal{N}_{n,r} \cong \mathrm{Spf}\, \mathcal{O}_{\breve{F}}[\![t_1, \ldots, t_{(n-r)r}]\!]$ (again by Grothendieck–Messing theory).

Lemma 5.6.1. *Given any $(X, \rho) \in \mathcal{N}_{2,1}(S)$, any principal polarization $\lambda_{\mathbf{X}_0}$ of $\mathbf{X}_0$ lifts uniquely to a principal polarization on X.*

Proof. Uniqueness follows from Drinfeld rigidity. Any two principal polarizations on X differ by $\mathbb{Z}_p^\times$ scalar (since this holds for $\mathbf{X}_0$), so it is enough to show existence of a principal polarization on X. Since $\mathcal{N}_{2,1} \cong \operatorname{Spf} \mathcal{O}_{\breve{F}}[\![t]\!]$, it is enough to check the case where the scheme S is a finite order thickening of $\operatorname{Spec} \overline{k}$. By Serre–Tate, we can view X as the p-divisible group of an elliptic curve over S (deforming an elliptic curve over $\operatorname{Spec} \overline{k}$ with p-divisible group $\mathbf{X}_0$). Any elliptic curve admits a (unique) principal polarization. $\square$

The preceding (possibly standard) argument also appeared in the proof of [RSZ17, Proposition 6.3] (for the same purpose), there in the supersingular case.

Recall the triple $(\mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F, \iota, \lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}})$ described in Section 5.1, arising from the Serre tensor construction (fixing some choice of $\lambda_{\mathbf{X}_0}$). For any $(X, \rho) \in \mathcal{N}_{2,1}(S)$, the same construction gives a tuple $(X \otimes_{\mathbb{Z}_p} \mathcal{O}_F, \iota, \lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}}, \rho \otimes 1)$ where $\lambda_{\mathbf{X}_0}$ denotes the unique lift to X as in Lemma 5.6.1 (by abuse of notation), and where $\rho \otimes 1 \colon X_{\overline{S}} \otimes_{\mathbb{Z}_p} \mathcal{O}_F \to \mathbf{X}_{0,\overline{S}} \otimes_{\mathbb{Z}_p} \mathcal{O}_F$.

Lemma 5.6.2 (Serre tensor isomorphism)**.** *For any F-linear quasi-isogeny $\phi \colon \mathbf{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_F \to \mathbf{X}$ preserving polarizations exactly, the induced map*

$$\begin{aligned} \mathcal{N}_{2,1} &\longrightarrow \mathcal{N}(1,1) \\ (X, \rho) &\longmapsto (X \otimes_{\mathbb{Z}_p} \mathcal{O}_F, \iota, -\mathfrak{d}^2 \cdot (\lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}}), \phi_{\overline{S}} \circ (\rho \otimes 1)) \end{aligned} \tag{5.6.2}$$

(defined on S-points for schemes S over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$) is an open and closed immersion whose set-theoretic image is a single point.

If $F/\mathbb{Q}_p$ is split, the inverse is given by restricting $(X, \iota, \lambda, \rho) \mapsto (X^-, (\phi^-_{\overline{S}})^{-1} \circ \rho^-)$ to the appropriate component of $\mathcal{N}(1,1)$.

Proof. For the ramified case, we refer to [RSZ17, Proposition 6.3]. In the unramified case, the lemma follows by identifying the deformation theory of $\mathcal{N}_{2,1}$ and $\mathcal{N}(1,1)$ using Grothendieck–Messing theory, using the eigenspace decomposition for the $\mathcal{O}_F$-action on the Dieudonné crystals of objects in $\mathcal{N}(1,1)(S)$ as in (3.5.4) and surrounding discussion (so the deformation problem for $\mathcal{N}_{2,1}$ identifies with the deformation problem of the "$-$ eigenspace" of the Hodge filtration for objects in $\mathcal{N}(1,1)(S)$ in the notation of loc. cit..). This is essentially how we verified generic formal smoothness of special cycles in loc. cit.. $\square$

When $F/\mathbb{Q}_p$ is split and the signature $(n-r, r)$ is arbitrary, recall that any $(X, \iota, \lambda, \rho) \in \mathcal{N}(n-r, r)$ admits a decomposition $X = X^+ \times X^-$ and $\rho = \rho^+ \times \rho^-$ where $\rho^\pm \colon X^\pm \to \mathbf{X}^\pm$ using the nontrivial idempotents $e^\pm \in \mathcal{O}_F$.

Lemma 5.6.3. *Suppose $F/\mathbb{Q}_p$ is split, and consider arbitrary signature $(n-r,r)$. For any formal scheme S over* $\operatorname{Spf} \mathcal{O}_{\breve{F}}$*, the forgetful functor*

$$\begin{Bmatrix} \textit{groupoid of principally polarized} \\ \textit{Hermitian } p\textit{-divisible groups } (X,\iota,\lambda) \\ \textit{over } S \textit{ of signature } (n-r,r) \end{Bmatrix} \longrightarrow \begin{Bmatrix} \textit{groupoid of ordinary } p\textit{-divisible groups} \\ \textit{over } S \textit{ of height } n \textit{ and dimension } r \end{Bmatrix}$$
$$(X,\iota,\lambda) \longmapsto X^- \tag{5.6.3}$$

is an equivalence of categories. The same holds if we consider the groupoids with morphisms being quasi-isogenies (rather than isomorphisms).

Proof. An explicit quasi-inverse is given by $(X^-) \mapsto (X,\iota,\lambda)$ (over a scheme S) with

$$\begin{aligned} & X = (X^-)^\vee \times X^- \\ & \iota(e^+) \colon X \to (X^-)^\vee \qquad \iota(e^-) \colon X \to X^- \qquad \text{projections} \\ & \lambda = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} : (X^-)^\vee \times X^- \to X^- \times (X^-)^\vee. \end{aligned} \tag{5.6.4}$$

This is analogous to the following phenomenon: if L is a free $\mathcal{O}_F$-module of rank n equipped with a perfect Hermitian pairing, then $U(L) \cong \mathrm{GL}_n(L^-)$ (and similarly with F instead of $\mathcal{O}_F$). $\square$

Remark 5.6.4. Suppose $F/\mathbb{Q}_p$ is split, and suppose R is a complete Noetherian local ring with algebraically closed residue field κ. If $n \geq 2$ and if $\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-2} \times X^-$ is an ordinary p-divisible group of height n and dimension r, then $(\mathfrak{X}_0)^{n-2} \times (X^- \otimes_{\mathbb{Z}_p} \mathcal{O}_F)$ (with the product $\mathcal{O}_F$-action and product polarization $(\lambda_{\mathfrak{X}_0})^{n-2} \times (\lambda_{\mathbf{X}_0} \otimes \lambda_{\mathrm{tr}})$, for some choice of isomorphism $X^-_\kappa \cong \mathbf{X}_0$) is a pre-image under the equivalence in (5.6.3). Here $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \lambda_{\mathfrak{X}_0})$ is the canonical lift (over S) as in Definition 5.1.7 (forgetting the framing).

Remark 5.6.5. If we drop the ordinary hypothesis on both sides of Lemma 5.6.3, the lemma still holds (by the same proof).

Lemma 5.6.6. *Suppose $F/\mathbb{Q}_p$ is split, and form $\widetilde{\mathcal{N}}_{n,r}$ using the framing object $\mathbf{X}^-$. For arbitrary signature $(n-r,r)$, the forgetful map*

$$\begin{aligned} \mathcal{N}(n-r,r) &\longrightarrow \widetilde{\mathcal{N}}_{n,r} \\ (X,\iota,\lambda,\rho) &\longmapsto (X^-,\rho^-) \end{aligned} \tag{5.6.5}$$

is an isomorphism.

Proof. This is immediate from Lemma 5.6.3. Alternative (less elementary) proof: first observe that the forgetful map is an isomorphism on κ-points for any algebraically closed field κ over $\overline{k}$ (see Lemma 5.4.1 and above discussion). As in the proof of Lemma 5.6.2, the claim now follows from Grothendieck–Messing theory. □

Remark 5.6.7. In the situation of Lemma 5.6.6, the open and closed subfunctor $\mathcal{N}_{n,r} \subseteq \widetilde{\mathcal{N}}_{n,r} \cong \mathcal{N}(n-r,r)$ has $\mathcal{N}_{n,r}(\overline{k})$ being a singleton set, corresponding (via Lemma 5.4.1) to the lattice $\mathbf{L} \oplus \mathbf{L}^{\perp} \subseteq \mathbf{W} \oplus \mathbf{W}^{\perp}$ (i.e. the locus where the framing ρ is a fiberwise isomorphism).

For $F/\mathbb{Q}_p$ in all cases (inert, ramified, split) and for any $x \in \mathbf{W}$, the local special cycle $\mathcal{Z}(x) \to \mathcal{N}(1,1)$ pulls back along the Serre tensor isomorphism (Lemma 5.6.2) to a certain local special cycle on $\mathcal{N}_{2,1}$ associated with an element $x' \in \mathrm{Hom}^0(\mathbf{X}_0, \mathbf{X}_0)$ (arising from adjunction in the Serre tensor construction). The ramified case is explained in [RSZ17, §6.2]. The inert and split cases may be formulated in a similar way (we omit a more detailed statement, which we will not need). This may be viewed as a local version of [KR14, Proposition 14.5] (see also Section 22.2).

For $F/\mathbb{Q}_p$ nonsplit (at least if $p \neq 2$), Kudla–Rapoport [KR11, Proposition 8.1] and Rapoport–Smithling–Zhang [RSZ17, Proposition 7.1] use this to describe $\mathcal{Z}(x)$ in terms of certain *quasi-canonical lifting cycles* on $\mathcal{N}_{2,1}$, corresponding to closed immersions $\mathrm{Spf}\, \mathcal{O}_{\breve{E}_s} \to \mathcal{N}_{2,1}$ associated with $(\mathfrak{X}_s, \rho) \in \mathrm{Spf}\, \mathcal{O}_{\breve{E}_s}$ where $(\mathfrak{X}_s, \rho)$ arises from a *quasi-canonical lifting* of $(\mathbf{X}_0, j)$ for suitable $j \colon \mathcal{O}_F \to \mathrm{End}(\mathbf{X}_0)$ (in the notation and sense of Section 7.2 below). This was extended by Li–Zhang [LZ22a] (inert) and Li–Liu [LL22] (ramified) to flat parts of 1-cycles in signature $(n-1,1)$, for arbitrary n in the inert case and even n in the ramified case. We will need this result, which we recall in Section 7.3 below (to the precision we need).

We will need an analogue of the previous paragraph when $F/\mathbb{Q}_p$ is split (allowing $p=2$). This is accomplished in Section 6 below (statement given in Section 7.3). Our method in the split case is somewhat different from the proofs cited above.

6 More on moduli of p-divisible groups: split

Retain notation from Section 5. Throughout Section 6, we assume $F/\mathbb{Q}_p$ is split.

6.1 Lifting theory for ordinary p-divisible groups

We discuss lifting theory for ordinary p-divisible groups over an algebraically closed field κ of characteristic p. The case of height 2 dimension 1 ordinary p-divisible groups is discussed in

[Mes72, Appendix]. We spell out the case of general height and dimension (which reduces to the results in [Mes72, Appendix]). See also the exposition in [Meu07] (or the sketch in [Gro86, §6]), though we will need some additional material on homomorphisms between liftings.

Take integers $r_1, r_2 \geq 0$. The unique ordinary p-divisible group X over κ of height $r_1 + r_2$ and dimension r_1 is $X = \boldsymbol{\mu}_{p^\infty}^{r_1} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}$.

Next, let R be an adic Noetherian local ring (with maximal ideal being an ideal of definition) with residue field κ. The p-divisible groups $\boldsymbol{\mu}_{p^\infty}$ and $\underline{\mathbb{Q}_p/\mathbb{Z}_p}$ lift uniquely to $\operatorname{Spf} R$ (e.g. by Grothendieck–Messing deformation theory), which we still notate as $\boldsymbol{\mu}_{p^\infty}$ and $\underline{\mathbb{Q}_p/\mathbb{Z}_p}$. If $\mathfrak{X}$ is a lift of X over $\operatorname{Spf} R$, its connected-étale exact sequence must be $0 \to \boldsymbol{\mu}_{p^\infty}^{r_1} \to \mathfrak{X} \to \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2} \to 0$. Classifying lifts $\mathfrak{X}$ is thus the same as classifying such extensions, which are in canonical bijection with $\operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1})$ (using also Drinfeld rigidity, as well as the fact $\operatorname{Hom}(\boldsymbol{\mu}_{p^\infty}, \underline{\mathbb{Q}_p/\mathbb{Z}_p}) = \operatorname{Hom}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}, \boldsymbol{\mu}_{p^\infty}) = 0$). Here, $\operatorname{Ext}^1_{\operatorname{Spf} R}$ is calculated in the abelian category of fppf sheaves of abelian groups over $\operatorname{Spf} R$ (this is also [Mes72, Appendix, Corollary (2.3)]). We typically suppress the R-dependence in $\operatorname{Hom}(-,-)$.

Applying $\operatorname{Hom}(-, \boldsymbol{\mu}_{p^\infty})$ to the short exact sequence of sheaves (fppf sheaves over $\operatorname{Spf} R$)

$$0 \longrightarrow \underline{\mathbb{Z}} \longrightarrow \underline{\mathbb{Z}[1/p]} \longrightarrow \underline{\mathbb{Q}_p/\mathbb{Z}_p} \longrightarrow 0 \tag{6.1.1}$$

gives a boundary morphism $\delta \colon \operatorname{Hom}(\underline{\mathbb{Z}}, \boldsymbol{\mu}_{p^\infty}) \to \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}, \boldsymbol{\mu}_{p^\infty})$ in the associated long exact sequence. This map δ is an isomorphism [Mes72, Appendix, Proposition (2.5)].[28] By compatibility of Ext with finite direct sums, it follows that the boundary morphism $\delta \colon \operatorname{Hom}(\underline{\mathbb{Z}}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1}) \to \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1})$ is also an isomorphism.

Given an element $\alpha \in \operatorname{Hom}(\underline{\mathbb{Z}}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1})$, we can identify the extension corresponding to $\delta(\alpha)$ with the bottom row of the diagram

$$\begin{tikzcd}
0 \arrow[r] & \underline{\mathbb{Z}}^{r_2} \arrow[r] \arrow[d, "\alpha"] & \underline{\mathbb{Z}[1/p]}^{r_2} \arrow[r] \arrow[d] \arrow[dl, phantom, "\ulcorner"] & \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \boldsymbol{\mu}_{p^\infty}^{r_1} \arrow[r] & \mathfrak{X} \arrow[r] & \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2} \arrow[r] & 0
\end{tikzcd} \tag{6.1.2}$$

where the rows are exact and the left square is a pushout. This follows from general homological algebra valid in any abelian category (e.g. [SProject, Section 010I] and [SProject, Section 06XP]).

[28] In loc. cit. this is stated for Artinian local rings R, but one can pass to the limit and obtain the statement here (compare [Mes72, Appendix, Remark (2.2)]).

Given $r_1, r_1', r_2, r_2' \in \mathbb{Z}_{\geq 0}$, we have

$$\begin{aligned}\mathrm{Hom}(\boldsymbol{\mu}_{p^\infty}^{r_1} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1'} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2'}) &= \mathrm{Hom}(\boldsymbol{\mu}_{p^\infty}^{r_1}, \boldsymbol{\mu}_{p^\infty}^{r_1'}) \times \mathrm{Hom}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}, \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2'}) \\ &\cong M_{r_1', r_1}(\mathbb{Z}_p) \times M_{r_2', r_2}(\mathbb{Z}_p), \end{aligned} \tag{6.1.3}$$

since any p-divisible group over κ of height 1 has endomorphism ring $\mathbb{Z}_p$. Here $M_{s,t}(\mathbb{Z}_p)$ denotes $s \times t$ matrices with entries in $\mathbb{Z}_p$. Given

$$\alpha \in \mathrm{Hom}(\underline{\mathbb{Z}}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1}) = \mathrm{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{r_2}, \boldsymbol{\mu}_{p^\infty}^{r_1}) \qquad \alpha' \in \mathrm{Hom}(\underline{\mathbb{Z}}^{r_2'}, \boldsymbol{\mu}_{p^\infty}^{r_1'}) = \mathrm{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{r_2'}, \boldsymbol{\mu}_{p^\infty}^{r_1'}) \tag{6.1.4}$$

with corresponding lifts $\mathfrak{X}$ and $\mathfrak{X}'$ (of $\boldsymbol{\mu}_{p^\infty}^{r_1} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2}$ and $\boldsymbol{\mu}_{p^\infty}^{r_1'} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{r_2'}$ respectively) over $\mathrm{Spf}\, R$, a morphism $(f_1, f_2) \in M_{r_1', r_1}(\mathbb{Z}_p) \times M_{r_2', r_2}(\mathbb{Z}_p)$ lifts to a map $f \colon \mathfrak{X} \to \mathfrak{X}'$ if and only if

$$f_1 \circ \alpha = \alpha' \circ f_2, \tag{6.1.5}$$

again by general facts about Ext in abelian categories (compare with the proof of [Mes72, Appendix, Proposition (3.3)], which discusses the case $r_1 = r_2 = r_1' = r_2' = 1$). We will repeatedly use this criterion for lifting to maps $f \colon \mathfrak{X} \to \mathfrak{X}'$. In (6.1.4), the subscripts $\mathbb{Z}_p$ indicate $\mathbb{Z}_p$-linearity (not the base $\mathrm{Spf}\, \mathbb{Z}_p$).

6.2 Quasi-canonical lifting cycles: split

Throughout Section 6.2, we write R for an adic Noetherian local ring (with maximal ideal being an ideal of definition) equipped with a morphism $\mathrm{Spf}\, R \to \mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ inducing an isomorphism on residue fields.

Allowing arbitrary signature $(n - r, r)$ for the moment, form the Rapoport–Zink spaces $\mathcal{N}(n-r, r)$, $\mathcal{N}_{n,r}$, and $\widetilde{\mathcal{N}}_{n,r}$ as in Section 5.6. With $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ denoting the framing object for $\mathcal{N}(n-r, r)$, we take $\mathbf{X}^-$ to be the framing object used to define $\mathcal{N}_{n,r}$ and $\widetilde{\mathcal{N}}_{n,r}$. There are non-canonical isomorphisms $\mathbf{X}^- \cong \boldsymbol{\mu}_{p^\infty}^r \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}$ and $\mathbf{X}_0^- \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}$.

Definition 6.2.1. Given a subset $L^- \subseteq \mathbf{W}^- = \mathrm{Hom}^0(\mathbf{X}_0^-, \mathbf{X}^-)$, consider the associated *local special cycle*

$$\mathcal{Y}(L^-) \subseteq \mathcal{N}_{n,r} \tag{6.2.1}$$

which is the subfunctor consisting of pairs (X, ρ) over schemes S over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ such that, for all $x^- \in L^-$, the quasi-homomorphism

$$\rho^{-1} \circ x'_{\overline{S}} \circ \rho^-_{\mathfrak{X}_0, \overline{S}} \colon \mathfrak{X}_0^- \to X_{\overline{S}} \tag{6.2.2}$$

lifts to a homomorphism $\mathfrak{X}_0^- \to X$.

As in Definition 5.1.7 (also Section 7.1), the notation $\mathfrak{X}_0$ refers to the canonical lifting of $\mathbf{X}_0$ (and $\mathfrak{X}_0 = \mathfrak{X}_0^+ \times \mathfrak{X}_0^-$ is the decomposition via the nontrivial idempotents $e^{\pm} \in \mathcal{O}_F$, with $\mathbf{X}_0^+ \cong \boldsymbol{\mu}_{p^\infty}$ and $\mathfrak{X}_0^- \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}$). Again, $\mathcal{Y}(L^-) \subseteq \mathcal{N}_{n,r}$ is a closed subfunctor (hence a locally Noetherian formal scheme) by [RZ96, Proposition 2.9] for quasi-homomorphisms.

Lemma 6.2.2. *Suppose $L \subseteq \mathbf{W} = \operatorname{Hom}^0_F(\mathbf{X}_0, \mathbf{X})$ is a subset with $L^+ \subseteq \mathbf{L}^+ = \operatorname{Hom}(\mathbf{X}_0^+, \mathbf{X}^+)$. The natural commutative diagram*

$$\begin{tikzcd}
\mathcal{Y}(L^-) \arrow[rr] \arrow[d] \arrow[dr, phantom, very near start, "\lrcorner"] & & \mathcal{Z}(L) \arrow[d] \\
\mathcal{N}_{n,r} \arrow[r, hook] & \widetilde{\mathcal{N}}_{n,r} \arrow[r, "\sim"] & \mathcal{N}(n-r, r)
\end{tikzcd} \tag{6.2.3}$$

is Cartesian.

Proof. The lower horizontal arrows are as described in Section 5.6 and Lemma 5.6.6 (the composite is an open and closed immersion). The lemma amounts to the claim that, for any $(X, \iota, \lambda, \rho) \in X(S)$ (for any scheme S over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$), if $x = x^+ \times x^- \in \operatorname{Hom}^0_F(\mathbf{X}_0, \mathbf{X})$ with $x^+ \in \operatorname{Hom}_{\mathcal{O}_F}(\mathbf{X}_0^+, \mathbf{X}^+)$, then x lifts to a homomorphism $\mathfrak{X}_0 \to X$ if and only if x^- lifts to a homomorphism $\mathfrak{X}_0^- \to X^-$. Stated alternatively, this is the claim that x^+ always lifts to a homomorphism $\mathfrak{X}_0^+ \to X^+$. Since $\mathcal{N}_{n,r} \cong \operatorname{Spf} \mathcal{O}_{\breve{F}}[\![t_1, \ldots, t_{(n-r)r}]\!]$, this is clear because $\mathfrak{X}_0^+ \to X^+$ automatically factors through a homomorphism to the connected part $(X^+)^0 \cong \boldsymbol{\mu}_{p^\infty}^{n-r}$ of X^+, over any base $\operatorname{Spf} R$ where R is Noetherian Henselian local ring (alternative proof: apply (6.1.5)). $\square$

Choose isomorphisms $(\mathbf{X}^-)^0 \cong \boldsymbol{\mu}^r_{p^\infty}$ and $(\mathbf{X}^-)^{\text{ét}} \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}$ for the connected and étale parts of $\mathbf{X}^-$ respectively. Any element $(X, \rho) \in \mathcal{N}_{n,r}(\operatorname{Spf} R)$ (i.e. a morphism $\operatorname{Spf} R \to \mathcal{N}_{n,r}$) then corresponds to a class $\alpha \in \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}, \boldsymbol{\mu}^r_{p^\infty}) = \operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-r}, \boldsymbol{\mu}^r_{p^\infty})$ via the lifting theory in Section 6.1.

Lemma 6.2.3. *Fix any isomorphism $\mathbf{X}_0^- \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}$. Consider $\varphi \colon \operatorname{Spf} R \to \mathcal{N}_{n,r}$, corresponding to $(X, \rho) \in \mathcal{N}_{n,r}(\operatorname{Spf} R)$ and hence a class $\alpha' \in \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}, \boldsymbol{\mu}^r_{p^\infty})$. Given any subset $L^- \subseteq \mathbf{L}^-$, the morphism φ factors through $\mathcal{Y}(L^-) \subseteq \mathcal{N}_{n,r}$ if and only if the map*

$$x^* \colon \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}, \boldsymbol{\mu}^r_{p^\infty}) \to \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}, \boldsymbol{\mu}^r_{p^\infty}) \tag{6.2.4}$$

satisfies $x^*(\alpha') = 0$ for all $x \in L^-$.

Proof. In the lemma statement, we have viewed $x \in L^-$ as a morphism $\underline{\mathbb{Q}_p/\mathbb{Z}_p} \to \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-r}$ via the various identifications. The lemma follows from the lifting criterion in (6.1.5) (in the notation of loc. cit., take $\alpha = 0$). $\square$

Next, we restrict to the case of signature $(n-1,1)$.

Lemma 6.2.4. *Assume that R is moreover a domain and that* $\operatorname{Spf} R \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$ *is flat. There is a natural map*

$$\left\{ \begin{array}{l} \textit{cyclic subgroups of order } p^s \\ \textit{in } \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty}) \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} \textit{Full rank integral lattices } M \subseteq \mathbf{L} \\ \textit{such that } t(M) \leq 1 \textit{ and } \operatorname{val}(M) = s \\ \textit{and } M^+ = \mathbf{L}^+ \end{array} \right\} \tag{6.2.5}$$

(functorial in R on the left). If R contains a primitive p^s-th root of unity, then the map is a bijection. Otherwise, the left-hand side is empty.

Proof. Recall the identification $\operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty}) \cong \operatorname{Ext}^1_{\operatorname{Spf} R}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ from Section 6.1. Suppose $\alpha' \in \operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ generates a cyclic subgroup of order p^s (possible if and only if R contains a primitive p^s-th root of unity). Let $M_{n-1,n-1}(\mathbb{Z}_p)$ act on $\operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ by pre-composition. The annihilator of α' is generated (as a one-sided ideal) by an element $f_2 \in M_{n-1,n-1}(\mathbb{Z}_p)$ which has Smith normal form $\operatorname{diag}(1, \ldots, 1, p^s)$.

We have a canonical identification $\mathbf{L}^- \cong \operatorname{Hom}(\mathbf{X}_0^-, (\mathbf{X}^-)^{\text{ét}})$. Via the identification $(\mathbf{X}^-)^{\text{ét}} \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}$, we obtain an action of $M_{n-1,n-1}(\mathbb{Z}_p)$ on $\mathbf{L}^-$ (post-composition). We then set $M^- = f_2(\mathbf{L}^-)$, and let $M = \mathbf{L}^+ \oplus M^-$. Note that M^- does not depend on the choice of generator f_2.

Conversely, given a lattice $M \subseteq \mathbf{L}$ as in the lemma statement, select any $f_2 \in M_{n-1,n-1}(\mathbb{Z}_p)$ satisfying $M^- = f_2(\mathbf{L}^-)$, and note that f_2 necessarily has Smith normal form $\operatorname{diag}(1, \ldots, 1, p^s)$. If R contains a primitive p^s-th root of unity, then f_2 acting on $\operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ has kernel which is cyclic of order p^s. This gives the inverse map. $\square$

For $s \in \mathbb{Z}_{\geq 0}$, set $\breve{E}_s \coloneqq \breve{F}[\zeta_{p^s}]$ with ring of integers $\mathcal{O}_{\breve{E}_s} = \mathcal{O}_{\breve{F}}[\zeta_{p^s}]$, where ζ_{p^s} is a primitive p^s-th root of unity. Suppose $M \subseteq \mathbf{L}$ is an integral full rank $\mathcal{O}_F$-lattice satisfying $M^+ = \mathbf{L}^+$, with type $t(M) \leq 1$ and $\operatorname{val}(M) = s$. By Lemma 6.2.4, there is an associated cyclic subgroup of $\operatorname{Ext}^1_{\operatorname{Spf} \mathcal{O}_{\breve{E}_s}}(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$. Any generator of this cyclic subgroup defines a morphism $\operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}_{n,1}$ (via the lifting theory from Section 6.1). Changing the choice of generator corresponds precisely to the action of $\operatorname{Gal}(\breve{E}_s/\breve{F})$ (by Lubin–Tate theory for $\boldsymbol{\mu}_{p^\infty}$). This morphism $\operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}_{n,1}$ must be a closed immersion: if the morphism factors through $\operatorname{Spf} R \to \mathcal{N}_{n,1}$ for some sub $\mathcal{O}_{\breve{F}}$-algebra $R \subseteq \mathcal{O}_{\breve{E}_s}$, then Lemma 6.2.4 implies that $R = \mathcal{O}_{\breve{E}_s}$.

We write $\operatorname{Spf} \mathcal{O}_{\breve{E}_s} \cong \mathcal{Z}(M)^\circ \subseteq \mathcal{N}_{n,1}$ for the resulting closed subfunctor, and call it a *quasi-canonical lifting cycle*. This closed subfunctor $\mathcal{Z}(M)^\circ$ does not depend on the choices of isomorphisms $(\mathbf{X}^-)^0 \cong \boldsymbol{\mu}_{p^\infty}$ and $(\mathbf{X}^-)^{\text{ét}} \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}$ appearing in the statement of Lemma 6.2.4.

Lemma 6.2.5. *With M as above, view $\mathcal{Z}(M)^\circ$ as a morphism* $\operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}_{n,1}$ *corresponding to $(X, \rho) \in \mathcal{N}_{n,1}(\operatorname{Spf} \mathcal{O}_{\breve{E}_s})$.*

If $n = 1$ then $X \cong \boldsymbol{\mu}_{p^\infty}$. If $n \geq 2$ then $X \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-2} \times \mathfrak{X}_s$ (forgetting ρ) where $\mathfrak{X}_s$ is a p-divisible group of height 2 and dimension 1 with $\operatorname{End}(\mathfrak{X}_s) = \mathbb{Z}_p + p^s \mathcal{O}_F$ *(a* quasi-canonical lifting *in the sense of Section 7.2).*

Proof. Let $\alpha' \in \operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ be the element corresponding to (X, ρ). If $n = 1$ then $\alpha' = 0$ and $X \cong \boldsymbol{\mu}_{p^\infty}$.

If $n \geq 2$, then (after replacing ρ by $\phi \circ \rho$ for some $\phi \in \mathrm{GL}_{n-1}(\mathbb{Z}_p)$), the lift (X, ρ) of $\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1} \times \boldsymbol{\mu}_{p^\infty}$ is associated with α' of the form $(0, \ldots, 0, \zeta_{p^s})$ for $\zeta_{p^s} \in \mathcal{O}_{\breve{E}_s}$ a primitive p^s-th root of unity. For some $\mathfrak{X}_s$ as in the lemma statement, we obtain a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \underline{\mathbb{Z}}^{n-1} & \longrightarrow & \underline{\mathbb{Z}[1/p]}^{n-1} & \longrightarrow & \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1} & \longrightarrow & 0 \\ & & {\scriptstyle (0,\ldots,0,\zeta_{p^s})}\Big\downarrow & & \Big\downarrow & & \Big\| & & \\ 0 & \longrightarrow & \boldsymbol{\mu}_{p^\infty} & \longrightarrow & \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-2} \times \mathfrak{X}_s & \longrightarrow & \underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1} & \longrightarrow & 0\,, \end{array} \tag{6.2.6}$$

using the lifting criterion of (6.1.5) again. □

Lemma 6.2.6. *Suppose $M \subseteq \mathbf{L}$ is an integral full rank $\mathcal{O}_F$-lattice satisfying $M^+ = \mathbf{L}^+$ with type $t(M) \leq 1$ and* $\operatorname{val}(M) = s$*. Let $L^- \subseteq \mathbf{W}^-$ be any subset. We have $\mathcal{Z}(M)^\circ \subseteq \mathcal{Y}(L^-)$ if and only if $L^- \subseteq M$.*

Proof. If $n = 1$ then $\mathbf{W}^- = 0$ and $\mathcal{Z}(M)^\circ = \mathcal{Y}(L^-) = \mathcal{N}_{n,1}$, so the lemma is trivial in this case. We thus assume $n \geq 2$ below.

It is enough to check the case where L^- consists of a single element, i.e. a quasi-homomorphism $x \colon (\mathbf{X}_0)^- \to \mathbf{X}^-$ (or equivalently, $x \colon (\mathbf{X}_0)^- \to (\mathbf{X}^-)^{\text{ét}}$ since $\mathbf{X}_0^- \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}$ is étale). If $x \notin \mathbf{L}^-$ then $\mathcal{Y}(L^-) = \emptyset$ (while $\mathcal{Z}(M)^\circ \neq \emptyset$), so we may assume $x \in \mathbf{L}^-$.

Pick any identification $\mathbf{X}_0^- \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}$. Set $s = \operatorname{val}(M)$. Use the setup and notation in the proof of Lemma 6.2.4.

View $\mathcal{Z}(M)^\circ$ as a closed immersion $\operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}_{n,1}$, corresponding to an element $\alpha' \in \operatorname{Hom}_{\mathbb{Z}_p}(\underline{\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$. By Lemma 6.2.3, our task is to show that $\alpha' \circ x = 0$ if and only if $\alpha' \in M^-$. Since f_2 generates (in $M_{n-1,n-1}(\mathbb{Z}_p)$) the annihilator of α' (as a one-sided ideal), we see that $\alpha' \circ x = 0$ if and only if $x \in f_2(\mathbf{L}^-) = M^-$ (for example, view x as a column vector and observe that $(x, 0, \ldots, 0) \in M_{n-1,n-1}(\mathbb{Z}_p)$ lies in the one-sided ideal generated by f_2). □

Definition 6.2.7. Let $M \subseteq \mathbf{W}$ be a full rank integral $\mathcal{O}_F$-lattice, with type $t(M) \leq 1$ and $\mathrm{val}(M) = s$. Select any $\gamma \in U(\mathbf{W})$ satisfying $\gamma(\mathbf{L}^+) = M^+$ (also write γ for $(\gamma, 1) \in U(\mathbf{W}) \times U(\mathbf{W}^\perp)$, by abuse of notation).

The *quasi-canonical lifting cycle* associated with M is the closed subfunctor

$$\mathrm{Spf}\, \mathcal{O}_{\breve{E}_s} \cong \mathcal{Z}(M)^\circ \coloneqq \gamma(\mathcal{Z}(\gamma^{-1}(M))^\circ) \subseteq \mathcal{N}(n-1,1) \tag{6.2.7}$$

where $\gamma \in U(\mathbf{W}) \times U(\mathbf{W}^\perp)$ acts on $\mathcal{N}(n-1,1)$ as in Section 5.3.

In the situation of Definition 6.2.7, the closed subfunctor $\mathcal{Z}(M)^\circ$ does not depend on the choice of γ. We have also viewed $\mathcal{N}_{n,1}$ as an open and closed subfunctor of $\mathcal{N}(n-1,1)$ (as in the lower horizontal arrows in Lemma 6.2.2).

Lemma 6.2.8. *If $L \subseteq \mathbf{W}$ is any subset and $M \subseteq \mathbf{W}$ is any full rank integral lattice with $t(M) \leq 1$, we have $\mathcal{Z}(M)^\circ \subseteq \mathcal{Z}(L)$ if and only if $L \subseteq M$.*

Proof. After acting by $U(\mathbf{W})$, it is enough to check the case where $M^+ = \mathbf{L}^+$. In this case, we have $\mathcal{Z}(M)^\circ \subseteq \mathcal{N}_{n,1}$. If $L \not\subseteq \mathbf{L}$, then $\mathcal{Z}(L) \cap \mathcal{Z}(M)^\circ = \emptyset$ by Lemma 5.4.1(5) and Remark 5.6.7 (and $\mathcal{Z}(M)^\circ$ is nonempty). So assume $L \subseteq \mathbf{L}$. Then $\mathcal{Z}(L) = \mathcal{Y}(L^-)$ (Lemma 6.2.2). This reduces to the case proved in Lemma 6.2.6. $\square$

Corollary 6.2.9. *Let $L \subseteq \mathbf{W}$ be any subset. Form the horizontal (flat) part of the local special cycle $\mathcal{Z}(L)$, which we denote as $\mathcal{Z}(L)_{\mathscr{H}}$. We have an inclusion of closed formal subschemes*

$$\bigcup_{\substack{L \subseteq M \subseteq M^* \\ t(M) \leq 1}} \mathcal{Z}(M)^\circ \subseteq \mathcal{Z}(L)_{\mathscr{H}} \tag{6.2.8}$$

in $\mathcal{N}(n-1,1)$.

Proof. The union is a scheme-theoretic union (i.e. intersect associated ideal sheaves). The claim follows from Lemma 6.2.8 because each $\mathcal{Z}(M)^\circ$ is flat over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}}$. $\square$

Lemma 6.2.10. *Let $M \subseteq \mathbf{W}$ and $M' \subseteq \mathbf{W}$ be integral full-rank $\mathcal{O}_F$-lattices with $t(M) \leq 1$ and $t(M') \leq 1$. If $M \neq M'$, then $\mathcal{Z}(M)^\circ \neq \mathcal{Z}(M')^\circ$.*

Proof. Let $N \subseteq \mathbf{W}$ (resp. $N' \subseteq \mathbf{W}$) be the unique self-dual full rank lattice such that $N^+ = M^+$ (resp. $N'^+ = M'^+$). On reduced subschemes, we have $\mathcal{Z}(M)^\circ_{\mathrm{red}} = \mathcal{Z}(M')^\circ_{\mathrm{red}}$ if and only if $N = N'$ by Lemma 5.4.1 (more precisely, Remark 5.6.7, Definition 6.2.1, and the action on special cycles in (5.3.6)). So we may assume $N = N'$. Using the $U(\mathbf{W})$ action on $\mathcal{N}(n-1,1)$, we also reduce to the case where $N = \mathbf{L}$.

Set $s = \mathrm{val}(M)$ and $s' = \mathrm{val}(M')$, and view $\mathcal{Z}(M)^\circ$ and $\mathcal{Z}(M')^\circ$ as closed immersions $\varphi\colon \operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}(n-1,1)$ and $\varphi'\colon \operatorname{Spf} \mathcal{O}_{\breve{E}_{s'}} \to \mathcal{N}(n-1,1)$. Lemma 6.2.4 implies that $M = M'$ if and only if both $s = s'$ and the morphisms φ, φ' are the same up to $\mathrm{Gal}(\breve{E}_s/\breve{F})$-action (this is equivalent to requiring that the corresponding elements of Ext^1 in that lemma generate the same subgroup). This is satisfied if and only if $\mathcal{Z}(M)^\circ = \mathcal{Z}(M')^\circ$. $\square$

Lemma 6.2.11. *Let $L \subseteq \mathbf{W}$ be a full rank $\mathcal{O}_F$-lattice. Assume that R is moreover a domain and* $\operatorname{Spf} R \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$ *is flat. Any morphism $\varphi\colon$* $\operatorname{Spf} R \to \mathcal{Z}(L)$ *factors through some quasi-canonical lifting cycle $\mathcal{Z}(M)^\circ$.*

Proof. Again, we may act by $U(\mathbf{W})$ on $\mathcal{N}(n-1,1)$ to assume that $\varphi\colon \operatorname{Spf} R \to \mathcal{Z}(L)$ factors through the open and closed component $\mathcal{N}_{n,1} \subseteq \mathcal{N}(n-1,1)$ described in Section 5.6 and above. This implies $L \subseteq \mathbf{L}$ (as $\mathcal{Z}(L) \cap \mathcal{N}_{n,1}$ is otherwise empty, see Lemma 5.4.1).

Fix isomorphisms as in the statement of Lemma 6.2.4. Then φ corresponds to some $(X, \rho) \in \mathcal{N}_{n,1}$, and this lift of $\mathbf{X}^-$ corresponds to a class $\alpha' \in \mathrm{Ext}^1(\underline{\mathbb{Q}_p/\mathbb{Z}_p}^{n-1}, \boldsymbol{\mu}_{p^\infty})$ via the lifting theory in Section 6.1.

By Lemma 6.2.4, it is enough to show that α' is p^s-torsion for some $s \in \mathbb{Z}_{\geq 0}$ (then φ must factor through $\mathcal{Z}(M)^\circ$ where M is the lattice associated with the cyclic subgroup generated by α). Select $s \geq 0$ such that $p^s \mathbf{L} \subseteq L$ (such s exists because L is full rank). Then Lemma 6.2.3 implies $p^s a' = 0$, since φ factors through $\mathcal{Z}(L)$ (and hence through $\mathcal{Y}(L^-)$). $\square$

7 Canonical and quasi-canonical liftings

We retain $F/\mathbb{Q}_p$ and accompanying notation as in Section 5. In Sections 7.1 and 7.2, we allow $p = 2$ even if $F/\mathbb{Q}_p$ is ramified. We collect some needed facts about canonical and quasi-canonical lifts in all cases (inert, ramified, split). See also [Gro86], [Wew07], [Meu07]. Our conventions differ slightly from [Wew07], due to the phenomenon explained in [KR11, Footnote 7] (there in the inert case, which we also modify to apply in the ramified case).

7.1 Canonical liftings

As in Section 5.1, let $\mathbf{X}_0$ be the unique supersingular (resp. ordinary) p-divisible group of height 2 dimension 1 over $\overline{k}$ if $F/\mathbb{Q}_p$ is nonsplit (resp. split). Let $j\colon \mathcal{O}_F \hookrightarrow \mathrm{End}(\mathbf{X}_0)$ be a ring homomorphism. We reserve the notation $\iota_{\mathbf{X}_0}$ to mean a signature $(1,0)$ action, and allow j to have either signature (i.e. $(1,0)$ or $(0,1)$) for its action on $\operatorname{Lie} \mathbf{X}_0$.

Let $\breve{E}$ be any finite degree field extension of $\breve{F}$, with ring of integers $\mathcal{O}_{\breve{E}}$. The pair $(\mathbf{X}_0, j)$ admits a lift $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ (i.e. $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0})$ is a p-divisible group over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ with $\mathcal{O}_F$-action $\iota_{\mathfrak{X}_0}$, and $\rho_{\mathfrak{X}_0} \colon \mathfrak{X}_{0,\overline{k}} \to \mathbf{X}_0$ is a $\mathcal{O}_F$-linear isomorphism with respect to ι and j).

In the supersingular case, the pair $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0})$ may be described via Lubin–Tate formal groups. In the ordinary case, we have $\mathfrak{X}_0 \cong \boldsymbol{\mu}_{p^\infty} \times \underline{\mathbb{Q}_p/\mathbb{Z}_p}$.

By the *signature* of $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ (or $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0})$), we mean the signature of $\iota_{\mathfrak{X}_0}$ acting on $\operatorname{Lie} \mathfrak{X}_0$ (either $(1,0)$ or $(0,1)$). If $F/\mathbb{Q}_p$ is unramified (resp. ramified), then $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ must have the same signature as $(\mathbf{X}_0, j)$ (resp. can have either signature).

After fixing a signature, the triple $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ is unique up to unique isomorphism, and we call it the *canonical lifting*[29] of $(\mathbf{X}_0, j)$. The canonical lifting over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ is defined over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$ (i.e. is the base change of the canonical lift over $\operatorname{Spf} \mathcal{O}_{\breve{F}}$).

The map $\iota_{\mathfrak{X}_0} \colon \mathcal{O}_F \to \operatorname{End}(\mathfrak{X}_0)$ is an isomorphism, since $\operatorname{End}(\mathfrak{X}_0)$ is commutative and $\mathcal{O}_F$ is self-centralizing in $\operatorname{End}(\mathbf{X}_0)$ (in the nonsplit case, note $\operatorname{End}(\mathfrak{X}_0) \hookrightarrow \operatorname{End}(\operatorname{Lie} \mathfrak{X}_0) = \mathcal{O}_{\breve{E}}$ so $\operatorname{End}(\mathfrak{X}_0)$ must be commutative).

If $(\mathfrak{X}_0^\sigma, \iota_{\mathfrak{X}_0}^\sigma)$ is as in (5.1.1), we have $\operatorname{Hom}_{\mathcal{O}_F}(\mathfrak{X}_0, \mathfrak{X}_0^\sigma) = 0$ because $\operatorname{End}(\mathfrak{X}_0) = \mathcal{O}_F$.

Example 7.1.1. Assume $F/\mathbb{Q}_p$ is nonsplit, and let $\mathfrak{X}_0$ be the canonical lifting over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ (of some fixed signature). Drinfeld rigidity for quasi-homomorphisms implies $\operatorname{End}^0(\mathfrak{X}_0) \cong \operatorname{End}^0(\mathbf{X}_0) \cong D$, where D is the quaternion division algebra over $\mathbb{Q}_p$. On the other hand, if $\mathfrak{X}_0'$ denotes the p-divisible group over $\operatorname{Spec} \mathcal{O}_{\breve{E}}$ associated with $\mathfrak{X}_0$ via Lemma B.3.1, we have $\operatorname{End}^0(\mathfrak{X}_0') \cong \operatorname{End}(\mathfrak{X}_0') \otimes_{\mathbb{Z}_p} \mathbb{Q}_p \cong F$. Thus, by our conventions (explained in Appendix B.1), quasi-homomorphisms do not necessarily lift along the equivalence of p-divisible groups over $\operatorname{Spec} \mathcal{O}_{\breve{E}}$ and $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ from Lemma B.3.1. See also Remark B.3.5.

7.2 Quasi-canonical liftings

Let $\breve{E}$ and $(\mathbf{X}_0, j)$ be as in Section 7.1. For integers $s \geq 0$, let $\mathcal{O}_{F,s} \coloneqq \mathbb{Z}_p + p^s \mathcal{O}_F$ be the order of index p^s in $\mathcal{O}_F$. When $F/\mathbb{Q}_p$ is nonsplit (resp. split) the subgroup $\mathcal{O}_{F,s}^\times \subseteq \mathcal{O}_F^\times$ (resp. $(1 + p^s \mathbb{Z}_p)^\times \subseteq \mathbb{Z}_p^\times$) has an associated finite totally ramified abelian extension $\breve{E}_s$ of $\breve{F}$ by local class field theory. The index is

$$[\breve{E}_s : \breve{F}] = \begin{cases} p^s(1 - \eta^*(p)p^{-1}) & \text{if } s \geq 1 \\ 1 & \text{if } s = 0 \end{cases} \tag{7.2.1}$$

[29] When j has signature $(1,0)$, what Gross [Gro86] calls a *canonical lifting* is what we call a *canonical lifting of signature* $(1,0)$. This change in terminology allows additional flexibility when discussing quasi-canonical liftings, to account for e.g. [KR11, Footnote 7].

where $\eta^*(p) \coloneqq -1, 0, 1$ in the inert, ramified, and split cases respectively. In the split case, we have $\mathcal{O}_{\breve{E}_s} = \mathcal{O}_{\breve{F}}[\zeta_{p^s}]$ where ζ_{p^s} is a primitive p^s-th root of unity.

In all cases, a *quasi-canonical lifting of level s* of $(\mathbf{X}_0, j)$ is a triple $(\mathfrak{X}_s, \iota_{\mathfrak{X}_s}, \rho_{\mathfrak{X}_s})$ where

$\mathfrak{X}_s$	is a p-divisible group over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$
$\iota_{\mathfrak{X}_s} \colon \mathcal{O}_{F,s} \xrightarrow{\sim} \operatorname{End}(\mathfrak{X}_s)$	is a ring isomorphism
$\rho_{\mathfrak{X},s} \colon \mathfrak{X}_{s,\overline{k}} \to \mathbf{X}_0$	is an $\mathcal{O}_{F,s}$-linear isomorphism of p-divisible groups over $\overline{k}$.

Note that a quasi-canonical lifting of level $s = 0$ is the same as a canonical lifting. As above, we speak of the *signature* of a quasi-canonical lifting, which means the signature of the action $\iota_{\mathfrak{X}_s}|_{\operatorname{Lie} \mathfrak{X}_s}$.

The signature of $(\mathbf{X}_0, j)$ and the signature of a level s quasi-canonical lifting must be

$$\begin{cases} \text{same} & \text{if } F/\mathbb{Q}_p \text{ is inert and } s \text{ is even, or } F/\mathbb{Q}_p \text{ is split} \\ \text{opposite} & \text{if } F/\mathbb{Q}_p \text{ is inert and } s \text{ is odd} \\ \text{either signature} & \text{if } F/\mathbb{Q}_p \text{ is ramified.} \end{cases} \tag{7.2.2}$$

Quasi-canonical liftings of level $s \geq 0$ exist in all such situations, and are defined over $\operatorname{Spf} \mathcal{O}_{\breve{E}_s}$. The property of being a level s quasi-canonical lifting is preserved under base change along $\operatorname{Spf} \mathcal{O}_{\breve{E}'} \to \operatorname{Spf} \mathcal{O}_{\breve{E}}$ for any finite degree field extension $\breve{E}'$ over $\breve{E}$. If $F/\mathbb{Q}_p$ is split, a choice of level s quasi-canonical lifting corresponds to a choice of morphism $\underline{\mathbb{Z}}_p \to \boldsymbol{\mu}_{p^\infty}$ over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ of exact order p^s (i.e. a choice of primitive p^s-th root of unity in $\breve{E}_s$) via the lifting theory in Section 6.1.

The group $\operatorname{Gal}(\breve{E}_s/\breve{F})$ acts simply transitively on the set of level s quasi-canonical liftings for any fixed signature (if such liftings exist). By Lubin–Tate theory, this action is compatible with the identification $\operatorname{Gal}(\breve{E}_s/\breve{F}) \cong \mathcal{O}_F^\times/\mathcal{O}_{F,s}^\times$ via local class field theory (normalized to send uniformizers to geometric Frobenius) where $a \in \mathcal{O}_F^\times$ acts on the set of quasi-canonical liftings as $(\mathfrak{X}_s, \iota_{\mathfrak{X}_s}, \rho_{\mathfrak{X}_s}) \mapsto (\mathfrak{X}_s, \iota_{\mathfrak{X}_s}, a\rho_{\mathfrak{X}_s})$. In the split case, we have used the isomorphism

$$\begin{aligned} \mathcal{O}_F^\times/\mathcal{O}_{F,s}^\times &\longrightarrow \mathbb{Z}_p^\times/(1 + p^s\mathbb{Z}_p)^\times \\ x &\longmapsto e^+(x)e^-(x^{-1}) \end{aligned} \tag{7.2.3}$$

if $(\mathbf{X}_0, j)$ has signature $(1, 0)$ and its reciprocal if $(\mathbf{X}_0, j)$ has signature $(0, 1)$. In particular, the quasi-canonical liftings of a fixed level s are all isomorphic if the framing $\rho_{\mathfrak{X}_s}$ is forgotten.

Let $(\mathfrak{X}_0, \iota_{\mathfrak{X}_0}, \rho_{\mathfrak{X}_0})$ and $(\mathfrak{X}_s, \iota_{\mathfrak{X}_s}, \rho_{\mathfrak{X}_s})$ be canonical and quasi-canonical lifts over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$, for some $(\mathbf{X}_0, j)$ and $(\mathbf{X}_0, j')$ respectively (possibly $j \neq j'$). Then

$$\operatorname{Hom}(\mathfrak{X}_0, \mathfrak{X}_s) \cong \psi_s \cdot \mathcal{O}_F \tag{7.2.4}$$

(no $\mathcal{O}_F$-linearity imposed) is a free $\mathcal{O}_F$-module of rank 1 (where $\mathcal{O}_F$ acts by pre-composition), generated by some isogeny ψ_s of degree p^s. The isogeny ψ_s is defined over $\operatorname{Spf} \mathcal{O}_{\breve{E}_s}$. If $\mathfrak{X}_0$ and $\mathfrak{X}_s$ have the same signature, then ψ_s is automatically $\mathcal{O}_{F,s}$-linear. When $F/\mathbb{Q}_p$ is split, we may take ψ_s to be the map inducing the map $\mathbf{X}_0 \to \mathbf{X}_0$ which is

$$\psi_s|_{\mathbf{X}_0^0} \colon \mathbf{X}_0^0 \xrightarrow{\mathrm{id}} \mathbf{X}_0^0 \qquad \psi_s|_{\mathbf{X}_0^{\text{ét}}} \colon \mathbf{X}_0^{\text{ét}} \xrightarrow{\times p^s} \mathbf{X}_0^{\text{ét}} \tag{7.2.5}$$

on the connected and étale parts, respectively. This follows from the lifting criterion in (6.1.5).

For any generator ψ_s of $\operatorname{Hom}(\mathfrak{X}_0, \mathfrak{X}_s)$, we have

$$\operatorname{length}_{\mathcal{O}_{\breve{E}}}(e^* \Omega^1_{\ker \psi_s / \operatorname{Spec} \mathcal{O}_{\breve{E}}}) = \frac{1}{2}[\breve{E} : \breve{\mathbb{Q}}_p] \frac{(1 - p^{-s})(1 - \eta^*(p))}{(1 - p^{-1})(p - \eta^*(p))} \tag{7.2.6}$$

where $\eta^*(p) \coloneqq -1, 0, 1$ in the inert, ramified, split cases respectively and where $e \colon \operatorname{Spec} \mathcal{O}_{\breve{E}} \to \ker \psi_s$ denotes the identity section.

The nonsplit case of (7.2.6) is essentially a computation of Nakkajima and Taguchi [NT91] (see also [KRY04, Proposition 10.3] and its proof). The split case follows from (7.2.5), which implies that $\ker \psi_s$ is étale over $\operatorname{Spec} \overline{k}$ (cf. the closely related [KRY04, Proposition 10.1]).

The following constant $\delta_{\mathrm{tau}}(s) \in \mathbb{Q}$ ("local change of tautological height") will be crucial for the formulation of our local main theorems. With notation as above, we define

$$\begin{aligned} \delta_{\mathrm{tau}}(s) &\coloneqq -\frac{1}{2} v_p(\deg \psi_s) + \frac{1}{[\breve{E} : \breve{\mathbb{Q}}_p]} \operatorname{length}_{\mathcal{O}_{\breve{E}}}(e^* \Omega^1_{\ker \psi_s / \operatorname{Spec} \mathcal{O}_{\breve{E}}}) && (7.2.7) \\ &= -\frac{1}{2} \left(s - \frac{(1 - p^{-s})(1 - \eta^*(p))}{(1 - p^{-1})(p - \eta^*(p))} \right) && (7.2.8) \end{aligned}$$

for integers $s \in \mathbb{Z}$. We used Eq. (7.2.6) for the second equality. We also set $\delta_{\mathrm{Fal}}(s) \coloneqq -2\delta_{\mathrm{tau}}(s)$.

The quantity $\delta_{\mathrm{tau}}(s)$ depends only on s and $\eta^*(p)$, and does not depend on the choice of ψ_s. In Part III below, we will explain the relation of $\delta_{\mathrm{tau}}(s)$ and $\delta_{\mathrm{Fal}}(s)$ with local decompositions of "tautological" and Faltings heights of special cycles.

7.3 Quasi-canonical lifting cycles

We state how certain local special cycles decompose into *quasi-canonical lifting cycles* (Section 6.2). We continue to use the notation in Section 5.2, now restricting to signature $(n-1, 1)$. We also assume $p \neq 2$ unless $F/\mathbb{Q}_p$ is split (in the inert case, this is so that we may cite [LZ22a, Theorem 4.2.1]).

Suppose $M^\flat \subseteq \mathbf{W}$ is an integral $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$. Set $s = \lfloor \mathrm{val}(M^\flat) \rfloor$ (notation as in Section 2.2). There is an associated *quasi-canonical lifting cycle* $\mathcal{Z}(M^\flat)^\circ \subseteq \mathcal{N}(n-1,1)$, which is a certain closed subfunctor such that

$$\mathcal{Z}(M^\flat)^\circ \cong \begin{cases} \operatorname{Spf} \mathcal{O}_{\breve{E}_s} & \text{if } F/\mathbb{Q}_p \text{ is unramified} \\ \operatorname{Spf} \mathcal{O}_{\breve{E}_s} \sqcup \operatorname{Spf} \mathcal{O}_{\breve{E}_s} & \text{if } F/\mathbb{Q}_p \text{ is ramified.} \end{cases} \tag{7.3.1}$$

Suppose $\varphi \colon \operatorname{Spf} \mathcal{O}_{\breve{E}_s} \to \mathcal{N}(n-1,1)$ is a morphism representing any component of $\mathcal{Z}(M^\flat)^\circ$, with corresponding tuple $(X, \iota, \lambda, \rho) \in \mathcal{N}(n-1,1)(\operatorname{Spf} \mathcal{O}_{\breve{E}_s})$. If $n = 1$, then $M^\flat = 0$ and $X \cong \mathfrak{X}_0^\sigma$. If $n \geq 2$, then there exists a polarization-preserving $\mathcal{O}_F$-linear isomorphism (forgetting ρ)

$$X \cong (\mathfrak{X}_0)^{n-2} \times (\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_F) \tag{7.3.2}$$

for some level s quasi-canonical lift $\mathfrak{X}_s$ (and $\mathfrak{X}_0$ being the canonical lift), where $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_F$ is equipped with the polarization as in (5.6.2), where $\mathfrak{X}_0^{n-2}$ has the diagonal polarization $\lambda_{\mathfrak{X}_0}^{n-2}$ for some principal polarization $\lambda_{\mathfrak{X}_0}$ on $\mathfrak{X}_0$ if $F/\mathbb{Q}_p$ is unramified, and where $\mathfrak{X}_0^{n-2}$ has a product polarization as in (5.1.6) (with respect to some principal polarization $\lambda_{\mathfrak{X}_0}$ on $\mathfrak{X}_0$) if $F/\mathbb{Q}_p$ is ramified.

For the inert case of the above assertions, see [LZ22a, §4.2] (we are using the same notation), and also [KR11, Proposition 8.1] (there for $n = 2$).

For the ramified case, see [RSZ17, Proposition 7.1] (there for $n = 2$) and also the proof of [LL22, Proposition 2.44] (also [LL22, Definition 2.45]; we are using their notation but with $\mathcal{N}$ replaced by $\mathcal{Z}$). In the ramified case, the two components $\mathcal{Z}(M^\flat)^\circ$ correspond to the two components of $\mathcal{N}(n-1,1)$ (as in Lemma 5.4.1, particularly part (5)), i.e. $\mathcal{Z}(M^\flat)^\circ \to \mathcal{N}(n-1,1)$ is surjective on underlying topological spaces.

For the split case, $\mathcal{Z}(M^\flat)^\circ$ was defined in Definition 6.2.7. The assertion $X \cong (\mathfrak{X}_0)^{n-2} \times (\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_F)$ follows from Lemma 6.2.5 (note that X in loc. cit. is X^- in the present notation) and Remark 5.6.4.

Proposition 7.3.1. *Let $L^\flat \subseteq \mathbf{W}$ be an $\mathcal{O}_F$-lattice of rank $n-1$. Form the horizontal (flat) part of the local special cycle $\mathcal{Z}(L^\flat)$, which we denote as $\mathcal{Z}(L^\flat)_{\mathscr{H}}$. We have an equality of closed formal subschemes*

$$\mathcal{Z}(L^\flat)_{\mathscr{H}} = \bigcup_{\substack{L^\flat \subseteq M^\flat \subseteq M^{\flat *} \\ t(M^\flat) \leq 1}} \mathcal{Z}(M^\flat)^\circ \tag{7.3.3}$$

in $\mathcal{N}(n-1,1)$, where the union runs over full rank lattices $M^\flat \subseteq L^\flat_F$.

Proof. The union is the scheme-theoretic union (i.e. intersect associated ideal sheaves).

The inert case is [LZ22a, Theorem 4.2.1]. The ramified case is [LL22, Lemma 2.54] (if $F/\mathbb{Q}_p$ is ramified, the condition $t(M^\flat) \leq 1$ implies $t(M^\flat) = 1$ since we have assumed n is even in the ramified case).

For the split case, the inclusion $\subseteq$ is Corollary 6.2.9. By Lemma 6.2.11, the inclusion $\supseteq$ will hold if we can verify that $\mathcal{Z}(L^\flat)_{\mathscr{H}} \cong \operatorname{Spf} R$ for some finite flat $\mathcal{O}_{\breve{F}}$-algebra R with $R \otimes_{\mathcal{O}_{\breve{F}}} \breve{F}$ reduced (with R not necessarily a domain). We will check this later by passing to global special cycles via uniformization (Lemma 11.7.4). □

For readers interested in Krämer integral models for $F/\mathbb{Q}_p$ ramified, we mention the analogous [HSY23, Theorem 4.2], which we will not need.

8 Hermitian symmetric domain

8.1 Setup

We recall/fix some notation, mostly as in [Liu11, §4B] (see also [GS19, §2.2.2]). Let $n \geq 1$ be an integer, and let V be the non-degenerate $\mathbb{C}/\mathbb{R}$ Hermitian space of signature $(n-1, 1)$. We write $(-,-)$ for the Hermitian pairing on V. Consider the Hermitian symmetric domain

$$\mathcal{D} = \{\text{maximal negative definite } \mathbb{C}\text{-linear subspaces of } V\}. \tag{8.1.1}$$

Choosing a basis $\{e_1, \ldots, e_n\}$ of V with Gram matrix $\operatorname{diag}(1_{n-1}, -1)$, we take the identification

$$\begin{aligned} \mathcal{D} &\xrightarrow{\quad\sim\quad} \{z \in \mathbb{C}^{n-1} : |z| < 1\} \\ (a_1 : \cdots : a_n) &\longmapsto (a_1/a_n, \ldots, a_{n-1}/a_n) \end{aligned} \tag{8.1.2}$$

and write $z_i = a_i/a_n$. Here $(a_1 : \cdots : a_n)$ stands for the complex line spanned by $a_1 e_1 + \cdots + a_n e_n$. We implicitly use the (standard) orientation $i^{n-1} dz_1 \wedge d\overline{z}_1 \wedge \cdots \wedge dz_{n-1} \wedge d\overline{z}_{n-1}$ on $\mathcal{D}$.

We write $\mathcal{E}$ for the tautological line bundle over $\mathcal{D}$, whose fiber over a point $z \in \mathcal{D}$ is identified with the corresponding $\mathbb{C}$-line in V. We give $\mathcal{E}$ the following metric: if $w_z \in \mathcal{E}$ lies over $z \in \mathcal{D}$, set $\|w_z\|^2 = -(w_z, w_z)$. Let $\widehat{\mathcal{E}}$ be the resulting metrized line bundle. We write $c_1(\widehat{\mathcal{E}})$ for the corresponding Chern form, given locally by

$$c_1(\widehat{\mathcal{E}}) = \frac{1}{2\pi i} \partial\overline{\partial} \log \|s\|^2 \tag{8.1.3}$$

for local nowhere vanishing holomorphic sections s of $\mathcal{E}$.

8.2 Local special cycles

Given any tuple $\underline{x} = (x_1, \ldots, x_m)$ with $x_i \in V$, there is a *local special cycle*

$$\mathcal{D}(\underline{x}) \coloneqq \{z \in \mathcal{D} : z \perp x_i \text{ for all } i\} \subseteq \mathcal{D}. \tag{8.2.1}$$

This is a closed complex submanifold of $\mathcal{D}$.

Given $x \in V$, there is an associated global holomorphic section s_x of the dual metrized tautological bundle $\widehat{\mathcal{E}}^\vee$, given by $s_x(w_z) = (x, w_z)$. For $x \in V$ and $z \in \mathcal{D}$, we set $R(x, z) \coloneqq \|s_x(z)\|^2 = -(x_z, x_z)$ where $\|-\|$ is the norm on $\widehat{\mathcal{E}}^\vee$, and x_z is the orthogonal projection of x to the $\mathbb{C}$-line z.

We write $\operatorname{Ei}(u) \coloneqq -\int_1^\infty e^{ut} t^{-1}\, dt$ for the exponential integral function, where $u \in \mathbb{R}$ is negative. We will use the asymptotics

$$|\operatorname{Ei}(u)| \leq -u^{-1} e^u \qquad \lim_{u \to 0^-} (\operatorname{Ei}(u) - \log|u|) = \gamma, \tag{8.2.2}$$

where γ is the Euler–Mascheroni constant. These may be verified by brief computations (omitted, but see the integral representation for γ in [WW73, §12.2 Example 4]).

Given $x \in V$ nonzero, we set[30]

$$\xi(x) = -\operatorname{Ei}(-4\pi R(x, z)) \tag{8.2.3}$$

which is a smooth function of $z \in (\mathcal{D} \setminus \mathcal{D}(x))$ with singularity of log type along $\mathcal{D}(x)$ (in the sense of [GS90, (1.3.2.1)]).

For locally L^1-forms ξ on $\mathcal{D}$, we write $[\xi]$ for the associated current. With x as above, we have the Green current equation

$$-\frac{1}{2\pi i} \partial\overline{\partial}[\xi(x)] + \delta_{\mathcal{D}(x)} = [\omega(x)] \tag{8.2.4}$$

where $\omega(x)$ is a smooth $(1,1)$-form on $\mathcal{D}$ coinciding with the Kudla–Millson form up to a normalization [Liu11, Proposition 4.9]. Given a linearly independent tuple $\underline{x} = (x_1, \ldots, x_m) \in V^m$, we consider the current

$$[\xi(\underline{x})] \coloneqq [\xi(x_1)] * ([\xi(x_2)] * \cdots ([\xi(x_{m-1})] * [\xi(x_m)])) \tag{8.2.5}$$

defined via star product (compare [GS90, §2.1.3]), e.g.

$$[\xi(x_1)] * ([\xi(x_2)] * [\xi(x_3)]) = \xi(x_1) \wedge \delta_{\mathcal{D}(x_2) \cap \mathcal{D}(x_3)} + \omega(x_1) \wedge \xi(x_2) \wedge \delta_{\mathcal{D}(x_3)} + \omega(x_1) \wedge \omega(x_2) \wedge \xi(x_3). \tag{8.2.6}$$

[30]Note that Liu instead uses $-\operatorname{Ei}(-2\pi R(x, z))$ [Liu11, §4B]. This is because he considers Gram matrices $T = \frac{1}{2}(\underline{x}, \underline{x})$ while we consider Gram matrices $T = (\underline{x}, \underline{x})$ (to match our global and non-Archimedean conventions). This also affects other normalizations, e.g. our $\omega(x)$ is Liu's $\omega(\sqrt{2}x)$.

We then have the Green current equation

$$-\frac{1}{2\pi i}\partial\overline{\partial}[\xi(\underline{x})] + \delta_{\mathcal{D}(\underline{x})} = [\omega(\underline{x})] \tag{8.2.7}$$

where $\omega(\underline{x}) \coloneqq \omega(x_1) \wedge \cdots \wedge \omega(x_m)$ (follows from (8.2.7) as in the proof of [GS90, Theorem 2.4.1(i)]).

For any nonzero $x \in V$ and $a \in \mathbb{C}^\times$, we have

$$\lim_{a \to 0} \omega(ax) = c_1(\widehat{\mathcal{E}}^\vee) \tag{8.2.8}$$

where the convergence is pointwise and uniform on compact subsets of $\mathcal{D}$ (the derivatives also converge uniformly on compact subsets). This limiting statement follows upon inspecting [GS19, (2.40)] (see also (8.3.1) and (8.3.3)). For convenience, we set $\omega(x) \coloneqq c_1(\widehat{\mathcal{E}}^\vee)$ when $x = 0$.

The group $U(V)$ acts on $\mathcal{D}$ via the moduli description. For any $g \in U(V)$, we have

$$g(\mathcal{D}(\underline{w})) = \mathcal{D}(g \cdot \underline{w}) \qquad g_*[\xi(\underline{x})] = [\xi(g \cdot \underline{x})] \tag{8.2.9}$$

where $\underline{w} \in V^m$ is any tuple and $\underline{x} \in V^m$ is any linearly independent tuple.

8.3 Green current convergence

We record some convergence estimates for the integrals appearing in our main Archimedean local identities (Section 19.1). We work with the explicit coordinates $z = (z_1, \ldots, z_{n-1})$ on $\mathcal{D}$ from Section 8.1 above (via the choice of basis $\{e_1, \ldots, e_n\}$ for V). For any nonzero $x \in V$, we have

$$c_1(\widehat{\mathcal{E}}^\vee) = \frac{1}{2\pi i}\partial\overline{\partial}\log R = \frac{1}{2\pi i}\frac{R\partial\overline{\partial}R - \partial R \wedge \overline{\partial}R}{R^2} \tag{8.3.1}$$

$$= \frac{1}{2\pi i}\left(\frac{\sum dz_j \wedge d\overline{z}_j}{1 - z\overline{z}} + \frac{(\sum \overline{z}_j dz_j) \wedge (\sum z_j d\overline{z}_j)}{(1 - z\overline{z})^2}\right). \tag{8.3.2}$$

and

$$\omega(x) = -\frac{1}{2\pi i}\partial\overline{\partial}\xi(x) = \frac{1}{2\pi i}e^{-4\pi R}\left(\frac{-4\pi\partial R \wedge \overline{\partial}R}{R} + \frac{\partial\overline{\partial}R}{R} - \frac{\partial R \wedge \overline{\partial}R}{R^2}\right) \tag{8.3.3}$$

on $\mathcal{D} \setminus \mathcal{D}(x)$, where $R \coloneqq R(x, z)$ for short.

Lemma 8.3.1. *For any fixed $x \in V$ (possibly $x = 0$) with $\omega(x) = \sum_{i,j} \omega(x)_{i,j} dz_i \wedge d\overline{z}_j$, the functions $(1 - z\overline{z})^3\omega(x)_{i,j}$ are bounded on $\mathcal{D}$.*

Proof. If $x = \sum a_j e_j$, we have

$$R(x,z) = \frac{(a_1\overline{z}_1 + \cdots + a_{n-1}\overline{z}_{n-1} - a_n)(\overline{a}_1 z_1 + \cdots + \overline{a}_{n-1} z_{n-1} - \overline{a}_n)}{(1 - z\overline{z})}. \tag{8.3.4}$$

This expression and the formulas for $\omega(x)$ (see above) yield the lemma via straightforward computation (omitted). □

Lemma 8.3.2. *Let $\underline{x} = (x_1, \ldots, x_m) \in V^m$ be an m-tuple with nonsingular Gram matrix $(\underline{x}, \underline{x})$. Assume either that $m \geq n-1$ or that $(\underline{x}, \underline{x})$ is not positive definite. Then exists $\epsilon > 0$ such that*

$$\sum_{i=1}^{n-1} R(x_i, z) > \frac{\epsilon}{1 - z\overline{z}} \tag{8.3.5}$$

for all $z \in \mathcal{D}$ with $|z| \gg 0$.

Proof. Given $x = \sum_j a_j e_j \in V$, we use the temporary notation $x \cdot z \coloneqq a_1\overline{z}_1 + \cdots + a_{n-1}\overline{z}_{n-1} - a_n$ for $z = (z_1, \ldots, z_{n-1}) \in \mathbb{C}^{n-1}$. Note $R(x,z) = |x \cdot z|^2 (1 - z\overline{z})^{-1}$ for $z \in \mathcal{D}$. View $\mathbb{C}^{n-1}$ as a standard coordinate chart in the projective space of lines in V (i.e. the lines which are not orthogonal to e_n). The zeros of $\sum_i |x_i \cdot z|^2$ on $\mathbb{C}^{n-1}$ correspond to those lines in V (in the given chart) which are orthogonal to $\operatorname{span}(\underline{x})$. This (closed) set of zeros is disjoint from the set $\{z \in \mathbb{C}^{n-1} : |z| = 1\}$, which corresponds to isotropic lines in V (i.e. no isotropic lines in V are orthogonal to $\operatorname{span}(\underline{x})$). Hence $\sum_i R(x_i, z)(1 - z\overline{z})$ is bounded below (as a function of $z \in \mathcal{D}$) by a positive constant as $|z| \to 1$. □

Lemma 8.3.3. *Let $\underline{x} = (x_1, \ldots, x_m) \in V^m$ be an m-tuple with nonsingular Gram matrix $(\underline{x}, \underline{x})$. Assume either that $m \geq n-1$ or that $(\underline{x}, \underline{x})$ is not positive definite.*

Let $\omega = \sum \omega_{I,J} dz_I \wedge d\overline{z}_J$ (multi-indices) be any smooth complex differential form on $\mathcal{D}$ such that each $(1 - z\overline{z})^b \omega_{I,J}$ is bounded on $\mathcal{D}$ for some real constant $b \gg 0$. Then the integral

$$\int_{\mathcal{D}} \xi(x_1)\omega(x_2) \wedge \cdots \wedge \omega(x_m) \wedge \omega \tag{8.3.6}$$

is absolutely convergent.

Proof. After making a unitary change of basis for V, we may assume

$$x_1 = \begin{cases} a e_n & \text{if } (x_1, x_1) < 0 \\ a e_1 & \text{if } (x_1, x_1) > 0 \\ e_{n-1} + e_n & \text{if } (x_1, x_1) = 0 \end{cases} \tag{8.3.7}$$

for some nonzero $a \in \mathbb{R}$ (where $(e_1, \ldots, e_n)$ is the basis of V used to define the coordinates $(z_1, \ldots, z_{n-1})$ in Section 8.1). This will aid calculation in coordinates.

Lemma 8.3.1 shows that it is enough to check (absolute) convergence of

$$\int_{\mathcal{D}} \xi(x_1) e^{-4\pi(R(x_2,z)+\cdots+R(x_m,z))}(1 - z\overline{z})^{-b} \tag{8.3.8}$$

for any $b \in \mathbb{R}$ (for the Euclidean measure on $\mathcal{D}$). It is enough to check convergence when $b \gg 0$, so we assume $b \geq n$ for convenience.

Set

$$u_j \coloneqq \frac{\mathrm{Re}(z_j)}{\sqrt{1 - z\overline{z}}} \qquad v_j \coloneqq \frac{\mathrm{Im}(z_j)}{\sqrt{1 - z\overline{z}}} \tag{8.3.9}$$

for $j = 1, \ldots, n-1$. A change of variables gives

$$\begin{aligned} &\int_{\mathcal{D}} \xi(x_1) e^{-4\pi(R(x_2,z)+\cdots+R(x_m,z))}(1 - z\overline{z})^{-b} \\ &= \int_{\mathbb{R}^{2(n-1)}} \xi(x_1) e^{-4\pi(R(x_2,z)+\cdots+R(x_m,z))}(1 + |u|^2 + |v|^2)^{b-n}, \end{aligned} \tag{8.3.10}$$

where $|u|^2 \coloneqq \sum_j u_j^2$ and $|v|^2 \coloneqq \sum_j v_j^2$, with $R(x_i, z)$ a function of u, v via (8.3.10), and with the Euclidean measure $du_1\, dv_1 \cdots du_{n-1} dv_{n-1}$ understood on the right-hand side.

The asymptotics for $\mathrm{Ei}(u)$ as in (8.2.2) show it is enough to check convergence of the integrals

$$\int_{\mathbb{R}^{2(n-1)}} e^{-4\pi(R(x_1,z)+R(x_2,z)+\cdots R(x_m,z))}(1 + |u|^2 + |v|^2)^{b-n} \tag{8.3.11}$$

$$\text{and} \qquad \int_{\substack{\mathbb{R}^{2(n-1)} \\ R(x_1,z) \leq 1/(8\pi)}} \log(4\pi R(x_1, z)) e^{-4\pi(R(x_2,z)+\cdots+R(x_m,z))}(1 + |u|^2 + |v|^2)^{b-n} \tag{8.3.12}$$

(where the second integral is over the set of $(u, v) \in \mathbb{R}^{2(n-1)}$ satisfying $R(x_1, z) \leq 1/(8\pi)$).

Since we have $(1 - z\overline{z})^{-1} = 1 + |u|^2 + |v|^2$, Lemma 8.3.2 implies that (8.3.11) is absolutely convergent (by exponential decay of the integrand as $|u|^2 + |v|^2 \to \infty$).

For convergence of (8.3.12), the same lemma shows that it is enough to check convergence of the integral

$$\int_{\substack{\mathbb{R}^{2(n-1)} \\ R(x_1,z) \leq 1/(8\pi)}} \log(8\pi R(x_1, z)) e^{-4\pi\epsilon(1+|u|^2+|v|^2)}(1 + |u|^2 + |v|^2)^{b-n} \tag{8.3.13}$$

for all $\epsilon > 0$ (using also $R(x_1, z) \leq 1/(8\pi)$). We check this convergence via casework.

Case when $(x_1, x_1) < 0$*:* In this case, we have $R(x_1, z) = a^2(1 + |u|^2 + |v|^2)$. The integrand in (8.3.12) is bounded on the compact set $\{(u, v) \in \mathbb{R}^{2(n-1)} : R(x_1, z) \leq 1/(8\pi)\}$, hence the integral is convergent.

Case when $(x_1, x_1) > 0$*:* In this case, we have $R(x_1, z) = a^2(u_1^2 + v_1^2)$. To check convergence of (8.3.13), it is enough to check that

$$\int_{\substack{\mathbb{R}^{2(n-1)} \\ a^2(u_1^2+v_1^2) \leq 1/(8\pi)}} \log(4\pi a^2(u_1^2+v_1^2))e^{-4\pi\epsilon(1+u_2^2+v_2^2+\cdots+u_{n-1}^2+v_{n-1}^2)}(1+u_2^2+v_2^2+\cdots+u_{n-1}^2+v_{n-1}^2)^{b-n} \tag{8.3.14}$$

is convergent (using $R(x_1, z) \leq 1/(8\pi)$). The integral over (u_1, v_1) converges because the singularity at $u_1 = v_1 = 0$ is logarithmic, and the integral over $(u_2, v_2, \ldots, u_{n-1}, v_{n-1})$ converges because of the exponential decay.

Case when $(x_1, x_1) = 0$*:* In this case, we have $R(x_1, z) = (u_1 - \sqrt{1 + |u|^2 + |v|^2})^2 + v_1^2$. Under the condition $R(x_1, z) \leq 1/(8\pi)$, we may bound $|\log R(x_1, z)| \leq C \cdot (1 + |u_1|)$ for some constant $C > 0$. To check convergence of (8.3.13), it is thus enough to check that

$$\int_{\mathbb{R}^{2(n-1)}} (1 + |u_1|)e^{-4\pi\epsilon(1+|u|^2+|v|^2)}(1 + |u|^2 + |v|^2)^{b-n} \tag{8.3.15}$$

is convergent, which follows from exponential decay of the integrand. $\square$

Remark 8.3.4. The convergence result of Lemma 8.3.3 fails in general if $m < n - 1$ and $(\underline{x}, \underline{x})$ is positive definite. For example, if $n = 3$, if $m = 1$, and if $x \in V$ with $(x, x) > 0$, the integral

$$\int_{\mathcal{D}} \xi(x) \wedge \omega(0)^2 \tag{8.3.16}$$

is not absolutely convergent.

Part III

Local change of heights

Fix an imaginary quadratic field $F/\mathbb{Q}$. Write Δ for the discriminant and σ for the nontrivial involution. We allow $2 \mid \Delta$ in Part III unless otherwise specified. We set $F_p \coloneqq F \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and $\mathcal{O}_{F_p} \coloneqq \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$.

Throughout Part III, we write E for a number field, with ring of integers $\mathcal{O}_E$. Given a prime p, we set $\mathcal{O}_{E,(p)} \coloneqq \mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$. We use $\breve{E}$ to denote a finite degree field extension of $\breve{\mathbb{Q}}_p$, with ring of integers $\mathcal{O}_{\breve{E}}$. We write $\mathfrak{d}_p \subseteq \mathcal{O}_{F_p}$ for the different ideal of $\mathcal{O}_{F_p/\mathbb{Q}_p}$. We abuse notation and also mean $\mathfrak{d}_p \coloneqq \sqrt{\Delta}$, which is a generator of the different ideal.

By a *place* $\breve{w}$ of $E \otimes_{\mathbb{Q}} \breve{\mathbb{Q}}_p$, we mean a prime ideal of $E \otimes_{\mathbb{Q}} \breve{\mathbb{Q}}_p$ (equivalently, an element of $\mathrm{Hom}_{\breve{\mathbb{Q}}_p}(E \otimes_{\mathbb{Q}} \breve{\mathbb{Q}}_p, \mathbb{C}_p)$ up to automorphisms of $\mathbb{C}_p$). We write $\breve{E}_{\breve{w}}$ for the residue field of $\breve{w}$, with ring of integers $\mathcal{O}_{\breve{E}_{\breve{w}}}$. We use the shorthand $\breve{w} \mid p$ to indicate a place of $E \otimes_{\mathbb{Q}} \breve{\mathbb{Q}}_p$, and may use subscripts (e.g. $X'_{\breve{w}}$ and $\phi_{\breve{w}}$ in Section 10.2) to indicate base-change from $\operatorname{Spec} \mathcal{O}_{E,(p)}$ to $\operatorname{Spf} \mathcal{O}_{\breve{E}_{\breve{w}}}$.

Whenever an $\mathcal{O}_F$-action or F-action is mentioned (e.g. on a sheaf of modules on $\operatorname{Spec} \mathcal{O}_E$), we assume that $\mathcal{O}_E$ (resp. $\mathcal{O}_{\breve{E}}$) is equipped with morphism $\mathcal{O}_F \to \mathcal{O}_E$ (resp. $\mathcal{O}_F \to \mathcal{O}_{\breve{E}}$).

We write $\mathfrak{X}_s$ for a level $s \geq 0$ quasi-canonical lifting of signature $(1,0)$ over $\operatorname{Spec} \mathcal{O}_{\breve{E}}$ with its $\mathcal{O}_{F_p}$-action $\iota_{\mathfrak{X}_s}$, as explained in Section 7. The framing $\rho_{\mathfrak{X}_s}$ of loc. cit. is unimportant in Part III (and will be omitted). As before, the notation $\mathfrak{X}_s^\sigma$ means $\mathfrak{X}_s$ but with $\mathcal{O}_{F_p}$-action pre-composed by σ.

Given a group scheme G over a base S, we typically write $e \colon S \to G$ for the identity section. We abuse notation and use "e" simultaneously for different group schemes.

9 Faltings and "tautological" heights

9.1 Heights

Suppose $A \to \operatorname{Spec} \mathcal{O}_E$ is a semi-abelian Néron model of an abelian variety over E. The *Faltings height* of A (or its generic fiber A_E) is

$$h_{\mathrm{Fal}}(A_E) \coloneqq h_{\mathrm{Fal}}(A) \coloneqq \frac{1}{[E : \mathbb{Q}]} \widehat{\deg}(\widehat{\omega}_A) \tag{9.1.1}$$

where $\widehat{\omega}_A = (\omega_A, \|-\|) = (e^* \bigwedge^n \Omega^1_{A/\mathcal{O}_E}, \|-\|)$ is the Hermitian line bundle with norm $\|-\|$ normalized as in (4.3.1). The usual arithmetic degree $\widehat{\deg}$ was recalled in Section 4.1. Any abelian variety over a number field has everywhere potentially semi-abelian reduction, and the *Faltings height* of any abelian variety B over $\operatorname{Spec} E$ is defined so that $h_{\mathrm{Fal}}(B)$ is remains constant under finite field extensions $E \to E'$. (We only consider stable Faltings height, as defined above.)

We also consider certain "tautological heights" to describe the arithmetic intersections appearing in Section 4.7. The terminology we introduce for this (e.g. "Krämer datum") is likely nonstandard.

Definition 9.1.1.

(1) Given a scheme S over $\operatorname{Spec} \mathcal{O}_F$, a *Krämer datum* (of signature $(n-1,1)$) is a tuple $(A, \iota, \mathcal{F})$ where $A \to S$ is an abelian scheme, where $\iota \colon \mathcal{O}_F \to \operatorname{End}(A)$ an action of signature $(n-1,1)$, and where $\mathcal{F} \subseteq \operatorname{Lie} A$ is a ι-stable local direct summand of rank $n-1$ such that the $\mathcal{O}_F$ action via ι on $\mathcal{F}$ (resp. $(\operatorname{Lie} A)/\mathcal{F}$) is $\mathcal{O}_F$-linear (resp. σ-linear). We say that $\mathcal{F}$ is the associated *Krämer hyperplane*.

(2) Given a formal scheme S over $\operatorname{Spf} \mathcal{O}_{F_p}$, a *local Krämer datum* (of signature $(n-1,1)$) is a tuple $(X, \iota, \mathcal{F})$ where X is a p-divisible group over S of height $2n$ and dimension n, where $\iota \colon \mathcal{O}_{F_p} \to \operatorname{End}(X)$ is an action of signature $(n-1,1)$, and where $\mathcal{F} \subseteq \operatorname{Lie} X$ is a ι-stable local direct summand of rank $n-1$ such that the $\mathcal{O}_F$ action via ι on $\mathcal{F}$ (resp. $(\operatorname{Lie} X)/\mathcal{F}$) is $\mathcal{O}_F$-linear (resp. σ-linear). We say that $\mathcal{F}$ is the associated *Krämer hyperplane*.

(3) A *quasi-polarized* Krämer datum (resp. *quasi-polarized* local Krämer datum) is a tuple $(A, \iota, \lambda, \mathcal{F})$ (resp. $(X, \iota, \lambda, \mathcal{F})$) where (A, ι, λ) is a Hermitian abelian scheme (Definition 3.1.1) (resp. (X, ι, λ) is a Hermitian p-divisible group (Definition 5.1.1, but we allow $p = 2$ even if $F/\mathbb{Q}_p$ is ramified)) and $(A, \iota, \mathcal{F})$ is a Krämer datum (resp. $(X, \iota, \mathcal{F})$ is a Krämer datum).

The name "Krämer datum" refers to the Krämer model mentioned in Remark 3.2.7. For an understood Krämer datum $(A, \iota, \lambda, \mathcal{F})$, we will use the shorthand $\mathscr{E}^\vee := (\operatorname{Lie} A)/\mathcal{F}$ (cf. the "tautological bundles" of Definition 3.1.7 and Definition 3.2.6). We use the same notation $\mathscr{E}^\vee := (\operatorname{Lie} X)/\mathcal{F}$ given an understood local Krämer datum $(X, \iota, \mathcal{F})$. In both cases, the sheaf $\mathscr{E}^\vee$ is locally free of rank 1, and we call it the associated *Krämer hyperplane quotient*.

Definition 9.1.2. A *morphism* (resp. *isogeny*) of Krämer data $(A_1, \iota_1, \mathcal{F}_1) \to (A_2, \iota_2, \mathcal{F}_2)$ is an $\mathcal{O}_F$-homomorphism (resp. isogeny) $A_1 \to A_2$ such that $\operatorname{im}(\mathcal{F}_1) \subseteq \mathcal{F}_2$, where $\operatorname{im}(\mathcal{F}_1)$ is the image of $\mathcal{F}_1$ under $\operatorname{Lie} A_1 \to \operatorname{Lie} A_2$. A *morphism* (resp. *isogeny*) of local Krämer data is defined in the same way.

Lemma 9.1.3. *Let S be a scheme over $\operatorname{Spec} \mathcal{O}_F$. Assume either that S is a scheme over $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$ or that $S = \operatorname{Spec} R$ where R is a Dedekind domain with fraction field of characteristic 0.*

(1) Suppose $A \to S$ is an abelian scheme with an action $\iota\colon \mathcal{O}_F \to \mathrm{End}(X)$ of signature $(n-1,1)$. Then the pair (A,ι) extends uniquely to a Krämer datum $(A,\iota,\mathcal{F})$ over S.

(2) Given pairs (A_1,ι_1) and (A_2,ι_2) as above, any $\mathcal{O}_F$-linear homomorphism (resp. isogeny) $A_1 \to A_2$ induces a morphism (resp. isogeny) of Krämer data.

(3) If S is a scheme over $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$, the exact sequence

$$0 \to \mathcal{F} \to \operatorname{Lie} A \to \mathscr{E}^\vee \to 0 \tag{9.1.2}$$

has a unique $\mathcal{O}_F$-linear splitting.

Proof. If S is a scheme over $\operatorname{Spec} \mathcal{O}_F[1/\Delta]$, the claims hold because there is a unique decomposition $\operatorname{Lie} A = (\operatorname{Lie} A)^+ \oplus (\operatorname{Lie} A)^-$ characterized by ι acting $\mathcal{O}_F$-linearly on the rank $n-1$ subbundle $(\operatorname{Lie} A)^+$ (resp. σ-linearly on the rank 1 bundle $(\operatorname{Lie} A)^-$).

Suppose instead that $S = \operatorname{Spec} R$ is a Dedekind domain with fraction field K of characteristic 0. By localizing, it is enough to verify the lemma when R is a discrete valuation ring. Then part (1) amounts to the following fact: given a finite free R-module M and any K-subspace $W \subseteq M \otimes K$, there is a unique summand $M' \subseteq M$ such that $W = M' \otimes K$ (namely $M' = M \cap W$; note that $M' \subseteq M$ is a saturated sublattice). We are applying this when $M = \operatorname{Lie} A$ and $W = (\operatorname{Lie} A \otimes K)^+$, in the notation above (and taking $\mathcal{F} = M'$). The signature $(n-1,1)$ condition forces the $\mathcal{O}_F$-action on $(\operatorname{Lie} A)/\mathcal{F}$ to be σ-linear. These considerations also verify the claim in part (2) (since it holds in the generic fiber). $\square$

Lemma 9.1.4. *Let S be a formal scheme over $\operatorname{Spf} \mathcal{O}_{F_p}$. Assume either that p is unramified in $\mathcal{O}_F$ or that $S = \operatorname{Spf} R$ for an adic ring which is a Dedekind domain with fraction field of characteristic 0. Then the following conclusions hold.*

(1) Suppose X is a p-divisible group of height $2n$ over S with an action $\iota\colon \mathcal{O}_{F_p} \to \mathrm{End}(X)$ of signature $(n-1,1)$. Then the pair (X,ι) extends uniquely to a local Krämer datum $(X,\iota,\mathcal{F})$ over S.

(2) Given pairs (X_1,ι_1) and (X_2,ι_2) as above, any $\mathcal{O}_{F_p}$-linear homomorphism (resp. isogeny) $X_1 \to X_2$ induces a morphism (resp. isogeny) of local Krämer data.

(3) If p is unramified, the exact sequence

$$0 \to \mathcal{F} \to \operatorname{Lie} A \to \mathscr{E}^\vee \to 0 \tag{9.1.3}$$

has a unique $\mathcal{O}_{F_p}$-linear splitting.

Proof. This may be proved in the same way as Lemma 9.1.3. If $S = \operatorname{Spf} R$ for R a Dedekind domain with fraction field of characteristic 0, note that R must be a complete discrete valuation ring. $\square$

In the situations of Lemma 9.1.3 and 9.1.3, we also use the alternative terminology *dual tautological bundle* for the Krämer hyperplane quotient $\mathscr{E}^\vee$.

If (A, ι, λ) is a Hermitian abelian scheme of signature $(n-1, 1)$ over $\operatorname{Spec} \mathcal{O}_E$ with associated quasi-polarized Krämer datum $(A, \iota, \lambda, \mathcal{F})$, we thus obtain a Hermitian line bundle $\widehat{\mathscr{E}^\vee} = (\mathscr{E}^\vee, \|-\|)$ on $\operatorname{Spec} \mathcal{O}_E$ as follows: the metric $\|-\|$ is given by restricting the metric on $\operatorname{Lie} A$ induced by λ (which we take to be normalized as in (4.3.3)) along the $\mathcal{O}_F$-linear splitting $\mathscr{E}^\vee[1/\Delta] \hookrightarrow (\operatorname{Lie} A)[1/\Delta]$ (where $(-)[1/\Delta]$ means restriction to $\operatorname{Spec} \mathcal{O}_E[1/\Delta]$). We say $\widehat{\mathscr{E}}^\vee$ is the associated *metrized dual tautological bundle*. We also make the same construction over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ and $\operatorname{Spec} E$.

Definition 9.1.5. Let (A, ι, λ) be a Hermitian abelian scheme of signature $(n-1, 1)$ over $\operatorname{Spec} \mathcal{O}_E$. The associated *tautological height* is

$$h_{\mathrm{tau}}(A_E) := h_{\mathrm{tau}}(A) := \frac{1}{[E : \mathbb{Q}]} \widehat{\deg}(\widehat{\mathscr{E}}^\vee). \tag{9.1.4}$$

The tautological height depends on the auxiliary data in the definition (not just A_E or A), which we have suppressed from notation. If (A, ι, λ) is a Hermitian abelian scheme of signature $(n-1, 1)$ over $\operatorname{Spec} E$ such that A has everywhere potentially good reduction, we define the *tautological height* $h_{\mathrm{tau}}(A)$ so that it is invariant under finite degree field extension $E \to E'$.

Remark 9.1.6. If we instead work over $\mathcal{O}_E[1/N]$ for some integer $N \geq 1$, we may define *Faltings height* and *tautological height* as above, but where $\widehat{\deg}$ now takes values in $\mathbb{R}_N := \mathbb{R}/\sum_{p|N} \mathbb{Q} \cdot \log p$ (as explained in Section 4.1).

9.2 Change along global isogenies

Let $A_1 \to \operatorname{Spec} \mathcal{O}_E$ and $A_2 \to \operatorname{Spec} \mathcal{O}_E$ be semi-abelian Néron models of abelian varieties over E. We have

$$[E : \mathbb{Q}](h_{\mathrm{Fal}}(A_2) - h_{\mathrm{Fal}}(A_1)) = \widehat{\deg}(\widehat{\omega}_{A_2}) - \widehat{\deg}(\widehat{\omega}_{A_1}) = -\widehat{\deg}(\underline{\mathrm{Hom}}(\widehat{\omega}_{A_2}, \widehat{\omega}_{A_1})). \tag{9.2.1}$$

Any isogeny $\phi\colon A_1 \to A_2$ defines a section ϕ of the Hermitian line bundle $\underline{\mathrm{Hom}}(\widehat{\omega}_{A_2}, \widehat{\omega}_{A_1})$, which gives

$$h_{\mathrm{Fal}}(A_2) - h_{\mathrm{Fal}}(A_1) = \frac{1}{[E:\mathbb{Q}]}\left(\log \|\phi\|_\infty + \sum_{v<\infty} \log \|\phi\|_v\right) \tag{9.2.2}$$

$$= \frac{1}{2}\log(\deg\phi) - \frac{1}{[E:\mathbb{Q}]}\log |e^*\Omega^1_{\ker\phi/\mathcal{O}_E}| \tag{9.2.3}$$

(sum is over places v of E) as in [Fal86, Lemma 5], where $|e^*\Omega^1_{\ker\phi/\mathcal{O}_E}|$ denotes the cardinality of the finite length $\mathcal{O}_E$-module $e^*\Omega^1_{\ker\phi/\mathcal{O}_E}$. Note $|e^*\Omega^1_{\ker\phi/\mathcal{O}_E}| = |\mathrm{coker}(\phi^*\colon \omega_{A_2} \to \omega_{A_1})| = |\mathrm{coker}(\phi_*\colon \operatorname{Lie} A_1 \to \operatorname{Lie} A_2)|$. Also note

$$h_{\mathrm{Fal}}(A_2) - h_{\mathrm{Fal}}(A_1) = \sum_p a_p \log p = \sum_{p \mid \deg\phi} a_p \log p \tag{9.2.4}$$

for some $a_p \in \mathbb{Q}$ independent of ϕ.

Given Hermitian abelian schemes $(A_1, \iota_1, \lambda_1)$ and $(A_2, \iota_2, \lambda_2)$ of signature $(n-1,1)$ over $\operatorname{Spec}\mathcal{O}_E$ with associated Hermitian line bundles $\widehat{\mathscr{E}}_1^\vee$ and $\widehat{\mathscr{E}}_2^\vee$, we similarly have

$$h_{\mathrm{tau}}(A_2) - h_{\mathrm{tau}}(A_1) = \frac{1}{[E:\mathbb{Q}]}\widehat{\deg}(\underline{\mathrm{Hom}}(\widehat{\mathscr{E}}_1^\vee, \widehat{\mathscr{E}}_2^\vee)). \tag{9.2.5}$$

Any $\mathcal{O}_F$-linear isogeny $\phi\colon A_1 \to A_2$ defines a section ϕ of the Hermitian line bundle $\underline{\mathrm{Hom}}(\widehat{\mathscr{E}}_1^\vee, \widehat{\mathscr{E}}_2^\vee)$, and we have

$$h_{\mathrm{tau}}(A_2) - h_{\mathrm{tau}}(A_1) = \frac{1}{[E:\mathbb{Q}]}\left(-\log \|\phi\|_\infty - \sum_{v<\infty} \log \|\phi\|_v\right) \tag{9.2.6}$$

$$= \frac{1}{[E:\mathbb{Q}]}\left(-\log\|\phi\|_\infty + \log|\mathrm{coker}(\phi_*\colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)|\right). \tag{9.2.7}$$

9.3 Change along local isogenies: Faltings

Given an isogeny $\phi\colon A_1 \to A_2$ of abelian schemes over $\operatorname{Spec}\mathcal{O}_{E,(p)}$, we define the *semi-global change of Faltings height*

$$\delta_{\mathrm{Fal},(p)}(\phi) \coloneqq -\frac{1}{2}\log|\deg\phi|_p - \frac{1}{[E:\mathbb{Q}]}\log|e^*\Omega^1_{\ker\phi/\mathcal{O}_{E,(p)}}|_p \tag{9.3.1}$$

where $|-|_p$ is the usual p-adic norm. We have $\delta_{\mathrm{Fal},(p)}(\phi) \in \mathbb{Q}\cdot\log p$. The formula for change of Faltings height (9.2.2) shows that $\delta_{\mathrm{Fal},(p)}(\phi) = a_p \log p$, in the notation of (9.2.4). In particular, $\delta_{\mathrm{Fal},(p)}(\phi)$ does not depend on the choice of isogeny ϕ (and depends only on A_1 and A_2). If A_1 and A_2 have everywhere potentially good reduction, we have

$$h_{\mathrm{Fal}}(A_{2,E}) - h_{\mathrm{Fal}}(A_{1,E}) = \sum_{\ell} \delta_{\mathrm{Fal},(\ell)}(\phi) = \sum_{\ell \mid \deg \phi} \delta_{\mathrm{Fal},(\ell)}(\phi) \tag{9.3.2}$$

where ϕ also denotes the induced isogeny on Néron models over $\operatorname{Spec} \mathcal{O}_{E,(\ell)}$ for each prime ℓ (after enlarging E if necessary).

Given any isogeny $\phi \colon X_1 \to X_2$ of p-divisible groups over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$, we have $(\operatorname{Lie} X_i)^\vee \cong e^* \Omega^1_{X_i[p^N]/\operatorname{Spec} \mathcal{O}_{\breve{E}}}$ (canonically) for $N \gg 0$ by [Mes72, Corollary II.3.3.17] (passing to the limit over $\mathcal{O}_{\breve{E}}/p^k \mathcal{O}_{\breve{E}}$ as $k \to \infty$), so there is a canonical exact sequence

$$0 \to (\operatorname{Lie} X_2)^\vee \xrightarrow{\phi^*} (\operatorname{Lie} X_1)^\vee \to e^* \Omega^1_{\ker \phi / \operatorname{Spec} \mathcal{O}_{\breve{E}}} \to 0 \tag{9.3.3}$$

of finite free $\mathcal{O}_{\breve{E}}$-modules (note that $\operatorname{Lie} X_1 \to \operatorname{Lie} X_2$ is injective, e.g. by Lemma B.2.2). If X_1 and X_2 are moreover height $2n$ and dimension n, we define the *local change of Faltings height*

$$\breve{\delta}_{\mathrm{Fal}}(\phi) := \frac{1}{2} \log(\deg \phi) - \frac{1}{[\breve{E} : \breve{\mathbb{Q}}_p]} \operatorname{length}_{\mathcal{O}_{\breve{E}}}(e^* \Omega^1_{\ker \phi / \operatorname{Spec} \mathcal{O}_{\breve{E}}}) \cdot \log p. \tag{9.3.4}$$

We have $\breve{\delta}_{\mathrm{Fal}}(\phi) = \mathbb{Q} \cdot \log p$, as well as

$$\breve{\delta}_{\mathrm{Fal}}(\phi' \circ \phi) = \breve{\delta}_{\mathrm{Fal}}(\phi') + \breve{\delta}_{\mathrm{Fal}}(\phi) \qquad\qquad \breve{\delta}_{\mathrm{Fal}}([N]) = 0 \tag{9.3.5}$$

where $\phi' \colon X_2 \to X_3$ is any isogeny of p-divisible groups and $[N] \colon X_1 \to X_1$ is the multiplication-by-N isogeny (follows from (9.3.3)). Unlike $\delta_{\mathrm{Fal},(p)}(-)$ from above, the quantity $\breve{\delta}_{\mathrm{Fal}}(\phi)$ may depend on the isogeny ϕ.

Given isogenous abelian schemes over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ and an isogeny $\phi \colon X_1 \to X_2$ of the associated p-divisible groups, set

$$\delta_{\mathrm{Fal},(p)}(\phi) := \frac{1}{[E : \mathbb{Q}]} \sum_{\breve{w} \mid p} [\breve{E}_{\breve{w}} : \breve{\mathbb{Q}}_p] \breve{\delta}_{\mathrm{Fal}}(\phi_{\breve{w}}) \tag{9.3.6}$$

where $\phi_{\breve{w}}$ denotes the base-change of ϕ to $\operatorname{Spf} \mathcal{O}_{\breve{E}_{\breve{w}}}$.

Lemma 9.3.1. *Let A_1, A_2 be isogenous abelian schemes over $\operatorname{Spec} \mathcal{O}_{E,(p)}$. Let X_i be the associated p-divisible groups. Given any isogenies $\tilde{\phi} \colon A_1 \to A_2$ and $\phi \colon X_1 \to X_2$, we have*

$$\delta_{\mathrm{Fal},(p)}(\tilde{\phi}) = \delta_{\mathrm{Fal},(p)}(\phi). \tag{9.3.7}$$

Proof. The lemma is clear if ϕ is the isogeny associated with $\tilde{\phi}$. If $\phi' \colon X_1 \to X_2$ is another isogeny, we have $[p^N] \circ \phi = \phi' \circ \phi''$ for some isogeny $\phi'' \colon X_1 \to X_1$ (Lemma B.2.2). By additivity of $\breve{\delta}_{\mathrm{Fal}}$ and since $\breve{\delta}_{\mathrm{Fal}}([p^N]) = 0$, it is enough to show $\delta_{\mathrm{Fal},(p)}(\phi) = 0$ if $X_1 = X_2$. For this purpose, we may also assume $A_1 = A_2$.

Write $A := A_1$ and $X := X_1$ to lighten notation. As usual, A_E and X_E denote the respective generic fibers (over $\operatorname{Spec} E$). We write $\operatorname{Isog}(A)$ and $\operatorname{Isog}(X)$ for the set of self-isogenies of A and X.

We have canonical identifications

$$\operatorname{End}(X) = \operatorname{End}(X_E) = \operatorname{End}(T_p(X_E)). \tag{9.3.8}$$

The first equality holds by a theorem of Tate [Tat67a, Theorem 4] (base-change along $\operatorname{Spec} E \to \operatorname{Spec} \mathcal{O}_{E,(p)}$ is fully faithful on p-divisible groups) and the second equality holds because X_E is an étale p-divisible group. Here, the notation $\operatorname{End}(T_p(X_E))$ means endomorphisms of $T_p(X_E)$ as a Galois module.

Equip the finite $\mathbb{Z}_p$-module $\operatorname{End}(T_p(X_E))$ with the p-adic topology, and give $\operatorname{Isog}(X)$ the subspace topology. We have $\delta_{\mathrm{Fal},(p)}(\phi \circ \phi') = \delta_{\mathrm{Fal},(p)}(\phi)$ for any $\phi \in \operatorname{Isog}(X)$ and $\phi' \in \operatorname{End}(X)$ with $\phi' \equiv 1 \pmod{p}$, since any such ϕ' is an automorphism of X. The map $\operatorname{Isog}(X) \to \mathbb{R}$ given by $\phi \mapsto \delta_{\mathrm{Fal},(p)}(\phi)$ is thus locally constant.

We also have canonical identifications

$$\operatorname{End}(A) \otimes_{\mathbb{Z}} \mathbb{Z}_p = \operatorname{End}(A_E) \otimes_{\mathbb{Z}} \mathbb{Z}_p = \operatorname{End}(T_p(A_E)) = \operatorname{End}(T_p(X_E)). \tag{9.3.9}$$

The first equality holds by the Néron mapping property, and the second equality holds by Faltings's theorem [Fal86, §5 Corollary 1]. Hence $\operatorname{Isog}(A) = \operatorname{Isog}(X) \cap \operatorname{End}(A)$ is a dense subset of $\operatorname{Isog}(X)$. Since $\delta_{\mathrm{Fal},(p)}(\phi) = 0$ for any $\phi \in \operatorname{Isog}(A)$, this proves the lemma. $\square$

Corollary 9.3.2. *In the situation of Lemma 9.3.1, the quantity $\delta_{\mathrm{Fal},(p)}(\phi)$ does not depend on the choice of isogeny $\phi \colon X_1 \to X_2$.*

Proof. In the notation of the lemma, this follows immediately from $\tilde{\phi}$ independence of $\delta_{\mathrm{Fal},(p)}(\tilde{\phi})$ (discussed above). $\square$

We will use Lemma 9.3.1 to compute Faltings heights without producing isogenies on abelian varieties, only isogenies on underlying p-divisible groups over $\operatorname{Spec} \mathcal{O}_{E,(p)}$. The analogous lemma for tautological height (Lemma 9.4.5) serves a similar purpose.

9.4 Change along local isogenies: tautological

To locally decompose the change of tautological height along an isogeny, we impose an additional condition.

Definition 9.4.1.

(1) A Hermitian abelian scheme (A, ι, λ) of signature $(n-1,1)$ over E is *special* if A is $\mathcal{O}_F$-linearly isogenous to a product of elliptic curves, each with $\mathcal{O}_F$-action. A Hermitian abelian scheme of signature $(n-1,1)$ over $\operatorname{Spec} \mathcal{O}_E$ or $\operatorname{Spec} \mathcal{O}_{E,(p)}$ is *special* if its generic fiber is special.

(2) A Hermitian p-divisible group (X, ι, λ) of signature $(n-1,1)$ over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$ is *special* if X is $\mathcal{O}_{F_p}$-linearly isogenous to $\mathfrak{X}_0^{n-1} \times \mathfrak{X}_0^{\sigma}$.

We only use the term "special" this way in Part III (but we have global special cycles in mind, cf. Lemma 4.7.1). The norm $\|-\|_{\infty}$ below is as in (9.2.6).

Lemma 9.4.2. *Let $(A_1, \iota_1, \lambda_1)$ and $(A_2, \iota_2, \lambda_2)$ be special Hermitian abelian schemes of signature $(n-1,1)$ over $\operatorname{Spec} E$. For any $\mathcal{O}_F$-linear isogeny $\phi \colon A_1 \to A_2$, we have $\|\phi\|_{\infty}^2 \in \mathbb{Q}_{>0}$.*

Proof. Given such ϕ, form a diagram

$$B_1 \times B_1^{\perp} \xrightarrow{\phi_1} A_1 \xrightarrow{\phi} A_2 \xrightarrow{\phi_2} B_2 \times B_2^{\perp} \tag{9.4.1}$$

where each ϕ_i is an $\mathcal{O}_F$-linear isogeny, each B_i is a product of $(n-1)$ elliptic curves each with $\mathcal{O}_F$-action of signature $(1,0)$, and each $B_i^{\perp}$ is an elliptic curve with $\mathcal{O}_F$-action of signature $(0,1)$. Signature incompatibility implies that λ_1 pulls back to a diagonal quasi-polarization $\lambda_{B_1} \times \lambda_{B_1^{\perp}}$ on $B_1 \times B_1^{\perp}$ (e.g. $\operatorname{Hom}_{\mathcal{O}_F}(B_1^{\sigma}, B_1^{\perp \vee}) = \operatorname{Hom}_{\mathcal{O}_F}(B_1^{\perp \sigma}, B_1^{\vee}) = 0$). Similarly, λ_2 pulls back along the quasi-isogeny ϕ_2^{-1} to a diagonal quasi-polarization $\lambda_{B_2} \times \lambda_{B_2^{\perp}}$.

With these quasi-polarizations, we have $\|\phi_2 \circ \phi \circ \phi_1\|_{\infty} = \|\phi_2\|_{\infty} \|\phi\|_{\infty} \|\phi_1\|_{\infty} = \|\phi\|_{\infty}$ since $\|\phi_1\|_{\infty} = \|\phi_2\|_{\infty} = 1$ (because ϕ_1 and ϕ_2 preserve quasi-polarizations, by construction). On the other hand, if $\phi' \colon B_1^{\perp} \to B_2^{\perp}$ is the induced isogeny (signature incompatibility again implies $\operatorname{Hom}_{\mathcal{O}_F}(B_1, B_2^{\perp}) = \operatorname{Hom}_{\mathcal{O}_F}(B_1^{\perp}, B_2) = 0$), we must have $\|\phi_2 \circ \phi \circ \phi_1\|_{\infty} = \|\phi'\|_{\infty}$ (the latter norm is taken with respect to $\lambda_{B_1^{\perp}}$ and $\lambda_{B_2^{\perp}}$). For each embedding $\tau \colon E \to \mathbb{C}$, the quantity $\|\phi'\|_{\tau}^2$ must be the element of $\mathbb{Q}_{>0}$ satisfying

$$\phi'^* \lambda_{B_2^{\perp}} = \|\phi'\|_{\tau}^2 \, \lambda_{B_1^{\perp}}, \tag{9.4.2}$$

(quasi-polarizations on elliptic curves are unique up to $\mathbb{Q}_{>0}$ scalar), so we have $\|\phi'\|_{\infty}^2 = \prod_{\tau \colon E \to \mathbb{C}} \|\phi'\|_{\tau}^2 \in \mathbb{Q}_{>0}$. $\square$

For the rest of Section 9.4, we let $(A_i, \iota_i, \lambda_i)$ for $i = 1, 2$ be special Hermitian abelian schemes of signature $(n-1, 1)$ over $\operatorname{Spec} \mathcal{O}_{E,(p)}$, with associated Krämer hyperplanes $\mathcal{F}_i$ and dual tautological bundles $\mathscr{E}_i^\vee$. We also let $(X_i, \iota_i, \lambda_i)$ for $i = 1, 2, 3$ be special Hermitian p-divisible groups of signature $(n-1, 1)$ over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$, and reuse the notation $\mathcal{F}_i$ and $\mathscr{E}_i^\vee$ for the respective Krämer hyperplanes and dual tautological bundles.

Given $(X_1, \iota_1, \lambda_1)$ and an $\mathcal{O}_{F_p}$-linear isogeny $Y_1 \times Y_1^\perp \to X_1$ with Y_1 being a product of $n-1$ canonical liftings of signature $(1, 0)$ and $Y_1^\perp$ being a canonical lifting of signature $(0, 1)$, there is an induced decomposition

$$T_p(X_1)^0 = T_p(Y_1)^0 \oplus T_p(Y_1^\perp)^0 \tag{9.4.3}$$

on rational Tate modules (of the generic fibers). Equip $Y_1 \times Y_1^\perp$ with the pullback of λ_1. This gives a product quasi-polarization $\lambda_{Y_1} \times \lambda_{Y_1^\perp}$ on $Y_1 \times Y_1^\perp$ (by signature incompatibility as in the abelian scheme case, i.e. $\operatorname{Hom}_{\mathcal{O}_{F_p}}(Y_1^\sigma, Y_1^{\perp\vee}) = \operatorname{Hom}_{\mathcal{O}_{F_p}}(Y_1^{\perp\sigma}, Y_1^\vee) = 0$). Hence the decomposition in (9.4.3) is orthogonal for the Hermitian pairing on $T_p(X_1)^0$.

Consider $(X_2, \iota_2, \lambda_2)$ with $\mathcal{O}_{F_p}$-linear isogeny $Y_2 \times Y_2^\perp \to X_2$ as above, and suppose $\phi \colon X_1 \to X_2$ is an $\mathcal{O}_{F_p}$-linear isogeny. Then the induced map $\phi_* \colon T_p(X_1)^0 \to T_p(X_2)^0$ sends $T_p(Y_1)^0$ to $T_p(Y_2)^0$ and similarly for $T_p(Y_i^\perp)^0$ (again by signature incompatibility, i.e. $\operatorname{Hom}_{\mathcal{O}_{F_p}}(Y_1, Y_2^\perp) = \operatorname{Hom}_{\mathcal{O}_{F_p}}(Y_1^\perp, Y_2) = 0$). In particular, the decomposition in (9.4.3) does not depend on the choice of $Y_1 \times Y_1^\perp \to X_1$.

Any $\mathcal{O}_{F_p}$-linear isogeny $\phi \colon X_1 \to X_2$ thus gives a nonzero element $\phi \in \operatorname{Hom}_{F_p}(T_p(Y_1^\perp)^0, T_p(Y_2^\perp)^0)$. We then set

$$\|\phi\|_{\infty,p} \coloneqq \|\phi\| \tag{9.4.4}$$

where $\|-\|$ on the right means the norm for the (one-dimensional and non-degenerate) F_p-Hermitian space $\operatorname{Hom}_{F_p}(T_p(Y_1^\perp)^0, T_p(Y_2^\perp)^0)$.

We may now proceed as in the Faltings height case. Given an $\mathcal{O}_F$-linear isogeny $\phi \colon A_1 \to A_2$, we define the *semi-global change of tautological height*

$$\delta_{\mathrm{tau},(p)}(\phi) \coloneqq \frac{1}{[E : \mathbb{Q}]} \left(\log |\, \|\phi\|_\infty \,|_p + \log |\operatorname{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)|\right) \tag{9.4.5}$$

where $| - |_p$ is the usual p-adic norm (well-defined by Lemma 9.4.2). We have $\delta_{\mathrm{tau},(p)}(\phi) \in \mathbb{Q} \cdot \log p$. Since A_1 and A_2 have everywhere potentially good reduction (implied by the special hypothesis: elliptic curves with $\mathcal{O}_F$-action over number fields have everywhere potentially good reduction) the formula for change of tautological height (9.2.6) implies

$$h_{\mathrm{tau}}(A_{2,E}) - h_{\mathrm{tau}}(A_{1,E}) = \sum_\ell \delta_{\mathrm{tau},(\ell)}(\phi) = \sum_{\ell \mid \deg \phi} \delta_{\mathrm{tau},(\ell)}(\phi) \tag{9.4.6}$$

where ϕ also denotes the induced isogeny on Néron models over $\operatorname{Spec} \mathcal{O}_{E,(\ell)}$ for each prime ℓ (after enlarging E if necessary). In particular, $\delta_{\mathrm{tau},(p)}(\phi)$ does not depend on the choice of isogeny ϕ.

Given any $\mathcal{O}_{F_p}$-linear isogeny $\phi \colon X_1 \to X_2$, we define the *local change of tautological height*

$$\breve{\delta}_{\mathrm{tau}}(\phi) \coloneqq \log \|\phi\|_{\infty,p} + \frac{1}{[\breve{E} : \breve{\mathbb{Q}}_p]} \mathrm{length}_{\mathcal{O}_{\breve{E}}}(\mathrm{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)) \cdot \log p. \tag{9.4.7}$$

We have $\breve{\delta}_{\mathrm{tau}}(\phi) \in \mathbb{Q} \cdot \log p$, as well as

$$\breve{\delta}_{\mathrm{tau}}(\phi' \circ \phi) = \breve{\delta}_{\mathrm{tau}}(\phi') + \breve{\delta}_{\mathrm{tau}}(\phi) \qquad\qquad \breve{\delta}_{\mathrm{tau}}([N]) = 0 \tag{9.4.8}$$

where $\phi' \colon X_2 \to X_2$ is any $\mathcal{O}_{F_p}$-linear isogeny and $[N] \colon X_1 \to X_1$ is the multiplication-by-N isogeny. For use in later calculations, we note the identity

$$\begin{aligned} &\mathrm{length}_{\mathcal{O}_{\breve{E}}}(\mathrm{coker}(\phi_* \colon \mathrm{Lie}(X_1) \to \mathrm{Lie}(X_2))) \\ &= \mathrm{length}_{\mathcal{O}_{\breve{E}}}(\mathrm{coker}(\phi_* \colon \mathcal{F}_1 \to \mathcal{F}_2)) + \mathrm{length}_{\mathcal{O}_{\breve{E}}}(\mathrm{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)) \end{aligned} \tag{9.4.9}$$

(by the snake lemma).

Lemma 9.4.3. *If $F_p/\mathbb{Q}_p$ is nonsplit, we have $\breve{\delta}_{\mathrm{tau}}(\phi) = \breve{\delta}_{\mathrm{tau}}(\phi')$ for any two $\mathcal{O}_F$-linear isogenies $\phi, \phi' \colon X_1 \to X_2$.*

Proof. Set $X = \mathfrak{X}_0^{n-1} \times \mathfrak{X}_0^\sigma$, and equip X with any $\mathcal{O}_F$-action-compatible quasi-polarization. Select any $\mathcal{O}_F$-linear isogeny $\phi'' \colon X \to X_1$. Using the additivity property $\breve{\delta}_{\mathrm{tau}}(\phi \circ \phi'') = \breve{\delta}_{\mathrm{tau}}(\phi) + \breve{\delta}_{\mathrm{tau}}(\phi'')$ and similarly for ϕ', this reduces us to the case where $X = X_1$.

As in the proof of Lemma 9.3.1, there exists an isogeny $\phi'' \colon X \to X$ such that $[p^N] \circ \phi = \phi' \circ \phi''$ for some $N \geq 0$, so the additivity properties of $\breve{\delta}_{\mathrm{tau}}$ reduce us to showing $\breve{\delta}_{\mathrm{tau}}(\phi) = 0$ when $(X_1, \iota_1, \lambda_1) = (X_2, \iota_2, \lambda_2)$.

Since $\mathrm{Hom}_{\mathcal{O}_F}(\mathfrak{X}_0, \mathfrak{X}_0^\sigma) = \mathrm{Hom}_{\mathcal{O}_F}(\mathfrak{X}_0^\sigma, \mathfrak{X}_0) = 0$, we must have $\phi = f \times f^\perp$ where $f \colon \mathfrak{X}_0^{n-1} \to \mathfrak{X}_0$ and $f^\perp \colon \mathfrak{X}_0^\sigma \to \mathfrak{X}_0^\sigma$. We find $\breve{\delta}_{\mathrm{tau}}(\phi) = \breve{\delta}_{\mathrm{tau}}(f^\perp) = 0$ since $f^\perp \colon \mathfrak{X}_0^\sigma \to \mathfrak{X}_0^\sigma$ is an automorphism times $[p^N]$ for some $N \geq 0$. $\square$

Remark 9.4.4. If $F_p/\mathbb{Q}_p$ is split, then Lemma 9.4.3 fails (consider multiplication by $(1, p)$ and $(p, 1)$ in $\mathcal{O}_{F_p} \cong \mathbb{Z}_p \times \mathbb{Z}_p$). This is the reason for Lemma 9.4.5 below, which allows us to uniformly treat all cases of $F_p/\mathbb{Q}_p$.

Continuing to allow $F_p/\mathbb{Q}_p$ inert/ramified/split, now suppose that $(X_i, \iota_i, \lambda_i)$ is the Hermitian p-divisible group associated with $(A_i, \iota_i, \lambda_i)$, for $i = 1, 2$. Since each $(A_i, \iota_i, \lambda_i)$ is special, there automatically exists an $\mathcal{O}_F$-linear isogeny $A_1 \to A_2$ after possibly replacing E by a finite extension (by the theory of complex multiplication for elliptic curves). Given any $\mathcal{O}_{F_p}$-linear isogeny $\phi\colon X_1 \to X_2$, set

$$\delta_{\mathrm{tau},(p)}(\phi) := \frac{1}{[E:\mathbb{Q}]} \sum_{\breve{w}|p} [\breve{E}_{\breve{w}} : \breve{\mathbb{Q}}_p] \breve{\delta}_{\mathrm{tau}}(\phi_{\breve{w}}) \tag{9.4.10}$$

where $\phi_{\breve{w}}$ denotes the base-change of ϕ to $\operatorname{Spf} \mathcal{O}_{\breve{E}_{\breve{w}}}$.

Lemma 9.4.5. *Suppose that $(X_i, \iota_i, \lambda_i)$ is the Hermitian p-divisible group associated with $(A_i, \iota_i, \lambda_i)$, for $i = 1, 2$. For any $\mathcal{O}_F$-linear isogenies $\tilde{\phi}\colon A_1 \to A_2$ and $\phi\colon X_1 \to X_2$, we have*

$$\delta_{\mathrm{tau},(p)}(\tilde{\phi}) = \delta_{\mathrm{tau},(p)}(\phi). \tag{9.4.11}$$

Proof. This may be proved exactly as in Lemma 9.3.1, now requiring isogenies and endomorphisms to be $\mathcal{O}_F$-linear. □

Corollary 9.4.6. *In the situation of Lemma 9.4.5, the quantity $\delta_{\mathrm{tau},(p)}(\phi)$ does not depend on the choice of isogeny $\phi\colon X_1 \to X_2$.*

Proof. In the notation of the lemma, this follows immediately from $\tilde{\phi}$ independence of $\delta_{\mathrm{tau},(p)}(\tilde{\phi})$ (discussed above). □

9.5 Serre tensor

We compute local changes of Faltings and tautological heights for isogenies involving the Serre tensor p-divisible groups $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$. These results will later be used to compute heights of arithmetic special 1-cycles.

Given $s \in \mathbb{Z}_{\geq 0}$ and a quasi-canonical lifting $\mathfrak{X}_s$ over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$, we write $\lambda_{\mathfrak{X}_s}$ for an understood principal polarization of $\mathfrak{X}_s$. Recall that $\lambda_{\mathfrak{X}_s}$ exists and is unique up to $\mathbb{Z}_p^\times$ scalar (Lemma 5.6.1 and its proof). As in Section 5.1, we consider the map $\lambda_{\mathrm{tr}}\colon \mathcal{O}_{F_p} \to \mathcal{O}_{F_p}^*$ determined by the $\mathbb{Z}_p$-bilinear pairing $\operatorname{tr}_{F_p/\mathbb{Q}_p}(x^\sigma y)$ on $\mathcal{O}_{F_p}$, where $\mathcal{O}_{F_p}^* := \operatorname{Hom}_{\mathbb{Z}_p}(\mathcal{O}_{F_p}, \mathbb{Z}_p)$.

We equip $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$ with its Serre tensor $\mathcal{O}_{F_p}$-action ι and the polarization $-\iota(\mathfrak{d}_p^2)^{-1} \circ (\lambda_{\mathfrak{X}_s} \otimes \lambda_{\mathrm{tr}})$. We equip $\mathfrak{X}_0 \times \mathfrak{X}_0^\sigma$ with its diagonal $\mathcal{O}_{F_p}$ action $\iota_{\mathfrak{X}_0} \times \iota_{\mathfrak{X}_0}^\sigma$ (of signature $(1,1)$) and the diagonal quasi-polarization $-\iota(\mathfrak{d}_p^2)^{-1} \circ (\lambda_{\mathfrak{X}_0} \times \lambda_{\mathfrak{X}_0})$.

Lemma 9.5.1. *For the $\mathcal{O}_{F_p}$-linear isogeny*

$$\begin{aligned} \mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} &\xrightarrow{\quad\phi\quad} \mathfrak{X}_0 \times \mathfrak{X}_0^\sigma \\ x \otimes a &\longmapsto (\iota_{\mathfrak{X}_0}(a)x, \iota_{\mathfrak{X}_0}(a^\sigma)x) \end{aligned} \tag{9.5.1}$$

we have $\breve{\delta}_{\mathrm{Fal}}(\phi) = 0$. Assuming $p \neq 2$ if $F_p/\mathbb{Q}_p$ is ramified, we also have $\breve{\delta}_{\mathrm{tau}}(\phi) = 0$.

Proof. We already know $\deg \phi = |\Delta|_p^{-1}$ (see (5.1.3) and surrounding discussion).

Pick any $\mathcal{O}_{F_p}$-linear isomorphism $\operatorname{Lie} \mathfrak{X}_0 \cong \mathcal{O}_{\breve{E}}$. Then the map $\phi_* \colon \operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \to \operatorname{Lie}(\mathfrak{X}_0 \times \mathfrak{X}_0^\sigma)$ may be identified with the map of $\mathcal{O}_{\breve{E}}$-modules $f \colon \mathcal{O}_{\breve{E}} \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} \to \mathcal{O}_{\breve{E}} \oplus \mathcal{O}_{\breve{E}}$ given by $f(x \otimes a) = (ax, a^\sigma x)$. Thus ϕ_* is given by the matrix in (5.1.4) (the same matrix describing ϕ after identifying $\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} \cong \mathfrak{X}_0^2$ using a $\mathbb{Z}_p$-basis of $\mathcal{O}_{F_p}$). That matrix has determinant which generates the different ideal $\mathfrak{d}_p$, hence

$$\operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\phi_* \operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \to \operatorname{Lie}(\mathfrak{X}_0 \times \mathfrak{X}_0^\sigma))) = \frac{1}{2}[\breve{E} : \breve{\mathbb{Q}}_p] v_p(\Delta). \tag{9.5.2}$$

This gives $2\breve{\delta}_{\mathrm{Fal}}(\phi) = \log \deg \phi - v_p(\Delta) \log p = 0$.

We also know that $\phi^*(\lambda_{\mathfrak{X}_0} \times \lambda_{\mathfrak{X}_0}) = \lambda_{\mathfrak{X}_0} \otimes \lambda_{\mathrm{tr}}$ (see discussion surrounding (5.1.3) again). Thus $\|\phi\|_{\infty,p} = 1$, in the notation of (9.4.4).

Let $\mathcal{F}_1 \subseteq \operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ and $\mathcal{F}_2 \subseteq \operatorname{Lie}(\mathfrak{X}_0 \times \mathfrak{X}_0^\sigma)$ be the (unique) associated Krämer hyperplanes, with associated Krämer hyperplane quotients $\mathscr{E}_1^\vee$ and $\mathscr{E}_2^\vee$. If $F_p/\mathbb{Q}_p$ is unramified, then ϕ is an isomorphism, hence $\operatorname{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee) = 0$. If $F_p/\mathbb{Q}_p$ is ramified, assume $p \neq 2$ and select a uniformizer $\varpi \in \mathcal{O}_{F_p}$ satisfying $\varpi^\sigma = -\varpi$. Then $(\varpi \otimes 1 + 1 \otimes \varpi) \in \mathcal{O}_{\breve{E}} \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$ is a generator of $\mathcal{F}_1$. We thus find $\operatorname{coker}(\phi_* \colon \mathcal{F}_1 \to \mathcal{F}_2) \cong \mathcal{O}_{\breve{E}}/\varpi\mathcal{O}_{\breve{E}}$.

By (9.4.9), the previous computations imply $\operatorname{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee) = 0$, and hence

$$\breve{\delta}_{\mathrm{tau}}(\phi) = \log \|\phi\|_{\infty,p} + \frac{1}{[\breve{E} : \breve{\mathbb{Q}}_p]} \operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)) \cdot \log p = 0. \tag{9.5.3}$$

$\square$

For any given integer $s \in \mathbb{Z}_{\geq 0}$, recall the constants $\delta_{\mathrm{tau}}(s), \delta_{\mathrm{Fal}}(s) \in \mathbb{Q}$ ("local change of 'tautological' and Faltings heights") as defined in Eq. (7.2.7) and surrounding text. Given any isogeny $\psi_s \colon \mathfrak{X}_0 \to \mathfrak{X}_s$ of minimal degree (i.e. degree p^s), the discussion surrounding Eq. (7.2.7) implies $\breve{\delta}_{\mathrm{tau}}(\psi_s) = \delta_{\mathrm{tau}}(s) \cdot \log p$.

Lemma 9.5.2. *Let $\psi_s \colon \mathfrak{X}_0 \to \mathfrak{X}_s$ be any isogeny of degree p^s. For the $\mathcal{O}_{F_p}$-linear isogeny*

$$\begin{aligned} \mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} &\xrightarrow{\phi} \mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} \\ x \otimes a &\longmapsto \psi_s(x) \otimes a \end{aligned} \tag{9.5.4}$$

we have

$$\breve{\delta}_{\mathrm{Fal}}(\phi) = -2\breve{\delta}_{\mathrm{tau}}(\phi) = -2\delta_{\mathrm{tau}}(s) \cdot \log p. \tag{9.5.5}$$

Proof. Recall that ψ_s is unique up to pre-composition by elements of $\mathcal{O}_{F_p}^\times$ (7.2.4). Write $\mathcal{F}_1$ and $\mathcal{F}_2$ (resp. $\mathscr{E}_1^\vee$ and $\mathscr{E}_2^\vee$) for the associated Krämer hyperplanes (resp. dual tautological bundles) of $\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$ and $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$ respectively.

We have $\deg \phi = (\deg \psi_s)^2 = p^{2s}$. Since quasi-polarizations on $\mathfrak{X}_0$ are unique up to $\mathbb{Q}_p^\times$ scalar (follows from Drinfeld rigidity and the corresponding statement for $\mathbf{X}_0$ in Section 5.1), we have $\psi_s^* \lambda_{\mathfrak{X}_s} = b \lambda_{\mathfrak{X}_0}$ for some $b \in p^s \mathbb{Z}_p^\times$. Hence we have $\phi^*(\lambda_{\mathfrak{X}_s} \otimes \lambda_{\mathrm{tr}}) = b(\lambda_{\mathfrak{X}_0} \otimes \lambda_{\mathrm{tr}})$, so $\|\phi\|_{\infty,p} = p^{-s/2}$.

Pick any identifications $\operatorname{Lie} \mathfrak{X}_0 \cong \operatorname{Lie} \mathfrak{X}_s \cong \mathcal{O}_{\breve{E}}$ of $\mathcal{O}_{\breve{E}}$-modules. With these identifications, the map $\psi_{s,*} \colon \operatorname{Lie} \mathfrak{X}_0 \to \operatorname{Lie} \mathfrak{X}_s$ is multiplication by some $c \in \mathcal{O}_{\breve{E}}$ satisfying $[\breve{E} : \breve{\mathbb{Q}}_p] v_p(c) = \operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\psi_{s,*} \colon \operatorname{Lie} \mathfrak{X}_0 \to \operatorname{Lie} \mathfrak{X}_s))$.

We also obtain identifications $\operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \cong \operatorname{Lie}(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ of $\mathcal{O}_{\breve{E}} \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$-modules, with induced identifications $\mathcal{F}_1 \cong \mathcal{F}_2$ and $\mathscr{E}_1^\vee \cong \mathscr{E}_2^\vee$. Then $\phi_* \colon \operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \to \operatorname{Lie}(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ is identified with multiplication by c, and hence $\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee$ must also be multiplication by c. Hence

$$\operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\phi_* \colon \operatorname{Lie}(\mathfrak{X}_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \to \operatorname{Lie}(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}))) = 2[\breve{E} : \breve{\mathbb{Q}}_p] v_p(c) \tag{9.5.6}$$

$$\operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\phi_* \colon \mathscr{E}_1^\vee \to \mathscr{E}_2^\vee)) = [\breve{E} : \breve{\mathbb{Q}}_p] v_p(c). \tag{9.5.7}$$

The lemma now follows from the formula for $\operatorname{length}_{\mathcal{O}_{\breve{E}}}(\operatorname{coker}(\psi_{s,*} \colon \operatorname{Lie} \mathfrak{X}_0 \to \operatorname{Lie} \mathfrak{X}_s))$ in (7.2.6). $\square$

10 Heights and quasi-canonical liftings

10.1 A descent lemma

To compute Faltings and tautological heights, we will produce isogenies of p-divisible groups over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ from isogenies over $\operatorname{Spf} \mathcal{O}_{\breve{E}_{\breve{w}}}$ for any choice of $\breve{w} \mid p$. We now explain this descent procedure, in a more general setup.

Lemma 10.1.1. *Let $S' \to S$ be a morphism of schemes whose scheme-theoretic image is all of S. Suppose X is a p-divisible group over S which satisfies $\operatorname{End}^0(X) = \operatorname{End}^0(X_{S'})$. Let Y and Z be p-divisible groups over S which are isogenous to X. The base-change maps*

$$\begin{aligned} &\operatorname{Hom}^0(Y, Z) \to \operatorname{Hom}^0(Y_{S'}, Z_{S'}) && \operatorname{Hom}(Y, Z) \to \operatorname{Hom}(Y_{S'}, Z_{S'}) \\ &\operatorname{Isog}^0(Y, Z) \to \operatorname{Isog}^0(Y_{S'}, Z_{S'}) && \operatorname{Isog}(Y, Z) \to \operatorname{Isog}(Y_{S'}, Z_{S'}) \end{aligned}$$

are bijections.

Proof. Choose isogenies $\phi_Y \colon X \to Y$ and $\phi_Z \colon X \to Z$. There is a commutative diagram

$$\begin{array}{ccc} \alpha & \mathrm{End}^0(X) \xrightarrow{\ \sim\ } \mathrm{End}^0(X_{S'}) & \alpha' \\ \downarrow & \wr\downarrow \qquad\qquad\qquad \downarrow\wr & \downarrow \\ \phi_Z \circ \alpha \circ \phi_Y^{-1} & \mathrm{Hom}^0(Y, Z) \longrightarrow \mathrm{Hom}^0(Y_{S'}, Z_{S'}) & \phi_{Z,S'} \circ \alpha' \circ \phi_{Y,S'}^{-1} \end{array}$$

where horizontal arrows are base-change. The vertical arrows are isomorphisms, and the upper horizontal arrow is an isomorphism by hypothesis. Hence the bottom arrow is an isomorphism. Suppose $\beta \in \mathrm{Hom}^0(Y, Z)$ is any quasi-homomorphism. The functor $T \mapsto \{\phi \in \mathrm{Hom}(T, S) : \phi^*\beta \text{ is a homomorphism}\}$ is represented by a closed subscheme of T, see [RZ96, Proposition 2.9]. If $\beta|_{S'}$ is a homomorphism, then β must also be a homomorphism, since the smallest closed subscheme of S through which S' factors is all of S (by hypothesis). Hence $\mathrm{Hom}(Y, Z) \to \mathrm{Hom}(Y_{S'}, Z_{S'})$ is an isomorphism. The statements about (quasi-)isogenies follow from an essentially identical argument, replacing End and Hom with Isog, and noting $\mathrm{Isog}^0(X) = (\mathrm{End}^0(X))^\times$ (e.g. by Lemma B.2.3). $\square$

Remark 10.1.2. We will be interested in the case where $S = \operatorname{Spec} \mathcal{O}_{E,(p)}$ and $S' = \operatorname{Spec} \mathcal{O}_{\breve{E}}$ for some finite extension $\breve{E}$ of $\breve{E}_{\breve{w}}$ for some $\breve{w} \mid p$. In this case, Lemma 10.1.1 admits an alternative proof: a quasi-homomorphism of p-divisible groups over $\operatorname{Spec} E$ is a homomorphism if and only if the map on rational Tate modules preserves (integral) Tate modules, and this can be checked after base-change to $\operatorname{Spec} \breve{E}$. Then apply the theorem of Tate [Tat67a, Theorem 4] which states that the generic fiber functor for p-divisible groups over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ (similarly, for $\operatorname{Spec} \mathcal{O}_{\breve{E}}$) is fully faithful.

Lemma 10.1.3. *Let X be a p-divisible group over a formal scheme S. Suppose there is a decomposition $X = X_1 \times X_2$ as fppf sheaves of abelian groups (on $(Sch/S)_{fppf}$). Then X_1 and X_2 are both p-divisible groups.*

Proof. Write $e_1, e_2 \in \mathrm{End}(X)$ for the projections to X_1 and X_2 respectively. As being a p-divisible group can be checked locally on $(Sch/S)_{fppf}$, assume S is a usual scheme.

It is clear that the multiplication by p map $[p] \colon X \to X$ is a surjection if and only if $[p] \colon X_1 \to X_1$ and $[p] \colon X_2 \to X_2$ are surjections. We also have $X[p^n] = X_1[p^n] \times X_2[p^n]$ for all $n \geq 1$. Thus the natural map $\varinjlim X[p^n] \to X$ is an isomorphism if and only if $\varinjlim X_1[p^n] \to X_1$ and $\varinjlim X_2[p^n] \to X_2$ are isomorphisms.

Next, note $X_1[p] = \ker(e_2 \colon X[p] \to X[p])$ and similarly $X_2[p] = \ker(e_1 \colon X[p] \to X[p])$. Since $X[p]$ is representable by a finite locally free scheme over S, we conclude that $X_1[p]$ and

$X_2[p]$ are represented by schemes which are finite and finitely presented over S. We also have short exact sequences

$$0 \to X_1[p] \to X[p] \to X_2[p] \to 0$$
$$0 \to X_2[p] \to X[p] \to X_1[p] \to 0$$

so Lemma B.2.1 implies that $X_1[p]$ and $X_2[p]$ are finite locally free over S. $\square$

Corollary 10.1.4. *Let $S' \to S$ and X be as in Lemma 10.1.1. Suppose Y and Z are p-divisible groups over S isogenous to X.*

If $Y_{S'} = Y_1' \times \cdots \times Y_r'$ and $Z_{S'} = Z_1' \times \cdots \times Z_r'$ for p-divisible groups Y_i' and Z_i' over S', then there are unique decompositions $Y = Y_1 \times \cdots \times Y_r$ and $Z = Z_1 \times \cdots \times Z_r$ such that $Y_i|_{S'} = Y_i'$ and $Z_i|_{S'} = Z_i'$ for all i. For any i, the base-change maps

$$\operatorname{Hom}^0(Y_i, Z_i) \to \operatorname{Hom}^0(Y_{i,S'}, Z_{i,S'}) \qquad \operatorname{Hom}(Y_i, Z_i) \to \operatorname{Hom}(Y_{i,S'}, Z_{i,S'})$$
$$\operatorname{Isog}^0(Y_i, Z_i) \to \operatorname{Isog}^0(Y_{i,S'}, Z_{i,S'}) \qquad \operatorname{Isog}(Y_i, Z_i) \to \operatorname{Isog}(Y_{i,S'}, Z_{i,S'})$$

are bijective.

Proof. The decomposition $Y_{S'} = Y_1' \times \cdots \times Y_r'$ corresponds to a system of orthogonal idempotents $d_1', \dots, d_r' \in \operatorname{End}(Y_{S'})$, i.e. $d_i'^2 = d_i'$ for all i and $d_i'd_j' = 0$ for all $i \neq j$. Lifting to a decomposition $Y = Y_1 \times \cdots \times Y_r$ is the same as lifting $\{d_i'\}_i$ to a system of orthogonal idempotents $\{d_i\}_i$ in $\operatorname{End}(Y)$. Such a lift exists and is unique by Lemma 10.1.1. The same applies for Z, and we write $\{e_i'\}_i$ and $\{e_i\}_i$ for the corresponding systems of idempotents. Using Lemma 10.1.1, we have

$$\operatorname{Hom}^0(Y_i, Z_i) = d_i \operatorname{Hom}^0(Y, Z)e_i = d_i' \operatorname{Hom}^0(Y_{S'}, Z_{S'})e_i' = \operatorname{Hom}^0(Y_{i,S'}, Z_{i,S'})$$
$$\operatorname{Hom}(Y_i, Z_i) = d_i \operatorname{Hom}(Y, Z)e_i = d_i' \operatorname{Hom}(Y_{S'}, Z_{S'})e_i' = \operatorname{Hom}(Y_{i,S'}, Z_{i,S'}).$$

The statement about Isog^0 then follows from Lemma B.2.3, and the statement about Isog follows from the relation $\operatorname{Isog}(-,-) = \operatorname{Isog}^0(-,-) \cap \operatorname{Hom}(-,-)$. $\square$

10.2 Minimal isogenies

Given any abelian scheme $A \to S$ over some base S, we can form the *Serre tensor* abelian scheme $A \otimes_{\mathbb{Z}} \mathcal{O}_F$ given by $(A \otimes_{\mathbb{Z}} \mathcal{O}_F)(T) := A(T) \otimes_{\mathbb{Z}} \mathcal{O}_F$ for S-schemes T. There is a natural action of $\mathcal{O}_F$ on $A \otimes_{\mathbb{Z}} \mathcal{O}_F$, as we have discussed for p-divisible groups (B.1.1). If $\lambda\colon A \to A^\vee$ is a quasi-polarization, then $\lambda \otimes \lambda_{\mathrm{tr}}\colon A \otimes_{\mathbb{Z}} \mathcal{O}_F \to A^\vee \otimes_{\mathbb{Z}} \mathcal{O}_F^\vee \cong (A \otimes_{\mathbb{Z}} \mathcal{O}_F)^\vee$ is a polarization, where $\lambda_{\mathrm{tr}}\colon \mathcal{O}_F \to \mathcal{O}_F^\vee$ is induced by the trace pairing, as above.

Let $A_0 \to \operatorname{Spec} \mathcal{O}_{E,(p)}$ be any (relative) elliptic curve with $\mathcal{O}_F$-action ι_0 of signature $(1,0)$, and let λ_0 be the unique principal polarization of A_0. For $n \geq 2$, set

$$A := A_0^{n-2} \times (A_0 \otimes_{\mathbb{Z}} \mathcal{O}_F) \tag{10.2.1}$$

with $\mathcal{O}_F$ action ι which is diagonal on A_0^{n-2} and the Serre tensor action on $A_0 \otimes_{\mathbb{Z}} \mathcal{O}_F$, and polarization $\lambda_0^{n-2} \times (|\Delta|^{-1}(\lambda_0 \otimes \lambda_{\mathrm{tr}}))$. Then (A, ι, λ) is a special Hermitian abelian scheme of signature $(n-1, 1)$. We write (X, ι, λ) for the associated special Hermitian p-divisible group of signature $(n-1,1)$, with

$$X = X_0^{n-2} \times (X_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \tag{10.2.2}$$

where X_0 is the p-divisible group of A_0. For any $\breve{w} \mid p$, the base-change $X_{0,\breve{w}}$ is a canonical lifting. The preceding notation (e.g. for A_0 and X_0) will be fixed for all of Section 10.2.

In Proposition 10.2.1 and Corollary 10.2.2 below, we equip $\mathfrak{X}_0^{n-2} \times (\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ with the diagonal $\mathcal{O}_{F_p}$ action (which is the Serre tensor action on $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$) and a product quasi-polarization, for some quasi-polarization of $\mathfrak{X}_0^{n-2}$ and the quasi-polarization $-\iota(\mathfrak{d}_p^2)^{-1} \circ (\lambda_{\mathfrak{X}_s} \otimes \lambda_{\mathrm{tr}})$ on $(\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$.

Proposition 10.2.1. *Let (A', ι', λ') be a special Hermitian abelian scheme of signature $(n-1,1)$ over $\operatorname{Spec} \mathcal{O}_{E,(p)}$, with associated Hermitian p-divisible group (X', ι', λ'). Replace E with a finite extension if necessary, so that A and A' are $\mathcal{O}_F$-linearly isogenous.*

Suppose there exists an $\mathcal{O}_{F_p}$-linear quasi-polarization preserving isomorphism

$$X'_{\operatorname{Spf} \mathcal{O}_{\breve{E}}} \cong \mathfrak{X}_0^{n-2} \times (\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \tag{10.2.3}$$

over $\operatorname{Spf} \mathcal{O}_{\breve{E}}$, where $\breve{E}$ is a finite extension of $\mathcal{O}_{\breve{E}_{\breve{w}'}}$, for some $\breve{w}' \mid p$ and $s \geq 0$. Fix an isomorphism $X_{0,\operatorname{Spf} \mathcal{O}_{\breve{E}}} \cong \mathfrak{X}_0$.

(1) Then there exists an $\mathcal{O}_{F_p}$-linear quasi-polarization preserving isomorphism

$$X' \cong X_0^{n-2} \times (X_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \tag{10.2.4}$$

over $\operatorname{Spec} \mathcal{O}_{E,(p)}$, for some p-divisible group X_s of height 2 and dimension 1 with fixed identification $X_{s,\operatorname{Spf} \mathcal{O}_{\breve{E}}} \cong \mathfrak{X}_s$, such that (10.2.4) recovers (10.2.3) upon base-change to $\operatorname{Spf} \mathcal{O}_{\breve{E}}$. On the right-hand side of (10.2.4), the polarization is the product of a polarization on X_0^{n-2} and a quasi-polarization $-(\mathfrak{d}_p^2)^{-1} \cdot (\lambda_s \otimes_{\mathbb{Z}_p} \lambda_{\mathrm{tr}})$ on $X_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}$, where λ_s is a principal polarization on X_s.

(2) For any $\breve{w} \mid p$, the base-change $X_{s,\breve{w}}$ is a quasi-canonical lifting of level s, and hence there is an identification as in (10.2.4) for all $\breve{w} \mid p$.

(3) *There exists an isogeny $\psi_s \colon X_0 \to X_s$ of degree p^s. The $\mathcal{O}_{F_p}$-linear product isogeny $\phi \colon X \to X'$ given by*

$$\phi \coloneqq \mathrm{id}_{X_0^{n-2}} \times (\psi_s \otimes 1) \colon X_0^{n-2} \times (X_0 \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \to X_0^{n-2} \times (X_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \tag{10.2.5}$$

over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ *satisfies*

$$\breve{\delta}_{\mathrm{Fal}}(\phi_{\breve{w}}) = -2\breve{\delta}_{\mathrm{tau}}(\phi_{\breve{w}}) = -2\delta_{\mathrm{tau}}(s) \cdot \log p \qquad \text{for all } \breve{w} \mid p \tag{10.2.6}$$

$$\delta_{\mathrm{Fal},(p)}(\phi) = -2\delta_{\mathrm{tau},(p)}(\phi) = -2\delta_{\mathrm{tau}}(s) \cdot \log p. \tag{10.2.7}$$

Proof. Note that X satisfies the hypotheses of Lemma 10.1.1 with $S = \operatorname{Spec} \mathcal{O}_{E,(p)}$ and $S' = \operatorname{Spec} \mathcal{O}_{\breve{E}_{\breve{w}}}$ for any $\breve{w} \mid p$, as $\operatorname{End}(X) \cong M_{n,n}(\mathcal{O}_{F_p})$ over both $\operatorname{Spec} \mathcal{O}_{E,(p)}$ and $\operatorname{Spf} \mathcal{O}_{\breve{w}}$ for any $\breve{w}$. The same holds for $S' = \operatorname{Spec} \mathcal{O}_{\breve{E}}$. Again, we pass between $\operatorname{Spf} \mathcal{O}_{\breve{E}_{\breve{w}}}$ and $\operatorname{Spec} \mathcal{O}_{\breve{E}_{\breve{w}}}$ as in Appendix B.3.

The proposition then follows from repeated applications of Lemma 10.1.1 and Corollary 10.1.4, as we now explain.

(1) and (2) Corollary 10.1.4 implies that (10.2.3) descends to a $\mathcal{O}_{F_p}$-linear product decomposition $X' \cong X_0'^{n-2} \times (X_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ for some X_0' descending $\mathfrak{X}_0$ (first pick any identification of p-divisible groups $\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p} \cong \mathfrak{X}_s^2$, then descend the $\mathcal{O}_{F_p}$-action), and the fully-faithfulness in Corollary 10.1.4 implies $\operatorname{End}(X_s) = \mathcal{O}_{F_p,s}$ (with $\mathcal{O}_{F_p,s} = \mathbb{Z}_p + p^s \mathcal{O}_{F_p}$ as in Section 7.2) over $\operatorname{Spec} \mathcal{O}_{E,(p)}$ and also over $\operatorname{Spf} \mathcal{O}_{\breve{w}}$ for any $\breve{w} \mid p$. The fully-faithfulness in Corollary 10.1.4 also implies that the fixed $\mathcal{O}_{F_p}$-linear isomorphism $X_{0,\breve{w}} \to \mathfrak{X}_0$ lifts to an isomorphism $X_0 \to X_0'$. The polarization on $\mathfrak{X}_0^{n-2} \times (\mathfrak{X}_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ descends to $X_0^{n-2} \times (X_s \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p})$ by Corollary 10.1.4 again (applied to X' and $X'^{\vee}$; note that the property of being a polarization is represented by a closed subfunctor of $\operatorname{Spec} \mathcal{O}_{E,(p)}$, hence can be checked in the generic fiber or over $\operatorname{Spec} \breve{E}$)

(3) If $\psi_s \colon \mathfrak{X}_0 \to \mathfrak{X}_s$ is any isogeny of degree p^s (exists and is unique up to precomposition by $\mathcal{O}_{F_p}^{\times}$, as discussed in Section 7.2), we apply Corollary 10.1.4 to descend to an isogeny $\psi_s \colon X_0 \to X_s$ of degree p^s. Equation (10.2.6) now follows from Lemma 9.5.2. Equation (10.2.7) follows from this (by the definitions in (9.3.6) and (9.4.10)). □

We will use the following reformulation (tailored to our intended application for global heights via local special cycles). In the corollary statement and proof, A_0^{σ} and $A_0^{n-1} \times A_0^{\sigma}$ are equipped with the product quasi-polarizations $-\Delta^{-1}\lambda_0$ and $-\Delta^{-1}(\lambda_0 \times \cdots \times \lambda_0)$ (where $A_0^{\sigma} = A_0$ but with $\mathcal{O}_F$-action $\iota \circ \sigma$, as above).

Corollary 10.2.2. *Let S be a reduced scheme which is finite flat over $\operatorname{Spec} \mathcal{O}_F$. Let $(A', \iota', \lambda', \mathcal{F}')$ be a quasi-polarized Krämer datum over S (of signature $(n-1,1)$) for $n \geq 2$, with associated metrized line bundles $\widehat{\omega}$ and $\widehat{\mathscr{E}}^\vee$ on S. Assume that (A', ι', λ') is special at all generic points of S. Let (X', ι', λ') be the associated Hermitian p-divisible group.*

Suppose we are given a finite étale surjection

$$\coprod_j \mathcal{Z}_j \to S \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \breve{\mathbb{Z}}_p \tag{10.2.8}$$

such that each restricted map $\Theta_j \colon \mathcal{Z}_j \to S \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \breve{\mathbb{Z}}_p$ has constant degree $\deg(j)$ onto its image. Assume that Θ_j and $\Theta_{j'}$ have disjoint images for $j \neq j'$.

For each irreducible component $\mathcal{Z} \hookrightarrow \coprod_j \mathcal{Z}_j$, write $\breve{E}_{\mathcal{Z}}$ for the residue field of its generic point. Assume there exists an isomorphism of Hermitian p-divisible groups

$$X'|_{\operatorname{Spf} \mathcal{O}_{\breve{E}_{\mathcal{Z}}}} \cong \mathfrak{X}_0^{n-2} \times (\mathfrak{X}_{s_{\mathcal{Z}}} \otimes_{\mathbb{Z}_p} \mathcal{O}_{F_p}) \tag{10.2.9}$$

for all $\mathcal{Z}$, where $s_{\mathcal{Z}} \in \mathbb{Z}_{\geq 0}$ is an integer depending on $\mathcal{Z}$.

We then have

$$\widehat{\deg}(\widehat{\mathscr{E}}^\vee) - (\deg_{\mathbb{Z}} S) \cdot h_{\mathrm{tau}}(A_0^\sigma) = \sum_{j \in J} \frac{1}{\deg(j)} \sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}_j} (\deg_{\breve{\mathbb{Z}}_p} \mathcal{Z}) \delta_{\mathrm{tau}}(s_{\mathcal{Z}}) \log p \tag{10.2.10}$$
$$\widehat{\deg}(\widehat{\omega}) - (\deg_{\mathbb{Z}} S) \cdot h_{\mathrm{Fal}}(A_0^{n-1} \times A_0^\sigma) = \sum_{j \in J} \frac{1}{\deg(j)} \sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}_j} (\deg_{\breve{\mathbb{Z}}_p} \mathcal{Z}) \delta_{\mathrm{Fal}}(s_{\mathcal{Z}}) \log p$$

modulo $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$, where the inner sums run over all irreducible components $\mathcal{Z} \hookrightarrow \mathcal{Z}_j$.

Proof. In the corollary statement, the expression "modulo $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$" means an equality of elements in the additive quotient $\mathbb{R}/(\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell)$. The notation $\deg_{\mathbb{Z}} S$ (resp. $\deg_{\breve{\mathbb{Z}}_p} \mathcal{Z}$) denotes the degree of $S \to \operatorname{Spec} \mathbb{Q}$ (resp. $\mathcal{Z} \to \operatorname{Spec} \breve{\mathbb{Z}}_p$) in the generic fiber.

By additivity, we immediately reduce to the case where S is irreducible. Then J consists of a single element j. By normalization, we may assume $S = \operatorname{Spec} \mathcal{O}_E$ for a number field E. We may also enlarge E as necessary so that (A_0, ι, λ_0) also extends to $\operatorname{Spec} \mathcal{O}_E$, and such that there exists an $\mathcal{O}_F$-linear isogeny $\phi \colon A_0^{n-2} \times (A_0 \otimes_{\mathbb{Z}} \mathcal{O}_F) \to A$. We also consider the $\mathcal{O}_F$-linear isogeny $\phi' \colon A_0^{n-2} \times (A_0 \otimes_{\mathbb{Z}} \mathcal{O}_F) \to A_0^{n-2} \times (A_0 \times A_0^\sigma)$ which is the identity on A_0^{n-2} and given by $(x \otimes a) \mapsto (ax, a^\sigma x)$ for $(A_0 \otimes_{\mathbb{Z}} \mathcal{O}_F) \to A_0 \times A_0^\sigma$.

Since $\delta_{\mathrm{tau}}(\phi') = 0$ (Lemma 9.5.1, along with the local decomposition (9.4.10), also Lemma 9.4.5), the decomposition in (9.4.6) shows

$$\widehat{\deg}(\widehat{\mathscr{E}}^\vee) - [E : \mathbb{Q}] \cdot h_{\mathrm{tau}}(A_0^{n-1} \times A_0^\sigma) = [E : \mathbb{Q}] \delta_{\mathrm{tau},(p)}(\phi) \mod \sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell. \tag{10.2.11}$$

We have $\sum_{\mathcal{Z} \hookrightarrow \mathcal{Z}_j} (\deg_{\breve{\mathbb{Z}}_p} \mathcal{Z}) = \deg(j) \cdot [E : \mathbb{Q}]$. Applying Proposition 10.2.1(3) (combined with the "isogeny independence" result of Lemma 9.4.5) now shows $\delta_{\mathrm{tau},(p)}(\phi) = \delta_{\mathrm{tau}}(s_{\mathcal{Z}}) \log p$ for any $\mathcal{Z} \hookrightarrow \mathcal{Z}_j$. This also shows that all $s_{\mathcal{Z}}$ are equal (when S is irreducible): the quantity $\delta_{\mathrm{tau}}(s)$ takes distinct values for distinct $s \in \mathbb{Z}_{\geq 0}$ (in the nonsplit cases, note $\delta_{\mathrm{tau}}(s)$ has strictly decreasing p-adic valuation as s increases, for $s > 2$). We also have $h_{\mathrm{tau}}(A_0^{n-1} \times A_0^{\sigma}) = h_{\mathrm{tau}}(A_0^{\sigma})$ (straightforward from the definition). This verifies (10.2.10) for $\widehat{\deg}(\widehat{\mathscr{E}}^{\vee})$ and the tautological height.

Since $\delta_{\mathrm{Fal}}(\phi') = 0$ (Lemma 9.5.1, along with the local decomposition (9.3.6), also Lemma 9.3.1), (9.4.6) similarly shows

$$\widehat{\deg}(\widehat{\omega}) - [E : \mathbb{Q}] \cdot h_{\mathrm{Fal}}(A_0^{n-1} \times A_0^{\sigma}) = [E : \mathbb{Q}] \delta_{\mathrm{Fal},(p)}(\phi) \quad \bmod \sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell. \tag{10.2.12}$$

Applying Proposition 10.2.1(3) (combined with the "isogeny independence" result of Lemma 9.3.1) verifies (10.2.10) for $\widehat{\deg}(\widehat{\omega})$ and the Faltings height, just as for tautological height above. □

In the situation above, we have

$$h_{\mathrm{tau}}(A_0^{\sigma}) = h_{\mathrm{tau}}^{\mathrm{CM}} \qquad h_{\mathrm{Fal}}(A_0^{n-1} \times A_0^{\sigma}) = n \cdot h_{\mathrm{Fal}}^{\mathrm{CM}} \tag{10.2.13}$$

in the notation of (4.3.6) and (4.3.5).

Remark 10.2.3. In Proposition 10.2.1(3), it was important that ψ_s was an isogeny of minimal degree p^s. If ψ_s were replaced by an arbitrary isogeny $f \colon \mathfrak{X}_0 \to \mathfrak{X}_s$, we would not be able to determine $\breve{\delta}_{\mathrm{Fal}}(f)$ or $\breve{\delta}_{\mathrm{tau}}(f)$ using only $\deg f$ in the case when $F_p/\mathbb{Q}_p$ is split (due to Remark 9.4.4).

Part IV

Uniformization

We use global notation as in Part I, e.g. F is an imaginary quadratic field extension of $\mathbb{Q}$ with nontrivial involution $a \mapsto a^\sigma$ and discriminant Δ. The notation $\mathbb{A}_f$ (resp. $\mathbb{A}_f^p$) will always denote the finite adèle ring (resp. finite adèle ring away from p) for $\mathbb{Q}$.

For all of Part IV, let $L_0 := \mathcal{O}_F$ be the rank one Hermitian $\mathcal{O}_F$-lattice with pairing $(x, y) := x^\sigma y$. Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n and signature $(n-r, r)$, with associated moduli stack $\mathcal{M}$ (Definition 3.1.2 and Section 3.2).

We fix some group-theoretic setup (common in the literature, e.g. [RSZ20; BHKRY20]). Set

$$V_0 := L_0 \otimes_{\mathcal{O}_F} F \qquad V := L \otimes_{\mathcal{O}_F} F$$
$$G' := \{(g_0, g) \in GU(V_0) \times GU(V) : c(g_0) = c(g)\} \subseteq GU(V_0) \times GU(V)$$

where $c \colon GU(V_0) \to \mathbb{G}_m$ and $c \colon GU(V) \to \mathbb{G}_m$ are the similitude characters. We use the shorthand

$$L_p := L \otimes_{\mathbb{Z}} \mathbb{Z}_p \qquad V_p := V \otimes_{\mathbb{Q}} \mathbb{Q}_p \qquad V_{\mathbb{R}} := V \otimes_{\mathbb{Q}} \mathbb{R}$$

and use similar notation for local versions of other Hermitian spaces. Given a tuple $\underline{x} \in V^m$, we write $\underline{x}_p \in V_p^m$ and $\underline{x}_\infty \in V_{\mathbb{R}}^m$ and $\underline{x}_f \in (V \otimes_{\mathbb{Q}} \mathbb{A}_f)^m$ and $\underline{x}^p \in (V \otimes_{\mathbb{Q}} \mathbb{A}_f^p)^m$ for the corresponding projections (and similarly for other Hermitian spaces).

There is an isomorphism

$$\begin{aligned} G' &\longrightarrow GU(V_0) \times U(V) \\ (g_0, g) &\longmapsto (g_0, g_0^{-1} g). \end{aligned} \tag{$*$}$$

To avoid potential confusion: whenever we write $(g_0, g) \in G'$, we mean $g_0 \in GU(V_0)$ and $g \in GU(V)$ with the same similitude factor.

We use factorizable open compact subgroups $K'_f = K_{0,f} \times K_f \subseteq G'(\mathbb{A}_f)$ as in Section 3.4, where $K_{0,f} \subseteq GU(V_0)(\mathbb{A}_f)$ and $K_f \subseteq U(V)(\mathbb{A}_f)$ (using also $(*)$).

Recall the moduli stack with level structure $\mathcal{M}_{K'_f}$ defined in Section 3.4. We do not require K'_f to be a small level, so $\mathcal{M}_{K'_f}$ is allowed to be a stack.

Notation. In Part IV, we implicitly fix an open compact subgroup $K'_f \subseteq G'(\mathbb{A}_f)$ as above. We abusively suppress K'_f from notation: we write

$$\mathcal{M} \qquad \mathcal{Z}(T) \qquad {}^{\mathbb{L}}\mathcal{Z}(T) \qquad {}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}$$

instead of $\mathcal{M}_{K'_f}$, $\mathcal{Z}(T)_{K'_f}$, ${}^{\mathbb{L}}\mathcal{Z}(T)_{K'_f}$, ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p,K'_f}$, etc..

For example, given a Hermitian matrix $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with entries in F) and an appropriate scheme S, our notation entails

$$\mathcal{Z}(T)(S) = \{(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \tilde{\eta}_0, \tilde{\eta}, \underline{x}) \text{ over } S\}$$

where $(\tilde{\eta}_0, \tilde{\eta})$ is a K'_f level structure and $\underline{x} \in \mathrm{Hom}_{\mathcal{O}_F}(A_0, A)^m$ satisfies $(\underline{x}, \underline{x}) = T$.

11 Non-Archimedean

Fix a prime p. If p is not inert, we assume the signature is $(n-r, r) = (n-1, 1)$. We assume that $L \otimes_{\mathbb{Z}} \mathbb{Z}_p$ is self-dual. If p is ramified, we assume n is even, L is self-dual (for the trace pairing), and $p \neq 2$.

In all cases, we assume the implicit level $K'_f = K_{0,f} \times K_f$ at p is

$$K_{0,p} = K_{L_0,p} \qquad K_p = K_{L,p}. \tag{11.0.1}$$

Recall that these denote the stabilizers of L_0 and L, respectively.

Set $\mathcal{O}_{F_p} := \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$ and $\mathcal{O}_{F,(p)} := \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$ and $F_p := F \otimes_{\mathbb{Q}} \mathbb{Q}_p$. As in Section 5, we write $\breve{F}_p$ for the completion of the maximal unramified extension of F_p is p is nonsplit (resp. $\breve{F}_p := \breve{\mathbb{Q}}_p$ if p is split, with a choice of morphism $F_p \to \breve{F}_p$). In all cases, $\mathcal{O}_{\breve{F}_p}$ (resp. $\overline{k}$) will denote the ring of integers (resp. residue field) of $\breve{F}_p$.

We discuss Rapoport–Zink uniformization [RZ96, §6], as applied to supersingular loci on special cycles by Kudla–Rapoport [KR14, §5, §6] (there in the inert case, $p \neq 2$). For p inert or ramified, the material in Sections 11.1 to 11.6 is essentially a repackaging of Rapoport–Zink and Kudla–Rapoport. However, we need modified arguments at split places: the abelian varieties will be ordinary. We give a mostly uniform treatment for inert/ramified/split places. We also allow $p = 2$ if p is inert, except in Section 11.9.

Section 11.9 is the newest part of Section 11. Here, we explain how to use uniformization to reduce (global) Faltings or "tautological" heights to quantities expressed in terms of local special cycles and the "local change of heights" from Part III (with the main input being Corollary 10.2.2).

Section 11.7 is the next newest part of Section 11. We use global special cycles and an "approximation" argument to prove certain properties of local special cycles. Some of these results are available or implicit in the literature (for p nonsplit, sometimes with $p \neq 2$ hypotheses and signature $(n-1, 1)$ hypotheses); we indicate this where relevant. Our methods of proof are different, based on the approximation argument mentioned above.

Section 11.8 is the next newest part of Section 11. We explain how to reduce global "vertical intersection numbers" to local "vertical intersection numbers".

Sections 11.7 to 11.9 will need detailed information on the construction of Rapoport–Zink uniformization. This is our other reason for giving an exposition of uniformization in Sections 11.1-11.6, as we need to explain the relevant maps (and fix notation) to give precise statements.

Sections 11.1 to 11.5 state the precise Rapoport–Zink uniformization map for special cycles. The proof of uniformization appears in Section 11.6 (and allows $p = 2$ inert). We differ slightly from [RZ96] by working directly with formal algebraic stacks (rather than requiring sufficient level structure) in the sense of [Eme20]. We occasionally need some notions on formal algebraic stacks which are not defined in [Eme20]; we will define these as needed.

Throughout Section 11, we freely use the relevant Rapoport–Zink spaces and their (Kudla–Rapoport) local special cycles as in Section 5.

11.1 Formal completion

Throughout Section 11, the notation T will always mean an $m \times m$ Hermitian matrix with F-coefficients, i.e. $T \in \mathrm{Herm}_m(\mathbb{Q})$. If p is split in $\mathcal{O}_F$, we assume $\mathrm{rank}(T) \geq n-1$. Form the special cycle $\mathcal{Z}(T) \to \mathcal{M}$.

Suppose p is nonsplit. The *supersingular locus* on $\mathcal{Z}(T)_{\overline{k}} := \mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \overline{k}$ is the subset $\mathcal{Z}(T)^{ss} \subseteq |\mathcal{Z}(T)_{\overline{k}}|$ of the underlying topological space[31] consisting of geometric points $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \tilde{\eta}_0, \tilde{\eta}, \underline{x})$ with A supersingular. The supersingular locus $\mathcal{Z}(T)^{ss}$ is a closed subset of $|\mathcal{Z}(T)_{\overline{k}}|$ (by the Katz–Grothendieck theorem on specialization for Newton polygons). The *formal completion of* $\mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}} := \mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathcal{O}_{\breve{F}_p}$ *along its supersingular locus* is the (strictly full) substack $\breve{\mathcal{Z}}(T) \subseteq \mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}}$ given by

$$\breve{\mathcal{Z}}(T) := \{\alpha \in \mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}}(S) : \alpha(|S|) \subseteq \mathcal{Z}(T)^{ss}\} \tag{11.1.1}$$

for schemes S over $\operatorname{Spec} \mathcal{O}_{\breve{F}_p}$, where the condition $\alpha(|S|) \subseteq \mathcal{Z}(T)^{ss}$ means that the associated map on underlying topological spaces $|S| \to |\mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}}|$ factors through $\mathcal{Z}(T)^{ss}$ (with $\alpha \in \mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}}(S)$ "viewed" as a morphism $S \to \mathcal{Z}(T)_{\operatorname{Spec} \mathcal{O}_{\breve{F}_p}}$ by the 2-Yoneda lemma).

If p is split, we define

$$\breve{\mathcal{Z}}(T) := \mathcal{Z}(T)_{\operatorname{Spf} \mathcal{O}_{\breve{F}_p}} := \mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spf} \mathcal{O}_{\breve{F}_p}. \tag{11.1.2}$$

[31] By the *underlying topological space* $|\mathcal{X}|$ of a formal algebraic stack $\mathcal{X}$, we mean the underlying topological space of its reduced substack $\mathcal{X}_{\mathrm{red}}$. As $\mathcal{X}_{\mathrm{red}}$ is an algebraic stack, it has an underlying topological space in the sense of [SProject, Section 04XE].

This is also the formal completion of $\mathcal{Z}(T)_{\operatorname{Spec}\mathcal{O}_{\breve{F}_p}}$ along its special fiber. For any geometric point $(A_0, \iota_0, \lambda_0, A, \iota, \lambda, \tilde{\eta}_0, \tilde{\eta}, \underline{x})$ of $\breve{\mathcal{Z}}(T)$, the abelian variety A is ordinary (because Lemma 4.7.1 implies A is isogenous to a product of elliptic curves with $\mathcal{O}_F$ action).

In all cases, $\breve{\mathcal{Z}}(T)$ is a locally Noetherian formal algebraic stack in the sense of [Eme20] (formal completion is discussed in [Eme20, Example 5.9]). The structure morphism $\breve{\mathcal{Z}}(T) \to \operatorname{Spf}\mathcal{O}_{\breve{F}_p}$ is formally smooth,[32] formally locally of finite type,[33] separated, and quasi-compact. If K'_f is a small level, then $\breve{\mathcal{Z}}(T)$ is a locally Noetherian formal scheme.

If $\mathcal{M} = \mathcal{Z}(T)$ (e.g. $T = \emptyset$ or $T = 0$), we set $\breve{\mathcal{M}} \coloneqq \breve{\mathcal{Z}}(T)$. If p is nonsplit, this is the *formal completion of* $\mathcal{M}_{\operatorname{Spec}\mathcal{O}_{\breve{F}_p}}$ *along its supersingular locus* $\mathcal{M}^{ss}$.

11.2 Local special cycles away from p

Given an m-tuple $\underline{x}^p = [x_1, \ldots, x_m] \in (V \otimes_{\mathbb{Q}} \mathbb{A}_f^p)^m$, we consider an *"away-from-p" local special cycle*

$$\mathcal{Z}'(\underline{x}^p) \coloneqq \{(g_0, g) : G'(\mathbb{A}_f^p)/K_f'^p : g^{-1} g_0 x_i \in L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \text{ for all } x_i \in \underline{x}^p\}. \tag{11.2.1}$$

We often view $\mathcal{Z}'(\underline{x}^p)$ and $G'(\mathbb{A}_f^p)/K_f'^p$ as constant formal schemes over $\operatorname{Spf}\mathcal{O}_{\breve{F}_p}$. We also define the *"away-from-p" local special cycle*

$$\mathcal{Z}(\underline{x}^p) \coloneqq \{g : U(V)(\mathbb{A}_f^p)/K_f^p : g^{-1} x_i \in L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \text{ for all } x_i \in \underline{x}^p\}. \tag{11.2.2}$$

The isomorphism $G'(\mathbb{A}_f^p)/K_f'^p \to GU(V_0)(\mathbb{A}_f^p)/K_{0,f}^p \times U(V)(\mathbb{A}_f^p)/K_f^p$ $(*)$ induces an isomorphism

$$\mathcal{Z}'(\underline{x}^p) \xrightarrow{\sim} GU(V_0)(\mathbb{A}_f^p)/K_{0,f}^p \times \mathcal{Z}(\underline{x}^p). \tag{11.2.3}$$

11.3 Framing objects

To define the uniformization map, we fix an object $(\mathbf{A}_0, \iota_{\mathbf{A}_0}, \lambda_{\mathbf{A}_0}, \mathbf{A}, \iota_{\mathbf{A}}, \lambda_{\mathbf{A}}, \tilde{\boldsymbol{\eta}}_0, \tilde{\boldsymbol{\eta}}) \in \mathcal{M}(\overline{k})$ ("basepoint of the uniformization"). If p is nonsplit (resp. split), we assume $\mathbf{A}$ is supersingular (resp. $\mathbf{A}$ is $\mathcal{O}_F$-linearly isogenous to $\mathbf{A}_0^{n-r} \times (\mathbf{A}_0^{\sigma})^r$); such data exists by Lemma 3.1.5 and

[32] Given a morphism f of categories fibered in groupoids over some base scheme, there is a category of dotted arrows [SProject, Definition 0H18] associated to the infinitesimal lifting problem along each square-zero thickening of affine schemes. We say that f is *formally smooth* (resp. *formally étale*) (resp. *formally unramified*) if each such category of dotted arrows is nonempty (resp. a setoid with exactly one isomorphism class) (resp. either empty or a setoid with exactly one isomorphism class).

[33] We say a morphism of locally Noetherian formal algebraic stacks is *formally locally of finite type* if it is locally of finite type on underlying reduced substacks.

Remark 3.2.4. Let $(\mathbf{X}_0, \iota_{\mathbf{X}_0}, \lambda_{\mathbf{X}_0}, \mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ be the tuple obtained by passing to p-divisible groups (e.g. $\mathbf{X}$ is the p-divisible group of $\mathbf{A}$). We use this as the framing object over $\overline{k}$ to define the Rapoport–Zink space $\mathcal{N}'$ (Definition 5.1.8). Set $\mathcal{N} \coloneqq \mathcal{N}(n-r, r)$ (Definition 5.1.3).

In the supersingular cases, the abelian variety $\mathbf{A}$ is automatically $\mathcal{O}_F$-linearly isogenous to $\mathbf{A}_0^{n-r} \times (\mathbf{A}_0^{\sigma})^r$, since

$$\mathrm{Hom}_F^0(\mathbf{A}_0^{n-r} \times (\mathbf{A}_0^{\sigma})^r, \mathbf{A}) \otimes_{\mathbb{Q}} \mathbb{Q}_p \xrightarrow{\sim} \mathrm{Hom}_F^0(\mathbf{X}_0^{n-r} \times (\mathbf{X}_0^{\sigma})^r, \mathbf{X}) \tag{11.3.1}$$

by Tate's isogeny theorem (for any supersingular abelian variety over a finite field, some power of Frobenius will be a power of p, e.g. by [RZ96, Lemma 6.28]); then use uniqueness of the framing object $(\mathbf{X}, \iota_{\mathbf{X}}, \lambda_{\mathbf{X}})$ up to isogeny (Section 5.1).

Since $\mathcal{M}_{\mathrm{Spec}\, \mathcal{O}_{\breve{F}_p}} \to \mathrm{Spec}\, \mathcal{O}_{\breve{F}_p}$ is smooth, this framing object $(\mathbf{A}_0, \ldots)$ admits a lift $(\mathfrak{A}_0, \iota_{\mathfrak{A}_0}, \lambda_{\mathfrak{A}_0}, \mathfrak{A}, \iota_{\mathfrak{A}}, \lambda_{\mathfrak{A}}, \tilde{\mathfrak{y}}_0, \tilde{\mathfrak{y}}) \in \mathcal{M}(\mathrm{Spf}\, \mathcal{O}_{\breve{F}_p})$, which we also fix. We fix representatives

$$\boldsymbol{\eta}_0 \colon T^p(\mathbf{A}_0) \xrightarrow{\sim} L_0 \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \qquad \boldsymbol{\eta} \colon \mathrm{Hom}_{\mathcal{O}_F \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p}(T^p(\mathbf{A}_0), T^p(\mathbf{A})) \xrightarrow{\sim} L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \tag{11.3.2}$$

for the K_0^p-orbit $\tilde{\boldsymbol{\eta}}_0$ and the K^p-orbit $\tilde{\boldsymbol{\eta}}$ (see Section 3.4). Recall that $\boldsymbol{\eta}$ preserves Hermitian pairings but $\boldsymbol{\eta}_0$ need not. We also write

$$\mathfrak{y}_0 \colon T^p(\mathfrak{A}_0) \xrightarrow{\sim} L_0 \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \qquad \mathfrak{y} \colon \mathrm{Hom}_{\mathcal{O}_F \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p}(T^p(\mathfrak{A}_0), T^p(\mathfrak{A})) \xrightarrow{\sim} L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \tag{11.3.3}$$

for the identifications induced by $\boldsymbol{\eta}_0$ and $\boldsymbol{\eta}$.

We define Hermitian F-modules

$$\mathbf{W} \coloneqq \mathrm{Hom}_F^0(\mathbf{A}_0, \mathbf{A}) \qquad \mathbf{W}^{\perp} \coloneqq \begin{cases} \mathrm{Hom}_F^0(\mathbf{A}_0^{\sigma}, \mathbf{A}) & \text{if } p \text{ is split} \\ 0 & \text{if } p \text{ is nonsplit} \end{cases} \tag{11.3.4}$$

$$\mathbf{V}_0 \coloneqq \mathrm{Hom}_F^0(\mathbf{A}_0, \mathbf{A}_0) \qquad \mathbf{V} \coloneqq \mathbf{W} \oplus \mathbf{W}^{\perp} \tag{11.3.5}$$

where the direct sum defining $\mathbf{V}$ is orthogonal. In all cases, the Hermitian pairing is $(x, y) \coloneqq x^{\dagger} y \in F$. All of these Hermitian spaces are positive definite (positivity of the Rosati involution).

The canonical maps

$$\mathbf{W} \otimes_{\mathbb{Q}} \mathbb{Q}_p \xrightarrow{\sim} \mathrm{Hom}_{F_p}^0(\mathbf{X}_0, \mathbf{X}) \qquad \mathbf{W}^{\perp} \otimes_{\mathbb{Q}} \mathbb{Q}_p \xrightarrow{\sim} \mathrm{Hom}_{F_p}^0(\mathbf{X}_0^{\sigma}, \mathbf{X}) \tag{11.3.6}$$

are isomorphisms of Hermitian spaces. In the nonsplit (hence supersingular) cases, this follows from Tate's isogeny theorem as above. In the split case, this follows because $\mathbf{A}$ is

$\mathcal{O}_F$-linearly isogenous to $\mathbf{A}_0^{n-r} \times (\mathbf{A}_0^\sigma)^r$. In particular, the local invariant is $\varepsilon(\mathbf{W}_p) = (-1)^r$ if p is nonsplit (resp. $\varepsilon(\mathbf{W}_p) = 1$ if p is split).

If p is nonsplit, the natural map

$$\mathbf{W} \otimes_{\mathbb{Q}} \mathbb{A}_f^p \to \operatorname{Hom}_{F \otimes_{\mathbb{Q}} \mathbb{A}_f^p}(T^p(\mathbf{A}_0)^0, T^p(\mathbf{A})^0) \tag{11.3.7}$$

is an isomorphism of Hermitian spaces, by similar reasoning.

If p is split, any $\mathcal{O}_F$-linear isogeny $\mathbf{A}_0^{n-1} \times \mathbf{A}_0^\sigma \to \mathbf{A}$ defines an F-linear orthogonal decomposition

$$T^p(\mathbf{A})^0 = T^p(\mathbf{A}_0^{n-1})^0 \oplus T^p(\mathbf{A}_0^\sigma)^0. \tag{11.3.8}$$

This decomposition is independent of the choice of isogeny because $\operatorname{Hom}_F^0(\mathbf{A}_0, \mathbf{A}_0^\sigma) = \operatorname{Hom}_F^0(\mathbf{A}_0^\sigma, \mathbf{A}_0) = 0$ (e.g. because $\operatorname{End}^0(\mathbf{A}_0) = F$). Then the natural map

$$\mathbf{W} \otimes_{\mathbb{Q}} \mathbb{A}_f^p \to \operatorname{Hom}_{F \otimes_{\mathbb{Q}} \mathbb{A}_f^p}(T^p(\mathbf{A}_0)^0, T^p(\mathbf{A}_0^{n-1})^0) \tag{11.3.9}$$

is an isomorphism of Hermitian spaces.

Given a tuple $\underline{\mathbf{x}} \in \mathbf{W}^m$, we write

$$\underline{\mathbf{x}}_p \in \mathbf{W}_p^m = \operatorname{Hom}_{F_p}^0(\mathbf{X}_0, \mathbf{X}) \tag{11.3.10}$$

$$\underline{\mathbf{x}}^p \in \mathbf{W}^m \otimes_{\mathbb{Q}} \mathbb{A}_f^p \subseteq \operatorname{Hom}_{F \otimes \mathbb{A}_f^p}(T^p(\mathbf{A}_0)^0, T^p(\mathbf{A})^0)^m = V^m \otimes_{\mathbb{Q}} \mathbb{A}_f^p \tag{11.3.11}$$

for the respective images of $\underline{\mathbf{x}}$ (using $\boldsymbol{\eta}$ for the identification with $V^m \otimes_{\mathbb{Q}} \mathbb{A}_f^p$ in the second line).

11.4 Framed stack

We consider the stack $\breve{\mathcal{Z}}(T)_{\text{framed}}$ over $\operatorname{Spf} \mathcal{O}_{\breve{F}_p}$, given by

$$\breve{\mathcal{Z}}(T)_{\text{framed}}(S) := \left\{ (\alpha, \phi_0, \phi) : \begin{array}{l} \alpha = (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \tilde{\eta}_0, \tilde{\eta}, \underline{x}) \in \breve{\mathcal{Z}}(T)(S) \\ \phi_0 \colon A_0 \to \mathfrak{A}_{0,S} \text{ and } \phi \colon A \to \mathfrak{A}_S \quad \text{quasi-isogenies} \\ \quad \text{such that } \phi_0^* \lambda_{\mathfrak{A}_{0,S}} = b\lambda_0 \text{ and } \phi^* \lambda_{\mathfrak{A},S} = b\lambda \\ \quad \text{for some } b \in \mathbb{Q}_{>0} \end{array} \right\}$$

for schemes S over $\operatorname{Spf} \mathcal{O}_{\breve{F}_p}$. The similitude factor b is allowed to vary (and is only required to be locally constant). If $\mathcal{M} = \mathcal{Z}(T)$, we set $\breve{\mathcal{M}}_{\text{framed}} := \breve{\mathcal{Z}}(T)_{\text{framed}}$. There is a canonical forgetful map

$$\Theta \colon \breve{\mathcal{Z}}(T)_{\text{framed}} \to \breve{\mathcal{Z}}(T) \tag{11.4.1}$$

sending $(\alpha, \phi_0, \phi) \mapsto \alpha$. This will be the uniformization map (Section 11.6).

There is a canonical isomorphism

$$\breve{\mathcal{Z}}(T)_{\mathrm{framed}} \xrightarrow{\sim} \coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p) \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \tag{11.4.2}$$

which we now describe. Here $\mathcal{Z}'(\underline{\mathbf{x}}_p)$ is the local special cycle at p from Section 5.2, and $\mathcal{Z}'(\underline{\mathbf{x}}^p)$ is the away-from-p local special cycle from Section 11.2.

Consider $(\alpha, \phi_0, \phi) \in \breve{\mathcal{Z}}(T)_{\mathrm{framed}}(S)$ as above. Passing to p-divisible groups gives a datum $(X_0, \iota_0, \lambda_0, X, \iota, \lambda)$ (e.g. X is the p-divisible group of A), along with a framing quasi-isogeny $\rho \colon X_{\overline{S}} \to \mathbf{X}_{\overline{S}}$ induced by ϕ (where $\overline{S} \coloneqq S_{\overline{k}}$) and similarly a framing ρ_0 induced by ϕ_0. We also obtain $g_0 \coloneqq \mathfrak{y}_0 \circ \phi_{0,*} \circ \tilde{\eta}_0^{-1} \in G_0(\mathbb{A}_f^p)/K_{0,f}^p$ and $g \coloneqq \mathfrak{y} \circ (\phi_0^{-1,*}\phi_*) \circ \tilde{\eta}^{-1} \in U(V)(\mathbb{A}_f^p)/K_f^p$ where $\phi_{0,*} \colon T^p(A_0)^0 \to T^p(\mathfrak{A}_0)^0$ and

$$\phi_0^{-1,*}\phi_* \colon \operatorname{Hom}_{F \otimes_{\mathbb{Q}} \mathbb{A}_f^p}(T^p(A_0)^0, T^p(A)^0) \to \operatorname{Hom}_{F \otimes_{\mathbb{Q}} \mathbb{A}_f^p}(T^p(\mathfrak{A}_0)^0, T^p(\mathfrak{A})^0) \tag{11.4.3}$$

is pre- and post-composition (when S is connected, pick any geometric point; there is no dependence on this choice). In general, g_0 and g will be locally constant elements. For any $\underline{x} \in \operatorname{Hom}_F(A_0, A)^m$ over a connected base S, we have $\phi \circ \underline{x} \circ \phi_0^{-1} \in \mathbf{W}^m$ (canonically), by Mumford's rigidity lemma for morphisms of abelian schemes [MFK94, Corollary 6.2]. In the not-necessarily connected case, we obtain a locally constant element of $\mathbf{W}^m$.

The above constructions give a map

$$\begin{aligned} \breve{\mathcal{Z}}(T)_{\mathrm{framed}} &\longrightarrow \mathcal{N}' \times G'(\mathbb{A}_f^p)/K_f'^p \times \mathbf{W}^m \\ (\alpha, \phi_0, \phi) &\longmapsto ((X_0, \iota_0, \lambda_0, \rho_0, X, \iota, \lambda, \rho), (g_0, g_0 g), \phi \circ \underline{x} \circ \phi_0^{-1}) \end{aligned} \tag{11.4.4}$$

which induces an isomorphism from $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ to the open and closed subfunctor

$$\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p) \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \hookrightarrow \mathcal{N}' \times G'(\mathbb{A}_f^p)/K_f'^p \times \mathbf{W}^m. \tag{11.4.5}$$

One can verify that the map in (11.4.2) is an isomorphism by decomposing the kernels of framing quasi-isogenies (rescale to obtain an isogeny) into their p-power and ℓ-power torsion subgroups. The isomorphism implies that $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ is a locally Noetherian formal scheme.

11.5 Quotient

Consider the algebraic groups

$$I_0 \coloneqq GU(\mathbf{V}_0) \qquad I_1 \coloneqq U(\mathbf{W}) \times U(\mathbf{W}^\perp) \qquad I' \coloneqq I_0 \times I_1 \tag{11.5.1}$$

over $\mathbb{Q}$. Unless specified otherwise, an element $(\gamma_0, \gamma) \in I'$ will mean $\gamma_0 \in I_0$ and $\gamma \in GU(\mathbf{W}) \times GU(\mathbf{W}^\perp)$ with $\gamma_0^{-1}\gamma \in I_1$.

Uniformization will involve the stack quotient $[I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]$ for an action of $I'(\mathbb{Q})$ on $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$, which we now describe. For $\mathbb{Q}$-algebras R, there are canonical identifications

$$I_0(R) = \{\gamma_0 \in \mathrm{End}^0_F(\mathbf{A}_0) \otimes_{\mathbb{Q}} R : \gamma_0^\dagger \gamma_0 \in R^\times\} \tag{11.5.2}$$

$$I_1(R) = \{\gamma \in \mathrm{End}^0_F(\mathbf{A}) \otimes_{\mathbb{Q}} R : \gamma^\dagger \gamma = 1\} \tag{11.5.3}$$

(act on $\mathbf{V}_0$ and $\mathbf{V}$ by post-composition). View $I'(\mathbb{Q})$ as a discrete group. Then $(\gamma_0, \gamma) \in I'(\mathbb{Q})$ acts on $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ as $(\alpha, \phi_0, \phi) \mapsto (\alpha, \gamma_0 \circ \phi_0, \gamma \circ \phi)$. We are abusing notation: the elements γ_0 and γ lift (uniquely, by Mumford's rigidity lemma or Drinfeld rigidity and Serre–Tate) to quasi-endomorphisms of $\mathfrak{A}_{0,S}$ and $\mathfrak{A}_S$ respectively.

In terms of the isomorphism in (11.4.2), the action of $I'(\mathbb{Q})$ on $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ admits the following (equivalent) description. By the isomorphism in (11.3.6), the group $I'(\mathbb{Q}_p)$ acts on $\mathcal{N}'$ (discussed in Section 5.3). By (11.5.3), we have a faithful action of $I(\mathbb{A}_f^p)$ on

$$\mathrm{Hom}_{F \otimes \mathbb{A}_f^p}(T^p(\mathbf{A}_0)^0, T^p(\mathbf{A})^0) = V \otimes_{\mathbb{Q}} \mathbb{A}_f^p \tag{11.5.4}$$

by post-composition. This induces a homomorphism $I_1(\mathbb{A}_f^p) \to U(V)(\mathbb{A}_f^p)$ and hence an action of $I'(\mathbb{A}_f^p)$ on $G'(\mathbb{A}_f^p)/K_f'^p$ (left multiplication). The group $I'(\mathbb{Q})$ also acts on $\mathbf{W}$ by the projection $I'(\mathbb{Q}) \to U(\mathbf{W})$.

Hence $I'(\mathbb{Q})$ acts on

$$\mathcal{N}' \times G'(\mathbb{A}_f^p)/K_f'^p \times \mathbf{W}^m. \tag{11.5.5}$$

Under the inclusion (11.4.5), this induces the same action on $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ described previously. Both descriptions will be useful for us.

We now form the (fppf) stack quotient

$$\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p) \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \to \left[I'(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p) \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \right) \right] \tag{11.5.6}$$

The left-hand side is a locally Noetherian formal scheme, and the right-hand side is a locally Noetherian formal algebraic stack which is formally locally of finite type over $\mathrm{Spf}\, \mathcal{O}_{\breve{F}_p}$. The right-hand side is also $[I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]$. The quotient map is representable by schemes, separated, étale, and surjective.

Using (5.2.5) (and (5.4.4)) and (11.2.3) (various incarnations of the isomorphism $G' \xrightarrow{\sim} GU(V_0) \times U(V)$) yields a canonical isomorphism from the left-hand side of (11.5.6) to

$$GU(V_0)/K_{0,f} \times \coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}(\underline{\mathbf{x}}_p) \times U(\mathbf{W}_p^\perp)/K_{1, \mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \tag{11.5.7}$$

where $K_{1,\mathbf{L}_p^\perp} \subseteq U(\mathbf{W}_p^\perp)$ is the unique maximal open compact subgroup (since $\mathbf{W}^\perp$ has rank 0 or 1). This is a disjoint union of various local special cycles $\mathcal{Z}(\underline{\mathbf{x}}_p)$, indexed by the (discrete) set

$$J_p(T) \coloneqq GU(V_0)/K_{0,f} \times \coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p). \tag{11.5.8}$$

In particular, every element $j \in J_p(T)$ defines a morphism

$$\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p) \to [I'(\mathbb{Q}) \backslash \mathcal{Z}(T)_{\text{framed}}] \tag{11.5.9}$$

which is étale, separated, and representable by schemes. Given $j \in J_p(T)$, we let $\operatorname{Aut}(j) \subseteq I'(\mathbb{Q})$ be the stabilizer for the action of $I'(\mathbb{Q})$ on $J_p(T)$.

The right-hand side of (11.5.6) is then identified with

$$\left[GU(V_0)(\mathbb{Q}) \backslash \left(GU(V_0)(\mathbb{A}_f)/K_{0,f} \right) \right] \times \left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}(\underline{\mathbf{x}}_p) \times U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right]. \tag{11.5.10}$$

We have

$$\deg[GU(V_0)(\mathbb{Q}) \backslash (GU(V_0)(\mathbb{A}_f)/K_{0,f})] = [K_{L_0,f} : K_{0,f}] \cdot h_F / |\mathcal{O}_F^\times| \tag{11.5.11}$$

where the left-hand side denotes (stacky) groupoid cardinality, where $[K_{L_0,f} : K_{0,f}]$ is the index of $K_{0,f}$ in $K_{L_0,f}$, and h_F is the class number of $\mathcal{O}_F$. In the case where $\operatorname{rank}(T) \geq n-1$ (we have already assumed $\operatorname{rank}(T) \geq n-1$ if p is split), the groupoid

$$\left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right] \tag{11.5.12}$$

has finite automorphism groups and finitely many isomorphism classes, and its groupoid degree is essentially a product of special values of local Whittaker functions away from p (Lemma 20.4.1).

In the case $\operatorname{rank}(T) \geq n-1$, the map Θ_j associated with any $j \in J_p(T)$ is thus representable by schemes and finite étale of constant degree $\deg \Theta_j = |\operatorname{Aut}(j)|$.

11.6 Uniformization

We explain how the uniformization morphism $\Theta \colon \breve{\mathcal{Z}}(T)_{\mathrm{framed}} \to \breve{\mathcal{Z}}(T)$ (11.4.1) descends to an isomorphism of locally Noetherian formal algebraic stacks

$$\tilde{\Theta} \colon \left[I'(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p) \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \right) \right] \xrightarrow{\sim} \breve{\mathcal{Z}}(T). \tag{11.6.1}$$

The main point is surjectivity on $\overline{k}$-points via the Hasse principle (Lemma 11.6.2).

When p is split, we will allow a change of choice of framing data $(\mathbf{A}_0, \iota_{\mathbf{A}_0}, \lambda_{\mathbf{A}_0}, \mathbf{A}, \iota_{\mathbf{A}}, \lambda_{\mathbf{A}}, \tilde{\boldsymbol{\eta}}_0, \tilde{\boldsymbol{\eta}})$, $\boldsymbol{\eta}_0$, and $\boldsymbol{\eta}$, possibly depending on T.

Lemma 11.6.1. *The map $\Theta \colon \breve{\mathcal{Z}}(T)_{\mathrm{framed}} \to \breve{\mathcal{Z}}(T)$ factors uniquely through a monomorphism*[34]

$$\tilde{\Theta} \colon [I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}] \to \breve{\mathcal{Z}}(T) \tag{11.6.2}$$

of formal algebraic stacks. The map $\tilde{\Theta}$ is formally locally of finite type and formally étale.

Proof. Suppose (α, ϕ_0, ϕ) and $(\alpha', \phi_0', \phi')$ are objects of $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}(S)$, and suppose $f' \colon \alpha \to \alpha'$ is an isomorphism of objects in the groupoid $\breve{\mathcal{Z}}(T)(S)$ (for some base scheme S). We claim there is a unique $\gamma' = (\gamma_0, \gamma) \in I'(\mathbb{Q})$ such that f' induces an isomorphism $\gamma' \cdot (\alpha, \phi_0, \phi) \xrightarrow{\sim} (\alpha', \phi_0', \phi')$ in the setoid $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}(S)$.

The map f is given by a pair of isomorphisms $f_0 \colon A_0 \to A_0'$ and $f \colon A \to A'$ (where $\alpha = (A_0, \ldots)$ and $\alpha' = (A_0', \ldots)$, with notation as above). Then we take $\gamma_0 = \phi_0' \circ f_0 \circ \phi_0^{-1}$ and $\gamma = \phi' \circ f \circ \phi^{-1}$. Hence $\tilde{\Theta}$ is a monomorphism.

The map $\tilde{\Theta}$ is a map between locally Noetherian formal algebraic stacks which are formally locally of finite type over $\operatorname{Spf} \mathcal{O}_{\breve{F}_p}$, so $\tilde{\Theta}$ is formally locally of finite type.

The property of being formally étale may be checked "formally étale locally on the source". The quotient map $\breve{\mathcal{Z}}(T)_{\mathrm{framed}} \to [I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]$ is representable by schemes and formally étale, so it is enough to check that $\Theta \colon \breve{\mathcal{Z}}(T)_{\mathrm{framed}} \to \breve{\mathcal{Z}}(T)$ is formally étale. This property amounts to the following rigidity statement for abelian schemes: given any first order thickening of schemes $T \to T'$ on which p is locally nilpotent, and given abelian schemes A_1 and A_2 over T', any quasi-homomorphism $A_{1,T} \to A_{2,T}$ lifts uniquely to a quasi-homomorphism $A_1 \to A_2$ (e.g. by Drinfeld rigidity and Serre–Tate). $\square$

[34]By a *monomorphism* of formal algebraic stacks, we mean a morphism which is fully faithful on underlying fibered categories.

Lemma 11.6.2. *The map $\Theta(\overline{k}) \colon \breve{\mathcal{Z}}(T)_{\mathrm{framed}}(\overline{k}) \to \breve{\mathcal{Z}}(T)(\overline{k})$ (on groupoids of $\overline{k}$-points) is surjective (resp. surjective for some choice of framing data) on isomorphism classes if p is nonsplit (resp. split).*

Proof. If $\breve{\mathcal{Z}}(T)$ is empty, there is nothing to show, so assume $\breve{\mathcal{Z}}(T)$ is nonempty. If p is split, we can change the framing object to assume it extends to $(\mathbf{A}_0, \iota_{\mathbf{A}_0}, \lambda_{\mathbf{A}_0}, \mathbf{A}, \iota_{\mathbf{A}}, \lambda_{\mathbf{A}}, \tilde{\boldsymbol{\eta}}_0, \tilde{\boldsymbol{\eta}}, \underline{\mathbf{x}}) \in \breve{\mathcal{Z}}(T)(\overline{k})$ (i.e. $\underline{\mathbf{x}} \in \mathbf{W}^m$ with $(\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T$). This implies that T has rank $n-1$ if p is split (we already assumed $\operatorname{rank}(T) \geq n-1$ if p is split, then see Remark 4.7.2). We still know that $\mathbf{A}$ is $\mathcal{O}_F$-linearly isogenous to $\mathbf{A}_0^{n-1} \times \mathbf{A}_0^{\sigma}$ (Lemma 4.7.1).

In all cases, the task is to show that any $(\mathbf{A}'_0, \iota_{\mathbf{A}'_0}, \lambda_{\mathbf{A}'_0}, \mathbf{A}', \iota_{\mathbf{A}'}, \lambda_{\mathbf{A}'}, \tilde{\boldsymbol{\eta}}_0, \tilde{\boldsymbol{\eta}}, \underline{\mathbf{x}}') \in \breve{\mathcal{Z}}(T)(\overline{k})$ admits a framing (ϕ_0, ϕ), i.e. quasi-isogenies $\phi_0 \colon \mathbf{A}'_0 \to \mathbf{A}_0$ and $\phi \colon \mathbf{A}' \to \mathbf{A}$ which preserve quasi-polarizations up to the same scalar in $\mathbb{Q}_{>0}$.

Fix any $\mathcal{O}_F$-linear isogeny $\phi_0 \colon \mathbf{A}'_0 \to \mathbf{A}_0$, which exists because $\mathbf{A}'_0$ and $\mathbf{A}_0$ are elliptic curves with $\mathcal{O}_F$-action of the same signature (see the proof of Lemma 4.7.1). Let $b \in \mathbb{Q}_{>0}$ be such that $\phi_0^* \lambda_{\mathbf{A}_0} = b \lambda_{\mathbf{A}'_0}$. Set $\mathbf{W}' \coloneqq \operatorname{Hom}^0_F(\mathbf{A}'_0, \mathbf{A}')$ with the Hermitian pairing $(x, y) = x^\dagger y$.

Case p is nonsplit: There is an isomorphism of F vector spaces

$$\begin{aligned} \operatorname{Hom}^0_F(\mathbf{A}', \mathbf{A}) &\longrightarrow \operatorname{Hom}_F(\mathbf{W}', \mathbf{W}) \\ \phi &\longmapsto (f \mapsto \phi \circ f \circ \phi_0^{-1}). \end{aligned} \tag{11.6.3}$$

An element $\phi \in \operatorname{Hom}^0_F(\mathbf{A}', \mathbf{A})$ satisfies $\phi^\dagger \phi = b$ if and only if ϕ corresponds to an isomorphism of Hermitian spaces $\mathbf{W}' \to \mathbf{W}$. But we have $\mathbf{W}' \otimes_{\mathbb{Q}} \mathbb{A}_f^p \cong \mathbf{W} \otimes_{\mathbb{Q}} \mathbb{A}_f^p \cong V \otimes_{\mathbb{Q}} \mathbb{A}_f^p$ as Hermitian spaces, we have $\varepsilon(\mathbf{W}') = \varepsilon(\mathbf{W}) = (-1)^r$, and we have $\mathbf{W}'_{\mathbf{R}} \cong \mathbf{W}_{\mathbb{R}}$ (both are positive definite of rank n). So we have $\mathbf{W}' \cong \mathbf{W}$ as Hermitian spaces, by the Hasse principle for Hermitian spaces (Landherr's theorem).

Case p is split: Fix $\mathcal{O}_F$-linear isogenies $\mathbf{B} \times \mathbf{B}^{\perp} \to \mathbf{A}$ and $\mathbf{B}' \times \mathbf{B}'^{\perp} \to \mathbf{A}'$, where $\mathbf{B} \cong \mathbf{A}_0^{n-1}$, $\mathbf{B}^{\perp} \cong \mathbf{A}_0^{\sigma}$, $\mathbf{B}' \cong \mathbf{A}_0'^{n-1}$, and $\mathbf{B}'^{\perp} \cong \mathbf{A}_0'^{\sigma}$. Equip $\mathbf{B} \times \mathbf{B}^{\perp}$ and $\mathbf{B}' \times \mathbf{B}'^{\perp}$ with the quasi-polarizations pulled back from $\lambda_{\mathbf{A}}$ and $\lambda_{\mathbf{A}'}$ on $\mathbf{A}$ and $\mathbf{A}'$, respectively.

Any F-linear quasi-isogeny $\phi \colon \mathbf{A}' \to \mathbf{A}$ decomposes as a product of quasi-isogenies $\mathbf{B}' \to \mathbf{B}$ and $\mathbf{B}'^{\perp} \to \mathbf{B}^{\perp}$, since $\operatorname{Hom}_F(\mathbf{B}', \mathbf{B}^{\perp}) = \operatorname{Hom}_F(\mathbf{B}'^{\perp}, \mathbf{B}) = 0$ (because of the opposite signatures). We write $\phi^{\perp} \colon \mathbf{B}'^{\perp} \to \mathbf{B}^{\perp}$ for the quasi-isogeny induced by ϕ. By similar reasoning, the quasi-polarization on $\mathbf{B} \times \mathbf{B}^{\perp}$ is the product of a quasi-polarization on $\mathbf{B}$ and a quasi-polarization on $\mathbf{B}^{\perp}$.

There is an isomorphism of F vector spaces

$$\begin{aligned} \operatorname{Hom}^0_F(\mathbf{A}', \mathbf{A}) &\longrightarrow \operatorname{Hom}_F(\mathbf{W}', \mathbf{W}) \times \operatorname{Hom}^0_F(\mathbf{B}'^{\perp}, \mathbf{B}^{\perp}) \\ \phi &\longmapsto (f \mapsto (\phi \circ f \circ \phi_0^{-1})), \phi^{\perp}. \end{aligned} \tag{11.6.4}$$

An element $\phi \in \mathrm{Hom}_F^0(\mathbf{A}', \mathbf{A})$ satisfies $\phi^\dagger \phi = b$ if and only if ϕ corresponds to an isomorphism of Hermitian spaces $\mathbf{W}' \to \mathbf{W}$ and with $\phi^{\perp\dagger}\phi^\perp = b$.

We have $\underline{\mathbf{x}}' \in \mathbf{W}'$ and $\underline{\mathbf{x}} \in \mathbf{W}$ with $(\underline{\mathbf{x}}', \underline{\mathbf{x}}') = (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T$. Since $\mathrm{rank}(T) = \mathrm{rank}(\mathbf{W}') = \mathrm{rank}(\mathbf{W}) = n - 1$, this implies $\mathbf{W}' \cong \mathbf{W}$ as Hermitian spaces.

For every prime $\ell \neq p$, the natural map

$$\mathrm{Hom}_F^0(\mathbf{B}'^\perp, \mathbf{B}^\perp) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \longrightarrow \mathrm{Hom}_{F_\ell}(T_\ell(\mathbf{B}'^\perp)^0, T_\ell(\mathbf{B}^\perp)^0) \tag{11.6.5}$$

is an isomorphism of (one-dimensional) Hermitian spaces. If we set $\mathbf{U}_\ell \coloneqq \mathrm{Hom}_{F_\ell}(T_\ell(\mathbf{A}_0)^0, T_\ell(\mathbf{B}^\perp)^0)$ and $\mathbf{U}'_\ell \coloneqq \mathrm{Hom}_{F_\ell}(T_\ell(\mathbf{A}'_0)^0, T_\ell(\mathbf{B}'^\perp)^0)$, there is an isomorphism of F_ℓ vector spaces

$$\begin{aligned} \mathrm{Hom}_{F_\ell}(T_\ell(\mathbf{B}'^\perp)^0, T_\ell(\mathbf{B}^\perp)^0) &\longrightarrow \mathrm{Hom}_{F_\ell}(\mathbf{U}_\ell, \mathbf{U}'_\ell) \\ \phi^\perp &\longmapsto (f \mapsto (\phi^\perp \circ f \circ \phi_0^{-1})). \end{aligned} \tag{11.6.6}$$

An element $\phi^\perp$ on the left satisfies $\phi^{\perp\dagger}\phi^\perp = b$ if and only if $\phi^\perp$ corresponds to an isomorphism of Hermitian spaces $\mathbf{U}_\ell \to \mathbf{U}'_\ell$. We have $V_\ell \cong \mathbf{W}'_\ell \oplus \mathbf{U}'_\ell \cong \mathbf{W}_\ell \oplus \mathbf{U}_\ell$ (orthogonal direct sum) as Hermitian spaces, for all $\ell \neq p$. Hence $\mathbf{U}'_\ell \cong \mathbf{U}_\ell$ for all $\ell \neq p$ (consider the Hermitian space local invariants (in $\{\pm 1\}$) via ε as in Section 2.2).

The preceding discussion produces an element $\phi_\ell^\perp \in \mathrm{Hom}_F^0(\mathbf{B}'^\perp, \mathbf{B}^\perp) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ satisfying $\phi_\ell^{\perp\dagger}\phi_\ell^\perp = b$ for all primes $\ell \neq p$. Since p is split in $\mathcal{O}_F$, such an element exists for $\ell = p$ as well (i.e. $N_{F_p/\mathbb{Q}_p}(F_p^\times) = \mathbb{Q}_p^\times$). Since $b > 0$, such an element also exists if $\mathbb{Q}_\ell$ is replaced by $\mathbb{R}$ (positivity of the Rosati involution). By the Hasse principle for Hermitian spaces (or Hasse norm theorem), we obtain $\phi^\perp \in \mathrm{Hom}_F^0(\mathbf{B}'^\perp, \mathbf{B}^\perp)$ satisfying $\phi^{\perp\dagger}\phi^\perp = b$. $\square$

For the rest of Section 11, we fix framing data as in Lemma 11.6.2 if p is split, so that $\Theta(\overline{k})$ is surjective.

For the supersingular cases, we use the following lifting result to prove surjectivity of $\tilde{\Theta}$ by bootstrapping from surjectivity on $\overline{k}$ points (as in the proof of [RZ96, Theorem 6.30]). Recall that a p-divisible group X over a base scheme S is said to be *isoclinic* if for any geometric point $\overline{s}$ of S, the isocrystal of $X_{\overline{s}}$ has constant slope independent of $\overline{s}$.

Proposition 11.6.3 (Isoclinic lifting theorem)**.** *For any integer h, there exists an integer c with the following property: Let R be a reduced Noetherian Henselian local ring with residue field κ, and assume that R is an $\mathbb{F}_p$-algebra. Let X and Y be isoclinic p-divisible groups of heights $\leq h$ over* $\mathrm{Spec}\, R$*. For any homomorphism $f \colon X_\kappa \to Y_\kappa$, the homomorphism $p^c f$ lifts to a unique homomorphism $X \to Y$.*

Proof. See [OZ02, Corollary 3.4]. For the statement when $R = \kappa[\![t]\!]$ for an algebraically closed field κ (which is enough for Lemma 11.6.4), see also [Kat79, Theorem 2.7.1] combined with Grothendieck–Messing theory as in [RZ96, pg. 295]. $\square$

Lemma 11.6.4. *The uniformization map Θ is a surjection*[35] *of formal algebraic stacks.*

Proof. The reduced substack $\breve{\mathcal{Z}}(T)_{\mathrm{red}} \subseteq \breve{\mathcal{Z}}(T)$ is Jacobson, Deligne–Mumford, with quasi-compact diagonal, and finite type over $\operatorname{Spec} \overline{k}$. This implies that the closed points of $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ are dense in every closed subset (e.g. [SProject, Lemma 06G2]; the finite type points are the same as closed points here), each closed point is the image of a map $\operatorname{Spec} \overline{k} \to \breve{\mathcal{Z}}(T)_{\mathrm{red}}$, and every such map has image being a closed point.

Case p is nonsplit: We already know that Θ is surjective on $\overline{k}$ points. It is thus enough to prove the following claim: suppose $\alpha' \rightsquigarrow \alpha$ is an immediate specialization of points in $|\breve{\mathcal{Z}}(T)|$ (in the sense of [SProject, Definition 02I9], i.e. α is a point of "codimension one" in the closure of α'). If α is in the image of Θ, we claim that α' is also in the image of Θ. (This specialization process eventually terminates with a $\overline{k}$ point.)

Let κ an algebraically closed field with a morphism $\operatorname{Spec} \kappa[\![t]\!] \to \breve{\mathcal{Z}}(T)$, which sends the closed point to α and the open point to α'.[36] Enlarging κ if necessary, we may lift α to a point $(\alpha, \phi_0, \phi) \in \breve{\mathcal{Z}}(T)_{\mathrm{framed}}$. The task is to lift the framing pair (ϕ_0, ϕ) to $\operatorname{Spec} k[\![t]\!]$, which is then a framing pair for α'. Serre–Tate (and formal GAGA as in [EGAIII1, Théorème 5.4.1]) implies that it is enough to lift the induced quasi-isogenies of p-divisible groups to $\operatorname{Spec} k[\![t]\!]$. This is possible by the isoclinic lifting theorem (Proposition 11.6.3).

Case p is split: By Lemma 5.4.2 (finiteness of local special cycles), and since the groupoid $[I'(\mathbb{Q}) \backslash J_p(T)]$ has finite automorphism groups and finitely many isomorphism classes (Section 11.5; we assumed $\operatorname{rank}(T) \geq n-1$ for p split), we know there is a surjection from finitely many copies of $\operatorname{Spec} \overline{k}$ to $[I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]$. Since Θ is surjective on $\overline{k}$-points, the previous considerations show that $|\breve{\mathcal{Z}}(T)_{\mathrm{red}}|$ is a finite discrete topological space, and that Θ is a surjection. $\square$

Lemma 11.6.5. *The map $\tilde{\Theta}$ is proper on underlying reduced substacks, and the reduced substack $[I'(\mathbb{Q}) \backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]_{\mathrm{red}}$ is proper over* $\operatorname{Spec} \overline{k}$.

[35] We say a morphism of formal algebraic stacks is a *surjection* if it is surjective on underlying topological spaces.

[36] The following procedure produces such a morphism $\operatorname{Spec} \kappa[\![t]\!] \to \breve{\mathcal{Z}}(T)$. First, take an étale cover of $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ by a scheme U and lift $x' \rightsquigarrow x$ to an immediate specialization $y' \rightsquigarrow y$ on U. Write Z for the normalization of the integral closed subscheme of U with generic point y'. Note the normalization map is finite, and lift $y' \rightsquigarrow y$ to an immediate specialization $z' \rightsquigarrow z$ on Z. Completion of the local ring at z on Z is a power series ring over a field.

Proof. Since the reduced substack $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ is separated over $\operatorname{Spec} \overline{k}$, it is enough to check that $[I'(\mathbb{Q})\backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]_{\mathrm{red}}$ is proper over $\operatorname{Spec} \overline{k}$, by [SProject, Lemma 0CPT].

We already saw that $\tilde{\Theta}$ is a monomorphism, hence separated. Since $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ is separated, we see that $[I'(\mathbb{Q})\backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]_{\mathrm{red}}$ is also separated over $\operatorname{Spec} \overline{k}$.

We use the description of $\breve{\mathcal{Z}}(T)_{\mathrm{framed}}$ in (11.4.2). We know that every irreducible component of the reduced subscheme $\mathcal{N}'_{\mathrm{red}}$ is projective over $\overline{k}$ (Section 5.1), hence the same holds for $\mathcal{Z}'(\underline{\mathbf{x}}_p)$ for any $\underline{\mathbf{x}} \in \mathbf{W}^m$ (and $\mathcal{Z}'(\underline{\mathbf{x}}^p)$ is discrete). Hence each irreducible component of $\breve{\mathcal{Z}}(T)_{\mathrm{framed,red}}$ has closed image in $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$. Since Θ is surjective, we conclude that finitely many irreducible components of $\breve{\mathcal{Z}}(T)_{\mathrm{framed,red}}$ cover $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ (by Noetherianity of the latter). Since $\tilde{\Theta}$ is a monomorphism, hence injective on underlying topological spaces, we conclude that those finitely many irreducible components cover $[I'(\mathbb{Q})\backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]_{\mathrm{red}}$ as well. Then $[I'(\mathbb{Q})\backslash \breve{\mathcal{Z}}(T)_{\mathrm{framed}}]_{\mathrm{red}}$ is proper over $\operatorname{Spec} \overline{k}$ by [SProject, Lemma 0CQK]. $\square$

Proposition 11.6.6. *The map $\tilde{\Theta}$ is an isomorphism.*

Proof. We have seen that the morphism $\tilde{\Theta}$ of locally Noetherian formal algebraic stacks is formally étale, surjective, and a monomorphism. The underlying map of reduced substacks is proper. These properties imply that $\tilde{\Theta}$ is an isomorphism. $\square$

11.7 Global and local

The next lemma (purely linear-algebraic) helps us use uniformization to deduce properties of local special cycles via "approximating" them by global special cycles.

Lemma 11.7.1. *Let $L \subseteq \mathbf{W}_p$ be any non-degenerate Hermitian $\mathcal{O}_{F_p}$-lattice (of any rank). There exists an element $g_p \in U(\mathbf{W}_p)$ such that $g_p(L)$ admits a basis consisting of elements in $\mathbf{W}$.*

Proof. Set $W = L \otimes_{\mathcal{O}_{F_p}} F_p$. It is enough to produce $g_p \in U(\mathbf{W}_p)$ such that $g_p(W)$ admits an F_p-basis consisting of elements in $\mathbf{W}$ (since this implies that every full rank $\mathcal{O}_{F_p}$-lattice in $g_p(W)$ admits a basis consisting of elements of $\mathbf{W}$).

Select any F_p-basis $\underline{e} = [e_1, \ldots, e_d]$ for W. Since $\mathbf{W}$ is dense in $\mathbf{W}_p$, we may select $\underline{\tilde{e}} = [\tilde{e}_1, \ldots, \tilde{e}_d]$ such that each $\|\tilde{e}_i - e_i\|_p \ll 1$ for all i (meaning $\tilde{e}_i - e_i$ lies in a small neighborhood of 0 for the p-adic topology on $\mathbf{W}_p$). Set $\tilde{W} \coloneqq \operatorname{span}_{F_p}\{\tilde{e}_1, \ldots, \tilde{e}_d\}$. When each $\tilde{e}_i - e_i$ lies in a sufficiently small neighborhood of 0, there exists a (non-canonical) isomorphism of Hermitian spaces $W \cong \tilde{W}$ (the associated Gram matrices $(\underline{e}, \underline{e})$ and $(\underline{\tilde{e}}, \underline{\tilde{e}})$ can be made arbitrarily p-adically close; hence the local invariants $\varepsilon((\underline{e}, \underline{e}))$ and $\varepsilon((\underline{\tilde{e}}, \underline{\tilde{e}}))$ will agree). By Witt's theorem for Hermitian spaces, any isometry $W \to \tilde{W}$ extends to an isometry $g_p\colon \mathbf{W}_p \to \mathbf{W}_p$. This element $g_p \in U(\mathbf{W}_p)$ satisfies the conditions in the lemma statement. $\square$

Corollary 11.7.2. *Consider any tuple $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$ which spans a non-degenerate Hermitian $\mathcal{O}_{F_p}$-lattice, and write $m^\flat$ for its rank. Assume $m^\flat = n-1$ if p is split.*

For some $T \in \mathrm{Herm}_m(\mathbb{Q})$ (still assuming $\operatorname{rank} T \geq n-1$ if p is split), and some $j \in J_p(T)$ with associated $\underline{\mathbf{w}} \in \mathbf{W}$, there exists $g_p \in U(\mathbf{W}_p)$ inducing an automorphism $\mathcal{N} \to \mathcal{N}$ which takes $\mathcal{Z}(\underline{\mathbf{x}}_p)$ isomorphically to $\mathcal{Z}(\underline{\mathbf{w}}_p)$. In particular, there is an induced morphism

$$\mathcal{Z}(\underline{\mathbf{x}}_p) \xrightarrow{\sim} \mathcal{Z}(\underline{\mathbf{w}}_p) \xrightarrow{\Theta_j} \breve{\mathcal{Z}}(T) \tag{11.7.1}$$

which is representable by schemes, separated, and étale. If $m^\flat \geq n-1$ (equivalently, $\operatorname{rank}(T) \geq n-1$), this map is finite étale.

Proof. By Lemma 11.7.1, we may pick an element $g_p \in U(\mathbf{W}_p)$ so that $\operatorname{span}_{\mathcal{O}_{F_p}}(g_p \cdot \underline{\mathbf{x}}_p)$ admits an $\mathcal{O}_{F_p}$-basis $\underline{\mathbf{w}}^\flat$ of elements in $\mathbf{W}$. Extend $\underline{\mathbf{w}}^\flat$ to any m-tuple $\underline{\mathbf{w}} \in \mathbf{W}^m$, and set $T \coloneqq (\underline{\mathbf{w}}, \underline{\mathbf{w}})$. Recall that $U(\mathbf{W}_p)$ acts on $\mathcal{N}$, and that g_p gives an automorphism of $\mathcal{N}$ sending $\mathcal{Z}(\underline{\mathbf{x}}_p) \mapsto \mathcal{Z}(\underline{\mathbf{w}}_p)$ (Section 5.3). By uniformization (Proposition 11.6.6), any $j \in J_p(T)$ whose associated tuple is $\underline{\mathbf{w}}$ will satisfy the conditions of the lemma. Replacing $\underline{\mathbf{w}}$ with $a \cdot \underline{\mathbf{w}}$ for suitable $a \in \mathbb{Z}$ with $p \nmid a$ ensures $\mathcal{Z}'(\underline{\mathbf{w}}^p) \neq \emptyset$. Then such $j \in J_p(T)$ will exist. In Section 11.5, we saw that Θ_j is finite étale if $\operatorname{rank}(T) \geq n-1$. $\square$

If $p \neq 2$ and in signature $(n-1,1)$, the quasi-compactness proved in the next lemma is also [LZ22a, Lemma 2.9.] (inert), proved via Bruhat–Tits stratification. In the exotic smooth ramified case, quasi-compactness should be implicit in [LL22], via Bruhat–Tits stratification as discussed in [LL22, §2.3]. In the case when $\underline{\mathbf{x}}_p$ spans a lattice of rank n and signature $(n-1,1)$, see [LZ22a, Lemma 5.1.1] (inert, $p \neq 2$) and [LL22, Remark 2.26] (ramified, exotic smooth).

Lemma 11.7.3. *Let $\underline{\mathbf{x}}_p \in \mathbf{W}_p^d$ be any tuple which spans a non-degenerate Hermitian $\mathcal{O}_{F_p}$-lattice of rank $\geq n-1$. Then the local special cycle $\mathcal{Z}(\underline{\mathbf{x}}_p)$ is quasi-compact and the structure map $\mathcal{Z}(\underline{\mathbf{x}}_p) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is adic and proper.*

Proof. By Corollary 11.7.2, we obtain $T \in \mathrm{Herm}_m(\mathbb{Q})$ and a map $\mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ which is representable by schemes, and finite étale. In particular, $\mathcal{Z}(\underline{\mathbf{x}}_p)$ is quasi-compact because the (base-changed) global special cycle $\breve{\mathcal{Z}}(T)$ is quasi-compact.

If p is nonsplit, then $\mathcal{Z}(T)_{\overline{k}} \to \mathcal{M}_{\overline{k}}$ automatically factors through the supersingular locus (Corollary 4.7.3), so we have $\breve{\mathcal{Z}}(T) = \mathcal{Z}(T)_{\operatorname{Spf} \mathcal{O}_{\breve{F}_p}}$. This formula holds for p split as well, by definition. Hence $\breve{\mathcal{Z}}(T) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is adic[37] and proper (Lemma 4.7.5), so $\mathcal{Z}(\underline{\mathbf{x}}_p) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is adic and proper. $\square$

We write $\breve{\mathcal{Z}}(T)_{\mathscr{H}}$ for the flat[38] part ("horizontal") of $\breve{\mathcal{Z}}(T)$, i.e. the largest closed substack which is flat over $\operatorname{Spf} \mathcal{O}_{\breve{F}_p}$. We use similar notation $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}}$ for the flat part of the local special cycle $\mathcal{Z}(\underline{\mathbf{x}}_p)$.

Formation of "flat part" is flat local on the source. The quotient map Θ (11.4.1) is representable by schemes and étale, hence flat. So the uniformization result (Proposition 11.6.6) implies that there is an induced uniformization morphism

$$\Theta \colon \coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p)_{\mathscr{H}} \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \to \breve{\mathcal{Z}}(T)_{\mathscr{H}} \tag{11.7.3}$$

where $\mathcal{Z}'(\underline{\mathbf{x}}_p)_{\mathscr{H}}$ is the flat part of $\mathcal{Z}'(\underline{\mathbf{x}}_p)$. The action of $I'(\mathbb{Q})$ must preserve the flat part, so

[37] We say a morphism of formal algebraic stacks is *adic* if the morphism is representable by algebraic stacks in the sense of [Eme20, §3].

[38] Flatness for morphisms of locally Noetherian formal algebraic stacks was defined in [Eme20, Definition 8.42]. We are using a different definition, since the definition of loc. cit. does not recover the usual notion of flatness for morphisms of schemes (in the situation of [Eme20, Lemma 8.41(1)], consider $\mathcal{X} = \mathcal{Y} = \operatorname{Spec} k$ for a field k and any non-flat morphism of Noetherian affine k-schemes $U \to V$).

We define flatness in the style of [SProject, Section 06FL] (there for algebraic stacks), which recovers usual flatness for morphisms of locally Noetherian formal schemes. Let $f \colon X \to Y$ be a morphism of locally Noetherian formal algebraic spaces. Consider commutative diagrams

$$\begin{tikzcd} U \arrow[r, "h"] \arrow[d, "a"] & V \arrow[d, "b"] \\ X \arrow[r, "f"] & Y \end{tikzcd} \tag{11.7.2}$$

where U and V are locally Noetherian formal schemes and the vertical arrows are representable by schemes, flat, and locally of finite presentation. We say that f is *flat* if it satisfies either of the following equivalent conditions.

(1) For any diagram as above such that in addition $U \to X \times_Y V$ is flat, the morphism h is flat.

(2) For some diagram as above with $a \colon U \to X$ surjective, the morphism h is flat.

Next, consider a morphism $f \colon X \to Y$ of locally Noetherian formal algebraic stacks. Consider diagrams as above, but assume instead that U and V are locally Noetherian formal algebraic spaces, and that the arrows a and b are representable by algebraic spaces, flat, and locally of finite presentation. We say that f is *flat* if either of the equivalent conditions (1) and (2) as above are satisfied. If the morphism f is adic, then this agrees with the notion of flatness for adic morphisms as in [Eme20, Definition 3.11].

generalities on stack quotients imply that Θ induces an isomorphism

$$\tilde{\Theta}\colon \left[I'(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}'(\underline{\mathbf{x}}_p)_{\mathscr{H}} \times \mathcal{Z}'(\underline{\mathbf{x}}^p) \right) \right] \xrightarrow{\sim} \breve{\mathcal{Z}}(T)_{\mathscr{H}} \tag{11.7.4}$$

of formal algebraic stacks. For each $j \in J_p(T)$, the map $\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ induces a map $\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$ (reusing the notation Θ_j). Since Θ_j is flat and since formation of flat part is flat local on the source, the "horizontal" Θ_j arises from the original Θ_j by base-change along $\breve{\mathcal{Z}}(T)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)$.

In the case $p \neq 2$ and for signature $(n-1, 1)$ and $m^\flat = n - 1$, the next lemma is a consequence of [LZ22a, Theorem 4.2.1] (decomposition into quasi-canonical lifting cycles via Breuil modules) and is explained in [LL22, Lemma 2.49(1)] (also via decomposition into quasi-canonical lifting cycles). In the case $p \neq 2$, signature $(n-1,1)$, and $m^\flat = n$, see again [LZ22a, Lemma 5.1.1] (inert, $p \neq 2$) and [LL22, Remark 2.26] (ramified, exotic smooth).

Lemma 11.7.4. *Let $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$ be a tuple which spans a non-degenerate Hermitian $\mathcal{O}_{F_p}$-lattice, whose rank we denote $m^\flat$. Assume $m^\flat = n-1$ if $F/\mathbb{Q}_p$ is split. Form the horizontal part $\mathcal{Z}(\underline{x})_{\mathscr{H}}$ of $\mathcal{Z}(\underline{x})$.*

1. *If $\mathcal{Z}(\underline{x})_{\mathscr{H}}$ is nonempty, then it is equidimensional of dimension $(n-r)r + 1 - m^\flat r$.*
2. *If $m^\flat = n-1$ and the signature is $(n-r, r) = (n-1, 1)$, then the structure morphism $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is a finite adic morphism of Noetherian formal schemes. The associated finite scheme over $\operatorname{Spec} \mathcal{O}_{\breve{F}_p}$ has reduced generic fiber.*
3. *If $m^\flat = n$ and the signature is $(n-r, r) = (n-1, 1)$, then $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} = \emptyset$.*

Proof.

(1) By Corollary 11.7.2, we can find $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with $\mathrm{rank}(T) \geq n-1$ if p is split) and a morphism $\mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ which is representable by schemes and étale. As formation of flat part is flat local on the source, we obtain a morphism $\mathcal{Z}(\underline{\mathbf{x}})_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$ which is still representable by schemes and étale. The claim now follows from the corresponding global result for $\mathcal{Z}(T)_{\mathscr{H}}$ (Lemma 3.5.5). Note that we may assume K'_f is a small level (deepen the level away from p) to reduce to the case when $\breve{\mathcal{Z}}(T)_{\mathscr{H}}$ is a formal scheme.

(2) In this case, the map $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$ from part (1) is finite étale. We know that $\breve{\mathcal{Z}}(T)_{\mathscr{H}} \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ (with T as in the proof of loc. cit.) is proper and quasi-finite (Lemma 4.7.4). Since proper and quasi-finite implies finite (for morphisms of schemes)

and since $\breve{\mathcal{Z}}(\underline{\mathbf{x}}_p) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is adic, (already proved in Lemma 11.7.3) i.e. representable by schemes, the claimed finiteness holds. The claim on reducedness in the generic fiber follows because $\mathcal{Z}(T)_{\mathscr{H}} \to \operatorname{Spec} \mathcal{O}_F$ is étale in the generic fiber (Lemma 3.5.5). We are passing from finite relative schemes over $\operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ and $\operatorname{Spec} \mathcal{O}_{\breve{F}_p}$ as in Appendix B.3 (i.e. $\operatorname{Spf} R \mapsto \operatorname{Spec} R$).

(3) If $m^\flat = n$, then $\breve{\mathcal{Z}}(T)_{\mathscr{H}} = \emptyset$ (by Lemma 3.5.5), so $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} = \emptyset$ by existence of the map $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$. □

In the case of $p \neq 2$, signature $(n-1, 1)$, and $m^\flat = n-1$, the following lemma is [LZ22a, §2.9] (there proved differently, using their quasi-compactness result via Bruhat–Tits stratification).

Lemma 11.7.5 (Horizontal and vertical decomposition)**.** *Let $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$ be a tuple which span a non-degenerate Hermitian $\mathcal{O}_{F_p}$-lattice of rank $m^\flat$. Assume $m^\flat = n-1$ if $F/\mathbb{Q}_p$ is split. For $e \gg 0$, we have a scheme-theoretic union decomposition*

$$\mathcal{Z}(\underline{\mathbf{x}}_p) = \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \cup \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} \tag{11.7.5}$$

where $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} := \mathcal{Z}(\underline{\mathbf{x}}_p)_{\operatorname{Spf} \mathcal{O}_{\breve{F}_p}/p^e \mathcal{O}_{\breve{F}_p}}$.

Proof. If $\mathcal{I}$ denotes the ideal sheaf of $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}}$ as a closed subscheme of $\mathcal{Z}(\underline{\mathbf{x}}_p)$, it is enough to show that p^e annihilates $\mathcal{I}$ for $e \gg 0$. By Corollary 11.7.2, we can find $T \in \operatorname{Herm}_m(\mathbb{Q})$ (with $\operatorname{rank}(T) \geq n-1$ if p is split) and a morphism $f \colon \mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ which is representable by schemes and étale. We may assume K'_f is small (deepen the level away from p) so that $\breve{\mathcal{Z}}(T)$ is a formal scheme.

If $\mathscr{I}$ denotes the ideal sheaf of the flat part $\breve{\mathcal{Z}}(T)_{\mathscr{H}} \subseteq \breve{\mathcal{Z}}(T)$, then $f^*\mathscr{I} \to \mathcal{I}$ is surjective (by flatness of f, i.e. formation of flat part is flat local on the source). If p^e annihilates $\mathscr{I}$, then p^e also annihilates $\mathcal{I}$. We know that $\mathscr{I}$ consists (locally) of p-power torsion elements in the structure sheaf. Since $\breve{\mathcal{Z}}(T)$ is quasi-compact, we know that $\mathscr{I}$ is annihilated by p^e for $e \gg 0$. □

For the rest of Section 11, we restrict to signature $(n-1, 1)$ in all cases.

Lemma 11.7.6.

(1) If p is split, then $\breve{\mathcal{Z}}(T) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is proper and quasi-finite and we have ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} = 0$.

(2) Assume $n = 2$ and $\operatorname{rank}(T) \geq 1$. Then $\breve{\mathcal{Z}}(T) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is proper and quasi-finite. If $\operatorname{rank}(T) = 1$, then we have ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} = 0$.

Proof. (1) Recall our running assumption that $\operatorname{rank}(T) \geq n-1$ if p is split. Recall also $\breve{\mathcal{Z}}(T) \coloneqq \mathcal{Z}(T)_{\operatorname{Spf} \mathcal{O}_{\breve{F}_p}}$ in the split case, so the map $\breve{\mathcal{Z}}(T) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is representable by algebraic stacks and locally of finite type. This map is proper on reduced substacks by uniformization (Lemma 11.6.5 and Proposition 11.6.6), hence it is proper.

It remains to check that $\breve{\mathcal{Z}}(T) \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$ is quasi-finite in the sense of [SProject, Definition 0G2M]. It is enough to check that $\breve{\mathcal{Z}}(T)_{\mathrm{red}} \to \operatorname{Spec} \overline{k}$ is quasi-finite. This follows from the uniformization isomorphism, since $\breve{\mathcal{Z}}(T)_{\mathrm{red}}$ may be covered by finitely many copies of $\operatorname{Spec} \overline{k}$ (combine uniformization with the analogous result for local special cycles, which is Lemma 5.4.2; since we assume $\operatorname{rank}(T) \geq n-1$ when p is split, the groupoid $[I'(\mathbb{Q}) \backslash J_p(T)]$ has finitely many isomorphism classes, as discussed in Section 11.5).

The derived vertical special cycle class ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \operatorname{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}}$ was defined in Section 4.6. If $m \geq n$ then $\mathcal{Z}(T)$ is empty. If $m = n-1$, then $\operatorname{gr}^m_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} = 0$ because $\mathcal{Z}(T)_{\mathbb{F}_p}$ has dimension 0 (and $\mathcal{M}$ has dimension n).

(2) This may be proved as in part (1). We may assume p is nonsplit. We have $\breve{\mathcal{Z}}(T) = \mathcal{Z}(T)_{\operatorname{Spf} \mathcal{O}_{\breve{F}_p}}$ (Lemma 4.7.3). Then use quasi-compactness of local special cycles (Lemma 11.7.3), uniformization, and discreteness of $\mathcal{N}_{\mathrm{red}}$ (Section 5.4). Suppose $\operatorname{rank}(T) = 1$. First consider the case $m = 1$. Then ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \operatorname{gr}^1_{\mathcal{M}} K'_0(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}} = 0$ because $\mathcal{Z}(T)_{\mathbb{F}_p}$ has dimension 0 (and $\mathcal{M}$ has dimension 2). If $m = 2$, then ${}^{\mathbb{L}}\mathcal{Z}(t_i)_{\mathscr{V},p} = 0$ for any nonzero diagonal entry t_i of T by the preceding argument, so ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} = 0$ by construction (defined in Section 4.6 as the projection of a product against ${}^{\mathbb{L}}\mathcal{Z}(t_i)_{\mathscr{V},p} = 0$ for some i). □

Lemma 11.7.7. *Assume p is nonsplit. Assume that K'_f is a small level, so that $\mathcal{M}$ is a scheme. Fix any $j \in J_p(T)$ and consider the map*

$$\mathcal{Z}(\underline{\mathbf{x}}_p) \xrightarrow{\Theta_j} \breve{\mathcal{Z}}(T) \to \mathcal{Z}(T). \tag{11.7.6}$$

The class ${}^{\mathbb{L}}\mathcal{Z}(T) \in K'_0(\mathcal{Z}(T))_{\mathbb{Q}}$ pulls back to the class ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p) \in K'_0(\mathcal{Z}(\underline{\mathbf{x}}_p))_{\mathbb{Q}}$.

Proof. The maps $\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ and $\breve{\mathcal{Z}}(T) \to \mathcal{Z}(T)$ are flat maps of locally Noetherian formal schemes, so we may take the non-derived pullback. The lemma may be proved using the fact that the commutative diagrams

$$\begin{tikzcd} \breve{\mathcal{Z}}(T)_{\mathrm{framed}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & \breve{\mathcal{M}}_{\mathrm{framed}} \arrow[d] \\ \mathcal{Z}(T) \arrow[r] & \mathcal{M} \end{tikzcd} \qquad \begin{tikzcd} \breve{\mathcal{Z}}(t_i)_{\mathrm{framed}} \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & \breve{\mathcal{M}}_{\mathrm{framed}} \arrow[d] \\ \mathcal{Z}(t_i) \arrow[r] & \mathcal{M} \end{tikzcd} \tag{11.7.7}$$

are 2-Cartesian (where the t_i are the diagonal entries of T), and the fact that the tautological bundle $\mathcal{E}$ on $\mathcal{M}$ pulls back to the tautological bundle $\mathcal{E}$ on $\mathcal{N}$. □

Corollary 11.7.8. *Assume p is nonsplit. For any $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$, we have ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p) \in F_{\mathcal{N}}^m K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p))_{\mathbb{Q}}$.*

Proof. By Corollary 11.7.2, we can find $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with $\mathrm{rank}(T) \geq n-1$ if p is split) and a morphism $\mathcal{Z}(\underline{\mathbf{x}}_p) \to \breve{\mathcal{Z}}(T)$ which is representable by schemes and étale. We can deepen the level K_f' away from p to assume $\breve{\mathcal{Z}}(T)$ is a formal scheme. Since ${}^{\mathbb{L}}\mathcal{Z}(T) \in F_{\mathcal{N}}^m K_0'(\mathcal{Z}(T))_{\mathbb{Q}}$, the corollary follows from Lemma 11.7.7. □

We previously defined derived vertical (global) special cycles ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \mathrm{gr}_{\mathcal{M}}^m K_0'(\mathcal{Z}(T)_{\mathbb{F}_p})_{\mathbb{Q}}$ (Section 4.6). In the next lemma, we write $\mathcal{Z}(T)_{(p)} := \mathcal{Z}(T) \times_{\operatorname{Spec}\mathbb{Z}} \operatorname{Spec}\mathbb{Z}_{(p)}$. We also write $\mathcal{Z}(T)_{\mathscr{V},p} := \mathcal{Z}(T) \times_{\operatorname{Spec}\mathbb{Z}} \operatorname{Spec}\mathbb{Z}/p^e\mathbb{Z}$ and $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V},p} := \mathcal{Z}(\underline{\mathbf{x}}_p) \times_{\operatorname{Spf}\mathcal{O}_{\breve{F}_p}} \operatorname{Spec}\mathcal{O}_{\breve{F}_p}/p^e\mathcal{O}_{\breve{F}_p}$ for an understood integer $e \gg 0$. We also set $\breve{\mathcal{Z}}(T)_{\mathscr{V}} := \breve{\mathcal{Z}}(T) \times_{\operatorname{Spf}\mathcal{O}_{\breve{F}_p}} \operatorname{Spec}\mathcal{O}_{\breve{F}_p}/p^e\mathcal{O}_{\breve{F}_p}$.

Lemma 11.7.9. *Fix any $j \in J_p(T)$. Write $\underline{\mathbf{x}} \in \mathbf{W}^m$ for the associated m-tuple. Fix any $e \gg 0$ such that there are scheme-theoretic union decompositions*

$$\mathcal{Z}(T)_{(p)} = \mathcal{Z}(T)_{(p),\mathscr{H}} \cup \mathcal{Z}(T)_{\mathscr{V},p} \quad \mathcal{Z}(\underline{\mathbf{x}}_p) = \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \cup \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V},p} \tag{11.7.8}$$

("horizontal and vertical"). Pullback along the map

$$\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} \xrightarrow{\Theta_j} \breve{\mathcal{Z}}(T)_{\mathscr{V}} \to \mathcal{Z}(T)_{\mathscr{V},p} \tag{11.7.9}$$

sends ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \in \mathrm{gr}_{\mathcal{M}}^m K_0'(\mathcal{Z}(T)_{\mathscr{V},p})_{\mathbb{Q}}$ to ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} \in \mathrm{gr}_{\mathcal{N}}^m K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}})_{\mathbb{Q}}$.

Proof. If p is split, the derived vertical special cycles (global and local) are zero (Lemma 11.7.6 (global) and Section 5.5 (local)) and the lemma is trivial. We remind the reader of our running assumption that $\mathrm{rank}(T) \geq n-1$ if p is split.

We thus assume that p is nonsplit. By the local and global linear invariance results (Section 5.5 and (4.6.11)), it is enough to check the case where $T = \mathrm{diag}(0, T^\flat)$ where $\det T^\flat \neq 0$.

First consider the case where T is nonsingular, i.e. $T = T^\flat$. If K_f' is a small level, the lemma follows from Lemma 11.7.7, since the projections $\mathrm{gr}_{\mathcal{M}}^{m^\flat} K_0'(\mathcal{Z}(T^\flat))_{\mathbb{Q}} \to \mathrm{gr}_{\mathcal{M}}^{m^\flat} K_0'(\mathcal{Z}(T^\flat)_{\mathscr{V},p})_{\mathbb{Q}}$ and $\mathrm{gr}_{\mathcal{N}}^{m^\flat} K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p))_{\mathbb{Q}} \to \mathrm{gr}_{\mathcal{N}}^{m^\flat} K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}})_{\mathbb{Q}}$ are given by (non-derived) pullbacks of coherent sheaves, see Lemma A.1.5 (Deligne–Mumford stacks) and [Zha21, Lemma B.1] (locally Noetherian formal schemes). Note that the codimension graded pieces gr^m are preserved, by étale-ness of Θ_j. In general, we may reduce to the case where K_f' is a small level by compatibility of ${}^{\mathbb{L}}\mathcal{Z}(T^\flat)_{\mathscr{V},p}$ with (finite étale) pullback for varying levels (Section 4.6).

Next, consider the case where T is possibly singular with $T = \mathrm{diag}(0, T^\flat)$. If K_f' is a small level, this follows as in the proof of Lemma 11.7.7. That is, the class ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} := (\mathcal{E}^\vee)^{m-\mathrm{rank}(T)} \cdot \mathcal{Z}(T^\flat)_{\mathscr{V},p}$ pulls back to ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} := (\mathcal{E}^\vee)^{m-\mathrm{rank}(T)} \cdot \mathcal{Z}(\underline{\mathbf{x}}_p^\flat)_{\mathscr{V}}$ (use the result for $T^\flat$ just proved). In general, we may reduce to the case where K_f' is a small level (deepen level away from p) by compatibility of ${}^{\mathbb{L}}\mathcal{Z}(T^\flat)_{\mathscr{V},p}$ with (finite étale) pullback for varying levels (Section 4.6). □

11.8 Local intersection numbers: vertical

The main purpose of this section is to reduce "global vertical intersection numbers" to "local vertical intersection numbers" (see end of this section). We continue to assume signature $(n-1,1)$.

Consider $T' \in \mathrm{Herm}_m(\mathbb{Q}_p)$ (with F_p-coefficients) with $\mathrm{rank}(T') = n-1$, and either $m = n-1$ or $m = n$. For any tuple $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$ with Gram matrix T', we define the *local vertical intersection number*

$$\mathrm{Int}_{\mathscr{V},p}(T') := \begin{cases} 2[\breve{F}_p : \breve{\mathbb{Q}}_p]^{-1} \deg_{\overline{k}}({}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} \cdot \mathcal{E}^\vee) \log p & \text{if } m = n-1 \\ 2[\breve{F}_p : \breve{\mathbb{Q}}_p]^{-1} \deg_{\overline{k}}({}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}}) \log p & \text{if } m = n \end{cases} \tag{11.8.1}$$

Here, $\mathcal{E}^\vee$ stands for the class $[\mathcal{O}_{\mathcal{N}}] - [\mathcal{E}] \in K_0'(\mathcal{N})$. If no such $\underline{\mathbf{x}}_p$ exists, we set $\mathrm{Int}_{\mathscr{V},p}(T') := 0$. The definition of $\mathrm{Int}_{\mathscr{V},p}(T')$ does not depend on the choice of $\underline{\mathbf{x}}_p$ (by the action of $U(\mathbf{W}_p)$ on $\mathcal{N}(n-1,1)$, Section 5.3). The factor $2[\breve{F}_p : \breve{\mathbb{Q}}_p]^{-1}$ will account for total degree of $\mathrm{Spec}\,\mathcal{O}_{F_p} \to \mathrm{Spec}\,\mathbb{Z}_p$ on residue fields (e.g. we need to account for both primes in $\mathcal{O}_F$ over p in the split case). By local linear invariance (Section 5.5), we have

$$\mathrm{Int}_{\mathscr{V},p}(T') = \mathrm{Int}_{\mathscr{V},p}({}^t\overline{\gamma} T' \gamma) \tag{11.8.2}$$

for any $\gamma \in \mathrm{GL}_m(\mathcal{O}_{F_p})$.

Consider any $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with F-coefficients) with $\mathrm{rank}(T) = n-1$, and either $m = n-1$ or $m = n$. Pick any set of representatives $J \subseteq J_p(T)$ for the isomorphism classes of the groupoid $[I'(\mathbb{Q}) \backslash J_p(T)]$. By Lemma 11.7.9, we have

$$\mathrm{Int}_{\mathscr{V},p,\mathrm{global}}(T) := \deg_{\mathbb{F}_p}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p} \cdot (\mathcal{E}^\vee)^{n-m}) \log p \tag{11.8.3}$$

$$= \mathrm{Int}_{\mathscr{V},p}(T) \sum_{j \in J} \frac{1}{|\operatorname{Aut}(j)|} \tag{11.8.4}$$

$$= \mathrm{Int}_{\mathscr{V},p}(T) \frac{[K_{L_0,f} : K_{0,f}]}{|\mathcal{O}_F^\times|/h_F} \cdot \deg \left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right].$$

For later use in Remark 22.1.2, consider $T \in \mathrm{Herm}_n(\mathbb{Q})$ (with F-coefficients) with $\det T \neq 0$. We consider the *local intersection number*

$$\mathrm{Int}_p(T) := 2[\breve{F}_p : \breve{\mathbb{Q}}_p]^{-1} \deg_{\overline{k}}({}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)) \log p \tag{11.8.5}$$

where $\underline{\mathbf{x}}_p \in \mathbf{W}_p^n$ is any n-tuple with Gram matrix T (since $\mathrm{rank}\,\mathbf{W}_p = n-1$ when p is split, set $\mathrm{Int}_p(T) := 0$ in this case). Note ${}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{V}} = {}^{\mathbb{L}}\mathcal{Z}(\underline{\mathbf{x}}_p)$ by Lemmas 11.7.4 and 11.7.3 (under the dévissage pushforward identification $K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p)_{\overline{k}}) \xrightarrow{\sim} K_0'(\mathcal{Z}(\underline{\mathbf{x}}_p)))$. By Lemma 11.7.9, we have

$$\begin{aligned}\mathrm{Int}_{p,\mathrm{global}}(T) &\coloneqq \deg_{\mathbb{F}_p}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},p}) \log p && (11.8.6)\\ &= \mathrm{Int}_p(T) \frac{[K_{L_0,f} : K_{0,f}]}{|\mathcal{O}_F^\times|/h_F} \cdot \deg \left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^n \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right]. && (11.8.7)\end{aligned}$$

11.9 Local intersection numbers: horizontal

The main purpose of this section is to reduce "global horizontal intersection numbers" to "local horizontal intersection numbers" (see end of this section). We continue to assume signature $(n-1,1)$. In Section 11.9, we require $p \neq 2$ if p is inert (because we required this for our discussion of quasi-canonical lifting cycles, Section 7.3).

Consider $T' \in \mathrm{Herm}_m(\mathbb{Q}_p)$ (with F_p-coefficients) with $\mathrm{rank}(T') = n-1$, and either $m = n-1$ or $m = n$. Select any $\underline{\mathbf{x}}_p \in \mathbf{W}_p^m$ with Gram matrix T', and set $L_p^\flat \coloneqq \mathrm{span}_{\mathcal{O}_{F_p}}(\underline{\mathbf{x}}_p)$. We define the *local horizontal intersection number*

$$\mathrm{Int}_{\mathscr{H},p}(T') \coloneqq \sum_{\substack{L_p^\flat \subseteq M_p^\flat \subseteq M_p^{\flat *} \\ t(M_p^\flat) \leq 1}} \mathrm{Int}_{\mathscr{H},p}(M_p^\flat)^\circ \tag{11.9.1}$$

where the sum runs over lattices $M_p^\flat \subseteq L_p^\flat \otimes_{\mathcal{O}_{F_p}} F_p$, where

$$\mathrm{Int}_{\mathscr{H},p}(M_p^\flat)^\circ \coloneqq 2 \cdot \deg \mathcal{Z}(M_p^\flat)^\circ \cdot \delta_{\mathrm{tau}}(\mathrm{val}'(M_p^\flat)) \tag{11.9.2}$$

with $\mathrm{val}'(M_p^\flat) \coloneqq \lfloor \mathrm{val}(M_p^\flat) \rfloor$ and with $\delta_{\mathrm{tau}}(-)$ the "local change of tautological height" as defined in (7.2.7). Here $\mathcal{Z}(M_p^\flat)^\circ \subseteq \mathcal{N}(n-1,1)$ is the quasi-canonical lifting cycle associated with $M_p^\flat$ (Section 7.3). The local horizontal intersection number should be compared with the decomposition of horizontal local special cycles into quasi-canonical lifting cycles (Section 7.3). The notation $\deg \mathcal{Z}(M_p^\flat)^\circ$ means the degree of the finite flat adic morphism $\mathcal{Z}(M_p^\flat)^\circ \to \mathrm{Spf}\, \mathcal{O}_{\breve{F}_p}$. If no such $\underline{\mathbf{x}}_p$ exists, we set $\mathrm{Int}_{\mathscr{H},p}(T') \coloneqq 0$.

This definition of $\mathrm{Int}_{\mathscr{H},p}(T')$ does not depend on the choice of $\underline{\mathbf{x}}_p$ (again by the action of $U(\mathbf{W}_p)$ on $\mathcal{N}$ Section 5.3 and Witt's theorem). The formula for $\deg \mathcal{Z}(M_p^\flat)^\circ$ (combine (7.2.1) and (7.3.1)) shows $\mathrm{Int}_{\mathscr{H},p}(M_p^\flat)^\circ \in \mathbb{Z}$. The extra factor of 2 in (11.9.2) will account for the fact that $\mathrm{Spf}(\mathcal{O}_F \otimes_{\mathbb{Z}} \breve{\mathbb{Z}}_p) \to \mathrm{Spf}\, \breve{\mathbb{Z}}_p$ has degree 2.

In the above situation, we also set

$$\deg_{\mathscr{H},p}(T') \coloneqq \deg \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \tag{11.9.3}$$

where the right-hand side means the degree of the finite flat adic morphism $\mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \operatorname{Spf} \mathcal{O}_{\breve{F}_p}$. If no such $\underline{\mathbf{x}}_p$ exists, we set $\deg_{\mathscr{H},p}(T') = 0$. Again, the definition of $\deg_{\mathscr{H},p}(T')$ does not depend on the choice of $\underline{\mathbf{x}}_p$.

Suppose $T \in \operatorname{Herm}_m(\mathbb{Q})$ (with F-coefficients) with $\operatorname{rank}(T) = n-1$, and either $m = n-1$ or $m = n$. Then (in the notation of Sections 4.3 and 4.1) we have

$$\widehat{\deg}(\widehat{\mathcal{E}}^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}) = \widehat{\deg}(\widehat{\Omega}_0^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}) + \widehat{\deg}(\widehat{\mathscr{E}}^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}). \tag{11.9.4}$$

We have

$$\widehat{\deg}(\widehat{\Omega}_0^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}) = \deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}} \cdot (-h_{\mathrm{Fal}}^{\mathrm{CM}} - \frac{1}{4}\log|\Delta|) \tag{11.9.5}$$

where $\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}$ means the degree of $\mathcal{Z}(T)_{\mathscr{H}} \times_{\operatorname{Spec}\mathbb{Z}} \operatorname{Spec}\mathbb{Q} \to \operatorname{Spec}\mathbb{Q}$ (stacky degrees as in (A.1.10) and surrounding discussion).

Pick any set of representatives $J \subseteq J_p(T)$ for the isomorphism classes of the groupoid $[I'(\mathbb{Q}) \backslash J_p(T)]$. Using the finite étale maps $\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$ for $j \in J_p(T)$ (Section 11.5 and (11.7.3)) which cover $\breve{\mathcal{Z}}(T)_{\mathscr{H}}$ as j ranges over J, we find

$$\begin{aligned}\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}} &= \deg_{\mathscr{H},p}(T) \sum_{j \in J_p(T)} \frac{1}{|\operatorname{Aut}(j)|} \\ &= \deg_{\mathscr{H},p}(T) \frac{[K_{L_0,f} : K_{0,f}]}{|\mathcal{O}_F^\times|/h_F} \cdot \deg\left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}},\underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right].\end{aligned} \tag{11.9.6}$$

Combining the following: (1) the finite étale maps $\Theta_j \colon \mathcal{Z}(\underline{\mathbf{x}}_p)_{\mathscr{H}} \to \breve{\mathcal{Z}}(T)_{\mathscr{H}}$ for $j \in J_p(T)$ (Section 11.5 and (11.7.3)) which cover $\breve{\mathcal{Z}}(T)_{\mathscr{H}}$ as j ranges over J (2) Proposition 7.3.1 (decomposition of horizontal local special cycles into quasi-canonical liftings) and discussion surrounding (7.3.2), and (3) Corollary 10.2.2 (decomposition of global height into local "change of heights" for p-divisible groups), we find

$$\begin{aligned}\operatorname{Int}_{\mathscr{H},p,\mathrm{global}}(T) &\coloneqq \widehat{\deg}(\widehat{\mathscr{E}}^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}) - (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot h_{\mathrm{tau}}^{\mathrm{CM}} \mod \sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell \\ &= \operatorname{Int}_{\mathscr{H},p}(T) \sum_{j \in J_p(T)} \frac{1}{|\operatorname{Aut}(j)|} \\ &= \operatorname{Int}_{\mathscr{H},p}(T) \frac{[K_{L_0,f} : K_{0,f}]}{|\mathcal{O}_F^\times|/h_F} \cdot \deg\left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^m \\ (\underline{\mathbf{x}},\underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^\perp)/K_{1,\mathbf{L}_p^\perp} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right]\end{aligned} \tag{11.9.7}$$

with $h_{\mathrm{tau}}^{\mathrm{CM}}$ and $h_{\widehat{\mathcal{E}}^\vee}^{\mathrm{CM}}$ as in (4.3.6). The notation "mod $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$" means that equality holds as elements of the (additive) quotient $\mathbb{R}/(\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell)$. Note $\operatorname{Int}_{\mathscr{H},p,\mathrm{global}}(T) \in \mathbb{Q} \cdot \log p$. To apply Corollary 10.2.2, we first consider the case of small level K'_f so that $\mathcal{Z}(T)_{\mathscr{H}}$ is a scheme. This immediately implies the case of general (stacky) level, by compatibility of arithmetic degree with finite étale covers (see Section 4.1). We have

$$\widehat{\deg}(\widehat{\mathcal{E}}^\vee|_{\mathcal{Z}(T)_{\mathscr{H}}}) - (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^\vee} = \sum_\ell \mathrm{Int}_{\mathscr{H},\ell,\mathrm{global}}(T) \tag{11.9.8}$$

where the sum ranges over all primes ℓ, with all but finitely many terms equal to 0. The preceding expression should be understood modulo $\mathbb{Q} \cdot \log \ell$ for those primes ℓ for which $L \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ is not self-dual. If L is not self-dual, we also quotient by $\mathbb{Q} \cdot \log \ell$ for primes $\ell \mid \Delta$. We also quotient by $\mathbb{Q} \cdot \log 2$ unless 2 is split in $\mathcal{O}_F$.

We define *total "intersection numbers"*

$$\mathrm{Int}_p(T) := \mathrm{Int}_{\mathscr{H},p}(T) + \mathrm{Int}_{\mathscr{V},p}(T) \qquad \mathrm{Int}_{p,\mathrm{global}}(T) := \mathrm{Int}_{\mathscr{H},p,\mathrm{global}}(T) + \mathrm{Int}_{\mathscr{V},p,\mathrm{global}}(T) \tag{11.9.9}$$

(local and global) at p. These will feature in our main non-Archimedean local theorems (Section 18) and the proof of our main global theorem (Theorem 22.1.1) respectively.

For readers interested in Faltings heights, we record the relation

$$\widehat{\deg}(\widehat{\omega}|_{\mathcal{Z}(T)_{\mathscr{H}}}) - (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot n \cdot h^{\mathrm{CM}}_{\mathrm{Fal}} = -2 \sum_\ell \mathrm{Int}_{\mathscr{H},\ell,\mathrm{global}}(T) \tag{11.9.10}$$

where $\widehat{\omega}$ is the metrized Hodge determinant bundle (Section 4.3), the sum again runs over all primes ℓ, and where $\mathrm{Int}_{\mathscr{H},\ell,\mathrm{global}}(T)$ is the same quantity defined above. This follows by the same argument as above, using Corollary 10.2.2. The remarks following (11.9.8) (about quotienting by $\mathbb{Q} \cdot \log \ell$ for some primes ℓ) apply here verbatim.

12 Archimedean

We explain complex uniformization for special cycles and Green currents on $\mathcal{M}$. The only new part of Section 12 is our treatment of Green currents for singular T in Section 12.4, when $\mathrm{rank}(T) = n - 1$. The remaining material should be fairly standard, e.g. [KR14, §3] (uniformization of special cycles), [BHKRY20, §2] (including discussion of metrized tautological bundle), [Liu11, §4B] (Green currents via uniformization), etc.. Strictly speaking, however, the references [KR14; BHKRY20] restrict to principal polarizations. We will need non-principal polarizations (this slightly affects how we normalize the metric on the tautological bundle), so we explain the setup.

With notation as explained at the beginning of Part IV, we also assume L has signature $(n-1, 1)$. For technical convenience, we assume the implicit level K'_f is small so that $\mathcal{M}$ is a scheme (except at the very end of Section 12.4). Fix one of the two embeddings $F \to \mathbb{C}$, write $\mathcal{M}_{\mathbb{C}} := \mathcal{M} \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$, and let $\mathcal{M}_{\mathbb{C}}^{\mathrm{an}}$ be the analytification (outside of Section 12, we often abuse notation and drop the superscript an). This is a complex manifold of dimension $n-1$. Given any Hermitian matrix $T \in \operatorname{Herm}_m(\mathbb{Q})$ (with F-coefficients) with associated special cycle $\mathcal{Z}(T) \to \mathcal{M}$, we use similar notation $\mathcal{Z}(T)_{\mathbb{C}}$ and $\mathcal{Z}(T)_{\mathbb{C}}^{\mathrm{an}}$. Since $\mathcal{Z}(T)_{\mathbb{C}} \to \operatorname{Spec} \mathbb{C}$ is smooth (Lemma 3.5.5), we know that $\mathcal{Z}(T)_{\mathbb{C}}^{\mathrm{an}}$ is also a complex manifold.

We view $V_{\mathbb{R}}$ as a complex vector space via the identification $F \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathbb{C}$ (induced by the choice of $F \to \mathbb{C}$). We use notation from Section 8 on the Hermitian symmetric space $\mathcal{D}$ and its local special cycles $\mathcal{D}(\underline{x})$ for tuples $\underline{x} \in V_{\mathbb{R}}^m$, etc..

We set $V_{\mathbb{C}} := V \otimes_{\mathbb{Q}} \mathbb{C}$ and write $V_{\mathbb{C}} = V_{\mathbb{C}}^+ \oplus V_{\mathbb{C}}^-$ where the F-action on $V_{\mathbb{C}}^+$ (resp. $V_{\mathbb{C}}^-$) is F-linear (resp. σ-linear) with respect to the chosen map $F \to \mathbb{C}$. We use similar notation for other $F \otimes_{\mathbb{Q}} \mathbb{C}$-modules.

12.1 Local special cycles away from ∞

Given an m-tuple $\underline{x}_f = [x_1, \ldots, x_m] \in (V \otimes_{\mathbb{Q}} \mathbb{A}_f)^m$, we consider an "*away-from-*∞" local special cycle

$$\mathcal{D}'(\underline{x}_f) := \{(g_0, g) \in G'(\mathbb{A}_f)/K'_f : g^{-1} g_0 x_i \in L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \text{ for all } x_i \in \underline{x}_f\}. \tag{12.1.1}$$

We view $\mathcal{D}'(\underline{x}_f)$ as a discrete set. We also define the "*away-from-*∞" local special cycle

$$\mathcal{D}(\underline{x}_f) := \{g \in U(V)(\mathbb{A}_f)/K_f : g^{-1} x_i \in L \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}^p \text{ for all } x_i \in \underline{x}_f\}. \tag{12.1.2}$$

The isomorphism $G'(\mathbb{A}_f)/K'_f \to GU(V_0)(\mathbb{A}_f)/K_{0,f} \times U(V)(\mathbb{A}_f)/K_f$ $(*)$ induces a bijection

$$\mathcal{D}'(\underline{x}_f) \xrightarrow{\sim} GU(V_0)(\mathbb{A}_f)/K_{0,f} \times \mathcal{D}(\underline{x}_f). \tag{12.1.3}$$

12.2 Framing

Fix the isomorphism of Hermitian $\mathcal{O}_F$-lattices $\operatorname{Hom}_{\mathcal{O}_F}(L_0, L) \to L$ sending $x \mapsto x(1)$ (with $1 \in L_0$). This is analogous to $\boldsymbol{\eta}$ from (11.3.2).

Given $\alpha = (A_0, \iota_0, \lambda_0, A, \iota, \lambda, \tilde{\eta}_0, \tilde{\eta}) \in \mathcal{M}_{\mathbb{C}}^{\mathrm{an}}$, a *framing pair* (ϕ_0, ϕ) for α consists of isomorphisms of F vector spaces (singular homology)

$$\phi_0 \colon H_1(A_0, \mathbb{Q}) \to V_0 \qquad \phi \colon H_1(A, \mathbb{Q}) \to V \tag{12.2.1}$$

such that the induced map

$$\phi_0^{-1}\phi \colon \operatorname{Hom}_F(H_1(A_0, \mathbb{Q}), H_1(A, \mathbb{Q})) \to \operatorname{Hom}_F(V_0, V) = V \tag{12.2.2}$$

is an isomorphism of Hermitian spaces.

The Hodge structures of weight -1 on $H_1(A_0, \mathbb{Q})$ and $H_1(A, \mathbb{Q})$ induce a Hodge structure of weight 0 on V, with an associated complex line $F^1 V_{\mathbb{C}} \subseteq V_{\mathbb{C}}^+$. After pullback along the projection isomorphism $V_{\mathbb{R}} \to V_{\mathbb{C}}^+$ of $F \otimes_{\mathbb{Q}} \mathbb{R}$ vector spaces, the line $F^1 V_{\mathbb{C}} \subseteq V_{\mathbb{R}}$ is a negative definite subspace and hence defines a point $z \in \mathcal{D}$. There is a canonical isomorphism of $\mathbb{C}$ vector spaces

$$\mathrm{Hom}_{F \otimes_{\mathbb{Q}} \mathbb{R}}(\mathrm{Lie}\, A_0, F^0 H_1(A, \mathbb{Q})^+_{\mathbb{C}}) \cong F^1 V_{\mathbb{C}}. \tag{12.2.3}$$

We use the fixed choice of $\sqrt{\Delta}$ to pass between Hermitian/alternating/symmetric forms (Section 2.1). This makes $H_1(A_0, \mathbb{Q})$ and $H_1(A, \mathbb{Q})$ into Hermitian F-modules. Using the $\mathbb{C}$-bilinear extension of the symmetric $\mathbb{Q}$-bilinear trace pairing on $H_1(A, \mathbb{Q})$, we obtain an induced $\mathbb{C}$-linear identification $F^0 H_1(A, \mathbb{Q})^+_{\mathbb{C}} \cong \mathrm{Hom}_{\mathbb{C}}((\mathrm{Lie}\, A)^-, \mathbb{C})$.

We equip $F^1 V_{\mathbb{C}} \subseteq V_{\mathbb{R}}$ with the Hermitian metric obtained by restricting the metric on $V_{\mathbb{R}}$. Equip $\mathrm{Lie}\, A_0$ (resp. $\mathrm{Lie}\, A$) with the Hermitian metric as normalized in (4.3.2) (resp. (4.3.3)). Then $(\mathrm{Lie}\, A)^- \subseteq \mathrm{Lie}\, A$ inherits a Hermitian metric as well. Under the isomorphism

$$\mathrm{Hom}_{\mathbb{C}}(\mathrm{Lie}\, A_0, \mathbb{C}) \otimes \mathrm{Hom}_{\mathbb{C}}((\mathrm{Lie}\, A)^-, \mathbb{C}) \cong F^1 V_{\mathbb{C}} \tag{12.2.4}$$

induced by (12.2.3), the Hermitian metric on the left is $-(16\pi^3 e^{\gamma})^{-1}$ times the Hermitian pairing on the right.

To the datum (α, ϕ_0, ϕ), there are associated elements $g_0 \in GU(V_0)/K_{0,f}$ and $g \in U(V)(\mathbb{A}_f)/K_f$ given by $g_0 := \phi_0 \circ \tilde{\eta}_0^{-1}$ and $g := (\phi_0^{-1}\phi) \circ \tilde{\eta}^{-1}$ (strictly speaking, ϕ_0 and ϕ are tensored with $\mathbb{A}_f$ here, with $H_1(A, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}_f = T(A)^0$ (rational adèlic Tate module) and similarly for A_0).

12.3 Uniformization

For any Hermitian matrix $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with F-coefficients), define the set

$$\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} := \left\{ (\alpha, \underline{x}, \phi_0, \phi) : \begin{array}{l} \alpha \in \mathcal{M}^{\mathrm{an}}_{\mathbb{C}} \text{ with } (\alpha, \underline{x}) \in \mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C}} \\ \text{and } (\phi_0, \phi) \text{ a framing for } \alpha \end{array} \right\}. \tag{12.3.1}$$

There is a canonical injection of sets

$$\begin{aligned} \mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} &\longrightarrow \mathcal{D} \times G'(\mathbb{A}_f)/K'_f \times V^m \\ (\alpha, \underline{x}, \phi_0, \phi) &\longmapsto (z, (g_0, g_0 g), \phi \circ \underline{x} \circ \phi_0^{-1}) \end{aligned} \tag{12.3.2}$$

where the Hodge structure $z \in \mathcal{D}$ and the elements $g_0 \in GU(V_0)(\mathbb{A}_f)/K_{0,f}$ and $g \in U(V)(\mathbb{A}_f)/K_f$ are associated to (α, ϕ_0, ϕ) as in Section 12.2, and $\phi \circ \underline{x} \circ \phi_0^{-1} \in \mathrm{Hom}_F(V_0, V)^m =$

V^m (using the isomorphism $\mathrm{Hom}_{\mathcal{O}_F}(L_0, L) \cong L$ fixed above). This induces a bijection

$$\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} \xrightarrow{\sim} \coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_\infty) \times \mathcal{D}'(\underline{x}_f). \tag{12.3.3}$$

There is a forgetful map $\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} \to \mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C}}$ sending $(\alpha, \phi_0, \phi) \mapsto \alpha$. This is surjective, by the Hasse principle (Landherr's theorem) for Hermitian spaces, and factors through an isomorphism of complex manifolds

$$G'(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_\infty) \times \mathcal{D}'(\underline{x}_f) \right) \xrightarrow{\sim} \mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} \tag{12.3.4}$$

where $G'(\mathbb{Q})$ acts on $\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}}$ as $(\alpha, \phi_0, \phi) \mapsto (\alpha, \gamma_0 \circ \phi_0, \gamma \circ \phi)$ for $(\gamma_0, \gamma) \in G'(\mathbb{Q})$ with γ_0 and γ having the same similitude factor. The case $T = \emptyset$ (or $T = 0$) gives complex uniformization of $\mathcal{M}^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}}$.

The isomorphism $G' \xrightarrow{\sim} GU(V_0) \times U(V)$ (see $(*)$) induces an isomorphism

$$\begin{aligned} &G'(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_\infty) \times \mathcal{D}'(\underline{x}_f) \right) \\ &\xrightarrow{\sim} \Big(GU(V_0)(\mathbb{Q}) \backslash (GU(V_0)(\mathbb{A}_f)/K_{0,f}) \Big) \times \left(U(V)(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_\infty) \times \mathcal{D}(\underline{x}_f) \right) \right) \end{aligned} \tag{12.3.5}$$

where $U(V)(\mathbb{Q})$ acts on $\mathcal{D}$ via the $U(V)(\mathbb{R})$ action, and on $U(V)(\mathbb{A}_f)/K_f$ by left multiplication.

12.4 Local intersection numbers: Archimedean

Fix $T \in \mathrm{Herm}_m(\mathbb{Q})$ and $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$ (i.e. y is any positive definite complex Hermitian matrix). Throughout Section 12.4, we require $m \geq n - 1$ if T is positive definite. If T is singular, we also require $m = n$ and $\mathrm{rank}(T) = n - 1$.

For such T which are nonsingular, we recall Kudla's Green current $g_{T,y}$ for $\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C}}$ (i.e. the unitary analogue studied by Liu [Liu11, Proof of Theorem 4.20]), which is defined via uniformization and star products. For the case of singular T, we propose a definition of $g_{T,y}$ by a "linear invariance" method, which has some subtleties in the case where T is not $\mathrm{GL}_m(\mathcal{O}_F)$-equivalent to $\mathrm{diag}(0, T^\flat)$ for $\det T^\flat \neq 0$ ("not diagonalizable"). Our treatment of this non-diagonalizable case seems to be new.

Allowing T singular or not for the moment, define the set

$$J_\infty(T) := GU(V_0)(\mathbb{A}_f)/K_{0,f} \times \coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x})=T}} \mathcal{D}(\underline{x}_f) \tag{12.4.1}$$

We will see that the groupoid $[G'(\mathbb{Q})\backslash J_\infty(T)]$ has with finite stabilizers and finitely many isomorphism classes (Lemma 20.4.1). Given $j \in J_\infty(T)$, we let $\mathrm{Aut}(j) \subseteq G'(\mathbb{Q})$ be the stabilizer for the action of $G'(\mathbb{Q})$ on $J_\infty(T)$.

For any $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$, recall that there is an induced isomorphism $\mathcal{Z}(T) \xrightarrow{\sim} \mathcal{Z}({}^t\overline{\gamma}T\gamma)$ (i.e. send the tuple of special homomorphisms $\underline{x}$ to $\underline{x}\cdot\gamma$). Similarly, there is an induced isomorphism $\mathcal{Z}(T)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}} \to \mathcal{Z}({}^t\overline{\gamma}T\gamma)^{\mathrm{an}}_{\mathbb{C},\mathrm{framed}}$. There is corresponding a bijection $J_\infty(T) \to J_\infty({}^t\overline{\gamma}T\gamma)$ (which we denote $j \mapsto j\cdot\gamma$) sending $\underline{x} \mapsto \underline{x}\cdot\gamma$ for $\underline{x} \in V^m$ (acting trivially on the remaining data, i.e. view $J_\infty(T)$ as a subset of $G'(\mathbb{A}_f)/K'_f \times V^m$; note $\mathcal{D}(\underline{x}_f\cdot\gamma) = \mathcal{D}(\underline{x}_f)$). Note $\mathrm{Aut}(j) = \mathrm{Aut}(j\cdot\gamma)$.

For each $j \in J_\infty(T)$, there is a corresponding map

$$\Theta_j \colon \mathcal{D} \to \mathcal{M}^{\mathrm{an}}_{\mathbb{C}} \tag{12.4.2}$$

induced by the uniformization morphism $\mathcal{D} \times G'(\mathbb{A}_f)/K'_f \to \mathcal{M}^{\mathrm{an}}_{\mathbb{C}}$ (consider the projection $J_\infty(T) \to G'(\mathbb{A}_f)/K'_f$; by uniformization of $\mathcal{M}^{\mathrm{an}}_{\mathbb{C}}$, every element of $G'(\mathbb{A}_f)/K'_f$ determines a map $\mathcal{D} \to \mathcal{M}^{\mathrm{an}}_{\mathbb{C}}$). For any $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$, we have $\Theta_j = \Theta_{j\cdot\gamma}$.

If $\widehat{\mathcal{E}}_{\mathbb{C}}$ denotes the metrized tautological bundle on $\mathcal{M}^{\mathrm{an}}_{\mathbb{C}}$ (Section 4.3) we have $\Theta_j^*\widehat{\mathcal{E}}_{\mathbb{C}} \cong \widehat{\mathcal{E}}$, where $\widehat{\mathcal{E}}$ is the metrized tautological bundle on $\mathcal{D}$ (Section 8.2). By our normalizations, the metric on $\Theta_j^*\widehat{\mathcal{E}}_{\mathbb{C}}$ is $(16\pi^3 e^\gamma)^{-1}$ times the metric on $\widehat{\mathcal{E}}$ (this normalization constant does not change the Chern form $c_1(\widehat{\mathcal{E}})$).

Consider $\underline{x} \in V$ with $(\underline{x},\underline{x}) = T$. If T is nonsingular, set

$$[\xi(\underline{x}, y)] := [\xi(\underline{x}\cdot a)] \tag{12.4.3}$$

for a choice of $a \in \mathrm{GL}_m(\mathbb{C})$ satisfying $a^t\overline{a} = y$, with $[\xi(\underline{x}\cdot a)]$ the current from Section 8.2. We will not check that the current $[\xi(\underline{x}, y)]$ is independent of the choice of a, but the intersection numbers appearing in our main results will not depend on a (Remark 19.1.5, also the "linear invariance" from [Liu11, Proposition 4.10] when $m = n$).

Next, suppose that T is singular, with $m = n$ and $\mathrm{rank}(T) = n-1$. First consider the case when $T = \mathrm{diag}(0, T^\flat)$ where $T^\flat$ is nonsingular of rank $n-1$. If $(\underline{x},\underline{x}) = T$, we must have $\underline{x} = [0, x_2, \ldots, x_n] \in V^n$. Set $\underline{x}^\flat = [x_2, \ldots, x_n]$. There is a decomposition

$$y = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^\# & 0 \\ 0 & y^\flat \end{pmatrix} \begin{pmatrix} 1 & 0 \\ {}^t\overline{c} & 1 \end{pmatrix} \tag{12.4.4}$$

for uniquely determined $c \in M_{1,n-1}(\mathbb{C})$, $y^{\#} \in \mathbb{R}_{>0}$, and $y^{\flat} \in \mathrm{Herm}_{n-1}(\mathbb{R})_{>0}$. We then set

$$[\xi(\underline{x}, y)] \coloneqq c_1(\widehat{\mathcal{E}}^{\vee}) \wedge [\xi(\underline{x}^{\flat}, y^{\flat})] - \log(y^{\#}) \cdot \delta_{\mathcal{D}(\underline{x})}. \tag{12.4.5}$$

For T not necessarily block-diagonal, we define $[\xi(\underline{x}, y)]$ by the linear invariance requirement

$$[\xi(\underline{x}, y)] \coloneqq [\xi(\underline{x} \cdot \gamma^{-1}, \gamma y {}^t\overline{\gamma})] \quad \text{mod} \sum_{\substack{p \text{ such that} \\ \gamma \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)})}} \mathbb{Q} \cdot \log p \cdot \delta_{\mathcal{D}(\underline{x})} \tag{12.4.6}$$

for all $\gamma \in \mathrm{GL}_n(F)$, where ${}^t\overline{\gamma}$ means conjugate transpose, and where "mod" means that the equality (of currents) holds up to adding an element of the displayed sum.

Equivalently, suppose $\gamma \in \mathrm{GL}_n(F)$ is any element such that ${}^t\overline{\gamma}^{-1} T \gamma^{-1} = \mathrm{diag}(0, T^{\flat})$ is block diagonal with $T^{\flat}$ nonsingular. Write $\underline{x} \cdot \gamma^{-1} = [0, x_1^{\flat}, \ldots, x_{n-1}^{\flat}]$, set $\underline{x}_{\gamma}^{\flat} = [x_1^{\flat}, \ldots, x_{n-1}^{\flat}]$, and decompose

$$\gamma y {}^t\overline{\gamma} = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{\gamma}^{\#} & 0 \\ 0 & y_{\gamma}^{\flat} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ {}^t\overline{c} & 1 \end{pmatrix}, \tag{12.4.7}$$

as above (temporary notation). We then have

$$[\xi(\underline{x}, y)] = c_1(\widehat{\mathcal{E}}^{\vee}) \wedge [\xi(\underline{x}_{\gamma}^{\flat}, y_{\gamma}^{\flat})] - \log(\tilde{y}^{\#}) \cdot \delta_{\mathcal{D}(\underline{x})} \tag{12.4.8}$$

for a positive real number $\tilde{y}^{\#}$ uniquely determined by T and y. Indeed, we require

$$\log(\tilde{y}^{\#}) = \log(y_{\gamma}^{\#}) \quad \text{mod} \sum_{\substack{p \text{ such that} \\ \gamma \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)})}} \mathbb{Q} \cdot \log p. \tag{12.4.9}$$

For any fixed prime p, we can always find $\gamma \in \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)})$ such that ${}^t\overline{\gamma}^{-1} T \gamma^{-1}$ is block diagonal as above. The preceding expression thus characterizes $\tilde{y}^{\#}$ uniquely.[39] In all cases above (T singular or not), note $[\xi(\underline{x}, y)] = [\xi(\underline{x} \cdot \gamma^{-1}, \gamma y {}^t\overline{\gamma})]$ for all $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$ ("linear invariance").

Definition 12.4.1. For T as above (singular or not), we define the real current

$$g_{T,y} \coloneqq \sum_{j \in J} \frac{1}{|\mathrm{Aut}(j)|} \Theta_{j,*}[\xi(\underline{x}, y)] \tag{12.4.11}$$

on $\mathcal{M}_{\mathbb{C}}^{\mathrm{an}}$, where the sum runs over a set $J \subseteq J_{\infty}(T)$ of representatives for the isomorphism classes of $[G'(\mathbb{Q}) \backslash J_{\infty}(T)]$, where $\underline{x} \in V^m$ is the tuple associated with $j \in J_{\infty}(T)$.

[39] For any integer N, set $\mathbb{R}_N \coloneqq \mathbb{R}/(\sum_{p|N} \mathbb{Q} \cdot \log p)$. For any set of integers $\{N_i\}_{i \in I}$, the diagram

$$\mathbb{R}_{\gcd(\{N_i\}_{i \in I})} \longrightarrow \bigoplus_{i \in I} \mathbb{R}_{N_i} \rightrightarrows \bigoplus_{(i,i') \in I^2} \mathbb{R}_{N_i N_{i'}} \tag{12.4.10}$$

is an equalizer in the category of sets.

In the preceding definition, $\Theta_{j,*}$ denotes pushforward of currents along Θ_j (for singular T, see the convergence estimates in Section 8.3). The current $g_{T,y}$ does not depend on the choice of J, by compatibility of $\mathcal{D}(\underline{x})$ and $[\xi(\underline{x})]$ with the $U(V)(\mathbb{R})$ action on $\mathcal{D}$ (Section 8.2). It is also compatible with pullback of currents for varying (small) levels K'_f. When T is nonsingular, this $g_{T,y}$ agrees with the formulation in [Liu11, Proof of Theorem 4.20] (see also [LZ22a, §15.3]) up to our different normalization of the Green current (Footnote 30).

For any $\gamma \in \mathrm{GL}_m(\mathcal{O}_F)$, we have

$$g_{{}^t\overline{\gamma}T\gamma, \gamma^{-1}y{}^t\overline{\gamma}^{-1}} = g_{T,y} \tag{12.4.12}$$

("global linear invariance"). This follows from the definition of $g_{T,y}$, from local linear invariance of the currents on $\mathcal{D}$, and the formulas $\mathrm{Aut}(j) = \mathrm{Aut}(j \cdot \gamma)$ and $\Theta_j = \Theta_{j\cdot\gamma}$.

In all cases, we define the Archimedean intersection number

$$\mathrm{Int}_{\infty,\mathrm{global}}(T, y) := \int_{\mathcal{M}^{\mathrm{an}}_{\mathbb{C}}} g_{T,y} \wedge c_1(\widehat{\mathcal{E}}^\vee_{\mathbb{C}})^{n-m}. \tag{12.4.13}$$

This is a real number, and the integral is convergent by the estimates in Lemmas 8.3.3 and 8.3.1. It does not depend on the choice of embedding $F \to \mathbb{C}$. By the compatibility of $g_{T,y}$ with varying small levels K'_f, we can extend (12.4.13) to the case of not-necessarily small level by (4.5.2) (i.e. cover by a small level and divide by the degree of the cover). In the notation of loc. cit., the stack $\mathcal{M}$ implicitly has level $K'_{L,f}$ (while we are using the notation $\mathcal{M}$ to mean arbitrary level K'_f in Section 12.4).

In all cases (including possibly K'_f not necessarily small level), we have

$$\mathrm{Int}_{\infty,\mathrm{global}}(T, y) = \mathrm{Int}_\infty(T, y) \frac{[K_{L_0,f} : K_{0,f}]}{|\mathcal{O}_F^\times|/h_F} \cdot \deg\left[U(V)(\mathbb{Q}) \backslash \coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_f) \right] \tag{12.4.14}$$

by construction, where deg means (stacky) groupoid cardinality, where h_F is the class number of $\mathcal{O}_F$, and where

$$\mathrm{Int}_\infty(T, y) := \int_{\mathcal{D}} [\xi(\underline{x}, y)] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m}. \tag{12.4.15}$$

for any $\underline{x} \in V^m$ satisfying $(\underline{x}, \underline{x}) = T$. If there is no such $\underline{x}$, we set $\mathrm{Int}_\infty(T, y) := 0$.

Part V

Eisenstein series

13 Setup

13.1 The group $U(m,m)$

We fix notation for the unitary group $U(m,m)$.

Let $A \to B$ be a finite locally free morphism of (commutative) rings, and suppose B is given an involution $b \mapsto \overline{b}$ ("conjugation") over A. We are mostly interested in the case where F/F^+ is a CM extension of number fields (with F^+ the index 2 totally real subfield) and $B/A = \mathcal{O}_F/\mathcal{O}_{F^+}$ for the corresponding rings of integers (also the local analogues) etc..

Fix an integer $m \geq 0$. Write 1_m for the $m \times m$ identity matrix (sometimes we drop the subscript m), and let $H = U(m,m)$ be the unitary group

$$H = U(m,m) \coloneqq \left\{ h \in \operatorname{Res}_{B/A} \mathrm{GL}_{2m} : h \begin{pmatrix} 0 & 1_m \\ -1_m & 0 \end{pmatrix} {}^t\overline{h} = \begin{pmatrix} 0 & 1_m \\ -1_m & 0 \end{pmatrix} \right\} \tag{13.1.1}$$

where ${}^t\overline{h}$ denotes conjugate transpose (with H the trivial group if $m = 0$, by convention). Equivalently, H consists of block matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{satisfying} \quad {}^t\overline{a}c = {}^t\overline{c}a \qquad {}^t\overline{a}d - {}^t\overline{c}b = 1_m \qquad {}^t\overline{b}d = {}^t\overline{d}b \tag{13.1.2}$$

with $a, b, c, d \in \operatorname{Res}_{B/A} M_{m\times m}$. We refer to H as the group $U(m,m)$ (for signature reasons when B/A is $\mathbb{C}/\mathbb{R}$).

Given an integer j with $0 \leq j \leq m$, we consider the injection

$$\mu_j^m \colon U(j,j) \to U(m,m) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} 1_{m-j} & 0 & 0 & 0 \\ 0 & a & 0 & b \\ 0 & 0 & 1_{m-j} & 0 \\ 0 & c & 0 & d \end{pmatrix}. \tag{13.1.3}$$

Consider the subgroups

$$P \coloneqq \left\{ h = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \in H \right\} \tag{13.1.4}$$

$$M \coloneqq \left\{ m(a) = \begin{pmatrix} a & 0 \\ 0 & {}^t\overline{a}^{-1} \end{pmatrix} : a \in \operatorname{Res}_{B/A} \mathrm{GL}_m \right\} \tag{13.1.5}$$

$$N \coloneqq \left\{ n(b) = \begin{pmatrix} 1_m & b \\ 0 & 1_m \end{pmatrix} : b \in \mathrm{Herm}_m \right\} \tag{13.1.6}$$

of H. We have $P(R) = M(R)N(R)$ for all A-algebras R. We occasionally write P_m, M_m, N_m to emphasize dependence on m.

Set

$$w_j = \begin{pmatrix} 1_{m-j} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1_j \\ 0 & 0 & 1_{m-j} & 0 \\ 0 & -1_j & 0 & 0 \end{pmatrix} \tag{13.1.7}$$

for j with $0 \leq j \leq m$. We also write $w = w_m$ when $j = m$ and m is understood.

Let F_v be a finite étale algebra of degree 2 over a local field F_v^+. Consider $B/A = \mathcal{O}_{F_v}/\mathcal{O}_{F_v^+}$ for the respective rings of integers (with $\mathcal{O}_{F_v^+} \coloneqq F_v^+$ and $\mathcal{O}_{F_v} \coloneqq F_v$ if F_v^+ is Archimedean).

If $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$, we consider the standard maximal compact subgroup

$$K_v^\circ \coloneqq \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in H(\mathbb{R}) : a^t\overline{a} + b^t\overline{b} = 1_m \text{ and } a^t\overline{b} = b^t\overline{a} \right\} \subseteq H(\mathbb{R}) \tag{13.1.8}$$

We write $U(m) \subseteq \mathrm{GL}_m(\mathbb{C})$ for (the real points of) the unitary group for the usual positive definite rank m complex Hermitian space (specified by the Gram matrix 1_m). There is an isomorphism $K_v^\circ \to U(m) \times U(m)$ sending the displayed matrix to $(a+ib, a-ib) \in U(m)\times U(m)$ (see e.g. [GS19, §2.5.1]).

If F_v^+ is non-Archimedean, we consider the standard open compact subgroup

$$K_v^\circ \coloneqq H(\mathcal{O}_{F_v^+}) \subseteq H(F_v^+). \tag{13.1.9}$$

If $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$ or if F_v^+ is non-Archimedean, we have $H(F_v^+) = P(F_v^+)K_v^\circ$. If F_v^+ is non-Archimedean and

$$w^{-1}n(b)w = m(a)k \tag{13.1.10}$$

with $n(b) \in N(F_v^+)$ and $m(a) \in M(F_v^+)$ and $k \in K_v^\circ$, we have $|\det a|_{F_v} < 1$ and moreover $\det a \in F_v^+$ (see [Shi97, §13.4]).

If F/F^+ is a CM extension of number fields and $B/A = \mathcal{O}_F/\mathcal{O}_{F^+}$, we write

$$K^\circ = \prod_v K_v^\circ \qquad K_\infty^\circ = \prod_{v|\infty} K_v^\circ \qquad K_f^\circ = \prod_{v<\infty} K_v^\circ \tag{13.1.11}$$

where the products run over places v of F^+.

For places v of F^+, we use the notation $F_v \coloneqq \prod_{w|v} F_w$ where w runs over places of F, similarly $\mathcal{O}_{F_v} \coloneqq \prod_{w|v} \mathcal{O}_{F_w}$, as well as $F_\infty^+ = \prod_{v|\infty} F_v^+$ and $F_\infty = \prod_{w|\infty} F_w$, etc..

13.2 Adèlic and classical Eisenstein series

Characters are assumed continuous and unitary unless specified otherwise. Let F_v be a degree 2 étale algebra over a local field F_v^+, and form the corresponding unitary group $H = U(m,m)$ as in Section 13.1. If F_v^+ is Archimedean, we assume in Section 13.2 that F_v/F_v^+ is $\mathbb{C}/\mathbb{R}$.

Given $s \in \mathbb{C}$ and a character $\chi_v \colon F_v^\times \to \mathbb{C}^\times$, we may form the local *degenerate principal series*

$$I(s,\chi_v) \coloneqq \mathrm{Ind}_{P(F_v^+)}^{H(F_v^+)}(\chi_v|-|_{F_v}^{s+m/2}). \tag{13.2.1}$$

This is an unnormalized induction, consisting of smooth and K_v°-finite functions $\Phi_v \colon H(F_v^+) \to \mathbb{C}$ satisfying

$$\Phi_v(m(a)n(b)h,s) = \chi_v(a)|\det a|_{F_v}^{s+m/2} \tag{13.2.2}$$

for all $m(a) \in M(F_v^+)$ and $n(b) \in N(F_v^+)$ and $h \in H(F_v^+)$. Here we wrote $\chi_v(a) \coloneqq \chi_v(\det a)$ for short. A section $\Phi_v(h,s)$ of $I(s,\chi_v)$ is *standard* if $\Phi(k,s)$ is independent of s for any fixed $k \in K_v^\circ$. We say Φ_v is *spherical* if $\Phi_v(hk,s) = \Phi_v(h,s)$ for any $k \in K_v^\circ$. We write Φ_v° for the unique spherical standard section satisfying $\Phi_v^\circ(1,s) = 1$ for all s, and call Φ_v° the *normalized spherical section.*

Next, suppose F/F^+ is a CM extension of number fields. We write $\mathbb{A}_F$ for the adèle ring of F and $\mathbb{A}$ for the adèle ring of F^+. Given $s \in \mathbb{C}$ and a character $\chi \colon F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$ and $s \in \mathbb{C}$, we similarly form the global *degenerate principal series*

$$I(s,\chi) \coloneqq \mathrm{Ind}_{P(\mathbb{A})}^{H(\mathbb{A})}(\chi|-|_F^{s+m/2}) \tag{13.2.3}$$

which is an unnormalized induction, consisting of smooth and K°-finite functions $\Phi \colon H(\mathbb{A}) \to \mathbb{C}$ satisfying

$$\Phi(m(a)n(b)h,s) = \chi(a)|\det a|_F^{s+m/2} \tag{13.2.4}$$

for all $m(a) \in M(\mathbb{A})$ and $n(b) \in N(\mathbb{A})$ and $h \in H(\mathbb{A})$. Given characters $\chi_f \colon \mathbb{A}_{F,f}^\times \to C^\times$ and $\chi_\infty \colon \mathbb{A}_{F,\infty}^\times \to \mathbb{C}^\times$, we similarly form $I(s,\chi_f)$ and $I(s,\chi_\infty)$. We also speak of *spherical sections* and the *spherical standard section*, as above. We sometimes write $I_m(s,\chi)$ etc. to indicate dependence on m.

Given a standard section $\Phi(h,s)$ of the global degenerate principal series $I(s,\chi)$, we form the *Siegel Eisenstein series*

$$E(h,s,\Phi) = \sum_{\gamma \in P(F^+)\backslash H(F^+)} \Phi(\gamma h, s) \tag{13.2.5}$$

which is absolutely convergent for $\mathrm{Re}(s) > m/2$. We also form $E(h,\Phi,s)$ when Φ is a finite meromorphic linear combination of standard sections by extending linearly.

Define another character $\check{\chi}\colon F^\times\backslash\mathbb{A}_F^\times \to \mathbb{C}^\times$ as $\check{\chi}(a) := \chi(\overline{a})^{-1}$. There is a functional equation

$$E(h, -s, M(\chi, s)\Phi) = E(h, s, \Phi) \tag{13.2.6}$$

where $M(\chi, s)\colon I(s, \chi) \to I(-s, \check{\chi})$ is the intertwining operator

$$(M(s, \chi)\Phi)(h) = \int_{N(\mathbb{A})} \Phi(w^{-1}n(b)h, s)\; dn(b) \tag{13.2.7}$$

for $\mathrm{Re}(s) > m/2$ (see e.g. [Tan99]). We occasionally write $M_m(s, \chi)$ to emphasize the understood m (in $U(m, m)$).

Fix an identification of F_v^+-algebras $F_v \cong \mathbb{C}$ for each Archimedean place v of F^+. We consider classical Eisenstein series on the *Hermitian upper-half space*

$$\begin{aligned} \mathcal{H}_m &:= \{z \in M_{m,m}(F_\infty) : (2i)^{-1}(z - {}^t\overline{z}) > 0\} && (13.2.8)\\ &= \{z = x + iy : x, y \in \mathrm{Herm}_m(F_\infty^+) \text{ with } y > 0\}, && (13.2.9) \end{aligned}$$

where the latter expression means that x and y are $m \times m$ Hermitian matrices with y positive definite (at every place $v \mid \infty$ of F_v^+). Given $z = x + iy \in \mathcal{H}_m$, we write $h_z \in H(F_\infty^+) \subseteq H(\mathbb{A})$ for any element $h_z = n(x)m(a)$ where $a \in \mathrm{GL}_m(F_\infty)$ satisfies $a^t\overline{a} = y$. Note $h_z \cdot i1_m = z$.

We restrict to $\Phi = \Phi_\infty \otimes \Phi_f$ for standard sections $\Phi_\infty \in I(s, \chi_\infty)$ and $\Phi_f \in I(s, \chi_f)$. Fix an integer n_v for each place $v \mid \infty$ of F_v^+, and assume $\chi_v|_{F_v^{+\times}} = \mathrm{sgn}(-)^{n_v}$ for every $v \mid \infty$. We also let $k(\chi_v) \in \mathbb{Z}$ be the integer satisfying

$$\chi_v(z) = (z/|z|_{F_v}^{1/2})^{k(\chi_v)} \qquad \text{where } z \in F_v, \tag{13.2.10}$$

for each place $v \mid \infty$ of F_v^+. For such v, we let $\Phi_v = \Phi_v^{(n_v)}$ be the unique standard section of $I(s, \chi_v)$ of scalar weight

$$(n_1, n_2) \qquad \text{where} \qquad n_1 = \frac{n + k(\chi)}{2} \quad \text{and} \quad n_2 = \frac{-n + k(\chi)}{2} \tag{13.2.11}$$

such that $\Phi_v^{(n_v)}(1, s) = 1$ (as in [GS19, §3.2, §3.3]). The scalar weight condition means that $\Phi_v^{(n_v)}(hk, s) = \det(k_1)^{n_1}\det(k_2)^{n_2}\Phi_v^{(n_v)}(h, s)$ for all $h \in H(F_v^+)$ and $k \in K_v^\circ$ where $n_1 = (n_v + k(\chi_v))/2$ and $n_2 = (-n_v + k(\chi_v))/2$ and

$$k = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in K_v^\circ \qquad k_1 = a + ib \qquad k_2 = a - ib. \tag{13.2.12}$$

Note that $\Phi_v^{(n_v)}$ does not depend on the choice of identification $F_v \cong \mathbb{C}$.

If $y = a^t\overline{a}$ for some $a \in \mathrm{GL}_m(F_v)$, a computation (omitted) shows

$$\chi_v(a)^{-1}(\det y)^{-n_v/2}\Phi_v^{(n_v)}(w^{-1}n(b)m(a)) = (\det y)^{s-s_0}\det(-iy+b)^{-(s-s_0)}\det(iy+b)^{-(s-s_0)-n_v} \tag{13.2.13}$$

for any $b \in \mathrm{Herm}_m(F_v^+)$, where $s_0 = (n_v - m)/2$ (reduce to the case $a = 1_m$ and write $w^{-1}n(b) = n(-b(1_m + b^2)^{-1})m(b + i1_m)^{-1}k$ for $k \in K_v^\circ$). Equation (13.2.13) may be used to translate various statements from [Shi82] to statements about Archimedean Whittaker functions, etc. (see Section 19.3 for more on this).

Remark 13.2.1. Given $g = x_g + iy_g \in M_{m,m}(\mathbb{C})$ with x_g, y_g Hermitian and x_g positive definite, we define $\log\det(g)$ by the "principal branch" (such that $g \mapsto \log\det g$ is holomorphic, and $\log\det g \in \mathbb{R}$ if $y_g = 0$) as in [Shi82, (1.11)] and the surrounding discussion of loc. cit.. If y_g is positive definite and x_g is only assumed Hermitian, we also take

$$\log\det g = \log\det(-ig) + m\log i \qquad \log\det\overline{g} = \log\det(i\overline{g}) - m\log i \tag{13.2.14}$$

where $\log i \coloneqq \pi i/2$ (as in [Shi82, (1.11)]). This convention is implicit in (13.2.13).

We take $\Phi_\infty = \otimes_{v|\infty}\Phi_v^{(n_v)}$. We write $n = (n_v)_{v|\infty}$ for the collection of Archimedean weights (and will eventually focus on the case where all n_v are equal to some fixed integer n). In the above situation, we write $E(h,s,\Phi)_n \coloneqq E(h,s,\Phi)$ and consider an associated *classical Eisenstein series*

$$E(z,s,\Phi)_n \coloneqq E(z,s,\Phi) \coloneqq \chi_\infty(a)^{-1}\det(y)^{-n/2}E(h_z,s,\Phi)_n \tag{13.2.15}$$

where $z = x + iy$ and $h_z = n(x)m(a)$ with $a^t\overline{a} = y$ as above, and where $\det(y)^{-n/2}$ stands for $\prod_{v|\infty}\det(y_v)^{-n_v/2}$. This does not depend on the choice of h_z, i.e. $E(h_z k_\infty, s, \Phi)_n = E(h_z, s, \Phi)_n$ for any $k_\infty \in K_\infty^\circ$.

When $F^+ = \mathbb{Q}$ and $s_0 \coloneqq (n-m)/2$ (setting $n = n_\infty$ and $k(\chi) = k(\chi_\infty)$), a computation (omitted) gives the more classical form

$$E(z,s,\Phi)_n = \sum_{\gamma \in P(\mathbb{Z})\backslash U(m,m)(\mathbb{Z})} \frac{\det(y)^{s-s_0}\det(\gamma)^{(n+k(\chi))/2}}{\det(cz+d)^n|\det(cz+d)|^{2(s-s_0)}}\Phi_f(\gamma,s) \tag{13.2.16}$$

$$= \sum_{\gamma \in P_1(\mathbb{Z})\backslash SU(m,m)(\mathbb{Z})} \frac{\det(y)^{s-s_0}}{\det(cz+d)^n|\det(cz+d)|^{2(s-s_0)}}\Phi_f(\gamma,s) \tag{13.2.17}$$

where

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{13.2.18}$$

where $SU(m,m) \subseteq U(m,m)$ is the determinant 1 subgroup, and $P_1 := SU(m,m) \cap P$. We have $P(\mathbb{Q})\backslash H(\mathbb{Q}) = P(\mathbb{Z})\backslash H(\mathbb{Z}) = P_1(\mathbb{Q})\backslash H_1(\mathbb{Q}) = P_1(\mathbb{Z})\backslash H_1(\mathbb{Z})$ (e.g. [Ike08, Proposition 12.6]). When $m = 1$, the exceptional isomorphism $\mathrm{SL}_2 \to SU(1,1)$ (over $\operatorname{Spec}\mathbb{Q}$) implies that the above expression is a classical Eisenstein series for SL_2 on the upper-half plane.

Our main theorem (Theorem 22.1.1) concerns Fourier coefficients of $E(z,s,\Phi)_n$ (normalized as in Section 17.1), but the variant

$$\tilde{E}(a,s,\Phi)_n := \chi(a)^{-1}|\det(a)|_F^{-n/2} E(m(a),s,\Phi)_n \quad \text{for } a \in \mathrm{GL}_n(\mathbb{A}_F) \tag{13.2.19}$$

will be useful for studying Fourier coefficients of $E(z,s,\Phi)_n$ for singular T (see below). If $a \in \mathrm{GL}_m(F_\infty)$ is any element satisfying $a^t\overline{a}$ for $y \in \mathrm{Herm}_m(F_\infty^+)_{>0}$, we have

$$E(iy,s,\Phi)_n = \tilde{E}(a,s,\Phi)_n. \tag{13.2.20}$$

13.3 Fourier expansion and local Whittaker functions

Take notation as in Section 13.2, e.g. F/F^+ is a CM extension of number fields. Choose a nontrivial additive character $\psi \colon F^+\backslash\mathbb{A} \to \mathbb{C}^\times$. We have a Fourier expansion

$$E(h,s,\Phi) = \sum_{T \in \mathrm{Herm}_m(F^+)} E_T(h,s,\Phi) \tag{13.3.1}$$

where

$$E_T(h,s,\Phi) = \int_{N(F^+)\backslash N(\mathbb{A})} E(n(b)h,s,\Phi)\psi(-\mathrm{tr}(Tb))\, dn(b) \tag{13.3.2}$$

for $\mathrm{Re}(s) > m/2$, and where $dn(b)$ is the Haar measure on $N(\mathbb{A})$ which is self-dual with respect to the pairing $(b,b') \mapsto \psi(\mathrm{tr}(bb'))$. We refer to $E_T(h,s,\Phi)$ as the *T-th Fourier term.*

For any $a \in \mathrm{GL}_m(F)$, a change of variables gives

$$E_T(m(a)h,s,\Phi) = E_{{}^t\overline{a}Ta}(h,s,\Phi). \tag{13.3.3}$$

We also have

$$E_T(m(a)h,s,\Phi) = E_T(h,s,\Phi) \quad \text{for any} \quad \begin{pmatrix} 1_{m-m^\flat} & * \\ 0 & 1_{m^\flat} \end{pmatrix} \in \mathrm{GL}_m(\mathbb{A}_F) \quad \text{if} \quad T = \begin{pmatrix} 0 & 0 \\ 0 & T^\flat \end{pmatrix} \tag{13.3.4}$$

with the block matrix $T^\flat \in \mathrm{Herm}_{m^\flat}(F^+)$ having $\det T^\flat \neq 0$ (here $m^\flat$ is arbitrary) (follows from [GS19, Lemma 5.4, (5.56)]).

Allowing arbitrary T again, assume there is a factorization $\Phi = (\otimes_{v|\infty}\Phi_v) \otimes \Phi_f$. For each $v \mid \infty$, assume $\Phi_v = \Phi_v^{(n_v)}$ is the scalar weight standard section as in Section 13.2, for some $n_v \in \mathbb{Z}$. Write $n = (n_v)_{v|\infty}$ for the resulting tuple of integers.

Consider $a = a_\infty a_f \in \mathrm{GL}_m(\mathbb{A}_F)$, with $a_\infty \in \mathrm{GL}_m(F_\infty)$ and $a_f \in \mathrm{GL}_m(\mathbb{A}_{F,f})$. Set $y = a_\infty {}^t\overline{a}_\infty$ (temporary). We then have T-th *Fourier coefficients* $E_T(y, s, \Phi)_n$ and $\tilde{E}_T(a, s, \Phi)_n$ characterized by the relations

$$E_T(y, s, \Phi)_n q^T = \chi_\infty(a_\infty)^{-1} \det(y)^{-n/2} E_T(n(x)m(a), s, \Phi) \tag{13.3.5}$$

$$\tilde{E}_T(a, s, \Phi)_n \psi_f(\mathrm{tr}(Tb)) q^T = \chi(a)^{-1} |\det a|_F^{-n/2} E_T(n(x + b)m(a), s, \Phi) \tag{13.3.6}$$

for any $x \in \mathrm{Herm}_m(F_\infty)$ and $b \in \mathrm{Herm}_m(\mathbb{A}_f)$, with $z \coloneqq x + iy$, and with $q^T \coloneqq \psi_\infty(\mathrm{tr}(Tz))$. These correspond to the classical Eisenstein series and its variant in (13.2.15) and (13.2.19).

When $\det T \neq 0$ and $\Phi = \otimes_v \Phi_v$ is factorizable over all places, we have a factorization

$$E_T(h, s, \Phi) = \prod_v W_{T,v}(h_v, s, \Phi_v) \tag{13.3.7}$$

into *local Whittaker functions* defined below (13.3.8).

We switch to local notation: let F_v be a degree 2 étale algebra over a local field F_v^+, with nontrivial involution $a \mapsto \overline{a}$. We assume F_v^+ has characteristic 0 (because Karel assumes this [Kar79]). If F_v^+ is Archimedean, we also assume $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$.

Let $\chi_v \colon F_v^\times \to \mathbb{C}^\times$ and $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ be characters with ψ_v nontrivial, and suppose $\Phi_v \in I(s, \chi_v)$ is a standard section. Given $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$, there is a *local Whittaker function* defined by the absolutely convergent integral

$$W_{T,v}(h, s, \Phi_v) \coloneqq \int_{N(F_v^+)} \Phi_v(w^{-1} n(b) h, s) \psi_v(-\mathrm{tr}(Tb)) \, dn(b) \tag{13.3.8}$$

for $h \in H(F_v^+)$ and $s \in \mathbb{C}$ with $\mathrm{Re}(s) > m/2$, where $dn(b)$ is the Haar measure which is self-dual with respect to the pairing $(b, b') \mapsto \psi_v(\mathrm{tr}(bb'))$ on $\mathrm{Herm}_m(F_v^+) \cong N(F_v^+)$. For each fixed h, the function $W_{T,v}(h, s, \Phi_v)$ admits holomorphic continuation to $s \in \mathbb{C}$ [Kar79, Corollary 3.6.1][KS97][Ich04, §6]. Extending linearly defines $W_{T,v}(h, s, \Phi_v)$ whenever Φ_v is a finite meromorphic linear combination of standard sections. For any $a \in \mathrm{GL}_m(F_v)$, a change of variables shows

$$W_{T,v}(m(a)h, s, \Phi_v) = \check{\chi}_v(a) |\det a|_{F_v}^{-s+m/2} W_{{}^t\overline{a}Ta,v}(h, s, \Phi_v) \tag{13.3.9}$$

for $\check{\chi}_v(a) \coloneqq \chi_v(\overline{a})^{-1}$ as above. We use the shorthand $W_{T,v}(s, \Phi_v) \coloneqq W_{T,v}(1, s, \Phi_v)$.

Lemma 13.3.1. *With notation as above, assume that F_v^+ is non-Archimedean with residue field of cardinality q_v. Suppose $\Phi_v \in I(s, \chi_v)$ is a standard section and $h \in H(F_v^+)$ is a fixed element.*

(1) We have $W_{T,v}(h, s, \Phi_v) \in \mathbb{C}[q_v^{-s}, q_v^s]$.

(2) If $h \in K_v^\circ$, we have $W_{T,v}(h,s,\Phi_v) \in \mathbb{C}[q_v^{-2s}]$.

(3) Suppose $\chi'_v \colon F_v^\times \to \mathbb{C}^\times$ is another character satisfying $\chi'_v|_{F_v^{+\times}} = \xi_v \chi_v|_{F_v^{+\times}}$ for an unramified character $\xi_v \colon F_v^{+\times} \to \mathbb{C}^\times$. Assume $h \in K_v^\circ$, and suppose $\Psi_v \in I(s, \chi'_v)$ is a standard section satisfying $\Psi_v(w^{-1}h) = \Phi_v(w^{-1}h)$. If $f(X) \in \mathbb{C}[X]$ is the polynomial satisfying $f(q_v^{-2s}) = W_{T,v}(h,s,\Phi_v)$, then we have $f(\xi_v(\varpi_0)q^{-2s}) = W_{T,v}(h,s,\Psi_v)$, where $\varpi_0 \in F_v^+$ is a uniformizer.

Proof. A general result of Karel [Kar79, Corollary 3.6.1] states that $W_{T,v}(h,s,\Phi_v) \in \mathbb{C}[q_v^{-s}, q_v^s]$, and that $W_{T,v}(h,s,\Phi_v)$ may be computed for all s as the integral over a sufficiently large open compact subgroup of $N(F_v^+)$. Recall that we have $\Phi_v(m(a)h,s) = \chi_v(\det a)|\det a|_{F_v}^{s+m/2}\Phi_v(h,s)$ for all $a \in \mathrm{GL}_m(F_v)$ and all $h \in H(F_v^+)$. Then apply the discussion surrounding (13.1.10). $\square$

In the case where F_v^+ is non-Archimedean, consider the case where χ_v is unramified and $\chi_v|_{F_v^{+\times}} = \eta_v^n$ for some integer n, where $\eta_v \colon F_v^{+\times} \to \{\pm 1\}$ is the quadratic character associated to F_v/F_v^+. Consider the normalized spherical standard section $\Phi_v^\circ \in I(s,\chi_v)$. We temporarily write $W_{T,v}(h,s,\Phi_v^\circ)_n$ for the associated local Whittaker function, emphasizing the possible dependence on n. By Lemma 13.3.1(3), the implicit χ_v-dependence of $W_{T,v}(h,s,\Phi_v^\circ)_n$ is only on the restriction $\chi_v|_{F_v^{+\times}}$. If F_v/F_v^+ is not inert, then $W_{T,v}(h,s,\Phi_v^\circ)_n$ does not depend on n (note n must be even if F_v/F_v^+ is ramified). If F_v/F_v^+ is inert, then $W_{T,v}(h,s,\Phi_v^\circ)_n$ depends only on the parity of n. The $\mathbb{C}$-algebra endomorphism of $\mathbb{C}[q_v^{-2s}]$ sending $q_v^{-2s} \mapsto -q_v^{-2s}$ swaps $W_{T,v}(h,s,\Phi_v^\circ)_n$ and $W_{T,v}(h,s,\Phi_v^\circ)_{n+1}$, by Lemma 13.3.1(3).

13.4 Singular Fourier coefficients

Retain notation from Section 13.3 (switching back to global notation). The Fourier terms $E_T(h,s,\Phi)$ for singular $T \in \mathrm{Herm}_m(F^+)$ are known to be closely related with Fourier terms of Eisenstein series on smaller groups (e.g. [GS19, §5.2]). We focus on the case where $\operatorname{rank} T = m-1$ (assume this throughout Section 13.4). On account of (13.3.3), it will be enough to describe the case where T is block diagonal of the form

$$T = \begin{pmatrix} 0 & 0 \\ 0 & T^\flat \end{pmatrix} \tag{13.4.1}$$

with $\det T^\flat \neq 0$.

Assume $m \geq 1$, and fix an integer $n \in \mathbb{Z}$. Let $\chi \colon F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$ be a character satisfying $\chi|_{\mathbb{A}^\times} = \eta^n$, where η is the quadratic character associated with F/F^+. Note $\check{\chi} = \chi$ in this case.

Take a factorizable standard section $\Phi = \otimes_v \Phi_v \in I(s,\chi)$, and assume $\Phi_v = \Phi_v^{(n)}$ is the normalized scalar weight standard section (Section 13.2) for every Archimedean place v, with n the fixed integer from above (same for every $v \mid \infty$).

Take T as in (13.4.1). Given $a \in \mathrm{GL}_m(\mathbb{A}_F)$, we study the Fourier coefficient $\tilde{E}_T(a,s,\Phi)_n$. By the Iwasawa decomposition, every $a \in \mathrm{GL}_m(\mathbb{A}_F)$ admits a decomposition

$$a = \begin{pmatrix} 1_1 & * \\ 0 & 1_{m-1} \end{pmatrix} \begin{pmatrix} a^{\#} & 0 \\ 0 & a^{\flat} \end{pmatrix} k \tag{13.4.2}$$

with $a^{\#} \in \mathrm{GL}_1(\mathbb{A}_F)$, with $a^{\flat} \in \mathrm{GL}_{m-1}(\mathbb{A}_F)$, and with $k \in \prod_{v|\infty} U(m) \times \prod_{v<\infty} \mathrm{GL}_m(\mathcal{O}_{F_v})$. We will be eventually interested in the case when Φ_f is spherical, which implies $\tilde{E}_T(ak,s,\Phi)_n = \tilde{E}_T(a,s,\Phi)$ for any $k \in \prod_{v|\infty} U(m) \times \prod_{v<\infty} \mathrm{GL}_m(\mathcal{O}_{F_v})$ (also using the fact that Φ_v is a scalar weight standard section for each $v \mid \infty$). In light of the invariance property in (13.3.4), it is thus harmless to restrict to the case of block diagonal $a = \mathrm{diag}(a^{\#}, a^{\flat})$. Assume this for the rest of Section 13.4 (but we do not assume Φ_f is spherical for now).

Set $m^{\flat} \coloneqq m-1$. Arguing as in the proof of [KR88, Lemma 2.4] (see also [GS19, Lemma 5.4] and [HSY21, Theorem 2.2]) gives

$$\begin{aligned} \tilde{E}_T(a,s,\Phi)_n = {} & |\det a^{\#}|_F^{s-s_0} \tilde{E}_{T^{\flat}}(a^{\flat}, s+1/2, \mu_{m^{\flat}}^{m*}(s,\chi)\Phi)_n \\ & + |\det a^{\#}|_F^{-s-s_0} \tilde{E}_{T^{\flat}}(a^{\flat}, s-1/2, U_{m^{\flat}}^{m}(s,\chi)\Phi)_n \end{aligned} \tag{13.4.3}$$

where $s_0 \coloneqq (n-m)/2$, where

$$\begin{array}{ccc} I_m(s,\chi) & \xrightarrow{\mu_{m^{\flat}}^{m*}(s,\chi)} & I_{m^{\flat}}(s+1/2,\chi) \\ \Psi & \longmapsto & \Psi \circ \mu_{m^{\flat}}^{m} \end{array} \tag{13.4.4}$$

(with $\mu_{m^{\flat}}^{m} \colon U(m^{\flat}, m^{\flat}) \to U(m,m)$ as in Section 13.1), and where

$$\begin{array}{ccc} I_m(s,\chi) & \xrightarrow{U_{m^{\flat}}^{m}(s,\chi)} & I_{m^{\flat}}(s-1/2,\chi) \\ \Psi & \longmapsto & \left(h \mapsto \displaystyle\int_{\substack{b_1 \in \mathrm{Herm}_{m-m^{\flat}}(\mathbb{A}) \\ b_{12} \in M_{m-m^{\flat}, m^{\flat}}(\mathbb{A}_F)}} \Psi \left(w_m^{-1} \cdot n \begin{pmatrix} b_1 & b_{12} \\ {}^t\overline{b}_{12} & 0 \end{pmatrix} w_{m^{\flat}} \mu_{m^{\flat}}^{m}(h), s \right) \, db_1 \, db_{12} \right) \end{array} \tag{13.4.5}$$

for $\mathrm{Re}(s) > m/2$ (with meromorphic continuation to $s \in \mathbb{C}$). A calculation shows

$$M_{m^{\flat}}(s-1/2,\chi) \circ U_{m^{\flat}}^{m}(s,\chi) = \mu_{m^{\flat}}^{m*}(-s,\chi) \circ M_m(s,\chi), \tag{13.4.6}$$

compare [GS19, Lemma 5.5(iii)].

In Corollary 17.2.2, we rewrite (13.4.3) more explicitly when Φ_v is the normalized spherical standard section for every non-Archimedean v.

14 Weil representation

14.1 Weil index

We recall *Weil indices*, which are certain constants appearing in the Weil representation and other calculations below. We compute the instances which we need.

Suppose F_v^+ is a local field (arbitrary characteristic) with nontrivial additive character $\psi_v \colon F_v^+ \to \mathbb{C}^\times$, and suppose V_v is a (finite dimensional) F_v^+ vector space equipped with a non-degenerate quadratic form $Q(-)$. The map $V \to \mathbb{C}^\times$ given by $x \mapsto \psi_v(Q(x))$ is a "non-degenerate character of the second degree" in the sense of [Wei64] [Rao93, Appendix], so there is an associated *Weil index* $\gamma_{\psi_v}(V_v) \in \mathbb{C}^\times$ (which is an eighth root of unity). The quantity $\psi_{\psi_v}(V_v)$ depends only on ψ_v and the isomorphism class of V_v, and we have

$$\gamma_{\overline{\psi}_v}(V_v) = \overline{\gamma_{\psi_v}(V_v)} \qquad \gamma_{\psi_v}(V_v \oplus V_v') = \gamma_{\psi_v}(V_v)\gamma_{\psi_v}(V_v') \tag{14.1.1}$$

for orthogonal direct sums $V_v \oplus V_v'$ (follows from the definition, see [Rao93, Theorem A.2]). The Weil index also satisfies a global product formula [Wei64, Proposition 5].

When F_v^+ has characteristic $\neq 2$ and V_v has a bilinear pairing $(-,-)$, our convention is that $x \mapsto (x,x)$ is the associated quadratic form (and vice-versa).

Lemma 14.1.1. *Let F_v^+ be a local field of characteristic $\neq 2$, let $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ be a nontrivial additive character, and let V_v be a finite dimensional F_v^+ vector space with non-degenerate bilinear pairing. Assume any of the following situations hold.*

(1) The bilinear pairing on V_v is given by

$$\begin{pmatrix} 0 & 1_d \\ 1_d & 0 \end{pmatrix}. \tag{14.1.2}$$

(2) The field F_v^+ is non-Archimedean with residue characteristic $\neq 2$, there exists a self-dual lattice in V_v, and ψ_v is unramified.

Then the Weil index is $\gamma_{\psi_v}(V_v) = 1$.

Proof. (1) By compatibility with orthogonal direct sums, we reduce to the case $d = 1$. Given a nonzero element $a \in F_v^{+\times}$, we use the temporary notation $\gamma_{\psi_v}(a)$ for the Weil index of the one-dimensional quadratic space containing an element x with $(x,x) = a$. We have $\gamma_{\psi_v}(V_v) = \gamma_{\psi_v}(a)\gamma_{\psi_v}(-a^{-1})$ for some $a \in F_v^{+\times}$. We have $\gamma_{\psi_v}(a)\gamma_{\psi_v}(-a^{-1}) = 1$ (follows from [Rao93, Theorem A.4], which relates Weil indices and the Hilbert symbol).

(2) By compatibility with orthogonal direct sums, it is enough to show $\gamma_{\psi_v}(a) = 1$ for $a \in \mathcal{O}_{F_v^+}^\times$. This follows from [Rao93, Proposition A.11]. $\square$

Remark 14.1.2. The explicit formula of [Rao93, Proposition A.12] shows that Lemma 14.1.1(2) is false if $F_v^+ = \mathbb{Q}_2$ (e.g. if V_v has rank one).

Next, let F_v be an étale algebra of degree 2 over a local field F_v^+ of characteristic $\neq 2$ (residue characteristic 2 allowed). Write $\eta_v \colon F_v^{+\times} \to \{\pm 1\}$ for the quadratic character associated to F_v/F_v^+ (trivial if F_v/F_v^+ is split), and write $a \mapsto \overline{a}$ for the nontrivial involution of F_v over F_v^+. If F_v^+ is non-Archimedean, we write $\mathfrak{d}$ (resp. Δ) for the different (resp. discriminant) ideal for the extension F_v/F_v^+ (where $\mathfrak{d} = \mathcal{O}_{F_v}$ and $\Delta = \mathcal{O}_{F_v^+}$ in the split case). We sometimes abuse notation and write $\mathfrak{d}$ and Δ for understood/chosen generators of these ideals. We write q_v for the residue cardinality of F_v^+ if F_v^+ is non-Archimedean.

Any non-degenerate F_v/F_v^+ Hermitian space V_v has an associated F_v^+-bilinear pairing $\frac{1}{2}\mathrm{tr}_{F_v/F_v^+}(-,-)$ and quadratic form $x \mapsto \frac{1}{2}\mathrm{tr}_{F_v/F_v^+}(x,x)$. (Elsewhere, we typically normalize the trace bilinear pairing without the factor of $\frac{1}{2}$.) We write $\gamma_{\psi_v}(V_v)$ for the Weil index of this quadratic space with respect to a nontrivial additive character $\psi_v \colon F_v^+ \to \mathbb{C}^\times$. We know $\gamma_{\psi_v}(V_v)^4 = 1$ (see e.g. [Rao93, Corollary A.5(4)] and [Kud94, Theorem 3.1]).

We write $\gamma_{\psi_v}(F_v)$ for the Weil index associated to the one-dimensional Hermitian space F_v with pairing $(x,y) = \overline{x}y$. We write $\epsilon_v(s, \xi_v, \psi_v)$ for the local epsilon factor associated to a quasi-character $\xi_v \colon F_v^{+\times} \to \mathbb{C}^\times$ (as in [Tat79, §3][Tat67b], for the quasi-character $\xi_v|-|^s$ and the self-dual Haar measure for ψ_v).

If F_v^+ is non-Archimedean with uniformizer ϖ_0, we have

$$\epsilon_v(s, \eta_v, \psi_v) = |\varpi_0^{c(\psi_v)}\Delta|_{F_v^+}^{s-1/2}\gamma_{\psi_v}(F_v) \tag{14.1.3}$$

where

$$c(\psi_v) = \max\{j \in \mathbb{Z} : \psi_v|_{\varpi_0^{-j}\mathcal{O}_{F_v^+}} \text{ is trivial}\}. \tag{14.1.4}$$

If F_v^+ is Archimedean, we have

$$\epsilon_v(s, \eta_v, \psi_v) = |a|_{F_v^+}^{s-1/2}\gamma_{\psi_v}(F_v) \tag{14.1.5}$$

where $a \in F_v^{+\times}$ is such that

$$\psi_v(x) = e^{2\pi i a x} \quad \text{if } F_v^+ = \mathbb{R} \quad \text{and} \quad \psi_v(z) = e^{2\pi i \mathrm{tr}_{\mathbb{C}/\mathbb{R}}(az)} \quad \text{if } F_v^+ = \mathbb{C}. \tag{14.1.6}$$

These identities follow from [JL70, Lemma 1.2(iii),(iv)] and properties of epsilon factors. For the reader's convenience, we recall $\gamma_{\psi_v}(\mathbb{C}) = i$ if $F_v^+ = \mathbb{R}$ and $\psi_v(x) = e^{2\pi i x}$.

In all cases, we have

$$\gamma_{\psi_v}(F_v)^2 = \epsilon_v(1/2, \eta_v, \psi_v)^2 = \eta_v(-1). \tag{14.1.7}$$

If F_v^+ is non-Archimedean, recall that $\epsilon_v(s, \xi_v, \psi_v) = 1$ if ξ_v and ψ_v are unramified. If $F_v^+ = \mathbb{R}$ and $\psi_v(x) = e^{2\pi i x}$, recall $\epsilon_v(s, \mathrm{sgn}^j, \overline{\psi}_v) = 1$ (resp. $= -i$) if j is even (resp. odd) where $\mathrm{sgn}\colon \mathbb{R}^\times \to \{\pm 1\}$ is the sign character (these formulas will be used implicitly in Section 16.2). Recall our convention that self-duality for Hermitian lattices is understood with respect to the trace pairing (unless otherwise specified), Section 2.2.

For Hermitian lattices, we always use the term *self-dual* to mean self-dual with respect to the trace pairing (i.e. $L = L^\vee$) unless specified otherwise. If F_v/F_v^+ is ramified and L is a self-dual Hermitian lattice, then L must have even rank (see e.g. [Shi97, Lemma 13.3]).

Lemma 14.1.3. *Let F_v^+ be a local field of characteristic $\neq 2$, let $\psi_v\colon F_v^+ \to \mathbb{C}^\times$ be a nontrivial additive character, and let F_v/F_v^+ be a degree 2 étale algebra. Let V_v be a finite dimensional non-degenerate F_v/F_v^+ Hermitian space. Assume any of the following situations hold.*

(1) The Hermitian space V_v admits a basis with Gram matrix.

$$\begin{pmatrix} 0 & 1_d \\ 1_d & 0 \end{pmatrix}. \tag{14.1.8}$$

(2) We have $F_v = F_v^+ \times F_v^+$.

(3) The extension F_v/F_v^+ is unramified or F_v^+ has residue characteristic $\neq 2$. Moreover, the field F_v^+ is non-Archimedean, there exists a full-rank self-dual $\mathcal{O}_{F_v}$-lattice in V_v, and V_v has even rank.

(4) The field F_v^+ is non-Archimedean, the extension F_v/F_v^+ is unramified, there exists a full-rank self-dual lattice in V_v, and ψ_v is unramified.

Then the Weil index is $\gamma_{\psi_v}(V_v) = 1$.

Proof. We have (3) $\implies$ (1) (see [LL22, Lemma 2.12] for the ramified situation, in which case the even rank assumption is redundant). This implication is false if F_v/F_v^+ is ramified with F_v^+ of residue characteristic 2.

In situations (1) and (2) we may pick a basis $\{1, \alpha\}$ for F_v as an F_v^+ vector space where $\mathrm{tr}_{F_v/F_v^+}(\alpha) = 0$. Applying Lemma 14.1.1 proves the claims.

In situation (4), we may diagonalize the given self-dual lattice, hence reducing to the case where V_v has rank one. In this case, we have $\gamma_{\psi_v}(V_v) = \gamma_{\psi_v}(F_v) = \epsilon(1/2, \eta_v, \psi_v) = 1$. $\square$

14.2 Weil representation

Let F_v/F_v^+ and accompanying notation be as in Section 14.1. Assume $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$ if F_v^+ is Archimedean. We also assume F_v^+ has characteristic 0 (because [Kud94] assumes this).

Let V_v be a non-degenerate F_v/F_v^+ Hermitian space of dimension $n \geq 0$. Choose a nontrivial additive character $\psi_v \colon F_v^+ \to \mathbb{C}^\times$, and let $\chi_v \colon F_v^\times \to \mathbb{C}^\times$ be a character such that $\chi_v|_{F_v^{+\times}} = \eta_v^n$. There is a local *Weil representation* $\omega_v = \omega_{\chi_v,\psi_v}$ of $U(m,m)(F_v^+) \times U(V_v)(F_v^+)$ on the space of Schwartz functions $\mathcal{S}(V_v^m)$ (the Schrödinger model [Kud94]), which we normalize as

$$\begin{aligned}
(\omega_v(m(a))\varphi_v)(\underline{x}) &= \chi_v(\det a)|\det a|_{F_v}^{n/2}\varphi_v(\underline{x}\cdot a) && m(a) \in M(F_v^+)\\
(\omega_v(n(b))\varphi_v)(\underline{x}) &= \psi_v(\operatorname{tr} b(\underline{x},\underline{x}))\varphi_v(\underline{x}) && n(b) \in N(F_v^+)\\
(\omega_v(w)\varphi_v)(\underline{x}) &= \gamma_{\psi_v}(V_v)^m\widehat{\varphi}_v(\underline{x}) &&\\
(\omega_v(h)\varphi_v)(\underline{x}) &= \varphi_v(h^{-1}\cdot\underline{x}) && h \in U(m,m)(F_v^+)
\end{aligned}$$

for $\varphi_v \in \mathcal{S}(V_v^m)$ and $\underline{x} \in V_v^m$ (viewed as $n \times m$ matrices), where

$$\widehat{\varphi}_v(\underline{x}) = \int_{V_v^m} \varphi_v(\underline{y})\psi_v(\operatorname{tr}_{F_v/F_v^+}\operatorname{tr}(\underline{x},\underline{y}))\,d\underline{y} \tag{14.2.1}$$

is Fourier transform for the corresponding self-dual Haar measure on V_v^m. The constant $\gamma_{\psi_v}(V_v)$ is the Weil index from Section 14.1

With $s_0 \coloneqq (n-m)/2$, there is a map $\mathcal{S}(V_v^m) \to I(\chi_v, s_0)$ sending $\varphi_v \in \mathcal{S}(V_v^m)$ to the function $h \mapsto (\omega_v(h)\varphi_v)(0)$. The associated standard section $\Phi_{\varphi_v} \in I(\chi_v, s)$ is the *Siegel–Weil section* for φ_v [GS19, §5.1].

If F_v^+ is non-Archimedean, choose a generator $\mathfrak{d}$ of the different ideal of F_v/F_v^+, and let M_2° be the rank 2 Hermitian $\mathcal{O}_{F_v}$-lattice admitting a basis with Gram matrix

$$\begin{pmatrix} 0 & \mathfrak{d}^{-1} \\ \overline{\mathfrak{d}}^{-1} & 0 \end{pmatrix}. \tag{14.2.2}$$

Note that $M_2^\circ = M_2^{\circ *}$ is self-dual (with respect to the F_v^+-bilinear pairing $\operatorname{tr}_{F_v/F_v^+}(-,-)$).

Lemma 14.2.1. *In the situation above, assume moreover that χ_v and ψ_v are unramified, and that F_v^+ is non-Archimedean. Suppose $\varphi_v = \mathbf{1}_M^{\otimes m}$ where $\mathbf{1}_M$ is the characteristic function of a full rank $\mathcal{O}_{F_v}$-lattice $M \subseteq V_v$ in any of the following situations.*

(1) The lattice M is self-dual. Moreover, the extension F_v/F_v^+ is unramified, or F_v^+ has residue characteristic $\neq 2$.

(2) We have $M \cong (M_2^\circ)^{\oplus d}$ (orthogonal direct sum) for some $d \geq 0$.

Then the associated Siegel–Weil section Φ_{φ_v} is the normalized spherical section Φ_v°, i.e. K_v°-fixed with $\Phi_{\varphi_v}(1) = 1$.

Proof. Follows from the explicit formulas above, since w and $P(\mathcal{O}_{F_v^+})$ generate $K_v^\circ = U(m,m)(\mathcal{O}_{F_v^+})$ and since the Weil index $\gamma_{\psi_v}(V_v)$ is 1 (Lemma 14.1.3).

If M has even rank, then condition (1) implies condition (2) (the ramified case is [LL22, Lemma 2.12]). □

Next, consider the case where $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$. Suppose the n-dimensional Hermitian space V_v is positive definite, with Hermitian pairing $(-,-)$. If $\psi_v(x) = e^{2\pi i x}$, the Gaussian

$$\varphi_v(\underline{x}) = e^{-2\pi \operatorname{tr}(\underline{x},\underline{x})} \in \mathcal{S}(V_v^m) \tag{14.2.3}$$

for $\underline{x} = (x_1, \ldots, x_m) \in V_v^m$ (where $\operatorname{tr}(\underline{x},\underline{x}) = (x_1,x_1) + \cdots + (x_m,x_m)$) has associated Siegel–Weil section

$$\Phi_{\varphi_v} = \Phi_v^{(n)} \tag{14.2.4}$$

where $\Phi_v^{(n)}$ is the scalar weight standard section described surrounding (13.2.11), see [GS19, (2.68)].

Remark 14.2.2. Suppose F/F^+ is a CM extension of number fields with associated quadratic character η and accompanying notation as in Section 13.2. With m and n as above, choose any character $\chi\colon F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$ satisfying $\chi|_{\mathbb{A}^\times} = \eta^n$. Choose nontrivial additive characters $\psi_v\colon F_v^+ \to \mathbb{C}^\times$ for each place v (the ψ_v need not come from a global character). Suppose we are given a collection of local Weil representations ω_{χ_v,ψ_v} on some $\mathcal{S}(V_v^m)$ for each place v of F_v^+ (where the collection $(V_v)_v$ of local Hermitian spaces need not come from a global Hermitian space). Choose $\varphi_v \in \mathcal{S}(V_v^m)$ for each place v, and assume $\varphi_v = \mathbf{1}_{L_v}^m$ for some full-rank self-dual lattice $L_v \subseteq V_v$ for all but finitely many v. Set $\Phi \coloneqq \bigotimes_v \Phi_{\varphi_v}$.

In this situation, the Eisenstein series variant $\tilde{E}(a,s,\Phi)_n$ (13.2.19) does not depend on the choice of χ. This follows upon inspecting the Weil representation, particularly the action of $m(a)$.

This remark also has a local version, i.e. the Whittaker function variants $\tilde{W}_{T,v}^*(a,s)_n^\circ$ and $\tilde{W}_{T,v}^*(a,s,\Phi_{\varphi_v})_n$ (Sections 15.2 and 15.3) do not depend on the choice of χ_v.

15 Local Whittaker functions

Let F_v/F_v^+ and accompanying notation be as in Section 14.2. If F_v^+ has residue characteristic 2, we also assume F_v/F_v^+ is unramified. Let $\chi_v\colon F_v^\times \to \mathbb{C}^\times$ be a character satisfying $\chi_v|_{F_v^{+\times}} = \eta_v^n$

for some integer $n \in \mathbb{Z}$, with n even if F_v/F_v^+ is ramified. Assume χ_v is unramified if F_v^+ is non-Archimedean. Let $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ be an unramified nontrivial additive character. Assume $\psi_v(x) = e^{2\pi i x}$ if $F_v^+ = \mathbb{R}$. These are our default hypotheses, but weaker hypotheses often suffice (as will be indicated below).

Let $\Phi_v^\circ \in I(s, \chi_v)$ be the normalized spherical standard section if F_v^+ is non-Archimedean. Let $\Phi_v^{(n)} \in I(s, \chi_v)$ be the normalized scalar weight standard section from Section 13.2 if $F_v/F_v^+ = \mathbb{C}/\mathbb{R}$.

Given an integer $m \geq 0$ (we do not assume $m \leq n$, unless otherwise specified) and given $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$, we define *normalized local Whittaker functions*

$$W^*_{T,v}(h,s)^\circ_n \coloneqq \Lambda_{T,v}(s)^\circ_n W_{T,v}(h,s,\Phi_v^\circ) \qquad \text{for } F_v^+ \text{ non-Archimedean} \tag{15.0.1}$$

$$W^*_{T,v}(h,s)^\circ_n \coloneqq \Lambda_{T,v}(s)^\circ_n W_{T,v}(h,s,\Phi_v^{(n)}) \qquad \text{for } F_v^+ \text{ Archimedean} \tag{15.0.2}$$

for certain normalizing factors $\Lambda_{T,v}(s)^\circ_n$ (see (15.3.1) and (15.2.1) below).

The preceding normalization gives $W^*_{T,v}(h,s)^\circ_n$ a clean functional equation (Section 16). Moreover, the normalized function $W^*_{T,v}(h,s)^\circ_n$ (as opposed to the unnormalized versions) seem to correspond more naturally to local information about special cycles (e.g. local contributions to arithmetic degrees) in arithmetic (and non-arithmetic) Siegel–Weil formulas. For example, our main local theorems (Part VI) are proved in terms of the derivative of $W^*_{T,v}(1,s)^\circ_n$ and not $W_{T,v}(1,s,\Phi_v^\circ)$ or $W_{T,v}(1,s,\Phi_v^{(n)})$.

The normalizing factors $\Lambda_{T,v}(s)^\circ_n$ also carry geometric information. For example, consider an imaginary quadratic field $F/\mathbb{Q}$ of odd discriminant, suppose $m = n$ is even, and form the product $2\prod_v \Lambda_{T,v}(s)^\circ_n$ over all places v of $\mathbb{Q}$. If $n \equiv 0 \pmod 4$, evaluation at $s = 0$ returns the degree of a certain 0-dimensional unitary complex Shimura variety (stack), giving a case of a unitary analogue of the Siegel mass formula. If $n \equiv 2 \pmod 4$, evaluation at $s = 0$ returns the volume of a certain $(n-1)$-dimensional unitary complex Shimura variety (stack). These volume identities will be discussed in Section 21.2 (but are not needed for our main theorems on arithmetic Siegel–Weil).

15.1 Local L-factors

We use the following (standard) local factors as in [Tat79, §3].

If F_v^+ is a local field (allowing arbitrary characteristic in Section 15.1) and $\xi_v \colon F_v^{+\times} \to \mathbb{C}^\times$ is a quasi-character, we write $L_v(s, \xi_v)$ for the corresponding local L-factor (for the quasi-character $\xi_v| - |^s_{F_v^+}$). Given any nontrivial additive character $\psi_v \colon F_v^+ \to \mathbb{C}^\times$, we write $\epsilon_v(s, \xi_v, \psi_v)$ for the corresponding local epsilon factor (as appeared in Section 14.1) and

$\rho_v(s, \xi_v, \psi_v)$ for the local factor from Tate's thesis [Tat67b, Theorem 2.4.1], which satisfies

$$\rho_v(s, \xi_v, \psi_v) = \epsilon_v(s, \xi_v, \psi_v)^{-1} L_v(1 - s, \xi_v^{-1})^{-1} L_v(s, \xi_v). \tag{15.1.1}$$

If F^+ is a global field with a quasi-character $\xi \colon F^{+\times} \backslash \mathbb{A}_{F^+} \to \mathbb{C}^\times$ and nontrivial additive character $\psi \colon F^+ \backslash \mathbb{A}_{F^+} \to \mathbb{C}^\times$, we write

$$\Lambda(s, \xi) = \prod_v L_v(s, \xi_v) \qquad L(s, \xi) = \prod_{v < \infty} L_v(s, \xi) \qquad \epsilon(s, \xi) = \prod_v \epsilon_v(s, \xi_v, \psi_v) \tag{15.1.2}$$

and have $\Lambda(s, \xi) = \epsilon(s, \xi) \Lambda(1 - s, \xi^{-1})$. For the reader's convenience, we recall the formulas

$$L_v(s, \xi_v) = \begin{cases} (1 - \xi_v(\varpi_0) |\varpi_0|^s_{F_v^+})^{-1} & \text{if } \xi_v \text{ is unramified} \\ 1 & \text{if } \xi_v \text{ is ramified} \end{cases} \qquad \begin{array}{l} \text{if } F_v^+ \text{ is non-Archimedean} \\ \text{with uniformizer } \varpi_0 \in F_v^+ \end{array}$$

$$L_v(s, \operatorname{sgn}^j) = \begin{cases} \pi^{-s/2} \Gamma(s/2) & \text{if } j \text{ is even} \\ \pi^{-(s+1)/2} \Gamma((s+1)/2) & \text{if } j \text{ is odd} \end{cases} \qquad \begin{array}{l} \text{if } F_v^+ = \mathbb{R} \text{ and sgn denotes} \\ \text{the sign character.} \end{array}$$

15.2 Normalized Archimedean Whittaker functions

With notation as above, assume F_v / F_v^+ is $\mathbb{C}/\mathbb{R}$ and let $\psi_v \colon \mathbb{R} \to \mathbb{C}^\times$ be the standard additive character $x \mapsto e^{2\pi i x}$. The symbol h will denote an element of $U(m, m)(F_v^+)$. Fix integers n, m with $m \geq 0$.

Consider $T \in \operatorname{Herm}_m(F_v^+)$ with $\det T \neq 0$. With $s_0 \coloneqq (n - m)/2$ as above, we define the normalizing factor

$$\Lambda_{T,v}(s)^\circ_n \coloneqq \frac{(2\pi)^{m(m-1)/2}}{(-2\pi i)^{nm}} \pi^{m(-s+s_0)} \left(\prod_{j=0}^{m-1} \Gamma(s - s_0 + n - j) \right) |\det T|^{-s-s_0}_{F_v^+} \tag{15.2.1}$$

(compare [GS19, (3.54)], also Shimura [Shi82]) where Γ is the usual gamma function.

We define a *normalized Archimedean Whittaker function*

$$W^*_{T,v}(h, s)^\circ_n \coloneqq \Lambda_{T,v}(s)^\circ_n W_{T,v}(h, s, \Phi^{(n)}_v). \tag{15.2.2}$$

For $a \in \mathrm{GL}_m(F_v)$, we also consider the variant

$$\tilde{W}^*_{T,v}(a, s)^\circ_n \coloneqq \chi_v(a)^{-1} |\det a|^{-n/2}_{F_v} W^*_{T,v}(m(a), s)^\circ_n \cdot q^{-T} \qquad q^{-T} \coloneqq e^{-2\pi i \operatorname{tr}(iTy)} \tag{15.2.3}$$

with $y \coloneqq a^t \overline{a}$ (temporary notation). This is a (normalized) local analogue of (13.3.6). For any $a \in \mathrm{GL}_m(F_v)$ and $k \in U(m)$, we have the "linear invariance" properties

$$\tilde{W}^*_{T,v}(a, s)^\circ_n = \tilde{W}^*_{{}^t\overline{a}Ta}(1, s)^\circ_n \qquad \tilde{W}^*_{T,v}(1, s)^\circ_n = \tilde{W}^*_{T,v}(k, s)^\circ_n. \tag{15.2.4}$$

The first expression follows from (13.3.9), and the second expression follows from the scalar weight property of $\Phi_v^{(n)}$. Given $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$, we also set $W^*_{T,v}(y,s)^\circ_n := \tilde{W}^*_{T,v}(a,s)^\circ_n$ for any $a \in \mathrm{GL}_m(F_v)$ satisfying $a{}^t\overline{a} = y$ (does not depend on the choice of a).[40] We use the shorthand $W^*_{T,v}(s)^\circ_n := \tilde{W}^*_{T,v}(1,s)^\circ_n$.

For all $n \in \mathbb{Z}$, we have the functional equation

$$W^*_{T,v}(h,s)^\circ_n = \eta_v(\det T)^{n-m-1} W^*_{T,v}(h,-s)^\circ_n. \tag{15.2.5}$$

The case when T is positive definite follows from [Shi82, Theorem 3.1] (via (13.2.13), see also [GS19, (3.54)]). The case of general T (still with $\det T \neq 0$) should follow from [Shi82, Theorem 4.2, (4.34.K)], though we will give an alternative proof (Lemma 16.2.1). Here η_v is the sign character $\mathrm{sgn}(-)$.

Write (r_1, r_2) for the signature of T (temporary notation). If either $n \geq r_1$ or $r_2 = 0$, then the function $W^*_{T,v}(h,s)^\circ_n$ is holomorphic for all $s \in \mathbb{C}$, for fixed h (follows from [Shi82, Theorem 4.2, (4.34.K)]). For any $a \in \mathrm{GL}_m(F_v)$, we also have

$$\tilde{W}^*_{T,v}(a,s_0)^\circ_n = \begin{cases} 1 & \text{if } T \text{ is positive definite} \\ 0 & \text{if } m \leq n \text{ and } T \text{ is not positive definite.} \end{cases} \tag{15.2.6}$$

For the case when T is positive definite, see [Shi97, (3.15)] (also the proof of [GS19, Proposition 3.2]). The non positive definite case with $m \leq n$ follows from [Shi82, Theorem 4.2, (4.34.K)] (see also [GS19, Proposition 3.3(i)]).

15.3 Normalized non-Archimedean Whittaker functions

With n, χ_v, ψ_v, η_v, etc. as at the beginning of Section 15, assume F_v^+ is non-Archimedean. For the moment, we only assume F_v^+ has characteristic $\neq 2$, and allow χ_v possibly ramified. We can also allow F_v/F_v^+ to be ramified with F_v^+ of residue characteristic 2 in Section 15.3. The symbol h will denote an element of $U(m,m)(F_v^+)$.

Assume $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ is a nontrivial unramified additive character. Let ϖ_0 be a uniformizer of F_v^+, and let q_v be the residue cardinality of F_v^+. Consider $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$.

We define the *local normalizing factor*

$$\Lambda_{T,v}(s)^\circ_n := |\Delta|_{F_v^+}^{-m(m-1)/4} \left(\prod_{j=0}^{m-1} L_v(2s+m-j, \eta_v^{j+n}) \right) |(\det T)\Delta^{\lfloor m/2 \rfloor}|_{F_v^+}^{-s-s_0}. \tag{15.3.1}$$

[40]With this notation, there is possible ambiguity for the meaning of $W^*_{T,v}(1,s)^\circ_n$, which could refer to either $W^*_{T,v}(h,s)^\circ_n$ or $W^*_{T,v}(y,s)^\circ_n$ evaluated at $h = 1_{2m}$ or $y = 1_m$. To avoid confusion, we will avoid the symbol $W^*_{T,v}(1,s)^\circ_n$ when v is Archimedean.

The local L-factors appearing in $\Lambda_{T,v}(s)^{\circ}_n$ should be compared with e.g. [HKS96, §6].

Suppose V_v is an n-dimensional non-degenerate F_v/F_v^+ Hermitian space. Consider a full-rank lattice $L_v \subseteq V_v$, and take the Schwartz function $\varphi_v = \mathbf{1}_{L_v}^m \in \mathcal{S}(V_v^m)$. Form the associated Siegel–Weil standard section $\Phi_{\varphi_v} \in I(\chi_v, s)$. Let S be the Gram matrix of any basis for L_v.

We consider the *normalized local Whittaker function* $W^*_{T,v}$ and the variant $\tilde{W}^*_{T,v}$

$$W^*_{T,v}(h, s, \Phi_{\varphi_v})_n \coloneqq \gamma_{\psi_v}(V_v)^m \mathrm{vol}(L_v)^{-m} \Lambda_{T,v}(s)^{\circ}_n W_{T,v}(h, s, \Phi_{\varphi_v}) \tag{15.3.2}$$

$$\tilde{W}^*_{T,v}(a, s, \Phi_{\varphi_v})_n \coloneqq \chi_v(a)^{-1} |\det a|_{F_v}^{-n/2} W^*_{T,v}(m(a), s, \Phi_{\varphi_v}) \tag{15.3.3}$$

for $a \in \mathrm{GL}_m(F_v)$. The volume $\mathrm{vol}(L_v)$ is taken with respect to the self-dual Haar measure with respect to the pairing $x, y \mapsto \psi_v(\mathrm{tr}(x, y))$ on V_v (compare Lemma 15.4.2). The variant $\tilde{W}^*_{T,v}$ is a local analogue of (13.2.19). These will depend on n in general. For any $a \in \mathrm{GL}_m(F_v)$ and $k \in \mathrm{GL}_m(\mathcal{O}_{F_v^+})$, we have the "linear invariance" property

$$\tilde{W}^*_{T,v}(a, s, \Phi_{\varphi_v})_n = \tilde{W}^*_{{}^t\overline{a}Ta,v}(1, s, \Phi_{\varphi_v})_n \qquad W^*_{T,v}(1, s, \Phi_{\varphi_v})_n = \tilde{W}^*_{T,v}(1, s)_n = \tilde{W}^*_{T,v}(k, s, \Phi_{\varphi_v})_n. \tag{15.3.4}$$

The left expression follows from (13.3.9). The right expression follows as $\chi_v(k)^{-1}\omega_v(m(k))\varphi_v = \varphi_v$ holds for all $k \in \mathrm{GL}_m(\mathcal{O}_{F_v^+})$, where ω_v is the local Weil representation (Section 14.2).

Now assume χ_v is unramified, and recall the normalized spherical standard section $\Phi_v^{\circ} \in I(\chi_v, s)$. If L_v is self-dual, we have $\Phi_{\varphi_v} = \Phi_v^{\circ}$ (Section 14.2), at least outside the case of F_v/F_v^+ ramified with residue characteristic 2. If F_v/F_v^+ is ramified of residue characteristic 2, this still holds if $L_v = (M_2^{\circ})^{\oplus d}$ for some $d \geq 0$ (with M_2° the "standard" self-dual lattice from (14.2.2)). Note that $\gamma_{\psi_v}(V_v) = 1$ in these cases.

In the situation of the previous paragraph, we set

$$W^*_{T,v}(h, s)^{\circ}_n \coloneqq W^*_{T,v}(h, s, \Phi_{\varphi_v})_n \qquad \tilde{W}^*_{T,v}(a, s)^{\circ}_n \coloneqq \tilde{W}^*_{T,v}(a, s, \Phi_{\varphi_v})_n$$

for $h \in H(F_v^+)$ and $a \in \mathrm{GL}_m(F_v)$. Note $W^*_{T,v}(h, s)^{\circ}_n = \Lambda_{T,v}(s)^{\circ}_n W_{T,v}(h, s, \Phi_v^{\circ})$. The alternative normalization

$$W^{(*)}_{T,v}(h, s)^{\circ}_n \coloneqq |(\det T)\Delta^{\lfloor m/2 \rfloor}|_{F_v^+}^{s+s_0} W^*_{T,v}(h, s)^{\circ}_n \tag{15.3.5}$$

will also be useful.

We use the shorthand $W^*_{T,v}(s)^{\circ}_n \coloneqq W^*_{T,v}(1, s)^{\circ}_n$ and $W^{(*)}_{T,v}(s)^{\circ}_n \coloneqq W^{(*)}_{T,v}(1, s)^{\circ}_n$. We further describe these functions in the following sections (e.g. special values and functional equations). We are mostly interested in the spherical local Whittaker function $W^*_{T,v}(h, s)^{\circ}_n$, and the case of general φ_v plays a very limited role in the present work.

15.4 Local densities

We relate non-Archimedean Whittaker functions with local densities. This should be essentially known, but we restate the result for clarity (Lemma 15.4.2).[41] In Section 15.4, we do not need to assume χ_v is unramified (but still require $\chi_v|_{F_v^{+\times}} = \eta_v^n$).

Retain notation and assumptions from Section 15.3. In Section 15.4, we now require F_v^+ to have characteristic 0, exclude the case where F_v/F_v^+ is ramified with F_v^+ of residue characteristic 2, and take $n \geq 0$. We write

$$\begin{aligned}\mathrm{Herm}_m(\mathcal{O}_{F_v^+})^* &\coloneqq \{b \in \mathrm{Herm}_m(F_v^+) : \mathrm{tr}(bc) \in \mathcal{O}_{F_v^+} \text{ for all } c \in \mathrm{Herm}_m(\mathcal{O}_{F_v^+})\} \\ &= \{b \in \mathrm{Herm}_m(F_v^+) : b_{i,j} \in \mathcal{O}_{F_v^+} \text{ if } i = j \text{ and } b_{i,j} \in \mathfrak{d}^{-1}\mathcal{O}_{F_v} \text{ if } i \neq j\}.\end{aligned} \tag{15.4.1}$$

Given nonsingular Hermitian matrices $S \in \mathrm{Herm}_n(F_v^+)$ and $T \in \mathrm{Herm}_m(F_v^+)$, we consider the *local representation density* (or just *local density*)

$$\mathrm{Den}(S,T) \coloneqq \lim_{k \to \infty} \frac{\mathrm{vol}(\{x \in M_{n,m}(\mathcal{O}_{F_v}) : {}^t\overline{x}Sx - T \in \varpi_0^k \mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*\})}{q_v^{-km^2}} \tag{15.4.2}$$

where $M_{n,m}(\mathcal{O}_{F_v})$ is given the Haar measure of total volume 1. The limit argument stabilizes for $k \gg 0$ (follows from the proof of Lemma 15.4.2). The local density $\mathrm{Den}(S,T)$ depends only on the isomorphism classes of the Hermitian lattices defined by S and T. If $n < m$, then $\mathrm{Den}(S,T) = 0$.

If $S \in \mathrm{Herm}_n(\mathcal{O}_{F_v^+})^*$, we have

$$\mathrm{Den}(S,T) = \lim_{k \to \infty} \frac{\#\{x \in M_{n,m}(\mathcal{O}_{F_v}/\varpi_0^k \mathcal{O}_{F_v}) : {}^t\overline{x}Sx - T \in \varpi_0^k \mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*\}}{q_v^{k \cdot m(2n-m)}}. \tag{15.4.3}$$

If $S \in \mathrm{Herm}_n(\mathcal{O}_{F_v^+})^*$ and $T \notin \mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*$, we have $\mathrm{Den}(S,T) = 0$.

Remark 15.4.1. If $S \in \mathrm{Herm}_n(\mathcal{O}_{F_v^+})^*$ and $T \in \mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*$ with $m \leq n$, the local density $\mathrm{Den}(S,T)$ admits the following equivalent formulation. Suppose M (resp. L) is a Hermitian $\mathcal{O}_{F_v}$-lattice which admits a basis with Gram matrix S (resp. T). Write $\mathfrak{d}$ for any trace-zero generator of the different ideal $\mathfrak{d}$ of F_v/F_v^+, and let M' (resp. L') be the skew-Hermitian

[41] The proof is essentially as in [KR14, Proposition 10.1], with a few modifications. In the ramified situation, we should use M_2° (from Section 14.2) instead of $L_{1,1}$ (in the proof of loc. cit.); the proposition statement changes correspondingly, see [Shi22, Proposition 9.7]. Moreover, the quantity $\gamma_p(V)^n$ appearing before [KR14, (10.3)] should be $\gamma_p(V)^{-n}$ for consistency with the Schrödinger model of the Weil representation from [Kud94, Theorem 3.1 §3, §5] (and the same applies to [Shi22, Proposition 9.7]). The interpolation of $W_{T,v}(s, \Phi_{\varphi_v})$ in the two cited references should also be shifted by $s_0 = (n-m)/2$ in the s-variable. The cited results also restrict to the case $F_v^+ = \mathbb{Q}_p$, but the result and (modified) proof hold more generally.

lattice with pairing $\mathfrak{d}S$ (resp. $\mathfrak{d}T$). If $\mathrm{Herm}(M', L')$ is the *scheme of skew-Hermitian module homomorphisms* given by

$$\mathrm{Herm}(M', L')(R) \coloneqq \mathrm{Herm}(M' \otimes R, L' \otimes R) \tag{15.4.4}$$

for $\mathcal{O}_{F_v^+}$-algebras R (where the right-hand side means $\mathcal{O}_{F_v}$-linear homomorphisms preserving the skew-Hermitian pairing), we have

$$\#\mathrm{Herm}(M', L')(\mathcal{O}_{F_v^+}/\varpi_0^k \mathcal{O}_{F_v^+}) = \#\{x \in M_{n,m}(\mathcal{O}_{F_v}/\varpi_0^k \mathcal{O}_{F_v}) : {}^t\overline{x} S x - T \in \varpi_0^k \mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*\} \tag{15.4.5}$$

and also $m(2n-m) = \dim(\mathrm{Herm}(M', L') \times \operatorname{Spec} F_v^+)$ (and the right-hand side is nonempty). This recovers the formulations in [LZ22a, §3.1] (inert), [FYZ24, §2.3] (inert and split), and [HLSY23, §5.1] (ramified).

Return to the situation of general S and T (and possibly $m > n$). Fix characters $\chi_v \colon F_v^\times \to \mathbb{C}^\times$ and $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ as above, with ψ_v unramified. Let M be a Hermitian $\mathcal{O}_{F_v}$-lattice admitting a basis whose Gram matrix is S. Write $V_v = M \otimes_{\mathcal{O}_{F_v}} F_v$ for the associated F_v/F_v^+ Hermitian space, and let $\varphi_v \in \mathcal{S}(V_v^m)$ be the function $\varphi_v = \mathbf{1}_M^{\otimes m}$, where $\mathbf{1}_M$ is the characteristic function of M. Let $\Phi_{\varphi_v} \in I(s, \chi_v)$ be the associated Siegel–Weil section, and form the local Whittaker function $W_{T,v}(h, s, \Phi_{\varphi_v})$ as in Section 13.3. Set $W_{T,v}(s, \Phi_{\varphi_v}) \coloneqq W_{T,v}(1, s, \Phi_{\varphi_v})$.

With M_2° being the rank 2 self-dual Hermitian lattice from (14.2.2), let $S_{r,r}$ be the Gram matrix of a basis for $L_{v,r,r} \coloneqq M \oplus (M_2^\circ)^{\oplus r}$ (orthogonal direct sum). When F_v/F_v^+ is not ramified, we also let S_r be the Gram matrix of a basis for $L_{v,r} \coloneqq M \oplus \langle 1 \rangle^{\oplus r}$ (orthogonal direct sum), where $\langle 1 \rangle$ is a rank one self-dual lattice. The notations $L_{v,r,r}$ and $L_{v,r}$ will only be used in the proof of the next lemma.

Lemma 15.4.2. *With notation as above, there exists* $\mathrm{Den}(S, T, X) \in \mathbb{Q}[X]$ *(necessarily unique) such that*

$$W_{T,v}(s_0 + s, \Phi_{\varphi_v}) = \gamma_{\psi_v}(V_v)^{-m} |\det S|_{F_v^+}^m |\Delta|_{F_v^+}^e \mathrm{Den}(S, T, q_v^{-2s}) \quad \textit{for all } s \in \mathbb{C} \tag{15.4.6}$$

$$\mathrm{Den}(S_{r,r}, T) = \mathrm{Den}(S, T, q_v^{-2r}) \quad \textit{for all } r \in \mathbb{Z}_{\geq 0} \tag{15.4.7}$$

where $\gamma_{\psi_v}(V_v)$ *is the Weil index,* $s_0 = (n-m)/2$*, and* $e = nm/2 + m(m-1)/4$*. For all* $r \in \mathbb{Z}_{\geq 0}$*, we also have*

$$\mathrm{Den}(S_r, T) = \mathrm{Den}(S, T, (-q_v)^{-r}) \quad \textit{if } F_v/F_v^+ \textit{ is inert} \tag{15.4.8}$$

$$\mathrm{Den}(S_r, T) = \mathrm{Den}(S, T, q_v^{-r}) \quad \textit{if } F_v/F_v^+ \textit{ is split.} \tag{15.4.9}$$

Proof. As mentioned above (Footnote 41), this is a restatement of a result which should be essentially known [KR14, Proposition 10.1] [Shi22, Proposition 9.7], up to a few modifications. The modified version stated here may be proved by a similar interpolation argument, as explained below. For any $r \in \mathbb{Z}_{\geq 0}$, set $V_{v,r,r} := L_{v,r,r} \otimes_{\mathcal{O}_{F_v}} F_v$, and let $\varphi_{v,r,r} = \mathbf{1}^m_{L_{v,r,r}}$. Equip $\mathrm{Herm}_m(\mathcal{O}_{F_v^+})$ and $V_{v,r,r}$ with the self-dual Haar measures with respect to $(b,c) \mapsto \psi_v(\mathrm{tr}(bc))$ and $\psi_v(\mathrm{tr}_{F_v/F_v^+}(\mathrm{tr}(-,-)))$ respectively. Using the Weil representation, we compute

$$\begin{aligned}
&W_{T,v}(s_0 + r, \Phi_{\varphi_v}) \qquad (15.4.10)\\
&= \lim_{k\to\infty} \int_{\varpi_0^{-k}\mathrm{Herm}_m(\mathcal{O}_{F_v^+})} \psi_v(-\mathrm{tr}(Tb))\Phi_{\varphi_v}(w^{-1}n(b), s_0 + r)\ dn(b)\\
&= \gamma_{\psi_v}(V_v)^{-m} \lim_{k\to\infty} \int_{\varpi_0^{-k}\mathrm{Herm}_m(\mathcal{O}_{F_v^+})} \psi_v(-\mathrm{tr}(Tb)) \int_{V^m_{v,r,r}} \psi_v(\mathrm{tr}(b(\underline{x},\underline{x})))\varphi_{v,r,r}(\underline{x})\ d\underline{x}\ dn(b)\\
&= \gamma_{\psi_v}(V_v)^{-m} \lim_{k\to\infty} \mathrm{vol}(\varpi_0^{-k}\mathrm{Herm}_m(\mathcal{O}_{F_v^+})) \int_{\substack{\underline{x}\in V^m_{v,r,r}\\ (\underline{x},\underline{x})-T\in\varpi_0^k\mathrm{Herm}_m(\mathcal{O}_{F_v^+})^*}} \varphi_{v,r,r}(\underline{x})\ d\underline{x}\\
&= \gamma_{\psi_v}(V_v)^{-m}\,\mathrm{vol}(\mathrm{Herm}_m(\mathcal{O}_{F_v^+}))\,\mathrm{vol}(L^m_{v,r,r})\,\mathrm{Den}(S_{r,r},T).
\end{aligned}$$

We have the volume identities

$$\mathrm{vol}(\mathrm{Herm}_m(\mathcal{O}_{F_v^+})) = |\Delta|^{m(m-1)/4}_{F_v^+} \qquad \mathrm{vol}(L_{v,r,r}) = |\det S|_{F_v^+}|\Delta|^{n/2}_{F_v^+} \tag{15.4.11}$$

with respect to the self-dual Haar measures described above. We already know $W_{T,v}(s,\Phi_{\varphi_v}) \in \mathbb{C}[q_v^{-2s}]$ by Lemma 13.3.1. Since $\mathrm{Den}(S_{r,r},T) \in \mathbb{Q}$ for all $r \geq 0$, we conclude $W_{T,v}(s,\Phi_{\varphi_v}) \in \mathbb{Q}[q_v^{-2s}]$. The additional claims involving $\mathrm{Den}(S_r,T)$ in the unramified case may be proved similarly, using $L_{v,r}$ instead of $L_{v,r,r}$. $\square$

15.5 Local densities and spherical non-Archimedean Whittaker functions

Take F_v/F_v^+, ψ_v, and χ_v as in Section 15.4, and continue to assume $n \geq 0$ for the moment. Set $s_0 = (n-m)/2$. We assume χ_v is unramified.

Let M° be a self-dual Hermitian $\mathcal{O}_{F_v}$-lattice of rank n. This characterizes M° uniquely up to isomorphism, and forces n to be even if F_v/F_v^+ is ramified. We also have $\gamma_{\psi_v}(V_v) = 1$ (Lemma 14.1.3).

Set $V_v = M^\circ \otimes_{\mathcal{O}_{F_v}} F_v$, and let $\varphi_v \in \mathcal{S}(V_v^m)$ be the characteristic function of $M^{\circ m}$. Then the associated Siegel–Weil section $\Phi_{\varphi_v} \in I(s,\chi_v)$ coincides with the normalized spherical Whittaker function Φ_v° (Lemma 14.2.1).

Remark 15.5.1. Even if χ_v is possibly ramified, we still have $W_{T,v}(s, \Phi_{\varphi_v}) = W_{T,v}(s, \Phi_v^\circ)$ for any $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$ (by Lemma 13.3.1(3) or Lemma 15.4.2), where $\Phi_v^\circ \in I(s, \chi_v')$ is the standard normalized spherical section for an unramified χ_v'.

Suppose $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$. If S is the Gram matrix of any basis for M°, Lemma 15.4.2 gives

$$W_{T,v}(s_0 + s, \Phi_v^\circ) = |\Delta|_{F_v^+}^{m(m-1)/4} \mathrm{Den}(S, T, q_v^{-2s}) \tag{15.5.1}$$

for all $s \in \mathbb{C}$.

Suppose $M^{\circ\prime}$ is a rank m Hermitian $\mathcal{O}_{F_v}$-lattice such that

$$\begin{cases} M^{\circ\prime} \text{ is self-dual} & \text{if } F_v/F_v^+ \text{ is unramified or if } m \text{ is even} \\ M^{\circ\prime} \text{ is almost self-dual} & \text{if } F_v/F_v^+ \text{ is ramified and } m \text{ is odd.} \end{cases} \tag{15.5.2}$$

Let $S' \in \mathrm{Herm}(F_v^+)$ be the Gram matrix of a basis for $M^{\circ\prime}$. We have

$$\left(\prod_{j=0}^{m-1} L_v(2(s+s_0) + m - j, \eta_v^{j+n}) \right)^{-1} = \mathrm{Den}(S, S', X)|_{X = q_v^{-2s}}. \tag{15.5.3}$$

See [LZ22a, (3.2.0.1)] (inert), [FYZ24, Theorem 2.2] (split and inert), [LL22, Lemma 2.15] (ramified).

There is a (normalized) *local density polynomial* $\mathrm{Den}(X, T)_n \in \mathbb{Z}[1/q_v][X]$ such that

$$W_{T,v}^{(*)}(s + s_0)_n^\circ = \mathrm{Den}(q_v^{-2s}, T)_n \tag{15.5.4}$$

for all $s \in \mathbb{C}$ (with $W_{T,v}^{(*)}$ as in Section 15.3). See the "Cho–Yamauchi formulas" proved in [LZ22a, Theorem 3.5.1] (inert), [FYZ24, Theorem 2.2] (split and inert), and [LL22, Lemma 2.15] (ramified). Note that our convention differs slightly from [LL22] in the ramified case, where they consider polynomials in q_v^{-s} instead.

The polynomial $\mathrm{Den}(X, T)_n$ is nonzero if and only if $T \in \mathrm{Herm}(\mathcal{O}_{F_v^+})^*$, in which case $\mathrm{Den}(X, T)_n$ has constant term 1. When $m = n$, we have $\mathrm{Den}(X, T)_n \in \mathbb{Z}[X]$ for any T. More classically, see [Shi97, Theorem 13.6], which implies that $\mathrm{Den}(q_v^n X, T)_n \in \mathbb{Z}[X]$ with constant term 1.

We have

$$\begin{aligned} \mathrm{Den}(X, T)_{n+1} &= \mathrm{Den}(q_v^{-1} X, T)_n && \text{if } F_v/F_v^+ \text{ is split} \\ \mathrm{Den}(X, T)_{n+1} &= \mathrm{Den}(-q_v^{-1} X, T)_n && \text{if } F_v/F_v^+ \text{ is inert} \\ \mathrm{Den}(X, T)_{n+2} &= \mathrm{Den}(q_v^{-2} X, T)_n && \text{if } F_v/F_v^+ \text{ is ramified.} \end{aligned} \tag{15.5.5}$$

For $n < 0$, we define $\mathrm{Den}(X,T)_n$ using (15.5.5). Note that (15.5.4) continues to hold. For the rest of Section 15.5, we allow general $n \in \mathbb{Z}$ (assumed even if F_v/F_v^+ is ramified).

Similarly, there is a (normalized) *local density (Laurent) polynomial* $\mathrm{Den}^*(X,T)_n \in \mathbb{Z}[1/q_v][X, X^{-1/2}]$ such that

$$W^*_{T,v}(s+s_0)^\circ_n = \mathrm{Den}^*(q_v^{-2s}, T)_n \tag{15.5.6}$$

for all $s \in \mathbb{C}$ (with $W^*_{T,v}$ as in Section 15.3).

Remark 15.5.2. On the right-hand side of (15.5.6), we mean evaluating $\mathrm{Den}^*(X,T)_n$ at $X^{1/2} = q_v^{-s}$. We similarly abuse notation elsewhere. For example, $\mathrm{Den}^*(q_v X, T)_n \in \mathbb{Z}[1/q_v^{1/2}][X, X^{-1/2}]$ is obtained from $\mathrm{Den}^*(X,T)_n$ by replacing $X^{1/2}$ with $q_v^{1/2}X^{1/2}$. The notation $\frac{d}{dX} \colon \mathbb{Q}[X, X^{-1/2}] \to \mathbb{Q}[X, X^{-1/2}]$ means the $\mathbb{Q}$-linear map $X^{j/2} \mapsto (j/2)X^{j/2-1}$.

If T defines a self-dual Hermitian lattice when m is even or F_v/F_v^+ is unramified (resp. almost self-dual Hermitian lattice when m is odd and F_v/F_v^+ is ramified), we have

$$W^*_{T,v}(s)^\circ_n = W^{(*)}_{T,v}(s)^\circ_n = 1 \qquad \mathrm{Den}^*(X,T)_n = \mathrm{Den}(X,T)_n = 1 \tag{15.5.7}$$

(follows from (15.5.3)). For such T, an application of Lemma 13.3.1(3) also shows that

$$W_{T,v}(s, \Phi_v^\circ) = |\Delta|_{F_v^+}^{m(m-1)/4} \prod_{j=0}^{m-1} L_v(2s+m-j, \eta_v^j \chi'_v|_{F_v^{+\times}})^{-1} \tag{15.5.8}$$

if $\Phi_v^\circ \in I(s, \chi'_v)$ is the normalized spherical section for any unramified character $\chi'_v \colon F_v^\times \to \mathbb{C}^\times$ (not assuming $\chi'_v|_{F_v^{+\times}} = \eta_v^n$).

If L is a $\mathcal{O}_{F_v}$ Hermitian lattice of rank m, and if L admits a basis with Gram matrix T (allowing arbitrary $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$ again), we write $\mathrm{Den}(X,L)_n \coloneqq \mathrm{Den}(X,T)_n$ and similarly $\mathrm{Den}^*(X,L)_n \coloneqq \mathrm{Den}^*(X,T)_n$. We have

$$\mathrm{Den}^*(X,L)_n = (q_v^{2s_0} X^{-1/2})^{\mathrm{val}'(L)} \mathrm{Den}(X,L)_n \tag{15.5.9}$$

$$\mathrm{val}'(L) \coloneqq \lfloor \mathrm{val}(L) \rfloor = \begin{cases} \mathrm{val}(L) - 1/2 & \text{if } F_v/F_v^+ \text{ is ramified and } m \text{ is odd} \\ \mathrm{val}(L) & \text{else.} \end{cases} \tag{15.5.10}$$

The local densities satisfy a certain cancellation property (which we will use): if L° is a self-dual Hermitian lattice of rank n, then for any non-degenerate Hermitian lattice L and every integer $r \in \mathbb{Z}$ (assume r is even if F_v/F_v^+ is ramified), we have

$$\mathrm{Den}(X, L \oplus L^\circ)_{r+n} = \mathrm{Den}(X,L)_r \qquad \mathrm{Den}^*(X, L \oplus L^\circ)_{r+n} = \mathrm{Den}^*(X,L)_r \tag{15.5.11}$$

where $L \oplus L^\circ$ is the orthogonal direct sum. This follows from the Cho–Yamauchi type formulas cited above and the following linear algebra fact: every lattice $M' \subseteq (L \oplus L^\circ) \otimes_{\mathcal{O}_{F_v}} F_v$ satisfying $L^\circ \subseteq M' \subseteq M'^\vee$ admits an orthogonal direct sum decomposition $M' = L^\circ \oplus M''$ for some sublattice M''.

15.6 Limits of local Whittaker functions

Take integers m, n with $m \geq 1$, set $s_0 = (n-m)/2$, and set $m^\flat = m-1$. Take F_v/F_v^+ and other notation as in the beginning of Section 15 (allowing F_v^+ Archimedean or non-Archimedean).

We consider nonsingular $T \in \mathrm{Herm}_m(F_v^+)$ of the form $T = \mathrm{diag}(t, T^\flat)$ where $T^\flat \in \mathrm{Herm}_{m^\flat}(F_v^+)$ with $T^\flat$ nonsingular, and we study the local Whittaker function $W^*_{T,v}(s)^\circ_n$ as $t \to 0$. The following limiting identities will be crucial for the proofs of our main local theorems. We collect them here for easier comparison between the inert/ramified/split and Archimedean cases. Their proofs will appear in Part VI.

If F_v^+ is non-Archimedean and F_v/F_v^+ is inert, Proposition 18.5.2 implies

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,v}(s)^\circ_n = \lim_{t \to 0} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,v}(s)^\circ_n + (\log |t|_{F_v^+} - \log q_v) W^*_{T^\flat,v}(-1/2)^\circ_n \right) \tag{15.6.1}$$

if the limit is taken over nonzero $t \in F_v^+$ with $\varepsilon(\mathrm{diag}(t, T^\flat)) = -1$.

If F_v^+ is non-Archimedean and F_v/F_v^+ is split, Proposition 18.5.2 implies

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,v}(s)^\circ_n = \lim_{t \to 0} \left(\log q_v \cdot W^*_{T,v}(0)^\circ_n + (\log |t|_{F_v^+} - \log q_v) \cdot W^*_{T^\flat,v}(-1/2)^\circ_n \right) \tag{15.6.2}$$

if the limit is taken over nonzero $t \in F_v^+$.

If F_v^+ is non-Archimedean and F_v/F_v^+ is ramified, Proposition 18.5.2 implies

$$2\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,v}(s)^\circ_n = \lim_{t \to 0} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,v}(s)^\circ_n + (\log |t|_{F_v^+} - \log q_v) \cdot W^*_{T^\flat,v}(-1/2)^\circ_n \right) \tag{15.6.3}$$

if the limit is taken over nonzero $t \in F_v^+$ with $\varepsilon(\mathrm{diag}(t, T^\flat)) = -1$.

If F_v/F_v^+ is $\mathbb{C}/\mathbb{R}$, Proposition 19.1.2 gives

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,v}(s)^\circ_n = \lim_{t \to 0^\pm} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,v}(s)^\circ_n + (\log |t|_{F_v^+} + \log(4\pi) - \Gamma'(1)) W^*_{T^\flat,v}(-1/2)^\circ_n \right) \tag{15.6.4}$$

where the sign on $0^\pm$ is $-$ (resp. $+$) if $T^\flat$ is positive definite (resp. not positive definite). If $T^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{R})$ is not positive definite, Proposition 19.1.2 also proves a similar limiting statement for arbitrary $m^\flat$ (i.e. not necessarily $m^\flat = n-1$).

16 Local functional equations

Let F_v be a degree 2 étale algebra over a local field F_v^+ of characteristic $\neq 2$, with notation $\mathfrak{d}$, Δ, η_v, and $a \mapsto \overline{a}$ as above. If F_v^+ is Archimedean, we also assume F_v/F_v^+ is $\mathbb{C}/\mathbb{R}$. Fix an integer $m \geq 0$.

Consider a character $\chi_v \colon F_v^\times \to \mathbb{C}^\times$ and a nontrivial additive character $\psi_v \colon F^+ \to \mathbb{C}^\times$ (for the moment, we do not require $\chi_v|_{F_v^{+\times}} = \eta_v^n$, and allow χ_v and ψ_v to be ramified).

Set $\check{\chi}_v(a) \coloneqq \chi_v(\overline{a})^{-1}$. There is a local intertwining operator

$$M(s, \chi_v) \colon I(s, \chi_v) \to I(-s, \check{\chi}_v) \tag{16.0.1}$$

(where $I(s, \chi_v)$ and $I(-s, \check{\chi}_v)$ are degenerate local principal series for $U(m,m)$) defined by the integral

$$M(s, \chi_v)\Phi_v(h) = \int_{N(F_v^+)} \Phi_v(w^{-1} n(b) h, s)\ dn(b) \tag{16.0.2}$$

for $\mathrm{Re}(s) > m/2$, with meromorphic continuation to $\mathbb{C}$ (e.g. see [KS97] in the non-Archimedean case).

Given $T \in \mathrm{Herm}_m(F_v^+)$, we define the quantity

$$\begin{aligned}\kappa_T(s, \chi_v, \psi_v) = \chi_v(-1)^m \chi_v(\det T)^{-1} |\det T|_{F_v^+}^{-2s} \gamma_{\psi_v}(F_v)^{m(m-1)/2} \eta_v(\det T)^{m-1} \\ \cdot \prod_{j=0}^{m-1} \rho_v(2s + j - m + 1, \eta_v^j \cdot \chi_v|_{F_v^{+\times}}, \overline{\psi}_v)\end{aligned} \tag{16.0.3}$$

where $\gamma_{\psi_v}(F_v)$ is a Weil index (Section 14.1) and ρ_v is a local factor as in Tate's thesis (Section 15.1). This factor is taken from [KS97, §3][42] (see also [HKS96, Proposition 6.3]).

16.1 Non-Archimedean

Suppose F_v^+ is non-Archimedean (with notation as above). For any $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$ and any standard section Φ_v of $I(s, \chi_v)$, there is a functional equation

$$W_{T,v}(h, -s, M(s, \chi_v)\Phi_v) = \kappa_T(s, \chi_v, \psi_v) W_{T,v}(h, s, \Phi_v) \tag{16.1.1}$$

as in [KS97, §3, §7].

[42] The factor $\kappa_T(s, \chi_v, \psi_v)$ is given there in the non-Archimedean case, but we will use the same formula in the Archimedean case. For comparing formulas, note the different convention used to define $W_{T,v}$ and $M(s, \chi_v)$ (ψ_v versus $\overline{\psi}_v$ and w vs w^{-1}).

We next consider spherical Whittaker functions. Assume ψ_v and χ_v are unramified. We require F_v^+ to be characteristic 0 (because [Shi97, §13] assumes this). With Φ_v° denoting the normalized spherical sections of $I(s, \chi_v)$ and $I(s, \check{\chi}_v)$, we have

$$M(s, \chi_v)\Phi_v^\circ(s) = |\Delta|_{F_v^+}^{m(m-1)/4} \prod_{j=0}^{m-1} \frac{L_v(2s + j - m + 1, \eta_v^j \chi_v|_{F_v^{+\times}})}{L_v(2s + m - j, \eta_v^j \chi_v|_{F_v^{+\times}})} \Phi_v^\circ(-s), \tag{16.1.2}$$

see [Shi97, Theorem 13.6].[43]

Now, we further restrict to the situation where $\chi_v|_{F_v^{+\times}} = \eta_v^n$ for some $n \in \mathbb{Z}$, with n assumed even if F_v/F_v^+ is ramified. Note $\check{\chi}_v = \chi_v$. Combining (16.1.2) with the identities stated above (including the relation between Weil indices and epsilon factors in (14.1.3)), a straightforward computation (omitted) yields the functional equations

$$W_{T,v}^{(*)}(h, -s)_n^\circ = |(\det T)\Delta^{\lfloor m/2 \rfloor}|_{F_v^+}^{-2s} \eta_v((-1)^{m(m-1)/2} \det T)^{n-m-1} W_{T,v}^{(*)}(h, s)_n^\circ \tag{16.1.3}$$

$$W_{T,v}^{*}(h, -s)_n^\circ = \eta_v((-1)^{m(m-1)/2} \det T)^{n-m-1} W_{T,v}^{*}(h, s)_n^\circ \tag{16.1.4}$$

with $W_{T,v}^{(*)}(h, s)^\circ$ and $W_{T,v}^{*}(h, s)^\circ$ as in Section 15.3.

Next, assume that F_v/F_v^+ is unramified or that F_v^+ has residue characteristic $\neq 2$. If L is a Hermitian $\mathcal{O}_{F_v}$-lattice, we thus have

$$\mathrm{Den}(q_v^{2s_0} X^{-1}, L)_n = \varepsilon(L)^{n-m-1} X^{-\mathrm{val}'(L)} \mathrm{Den}(q_v^{2s_0} X, L)_n \tag{16.1.5}$$

$$\mathrm{Den}^*(q_v^{2s_0} X^{-1}, L)_n = \varepsilon(L)^{n-m-1} \mathrm{Den}^*(q_v^{2s_0} X, L)_n \tag{16.1.6}$$

with $\mathrm{val}'(L) \coloneqq \lfloor \mathrm{val}(L) \rfloor$ as in (15.5.10) (both $\varepsilon(L)$ and $\mathrm{val}(L)$ were defined in Section 2.2).

In the case where $\chi_v|_{F_v^{+\times}}$ is trivial, these functional equations are essentially [Ike08, Corollary 3.2].

16.2 Archimedean

Suppose F_v/F_v^+ is $\mathbb{C}/\mathbb{R}$ (with notation as above). For any $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$ and any standard section Φ_v of $I(s, \chi_v)$, we have

$$W_{T,v}(h, -s, M(s, \chi_v)\Phi_v) = \kappa_T(s, \chi_v, \psi_v) W_{T,v}(h, s, \Phi_v). \tag{16.2.1}$$

This may be deduced, e.g. by combining the non-Archimedean analogue (16.1.1) with the global functional equation (13.2.6).

[43] Take $\zeta = 0$ in the notation of loc. cit.. Strictly speaking, the statement there is only for $\chi_v|_{F_v^{+\times}}$ trivial, but the general case follows from this; see (13.1.10) and the proof of Lemma 13.3.1(3).

In the rest of Section 16.2, we require $\chi_v|_{F_v^{+\times}} = \eta_v^n$ for some $n \in \mathbb{Z}$, and let $\psi_v(x) = e^{2\pi i x}$. Recall that we have defined a normalized Archimedean Whittaker function $W^*_{T,v}(h,s)^\circ_n$ (Section 15.2).

Lemma 16.2.1. *For any $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$, we have the functional equation*

$$W^*_{T,v}(h,-s)^\circ_n = \eta_v(\det T)^{n-m-1} W^*_{T,v}(h,s)^\circ_n. \tag{16.2.2}$$

Proof. By (16.2.1), we must have $W^*_{T,v}(h,-s)^\circ_n = \eta_v(\det T)^{n-m-1} f(s) W^*_{T,v}(h,s)^\circ_n$ for some meromorphic factor $f(s)$ which is independent of T. When T is positive definite, we have $f(s) = 1$ (see Section 15.2), so we obtain the claimed functional equation for all $T \in \mathrm{Herm}_m(F_v^+)$ with $\det T \neq 0$. Note that η_v is simply the sign character $\mathrm{sgn}(-)$. $\square$

Recall that $\Phi_v^{(n)} \in I(s,\chi_v)$ is our notation for a certain scalar weight standard section, as in Section 13.2. For verifying the next lemma, it may be helpful to recall the relation between local epsilon factors $\epsilon_v(-)$ and Weil indices $\gamma_v(-)$ (Section 14.1).

Lemma 16.2.2. *We have*

$$\begin{aligned}
&M(s,\chi_v)\Phi_v^{(n)}(s) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (16.2.3)\\
&= \left(\prod_{j=0}^{m-1} \frac{L_v(2s+j-m+1, \eta_v^{n+j})\Gamma(-s-s_0+n-j)}{\epsilon_v(2s+j-m+1, \eta_v^{n+j}, \overline{\psi}_v) L_v(-2s-j+m, \eta_v^{n+j})\Gamma(s-s_0+n-j)} \right)\\
&\quad \cdot (-1)^{nm} i^{m(m-1)/2} \pi^{2ms} \Phi_v^{(n)}(-s)
\end{aligned}$$

with $s_0 = (n-m)/2$ as above.

Proof. A priori, the displayed identity holds up to some meromorphic scale factor. We may compute this scale factor by combining (16.2.1) and Lemma 16.2.1 (take $T = 1_m$). $\square$

Remark 16.2.3. Lemma 16.2.2 should be a reformulation (with alternative proof) of a case of [Shi82, (1.31)] (translating into Shimura's setup via (13.2.13)). Shimura's computation in loc. cit. implies

$$M(s,\chi_v)\Phi_v^{(n)}(s) = \left(\frac{i^{-mn}(2\pi)^{m^2}\pi^{-m(m-1)/2}}{2^{m(m-1)/2+2ms}} \prod_{j=0}^{m-1} \frac{\Gamma(2s-j)}{\Gamma(s-s_0+n-j)\Gamma(s-s_0-j)} \right) \Phi_v^{(n)}(-s). \tag{16.2.4}$$

Similarly, the functional equation in Lemma 16.2.1 should follow from [Shi82, Theorem 4.2, (4.34.K)] (alternative proof) after some rearranging.

For our later calculations, we prefer to use these results as stated in Lemmas 16.2.1 and 16.2.2.

17 Normalized Fourier coefficients

17.1 Global normalization

With notation as in Section 13.2 and Section 13.3, let F/F^+ be a CM extension of number field. For the moment, we allow 2-adic places of F^+ to ramify in F. Write $\mathfrak{d}$ (resp. Δ) for the different ideal (resp. discriminant ideal) of F/F^+. Let $\eta \colon F^{+\times}\backslash\mathbb{A}^\times \to \{\pm 1\}$ be the quadratic character associated with F/F^+.

Assume there exists a nontrivial additive character $\psi \colon F^+\backslash\mathbb{A} \to \mathbb{C}^\times$ such ψ_v is unramified for every non-Archimedean v and $\psi_v(x) = e^{2\pi i x}$ at every Archimedean place. Fix such a ψ. Fix integers m and n with $m \geq 0$, with $s_0 \coloneqq (n-m)/2$ as above. If any non-Archimedean places of F^+ are ramified in F, we assume n is even. Let $\chi \colon F^\times\backslash\mathbb{A}_F^\times \to \mathbb{C}^\times$ be a character satisfying $\chi|_{\mathbb{A}^\times} = \eta^n$. To simplify, we assume that χ is unramified at every non-Archimedean place (but see also Remark 15.5.1).

Take the standard section

$$\Phi^{(n)\circ} \coloneqq \left(\bigotimes_{v|\infty} \Phi_v^{(n)}\right) \otimes \left(\bigotimes_{v<\infty} \Phi_v^{\circ}\right) \in I(s,\chi) \tag{17.1.1}$$

(scalar weight at Archimedean places and spherical at non-Archimedean places). Form the associated Eisenstein series $E(h,s,\Phi^{(n)\circ})$ and its variants $E(z,s,\Phi^{(n)\circ})_n$ and $\tilde{E}(a,s,\Phi^{(n)\circ})_n$ as in Section 13.2. The Eisenstein series variant $\tilde{E}(a,s,\Phi^{(n)\circ})_n$ does not depend on the choice of χ (Remark 14.2.2).

Define the *global normalizing factor*

$$\begin{aligned}\Lambda_m(s)^\circ_n \coloneqq &\left(\frac{(2\pi)^{m(m-1)/2}}{(-2\pi i)^{nm}}\pi^{m(-s+s_0)}\right)^{[F^+:\mathbb{Q}]} |N_{F^+/\mathbb{Q}}(\Delta)|^{m(m-1)/4}|N_{F^+/\mathbb{Q}}(\Delta^{\lfloor m/2\rfloor})|^{s+s_0} \\ &\cdot \left(\prod_{j=0}^{m-1} \Gamma(s-s_0+n-j)^{[F^+:\mathbb{Q}]} \cdot L(2s+m-j,\eta^{j+n})\right).\end{aligned} \tag{17.1.2}$$

We define the *normalized Eisenstein series* and its variants

$$E^*(h,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n E(h,s,\Phi^{(n)\circ}) \tag{17.1.3}$$

$$E^*(z,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n E(z,s,\Phi^{(n)\circ})_n \qquad \tilde{E}^*(a,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n \tilde{E}(a,s,\Phi^{(n)\circ})_n \tag{17.1.4}$$

where $h \in U(m,m)(\mathbb{A})$ and $z \in \mathcal{H}_m$ and $a \in \mathrm{GL}_m(\mathbb{A}_F)$. For $T \in \mathrm{Herm}_m(F^+)$, we similarly define

$$E^*_T(h,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n E_T(h,s,\Phi^{(n)\circ}) \tag{17.1.5}$$

$$E^*_T(y,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n E_T(y,s,\Phi^{(n)\circ})_n \qquad \tilde{E}^*_T(a,s)^\circ_n \coloneqq \Lambda_m(s)^\circ_n \tilde{E}_T(a,s,\Phi^{(n)\circ})_n \tag{17.1.6}$$

The latter two are *normalized Fourier coefficients.*

Given any $T \in \mathrm{Herm}_m(F^+)$ with $\det T \neq 0$, the local normalizing factors from Sections 15.2 and 15.3 satisfy

$$\Lambda_m(s)_n^\circ := \prod_v \Lambda_{T,v}(s)_n^\circ \tag{17.1.7}$$

where the product (over all places v of F^+) is convergent for $\mathrm{Re}(s) > 0$. For such T, we have factorizations into (normalized) local Whitaker functions

$$E_T^*(h,s)_n^\circ = \prod_v W_{T,v}^*(h_v,s)_n^\circ \qquad \tilde{E}_T^*(a,s)_n^\circ = \prod_v \tilde{W}_{T,v}^*(a_v,s)_n^\circ \tag{17.1.8}$$

where all but finitely many factors are identically equal to 1 (as functions of s) for fixed T, h, and n.

Lemma 17.1.1. *We have*

$$E^*(h,-s)_n^\circ = (-1)^{m(m-1)(n-m-1)[F^+:\mathbb{Q}]/2} E^*(h,s)_n^\circ \tag{17.1.9}$$

Proof. Given $T \in \mathrm{Herm}_m(F^+)$ with $\det T \neq 0$, the local functional equations (Section 16) and the factorization from (17.1.8) imply

$$E_T^*(h,-s)_n^\circ = (-1)^{m(m-1)(n-m-1)[F^+:\mathbb{Q}]/2} E_T^*(h,s)_n^\circ. \tag{17.1.10}$$

The global functional equation (13.2.6) implies that $E^*(h,-s)_n^\circ = f(s)E^*(h,s)_n^\circ$ for some meromorphic function $f(s)$ (temporary notation) independent of T. There exists T with $\det T \neq 0$ and $E_T^*(h,s)_n^\circ$ not identically zero (e.g. $T = 1_m$; this follows from Section 15). So $f(s)$ is identically 1 and (17.1.10) holds for all $T \in \mathrm{Herm}_m(F^+)$. $\square$

17.2 Singular Fourier coefficients

Retain notation and assumptions from Section 17.1. In this section, the main result is Corollary 17.2.2 on singular Fourier terms of co-rank 1.

We use various subscripts to emphasize m-dependence (in the implicit $U(m,m)$). For example, we write $\Phi_{m,v}^\circ$ rather than just Φ_v° for non-Archimedean v (resp. $\Phi_{m,v}^{(n)}$ instead of $\Phi_v^{(n)}$ for Archimedean v), similarly $\Phi_m^{(n)\circ}$ instead of $\Phi^{(n)\circ}$ for the global standard section from Section 17.1, also $M_m(s,\chi)$ instead of $M(s,\chi)$ for the intertwining operator, etc..

Suppose $m \geq 1$ and set $m^\flat = m-1$. Recall the operators $\mu_{m^\flat}^{m*}(s,\chi)$, $M_m(s,\chi)$, $M_{m^\flat}(s,\chi)$ and $U_{m^\flat}^m(s,\chi)$ as in Section 13.4.

Lemma 17.2.1. *We have*

$$\mu^{m*}_{m^\flat}(s,\chi)\Phi^{(n)\circ}_m(s) = \Phi^{(n)\circ}_{m^\flat}(s+1/2) \tag{17.2.1}$$

$$U^m_{m^\flat}(s,\chi)\Phi^{(n)\circ}_m(s) = (-1)^e \frac{\Lambda_{m^\flat}(s-1/2)^\circ_n \Lambda_m(-s)^\circ_n}{\Lambda_m(s)^\circ_n \Lambda_{m^\flat}(-s+1/2)^\circ_n} \Phi^{(n)\circ}_{m^\flat}(s-1/2) \tag{17.2.2}$$

$$\begin{aligned} M_m(s,\chi)\Phi^{(n)\circ}_m(s) = {} & |N_{F^+/\mathbb{Q}}(\Delta)|^{-m(m-1)/4}((-1)^{nm} i^{m(m-1)/2}\pi^{2ms})^{[F^+:\mathbb{Q}]} \\ & \cdot \left(\prod_{j=0}^{m-1} \frac{L(2s+j-m+1,\eta^{n+j})}{L(2s+m-j,\eta^{n+j})} \right) \\ & \cdot \left(\prod_{v|\infty} \prod_{j=0}^{m-1} \frac{L_v(2s+j-m+1,\eta_v^{n+j})}{\epsilon_v(2s+j-m+1,\eta_v^{n+j},\overline{\psi}_v) L_v(-2s-j+m,\eta_v^{n+j})} \right) \\ & \cdot \left(\prod_{j=0}^{m-1} \frac{\Gamma(-s-s_0+n-j)}{\Gamma(s-s_0+n-j)} \right)^{[F^+:\mathbb{Q}]} \\ & \cdot \Phi^{(n)\circ}_m(-s), \end{aligned} \tag{17.2.3}$$

allowing $m = 0$ for in $M_m(s,\chi)$ formula, and where

$$e \coloneqq (m(m-1)(n-m-1)/2 - m^\flat(m^\flat-1)(n-m^\flat-1)/2)[F^+:\mathbb{Q}]$$

(temporary notation).

Proof. Each identity holds a priori up to a meromorphic scale factor. We may compute this scale factor by evaluating both sides at $1 \in U(m^\flat, m^\flat)$ or $1 \in U(m,m)$ as appropriate.

The identity for $\mu^{m*}_{m^\flat}(s,\chi)$ is then clear. For $M_m(s,\chi)$, the identity follows directly upon combining (16.1.2) and (16.2.3).

Define the temporary notation $\alpha_m(s)_n$ for the meromorphic function (in the lemma statement) satisfying $M_m(s,\chi)\Phi^{(n)\circ}_m(s) = \alpha_m(s)_n \Phi^{(n)\circ}_m(-s)$. By (13.4.6), proving the claimed identity for $U^m_{m^\flat}(s,\chi)$ is equivalent to showing

$$\frac{\alpha_m(s)_n}{\alpha_{m^\flat}(s-1/2)_n} = (-1)^e \frac{\Lambda_{m^\flat}(s-1/2)^\circ_n \Lambda_m(-s)^\circ_n}{\Lambda_m(s)^\circ_n \Lambda_{m^\flat}(-s+1/2)^\circ_n} \tag{17.2.4}$$

with e as in the lemma statement. This may be computed explicitly as follows. Some rearranging yields

$$\begin{aligned} \frac{\Lambda_{m^\flat}(s-1/2)^\circ_n \Lambda_m(-s)^\circ_n}{\Lambda_m(s)^\circ_n \Lambda_{m^\flat}(-s+1/2)^\circ_n} = {} & (\pi^{2ms}\pi^{(-2s+1)m^\flat})^{[F^+:\mathbb{Q}]} |N_{F^+/\mathbb{Q}}(\Delta^{\lfloor m/2 \rfloor})|^{-2s} |N_{F^+/\mathbb{Q}}(\Delta^{\lfloor m^\flat/2 \rfloor})|^{2s-1} \\ & \cdot \Gamma(s-s_0+n)^{-[F^+:\mathbb{Q}]} \Gamma(-s-s_0+n-m+1)^{[F^+:\mathbb{Q}]} \\ & \cdot L(2s,\eta^{m+n}) L(2s+m,\eta^n)^{-1} L(2s+m-1,\eta^{n+1})^{-1} \\ & \cdot L(-2s+1,\eta^{m+n+1}) \end{aligned}$$

and

$$\frac{\alpha_m(s)_n}{\alpha_{m^\flat}(s-1/2)_n} = |N_{F^+/\mathbb{Q}}(\Delta)|^{-(m-1)/2}((-1)^n i^{m-1}\pi^{2ms-2m^\flat(s-1/2)})^{[F^+:\mathbb{Q}]}$$
$$\cdot L(2s,\eta^{n+m+1})\cdot L(2s+m,\eta^n)^{-1}L(2s+m-1,\eta^{n+1})^{-1}L(2s,\eta^{n+m})$$
$$\cdot\left(\prod_{v\mid\infty}\frac{L_v(2s,\eta_v^{n+m+1})}{\epsilon_v(2s,\eta_v^{n+m+1},\overline{\psi}_v)L_v(-2s+1,\eta_v^{n+m+1})}\right)$$
$$\cdot\Gamma(-s-s_0+n-m+1)^{[F^+:\mathbb{Q}]}\Gamma(s-s_0+n)^{-[F^+:\mathbb{Q}]}.$$

We then use the global functional equation $\Lambda(s,\eta^{n+m+1}) = \epsilon(s,\eta^{n+m+1})\Lambda(1-s,\eta^{n+m+1})$ (notation as in Section 15.1). Recall the relation between Weil indices and epsilon factors (Section 14.1), the global product formula $\prod_v \gamma_{\overline{\psi}_v}(F_v) = 1$ for Weil indices, and the equality $\gamma_{\psi_v}(\mathbb{C}) = i$. Recall also that we have assumed n even if $\Delta \neq 1$. Combining these facts with some casework (which we omit) on m, n, Δ gives the claim. $\square$

Corollary 17.2.2. *Consider any $a = \operatorname{diag}(a^\#, a^\flat) \in \mathrm{GL}_m(\mathbb{A}_F)$ with $a^\# \in \mathrm{GL}_1(\mathbb{A}_F)$ and $a^\flat \in \mathrm{GL}_{m^\flat}(\mathbb{A}_F)$. For any $T \in \mathrm{Herm}_m(F^+)$ with $\operatorname{rank} T = m-1$ and $T = \operatorname{diag}(0, T^\flat)$ being block diagonal with $\det T^\flat \neq 0$, we have*

$$\tilde{E}^*_T(a,s)^\circ_n = |\det a^\#|_F^{s-s_0}\frac{\Lambda_m(s)^\circ_n}{\Lambda_{m^\flat}(s+1/2)^\circ_n}\tilde{E}^*_{T^\flat}(a^\flat, s+1/2)^\circ_n$$
$$+(-1)^e|\det a^\#|_F^{-s-s_0}\frac{\Lambda_m(-s)^\circ_n}{\Lambda_{m^\flat}(-s+1/2)^\circ_n}\tilde{E}^*_{T^\flat}(a^\flat, s-1/2)^\circ_n$$

with constant e as in Lemma 17.2.1.

Proof. This follows immediately from Lemma 17.2.1, (13.4.3), and the definition of the normalized Fourier coefficients $\tilde{E}^*_T(a,s)^\circ_n$ and $\tilde{E}^*_{T^\flat}(a^\flat,s)$ (Section 17.1). $\square$

Remark 17.2.3. In the situation of Corollary 17.2.2, the functional equation

$$\tilde{E}^*_T(a,s)^\circ_n = (-1)^{m(m-1)(n-m-1)[F^+:\mathbb{Q}]/2}\tilde{E}^*_T(a,-s)^\circ_n \tag{17.2.5}$$

is a visible consequence of the identity $\tilde{E}^*_{T^\flat}(a^\flat,s)^\circ_n = (-1)^{m^\flat(m^\flat-1)(n-m^\flat-1)[F^+:\mathbb{Q}]/2}\tilde{E}^*_{T^\flat}(a^\flat,-s)^\circ_n$.

Part VI

Local identities

18 Non-Archimedean local identity

Let F_0 be a non-Archimedean local field of characteristic 0, residue cardinality q, and residue characteristic p. Let F be a finite étale F_0-algebra of degree 2. We use notation $\breve{F}$ and $\breve{F}_0$ as in Part II (there with $F_0 = \mathbb{Q}_p$), so that $[\breve{F} : \breve{F}_0] = 1$ (resp. $[\breve{F} : \breve{F}_0] = 2$) if F/F_0 is unramified (resp. ramified).

Notation on Hermitian lattices from Section 2.2 will be used freely. For a non-degenerate Hermitian $\mathcal{O}_F$-lattice L, we use the shorthand $\mathrm{val}'(L) := \lfloor \mathrm{val}(L) \rfloor \in \mathbb{Z}_{\geq 0}$, as well as $\mathrm{val}'(x) := \lfloor \mathrm{val}(x) \rfloor$ for any $x \in L$ (i.e. $\mathrm{val}'(L) = \mathrm{val}(L) - 1/2$ if F/F_0 is ramified and L has odd rank, and $\mathrm{val}'(L) = \mathrm{val}(L)$ otherwise). Fix an integer $n \geq 1$, and assume n is even if F/F_0 is ramified.

If $F_0 = \mathbb{Q}_p$, we form the associated Rapoport–Zink space $\mathcal{N} := \mathcal{N}(n-1,1)$ (Section 5.1). Recall the space of local special quasi-homomorphisms $\mathbf{W} \subseteq \mathbf{V}$ (Section 5.2). Recall that $\mathbf{W}$ and $\mathbf{V}$ are non-degenerate Hermitian F-modules of rank n if $F/\mathbb{Q}_p$ is nonsplit (resp. rank $n-1$ and rank n is $F/\mathbb{Q}_p$ is split). Recall $\varepsilon(\mathbf{V}) = -1$ if $F/\mathbb{Q}_p$ is nonsplit (resp. $\varepsilon(\mathbf{V}) = 1$ if $F/\mathbb{Q}_p$ is split).

18.1 Statement of identity

We first define the geometric side of our main local identity, taking $F_0 = \mathbb{Q}_p$. We also assume $p \neq 2$ if $F/\mathbb{Q}_p$ is nonsplit. Let $L^\flat \subseteq \mathbf{W}$ be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank $n-1$. Form the associated local special cycle $\mathcal{Z}(L^\flat) \subseteq \mathcal{N}$. Recall that the flat part $\mathcal{Z}(L^\flat)_{\mathscr{H}} \subseteq \mathcal{Z}(L^\flat)$ decomposes into quasi-canonical lifting cycles $\mathcal{Z}(M^\flat)^\circ$ for certain lattices $M^\flat$ (Proposition 7.3.1). Recall also the derived vertical local special cycle ${}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \in \mathrm{gr}^{n-1}_{\mathcal{N}} K_0'(\mathcal{Z}(L^\flat))_{\mathbb{Q}}$ (Section 5.5).

Definition 18.1.1. Given a non-degenerate Hermitian $\mathcal{O}_F$-lattice $L^\flat \subseteq \mathbf{W}$ of rank $n-1$, the associated *local intersection number* is

$$\mathrm{Int}(L^\flat)_n := \mathrm{Int}_{\mathscr{H}}(L^\flat)_n + \mathrm{Int}_{\mathscr{V}}(L^\flat)_n \tag{18.1.1}$$

where

$$\mathrm{Int}_{\mathscr{H}}(L^\flat)_n := \sum_{\substack{L^\flat \subseteq M^\flat \subseteq M^{\flat *} \\ t(M^\flat) \leq 1}} \mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n \tag{18.1.2}$$

with the sum running over full rank lattices $M^\flat \subseteq L^\flat_F$, where

$$\mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n := 2 \deg \mathcal{Z}(M^\flat)^\circ \cdot \delta_{\mathrm{tau}}(\mathrm{val}'(M^\flat)) \tag{18.1.3}$$

for any non-degenerate integral lattice $M^\flat \subseteq \mathbf{W}$ with $t(M^\flat) \leq 1$, and where

$$\mathrm{Int}_{\mathscr{V}}(L^\flat)_n \coloneqq 2[\breve{F} : \breve{\mathbb{Q}}_p]^{-1} \deg_{\overline{k}}({}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \cdot \mathcal{E}^\vee). \tag{18.1.4}$$

We previously related these local intersection numbers with global intersection numbers (end of Sections 11.8 and 11.9). We are now using local notation, suppressing the p of loc. cit..

The quantity

$$\deg \mathcal{Z}(M^\flat)^\circ = \begin{cases} [\breve{F} : \breve{\mathbb{Q}}_p] p^{\mathrm{val}'(M^\flat)}(1 - \eta^*(p)p^{-1}) & \text{if } \mathrm{val}'(M^\flat) \geq 1 \\ [\breve{F} : \breve{\mathbb{Q}}_p] & \text{if } \mathrm{val}'(M^\flat) = 0 \end{cases} \tag{18.1.5}$$

is the degree of the adic finite flat morphism $\mathcal{Z}(M^\flat)^\circ \to \operatorname{Spf} \mathcal{O}_{\breve{F}}$, where $\eta^*(p) \coloneqq -1, 0, 1$ in the inert, ramified, and split cases respectively (see (7.2.1); the extra factor of $[\breve{F} : \breve{F}_0]$ accounts for the two components of $\mathcal{Z}(M^\flat)^\circ$ when F/F_0 is ramified, see (7.3.1)). Recall that $\delta_{\mathrm{tau}}(s)$ is the "local change of tautological height" defined in (7.2.7), and recall that $\mathcal{E}^\vee$ is the dual tautological bundle on $\mathcal{N}$ (Definition 5.1.9). In (18.1.4), we understand $\mathcal{E}^\vee = [\mathcal{O}_{\mathcal{N}}] - [\mathcal{E}] \in K_0'(\mathcal{N})$ so that ${}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \cdot \mathcal{E}^\vee \in F^n_{\mathcal{N}} K_0'(\mathcal{Z}(L^\flat)_{\overline{k}})_{\mathbb{Q}}$. For $L^\flat$ as above, recall that $\mathcal{Z}(L^\flat)_{\overline{k}}$ is a scheme proper over $\operatorname{Spec} \overline{k}$ (Lemma 11.7.3), so there is a degree map $\deg_{\overline{k}} \colon F^n_{\mathcal{N}} K_0'(\mathcal{Z}(L^\flat)_{\overline{k}}) \to \mathbb{Z}$.

We refer to $\mathrm{Int}_{\mathscr{H}}(L^\flat)_n$ as the "horizontal part" of the local intersection number (coming from the flat part $\mathcal{Z}(L^\flat)_{\mathscr{H}}$) and we refer to $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n$ as the "vertical part" of the local intersection number (coming from $\mathcal{Z}(L^\flat)_{\overline{k}}$, supported in positive characteristic).

We next define the automorphic side of our main local identity. For this, we allow F_0 to be an arbitrary finite extension of $\mathbb{Q}_p$ (allowing $p = 2$ if F/F_0 is unramified). If $L^\flat$ is a non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank $n-1$, we set

$$\partial\mathrm{Den}^*(L^\flat)_n \coloneqq -2[\breve{F} : \breve{F}_0] \frac{d}{dX}\bigg|_{X=1} \mathrm{Den}^*(q^2 X, L^\flat)_n \tag{18.1.6}$$

where $\mathrm{Den}^*(X, L^\flat)_n \in \mathbb{Z}[1/q][X, X^{-1/2}]$ is a normalized local density (15.5.6). We are abusing notation as in Remark 15.5.2, i.e. $\mathrm{Den}^*(q^2 X, L^\flat)_n$ means to evaluate $\mathrm{Den}^*(X, L^\flat)_n$ at $X^{1/2}$ being $qX^{1/2}$. We also set

$$\mathrm{Den}^*(L^\flat)_n \coloneqq [\breve{F} : \breve{F}_0] \cdot \mathrm{Den}^*(q^2, L^\flat)_n. \tag{18.1.7}$$

Suppose $M^\flat$ is a non-degenerate integral Hermitian $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$. If $M^\flat$ is maximal integral,[44] we set $\partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n \coloneqq \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)_n$. Otherwise, we define

[44]The symbol $\circ$ indicates "primitive" here (for quasi-canonical lifting cycles), while $\circ$ indices "spherical" in Part V (Eisenstein series). There is no notation clash as written, but we hope this remark helps to avoid confusion.

$\partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n$ inductively so that the relation

$$\partial\mathrm{Den}^*(M^\flat)_n = \sum_{M^\flat \subseteq N^\flat \subseteq N^{\flat *}} \partial\mathrm{Den}^*_{\mathscr{H}}(N^\flat)^\circ_n \tag{18.1.8}$$

is satisfied (induct on $\mathrm{val}(M^\flat)$), where the sum runs over lattices $N^\flat \subseteq M^\flat_F$. Given any non-degenerate integral Hermitian $\mathcal{O}_F$-lattice $L^\flat$ of rank $n-1$, we then define $\partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n$ so that the relation

$$\partial\mathrm{Den}^*(L^\flat) = \left(\sum_{\substack{L^\flat \subseteq M^\flat \subseteq M^{\flat *} \\ t(M^\flat) \leq 1}} \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n \right) + \partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n \tag{18.1.9}$$

is satisfied, where the sum runs over lattices $M^\flat \subseteq L^\flat_F$.

Theorem 18.1.2. *Suppose $F_0 = \mathbb{Q}_p$ and that $p \neq 2$ unless $F/\mathbb{Q}_p$ is split. For any non-degenerate Hermitian $\mathcal{O}_F$-lattice $L^\flat \subseteq \mathbf{W}$ of rank $n-1$, we have*

$$\mathrm{Int}(L^\flat)_n = \partial\mathrm{Den}^*(L^\flat)_n. \tag{18.1.10}$$

Moreover, we have

$$\mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n \qquad \mathrm{Int}_{\mathscr{V}}(L^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n \tag{18.1.11}$$

where $M^\flat \subseteq \mathbf{W}$ is any non-degenerate integral Hermitian $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$.

On account of the decompositions in (18.1.1), (18.1.2), and (18.1.9), it is clearly enough to prove the refined identities in (18.1.11). The theorem is also clear if $L^\flat$ is not integral, since both sides of (18.1.10) are zero in this case (the special cycle $\mathcal{Z}(L^\flat)$ will be empty, and $\mathrm{Den}(X, L^\flat)_n$ will be identically zero as discussed in Section 15.4).

We also record a special value formula (as observed in the inert case by Li and Zhang [LZ22a, Corollary 4.6.1]) for later use. Its proof will appear in Section 18.2.

Lemma 18.1.3. *Suppose $F_0 = \mathbb{Q}_p$ and that $p \neq 2$ unless $F/\mathbb{Q}_p$ is split. For any non-degenerate Hermitian $\mathcal{O}_F$-lattice $L^\flat \subseteq \mathbf{W}$ of rank $n-1$, we have*

$$\deg \mathcal{Z}(L^\flat)_{\mathscr{H}} = \mathrm{Den}^*(L^\flat)_n. \tag{18.1.12}$$

In the preceding lemma statement, $\deg \mathcal{Z}(L^\flat)_{\mathscr{H}}$ means the degree of the adic finite flat morphism $\mathcal{Z}(L^\flat)_{\mathscr{H}} \to \mathrm{Spf}\, \mathcal{O}_{\breve{F}}$ of formal schemes.

18.2 Horizontal identity

We will need Cho–Yamauchi formulas for local densities (unitary version, as proved in [LZ22a, Theorem 3.5.1] (inert) [FYZ24, Theorem 2.2(3)] (split) [LL22, Lemma 2.15] (ramified)). For this, we allow F_0 to be an arbitrary finite extension of $\mathbb{Q}_p$ (allowing $p = 2$ if F/F_0 is unramified). Then, if L is any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank n (still assuming n even if F/F_0 is ramified), we have

$$\mathrm{Den}(X, L)_n = \sum_{L \subseteq M \subseteq M^*} X^{\ell(M/L)} \mathrm{Den}(X, M)^\circ_n \tag{18.2.1}$$

$$\mathrm{Den}(X, M)^\circ_n := \prod_{i=0}^{t(M)-1} (1 - \eta^{*i}(\varpi_0) q^i X) \tag{18.2.2}$$

where $\eta^*(\varpi_0) := \eta^{*i}(\varpi_0) := -1, 0, 1$ if i is odd (resp. $\eta^{*i}(\varpi_0) := 1$ if i is even) in the inert, ramified, split cases respectively, and $\mathrm{Den}(X, L)_n \in \mathbb{Z}[X]$ is the local density polynomial normalized as in Section 15.5. The displayed sum runs over lattices $M \subseteq L_F$.

Suppose $L^\flat$ is a Hermitian $\mathcal{O}_F$-lattice of rank $n - 1$ (still assuming n even if F/F_0 is ramified). If F/F_0 is unramified, we have $\mathrm{Den}(X, L^\flat)_n = \mathrm{Den}(\eta^*(\varpi_0)q^{-1}X, L^\flat)_{n-1}$ (15.5.5) and we set $\mathrm{Den}(X, L^\flat)^\circ_n := \mathrm{Den}(\eta^*(\varpi_0)q^{-1}, L^\flat)^\circ_{n-1}$ if $L^\flat$ is also integral.

If F/F_0 is ramified, we have

$$\mathrm{Den}(X, L^\flat)_n = \sum_{L^\flat \subseteq M^\flat \subseteq M^{\flat *}} (q^{-1}X)^{\ell(M^\flat/L^\flat)} \mathrm{Den}(X, M^\flat)^\circ_n \tag{18.2.3}$$

$$\mathrm{Den}(X, M^\flat)^\circ_n := \prod_{i=0}^{\frac{t(M^\flat)-3}{2}} (1 - q^{2i} X) \tag{18.2.4}$$

where the sum runs over lattices $M^\flat \subseteq L^\flat_F$ (may be verified using [LL22, Lemma 2.15]).

If $M^\flat$ is a non-degenerate integral Hermitian $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$, set

$$\mathrm{Den}^*(M^\flat)^\circ_n := [\breve{F} : \breve{F}_0] q^{\mathrm{val}'(M^\flat)} \mathrm{Den}(1, M^\flat)^\circ_n. \tag{18.2.5}$$

We have

$$[\breve{F} : \breve{F}_0] \mathrm{Den}^*(q^2, L^\flat)_n = [\breve{F} : \breve{F}_0] \mathrm{Den}^*(1, L^\flat)_n = \sum_{L^\flat \subseteq M^\flat \subseteq M^{\flat *}} \mathrm{Den}^*(M^\flat)^\circ_n \tag{18.2.6}$$

where the sum runs over lattices $M^\flat \subseteq L^\flat_F$. The first equality follows from the functional equation (16.1.6), and the second equality follows from the Cho–Yamauchi formulas (and (15.5.9)). Note $\mathrm{Den}^*(M^\flat)_n = \mathrm{Den}^*(M^\flat)^\circ_n$ if $M^\flat$ is maximal integral.

Proof of Lemma 18.1.3. Follows from (18.2.6). Note $\mathrm{Den}^*(M^\flat)^\circ_n = \deg \mathcal{Z}(M^\flat)^\circ$ if $t(M^\flat) \leq 1$, and $\mathrm{Den}^*(M^\flat)^\circ_n = 0$ if $t(M^\flat) \geq 2$. □

Proposition 18.2.1. *Assume $F_0 = \mathbb{Q}_p$ and that $p \neq 2$ unless $F/\mathbb{Q}_p$ is split. For any rank $n-1$ non-degenerate Hermitian $\mathcal{O}_F$-lattice $M^\flat \subseteq \mathbf{W}$ with $t(M^\flat) \leq 1$, we have*

$$\mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n \tag{18.2.7}$$

Proof. By definition, the quantity $\mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n$ depends only on $\mathrm{val}(M^\flat)$. Since $t(M^\flat) \leq 1$, we may write $M^\flat = L^{\flat\prime} \oplus L^{\flat\prime\prime}$ (orthogonal direct sum) where $L^{\flat\prime}$ is self-dual of rank $n-2$ and $\mathrm{val}(L^{\flat\prime\prime}) = \mathrm{val}(M^\flat)$ (in the unramified case, this follows upon diagonalizing $M^\flat$; in the ramified case, this follows from picking a "standard basis" as in [LL22, Lemma 2.12]). Using the cancellation property of local densities explained in (15.5.11), we thus reduce to the case $n = 2$ (which we now assume).

By the inductive decompositions in (18.1.2) and (18.1.8), it is enough to show $\mathrm{Int}_{\mathscr{H}}(M^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)_n$ (induct on $\mathrm{val}(M^\flat)$). We have $\partial\mathrm{Den}^*(M^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)_n$ by construction, since $t(M^\flat) \leq 1$ (i.e. compare (18.1.8) and (18.1.9)).

Set $b = \mathrm{val}'(M^\flat)$. Using the Cho–Yamauchi formulas, we find

$$\mathrm{Den}^*(q^2X, M^\flat)_n = X^{-b/2} \sum_{j=0}^{b} (qX)^j \qquad \partial\mathrm{Den}^*(M^\flat)_n = [\breve{F} : \breve{F}_0] \sum_{j=0}^{b} (b-2j) q^j \tag{18.2.8}$$

in all cases. The preceding formulas are valid even if $F_0 \neq \mathbb{Q}_p$ (and also valid if $p = 2$ whenever F/F_0 is unramified), hence why we wrote q instead of p.

We have

$$\begin{aligned} \mathrm{Int}_{\mathscr{H}}(M^\flat)_n &= 2[\breve{F} : \breve{\mathbb{Q}}_p] \sum_{M^\flat \subseteq N^\flat \subseteq N^{\flat *}} p^s (1 - \eta^*(p) p^{-1}) \delta_{\mathrm{tau}}(s) \\ &= -[\breve{F} : \breve{\mathbb{Q}}_p] \sum_{M^\flat \subseteq N^\flat \subseteq N^{\flat *}} p^s (1 - \eta^*(p) p^{-1}) \left(s - \frac{(1-p^{-s})(1-\eta^*(p))}{(1-p^{-1})(p-\eta^*(p))} \right) \end{aligned} \tag{18.2.9}$$

where the sum runs over lattices $N^\flat \subseteq M^\flat_F$, where $s \coloneqq \mathrm{val}'(N^\flat)$, and where $\eta^*(p) \coloneqq -1, 0, 1$ in the inert, ramified, split cases respectively.

We prove the identity $\mathrm{Int}_{\mathscr{H}}(M^\flat)_n = \partial\mathrm{Den}^*(M^\flat)_n$ by induction on b. The case $b = 0$ is clear, as both quantities are 0.

Next suppose $b \geq 1$ and that $M^{\flat\prime}$ (resp. $M^{\flat\prime\prime}$) is a rank one non-degenerate lattice with $\mathrm{val}'(M^{\flat\prime}) = b - 1$ (resp. $\mathrm{val}'(M^{\flat\prime\prime}) = b - 2$). If $b - 2 \leq 0$, set $\mathrm{Int}_{\mathscr{H}}(M^{\flat\prime\prime})_n \coloneqq 0$ (in which case $\partial\mathrm{Den}^*(M^{\flat\prime\prime})_n = 0$ as well).

We have

$$\partial\mathrm{Den}^*(M^\flat)_n - \partial\mathrm{Den}^*(M^{\flat\prime})_n = [\breve{F} : \breve{F}_0](-bq^b + \sum_{j=0}^{b-1} q^j) \tag{18.2.10}$$

$$\partial\mathrm{Den}^*(M^\flat)_n - \partial\mathrm{Den}^*(M^{\flat\prime\prime})_n = [\breve{F} : \breve{F}_0](-bq^b - bq^{b-1} + 2\sum_{j=0}^{b-1} q^j). \tag{18.2.11}$$

If $F/\mathbb{Q}_p$ is inert, we find

$$\begin{aligned}\mathrm{Int}_{\mathscr{H}}(M^\flat)_n - \mathrm{Int}_{\mathscr{H}}(M^{\flat\prime\prime})_n &= -[\breve{F} : \breve{\mathbb{Q}}_p]p^b(1+p^{-1})\left(b - 2\frac{(1-p^{-b})}{(1-p^{-1})(p+1)}\right) \\ &= [\breve{F} : \breve{\mathbb{Q}}_p](-bp^b - bp^{b-1} + 2\sum_{j=0}^{b-1} p^j). \end{aligned} \tag{18.2.12}$$

If $F/\mathbb{Q}_p$ is ramified, we find

$$\mathrm{Int}_{\mathscr{H}}(M^\flat)_n - \mathrm{Int}_{\mathscr{H}}(M^{\flat\prime})_n = -[\breve{F} : \breve{\mathbb{Q}}_p]p^b\left(b - \frac{(1-p^{-b})}{(1-p^{-1})p}\right) \tag{18.2.13}$$

$$= [\breve{F} : \breve{\mathbb{Q}}_p](-bp^b + \sum_{j=0}^{b-1} p^j). \tag{18.2.14}$$

If $F/\mathbb{Q}_p$ is split, we find

$$\mathrm{Int}_{\mathscr{H}}(M^\flat)_n - \mathrm{Int}_{\mathscr{H}}(M^{\flat\prime})_n = -[\breve{F} : \breve{\mathbb{Q}}_p]\sum_{j=0}^{b} p^j(1-p^{-1})j \tag{18.2.15}$$

$$= [\breve{F} : \breve{\mathbb{Q}}_p](-bp^b + \sum_{j=0}^{b-1} p^j). \tag{18.2.16}$$

This proves the lemma in all cases, by induction on b. □

Corollary 18.2.2. *Theorem 18.1.2 holds when $n = 2$.*

Proof. If $n = 2$, Proposition 18.2.1 shows $\mathrm{Int}_{\mathscr{H}}(L^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{H}}(L^\flat)_n = \partial\mathrm{Den}^*(L^\flat)_n$. We have $\mathrm{gr}^1_{\mathcal{N}} K'_0(\mathcal{Z}(L^\flat)_{\overline{k}})_{\mathbb{Q}} = 0$ because $\mathcal{Z}(L^\flat)_{\overline{k}}$ is a scheme and because the reduced subscheme $\mathcal{N}_{\mathrm{red}} \subseteq \mathcal{N}$ is 0-dimensional (a disjoint union of copies of $\operatorname{Spec} \overline{k}$), see Lemma 5.4.1. Hence $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n = 0$ since ${}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \in \mathrm{gr}^1_{\mathcal{N}} K'_0(\mathcal{Z}(L^\flat)_{\overline{k}})_{\mathbb{Q}}$. □

18.3 Induction formula

Throughout Sections 18.3 and 18.5, we allow F_0 to be an arbitrary finite extension of $\mathbb{Q}_p$ (allowing $p = 2$ if F/F_0 is unramified). We take the following setup for the rest of Section 18 (i.e. the notations n, V, L, L', L'', $L^\flat$, x, x', and x'' are all reserved unless otherwise indicated).

Setup 18.3.1. Let V be a non-degenerate Hermitian F-module of rank n, with pairing $(-,-)$. Assume n is even if F/F_0 is ramified. Let $L^\flat \subseteq V$ be a non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank $n-1$. Let $x, x', x'' \in V$ be nonzero and orthogonal to $L^\flat$ with $\langle x \rangle \subseteq \langle x' \rangle \subseteq \langle x'' \rangle$ and $\mathrm{length}_{\mathcal{O}_F}(\langle x' \rangle / \langle x \rangle) = \mathrm{length}_{\mathcal{O}_F}(\langle x'' \rangle / \langle x' \rangle) = 1$. Set

$$L \coloneqq L^\flat \oplus \langle x \rangle \qquad L' \coloneqq L^\flat \oplus \langle x' \rangle \qquad L'' \coloneqq L^\flat \oplus \langle x'' \rangle \tag{18.3.1}$$

The notation L'' and x'' will only appear in our proof of the induction formula (Proposition 18.3.2) for F/F_0 split.

Proposition 18.3.2 (Induction formula)**.** *If* $\mathrm{val}(x) > a_{\max}(L^\flat)$ *in the nonsplit cases (resp. if* $\mathrm{val}(x) > 2a_{\max}(L^\flat)$ *in the case* F/F_0 *is split), we have*

$$\mathrm{Den}(X, L)_n = \begin{cases} X^2 \mathrm{Den}(X, L')_n + (1-X)\mathrm{Den}(q^2 X, L^\flat)_n & \text{if } F/F_0 \text{ is inert} \\ X \mathrm{Den}(X, L')_n + (1-X)\mathrm{Den}(q^2 X, L^\flat)_n & \text{if } F/F_0 \text{ is ramified} \\ X \mathrm{Den}(X, L')_n + \mathrm{Den}(q^2 X, L^\flat)_n & \text{if } F/F_0 \text{ is split.} \end{cases} \tag{18.3.2}$$

In the inert case, this is [Ter13, Theorem 5.1] (strictly speaking, there is a blanket $p \neq 2$ assumption there), which is a unitary analogue of [Kat99, Theorem 2.6(1)] (orthogonal groups); see also [LZ22a, Proposition 3.7.1] (there stated allowing $p = 2$) for a statement closer to ours.

Using the Cho–Yamauchi formulas, we give a uniform proof of the inert and ramified cases (Lemma 18.3.6). Our lower bounds on $\mathrm{val}(x)$ are possibly nonsharp (e.g. in the inert case, we only show the induction formula when $\mathrm{val}(x) > 2a_{\max}(L^\flat)$) but this makes no difference for the proof of Theorem 18.1.2, where we will take $\mathrm{val}(x) \to \infty$ (Proposition 18.5.2).

The case when F/F_0 is split is more difficult for us, and the same proof only shows a weaker version of the induction formula (stated in Lemma 18.3.6), which is insufficient for our purposes. Extracting the induction formula from this weak version is the subject of Section 18.4.

For the proof of Theorem 18.1.2, only the statement of the induction formula and the definitions in (18.3.3) and (18.3.4) will be needed.

We first record a few preparatory lemmas. As in Section 2.2, we fix a uniformizer $\varpi \in \mathcal{O}_F$ and a generator $u \in \mathcal{O}_F$ of the different ideal such that $\varpi^\sigma = -\varpi$ and $u^\sigma = -u$. We say a quantity *stabilizes*, e.g. for $\mathrm{val}(x) > C$ (for some constant C) if that quantity does not depend on x if $\mathrm{val}(x) > C$. When F/F_0 is nonsplit, given an $\mathcal{O}_F$-module M and $m \in M$, we say e.g. that m is *exact ϖ^e-torsion* for $e \geq 0$ if $\varpi^e m = 0$ but $\varpi^{e-1} m \neq 0$ (and if $e = 0$, the only exact ϖ^e-torsion element is 0). We use similar terminology for $\mathcal{O}_{F_0}$-modules and exact ϖ_0^e-torsion elements, etc..

Lemma 18.3.3. *Let M be a non-degenerate integral Hermitian $\mathcal{O}_F$-lattice of rank m. Suppose elements $w_1, \ldots, w_r \in M$ have $\mathcal{O}_F$-span M. Write T for the associated Gram matrix. Then $t(M) + \mathrm{rank}((uT) \otimes \mathcal{O}_F/\varpi) = m$.*

Proof. If F/F_0 is split, the rank continues to make sense because T is Hermitian (e.g. diagonalize the Hermitian form). In the unramified cases, the lemma follows by diagonalizing the Hermitian form. In the ramified case, the lemma follows by putting M into "standard form" (i.e. an orthogonal direct sum of rank one lattices and rank two hyperbolic lattices) as in [LL22, Definition 2.11]. We are allowing m even or odd. □

Lemma 18.3.4. *Let M be a non-degenerate integral Hermitian $\mathcal{O}_F$-lattice of rank m. Suppose $M_F = W' \oplus W''$ is an orthogonal decomposition with W'' of rank 1.*

(1) We have $t(M) - 1 \leq t(M \cap W') \leq t(M) + 1$.

(2) Let $M' \subseteq W'$ and $M'' \subseteq W''$ be the images of M under the projections $M_F \to W'$ and $M_F \to W''$. Assume that M' and M'' are integral and that $\mathrm{val}(M'') > 0$. Then we have $t(M) = t(M') + 1$.

Proof.

(1) The ramified case follows from [LL22, Lemma 2.23(2)]. The inert case when $t(M) = 0$ is [LZ22a, Lemma 4.5.1]. The same proof works in general for arbitrary F/F_0 in arbitrary characteristic: select any basis $(w_1, \ldots, w_{m-1})$ of $N \cap W'$, extend to a basis $(w_1, \ldots, w_m)$ of M with Gram matrix T, then use the formulas $t(M) + \mathrm{rank}((uT) \otimes \mathcal{O}_F/\varpi) = m$ and $t(M \cap W') + \mathrm{rank}((uT^\flat) \otimes \mathcal{O}_F/\varpi) = m - 1$.

(2) Let $\underline{w} = [w_1, \ldots, w_m]$ be any basis of M, and let $T = (\underline{w}, \underline{w})$ be the corresponding Gram matrix. Let $\underline{w}' = [w'_1, \ldots, w'_m]$ be the projection of $\underline{w}$ to W', with Gram matrix $T' = (\underline{w}', \underline{w}')$. Since $\mathrm{val}(M'') > 0$, we see $(uT) \otimes \mathcal{O}_F/\varpi = (uT') \otimes \mathcal{O}_F/\varpi$. Applying Lemma 18.3.3 twice (once for M and $r = m$ and once for M' and $r = m$) proves the claim. □

Set

$$\mathrm{Den}_{L^\flat,x}(X)^\circ_n \coloneqq \sum_{\substack{L\subseteq M\subseteq M^* \\ M\cap L^\flat_F=L^\flat}} X^{\ell(M/L)}\mathrm{Den}(X,M)^\circ_n \tag{18.3.3}$$

$$\mathrm{Den}_{L^\flat,x}(X)_n \coloneqq \mathrm{Den}(X,L)_n = \sum_{L^\flat\subseteq M^\flat\subseteq M^{\flat *}} \mathrm{Den}_{M^\flat,x}(X)^\circ_n \tag{18.3.4}$$

where the first sum runs over lattices $M \subseteq V$ and the second sum runs over lattices $M^\flat \subseteq L^\flat_F$. Note that the only dependence on x is on $\mathrm{val}(x)$ (since $\mathrm{Den}_{L^\flat,x}(X)^\circ_n$ only depends on the isomorphism class of the Hermitian lattice L).

Lemma 18.3.5. *The polynomial*

$$f_x(X) \coloneqq \begin{cases} \mathrm{Den}_{L^\flat,x}(X)^\circ_n - X^2\mathrm{Den}_{L^\flat,x'}(X)^\circ_n & \textit{if } F/F_0 \textit{ is inert} \\ \mathrm{Den}_{L^\flat,x}(X)^\circ_n - X\mathrm{Den}_{L^\flat,x'}(X)^\circ_n & \textit{if } F/F_0 \textit{ is ramified} \\ \mathrm{Den}_{L^\flat,x}(X)^\circ_n - 2X\mathrm{Den}_{L^\flat,x'}(X)^\circ_n + X^2\mathrm{Den}_{L^\flat,x''}(X)^\circ_n & \textit{if } F/F_0 \textit{ is split} \end{cases} \tag{18.3.5}$$

(an element of $\mathbb{Z}[X]$) stabilizes for $\mathrm{val}(x) > 2[\breve{F} : \breve{F}_0]^{-1}a_{\max}(L^\flat)$.

Proof. The notation $f_x(X)$ is temporary, used only for this lemma. If F/F_0 is split, write $\varpi = \varpi_1\varpi_2$ for $\varpi_i \in \mathcal{O}_F$ with $\mathrm{val}(\varpi_1 x) = \mathrm{val}(\varpi_2 x) = \mathrm{val}(x) + 1$. We may assume $x'' = \varpi^{-1}x$ in this case. Note $\mathrm{Den}_{L^\flat,x'}(X)^\circ_n = \mathrm{Den}_{L^\flat,\varpi_1^{-1}x}(X)^\circ_n = \mathrm{Den}_{L^\flat,\varpi_2^{-1}x}(X)^\circ_n$.

Inspecting (18.3.3) shows

$$f_x(X) = \sum_{\substack{L\subseteq M\subseteq M^* \\ M\cap L^\flat_F=L^\flat \\ \varpi^{-1}x\notin M}} X^{\ell(M/L)}\mathrm{Den}(X,M)^\circ_n \qquad \text{if } F/F_0 \text{ is nonsplit} \tag{18.3.6}$$

$$f_x(X) = \sum_{\substack{L\subseteq M\subseteq M^* \\ M\cap L^\flat_F=L^\flat \\ \varpi_1^{-1}x\notin M \\ \varpi_2^{-1}x\notin M}} X^{\ell(M/L)}\mathrm{Den}(X,M)^\circ_n \qquad \text{if } F/F_0 \text{ is split} \tag{18.3.7}$$

where the sums run over lattices $M \subseteq V$. For each such M, we know $L^\flat \subseteq M$ is a saturated sublattice, hence $M = L^\flat \oplus \langle \xi \rangle$ (not necessarily orthogonal direct sum) for some $\xi \in V$.

If $L^\flat$ is not integral, then the lemma is trivial as the polynomials of the lemma statement are 0. We assume $L^\flat$ is integral for the rest of the proof.

If F/F_0 is nonsplit, each lattice M appearing in (18.3.6) is of the form $M = L^\flat \oplus \langle y + \varpi^{-e}x\rangle$ for a uniquely determined element $y \in L^{\flat *}/L^\flat$, where $e \in \mathbb{Z}_{\geq 0}$ is such that $y \in L^{\flat *}/L^\flat$ is of exact ϖ^e-torsion. Conversely, an element $y \in L^{\flat *}/L^\flat$ gives rise to an M appearing in (18.3.6) if and only if $\mathrm{val}(y + \varpi^{-e}x) \geq 0$. If $\mathrm{val}(x) > 2[\breve{F} : \breve{F}_0]^{-1} a_{\max}(L^\flat)$, then $\mathrm{val}(\varpi^{-e}x) > 0$, so $\mathrm{val}(y + \varpi^{-e}x) \geq 0$ holds if and only if $\mathrm{val}(y) \geq 0$.

If F/F_0 is split, the preceding paragraph holds upon replacing ϖ^{-e} with $\varpi_1^{-e_1}\varpi_2^{-e_2}$ for $e_1, e_2 \in \mathbb{Z}_{\geq 0}$ such that $y \in L^{\flat *}/L^\flat$ is of exact $\varpi_1^{e_1}\varpi_2^{e_2}$-torsion (i.e. $\varpi_1^{e_1}\varpi_2^{e_2} y \in L^\flat$ but $\varpi_1^{e_1-1}\varpi_2^{e_2} \notin L^\flat$ and $\varpi_1^{e_1}\varpi_2^{e_2-1} \notin L^\flat$).

In the previous notation, we thus have

$$f_x(X) = \sum_{\substack{y \in L^{\flat *}/L^\flat \\ \mathrm{val}(y) \geq 0}} X^{\ell((L^\flat + \langle y \rangle)/L^\flat)} \mathrm{Den}(X, M)_n^\circ \tag{18.3.8}$$

where the sum runs over y, and $M = L^\flat \oplus \langle y + \varpi^{-e}x \rangle$ in the nonsplit case (resp. $M = L^\flat \oplus \langle y + \varpi_1^{-e_1}\varpi_2^{-e_2}x \rangle$ in the split case).

In the notation of (18.3.8), we have $t(M) = t(L^\flat + \langle y \rangle) + 1$ by Lemma 18.3.4(2) (using $\mathrm{val}(\varpi^{-e}x) > 0$ in the nonsplit case and $\mathrm{val}(\varpi_1^{-e_2}\varpi_2^{-e_2}x) > 0$ in the split case). Hence we have

$$\mathrm{Den}(X, M)_n^\circ = \prod_{i=0}^{t(M^\flat + \langle y \rangle)} (1 - \eta^{*i}(\varpi_0) q^i X) \tag{18.3.9}$$

(see definition in (18.2.2)), and now the right-hand side of (18.3.8) clearly depends only on $L^\flat$ (and not on x). □

Lemma 18.3.6. *With notation as above, assume* $\mathrm{val}(x) > 2[\breve{F} : \breve{F}_0]^{-1} a_{\max}(L^\flat)$. *We have*

$$(1-X)\mathrm{Den}(q^2X, L^\flat)_n = \begin{cases} \mathrm{Den}(X, L) - X^2 \mathrm{Den}(X, L') & \text{if } F/F_0 \text{ is inert} \\ \mathrm{Den}(X, L) - X \mathrm{Den}(X, L') & \text{if } F/F_0 \text{ is ramified} \\ \mathrm{Den}(X, L) - 2X \mathrm{Den}(X, L') + X^2 \mathrm{Den}(X, L'') & \text{if } F/F_0 \text{ is split.} \end{cases} \tag{18.3.10}$$

Proof. Combining (18.3.8) and (18.3.9), we find that the right-hand side of (18.3.10) is given by

$$\sum_{L^\flat \subseteq M^\flat \subseteq M^{\flat *}} X^{\ell(M^\flat/L^\flat)} \sum_{\substack{y \in M^{\flat *}/M^\flat \\ \mathrm{val}(y) \geq 0}} X^{\ell((M^\flat + \langle y \rangle)/M^\flat)} \prod_{i=0}^{t(M^\flat + \langle y \rangle)} (1 - \eta^{*i}(\varpi_0) q^i X) \tag{18.3.11}$$

in all cases, where the outer sum runs over lattices $M^\flat \subseteq L_F^\flat$.

Collecting the terms with $M^\flat + \langle y \rangle = N^\flat$ for fixed integral lattices $N^\flat \subseteq L_F^\flat$, we find that (18.3.11) is equal to

$$\sum_{L^\flat \subseteq N^\flat \subseteq N^{\flat *}} X^{\ell(N^\flat/L^\flat)} \prod_{i=0}^{t(N^\flat)} (1 - \eta^{*i}(\varpi_0) q^i X) \sum_{\substack{L^\flat \subseteq M^\flat \subseteq N^\flat \\ N^\flat/M^\flat \text{cyclic}}} (\text{number of generators of } N^\flat/M^\flat)$$

where the outer sum runs over lattices $N^\flat \subseteq L_F^\flat$ and the inner sum runs over lattices $M^\flat$. We have

$$\begin{aligned}\sum_{\substack{L^\flat \subseteq M^\flat \subseteq N^\flat \\ N^\flat/M^\flat \text{cyclic}}} (\text{number of generators of } N^\flat/M^\flat) &= \sum_{\substack{N^{\flat *} \subseteq M^{\flat *} \subseteq L^{\flat *} \\ M^{\flat *}/N^{\flat *} \text{cyclic}}} (\text{number of generators of } M^{\flat *}/N^{\flat *}) \\ &= |L^{\flat *}/N^{\flat *}| = |N^\flat/L^\flat| = q^{\ell(N^\flat/L^\flat)}\end{aligned}$$

so (18.3.11) is equal to

$$\sum_{L^\flat \subseteq N^\flat \subseteq N^{\flat *}} (qX)^{\ell(N^\flat/L^\flat)} \prod_{i=0}^{t(N^\flat)} (1 - \eta^{*i}(\varpi_0) q^i X). \tag{18.3.12}$$

Inspecting the Cho–Yamauchi formulas (and surrounding discussion) at the beginning of Section 18.2 shows that the displayed expression is equal to $(1-X)\mathrm{Den}(q^2X, L^\flat)_n$ in all cases (if F/F_0 is ramified, note that $t(N^\flat)$ is always odd because $N^\flat$ has rank $n-1$, which we have assumed is odd in the ramified case). $\square$

18.4 More on induction formula: split

Suppose F/F_0 is split. To prove the induction formula (Proposition 18.3.2), it remains only to show that $\mathrm{Den}(X, L) - X\mathrm{Den}(X, L')$ stabilizes for $\mathrm{val}(x) > 2a_{\max}(L^\flat)$, as Lemma 18.3.6 then shows $(1-X)(\mathrm{Den}(X, L) - X\mathrm{Den}(X, L')) = (1-X)\mathrm{Den}(q^2X, L^\flat)_n$.

We define some more notation (only used in Section 18.4). Fix a uniformizer ϖ_0 of $\mathcal{O}_{F_0}$, and consider the elements

$$\varpi_1 = (\varpi_0, 1) \qquad \varpi_2 = (1, -\varpi_0) \qquad e_1 = (1, 0) \qquad e_2 = (0, 1) \tag{18.4.1}$$

in $\mathcal{O}_F = \mathcal{O}_{F_0} \times \mathcal{O}_{F_0}$. Given an $\mathcal{O}_F$-module M, we set $M_1 \coloneqq e_1 M$ and $M_2 \coloneqq e_2 M$ (so $M = M_1 \oplus M_2$). We similarly write $y_1 \coloneqq e_1 y$ and $y_2 \coloneqq e_2 y$ for $y \in M$. If M is a non-degenerate Hermitian $\mathcal{O}_F$-lattice, we set $M_1^* \coloneqq e_2 M^*$ and $M_2^* \coloneqq e_1 M^*$. If M is moreover integral, the Hermitian pairing induces an identification $M_2^*/M_1 \cong \mathrm{Hom}_{\mathcal{O}_{F_0}}(M_1^*/M_2, F_0/\mathcal{O}_{F_0})$.

For integers $t \geq 0$, we set

$$\mathfrak{m}(t, X) \coloneqq \prod_{i=0}^{t-1} (1 - q^i X) \tag{18.4.2}$$

so that $\mathrm{Den}(X, M)^\circ_n = \mathfrak{m}(t(M), x)$ for any integral non-degenerate Hermitian $\mathcal{O}_F$-lattice M of rank n. If $\mathcal{T}$ is a finite length $\mathcal{O}_{F_0}$-module, we set

$$t_0(\mathcal{T}) \coloneqq \dim_{\mathbb{F}_q}(\mathcal{T} \otimes_{\mathcal{O}_{F_0}} \mathbb{F}_q) \qquad \ell_0(\mathcal{T}) \coloneqq \mathrm{length}_{\mathcal{O}_{F_0}}(\mathcal{T}). \tag{18.4.3}$$

Lemma 18.4.1. *Consider the polynomial*

$$h_{\mathrm{diff},x}(X) \coloneqq \sum_{\substack{L \subseteq M \subseteq M^* \\ M_1 \cap L^\flat_F = L^\flat_1 \\ M_2 \cap L^\flat_F \neq L^\flat_2 \\ M/L \text{ is cyclic} \\ \varpi_1^{-1} x \notin M}} X^{\ell(M/L)} \mathfrak{m}(t(M), X) \tag{18.4.4}$$

where the sum runs over lattices $M \subseteq V$ (satisfying the displayed conditions). This sum stabilizes for $\mathrm{val}(x) > 2a_{\max}(L^\flat)$.

Proof. Each lattice M in the sum is of the form $M = L + \langle \xi \rangle$ for a unique element $\xi = y + \varpi_1^{-e_1}\varpi_2^{-e_2} x \in L^*/L$ with $y \in L^{\flat *}$, such that $\mathrm{val}(\xi) \geq 0$, and with $e_1, e_2 \in \mathbb{Z}_{\geq 0}$.

Assume $\mathrm{val}(x) > 2a_{\max}(L^\flat)$. We claim that $\mathrm{val}(y) \geq 0$ (in the notation above). The additional conditions on M imply that $y_1 \in L^{\flat *}_2 / L^\flat_1$ is of exact $\varpi_1^{e_1}$-torsion and that $\varpi_2^{e_2} y_2 \notin L^\flat$. We thus have $e_1 \leq a_{\max}(L^\flat)$ and $e_2 < a_{\max}(L^\flat)$, so $\mathrm{val}(\varpi_1^{-e_1}\varpi_2^{-e_2} x) > 0$ when $\mathrm{val}(x) > 2a_{\max}(L^\flat)$. This implies that $\mathrm{val}(y) \geq 0$ as well.

Consider the F-linear (non-unitary) automorphism $\phi \colon V \to V$ which is the identity on $L^\flat_F$ and sends $x \mapsto \varpi_2 x$. Then $M \mapsto \phi(M)$ is a bijection from the set of lattices appearing in the sum for $h_{\mathrm{diff},x}(X)$ to the set of lattices appearing in the sum for $h_{\mathrm{diff},\varpi_2 x}(X)$ (we remind the reader that L depends on x as well).

In the above setup, an application of Lemma 18.3.4(2) shows $t(M) = t(\phi(M)) = t(L^\flat + \langle y \rangle) + 1$. We also find $\ell(M/L) = \ell(\phi(M)/(L^\flat \oplus \langle \varpi_2 x \rangle)) = \ell((L^\flat + \langle y \rangle)/L^\flat)$. This shows $h_{\mathrm{diff},x}(X) = h_{\mathrm{diff},\varpi_2 x}(X)$ (compare the M term and the $\phi(M)$ term). This proves the lemma, as the x-dependence of $h_{\mathrm{diff},x}(X)$ is only on $\mathrm{val}(x)$. $\square$

Lemma 18.4.2. *Let $\mathcal{T}$ be a finite length $\mathcal{O}_{F_0}$-module, and suppose $\mathcal{T}$ is ϖ_0^e-torsion. For any integer $b > e$, form the $\mathcal{O}_{F_0}$-module $A = \mathcal{T} \oplus (\varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0})$. Consider $u = t + w \in A$ with $t \in \mathcal{T}$ and $w \in \varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}$ both of exact ϖ_0^r-torsion. There is a (non-canonical) isomorphism*

$$A/(u) \cong (\mathcal{T}/(t)) \oplus (\varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}). \tag{18.4.5}$$

Proof. This follows from the structure theorem for finitely generated modules over the discrete valuation ring $\mathcal{O}_{F_0}$. For example, we can select elements $e_1, \ldots, e_m \in \mathcal{T}$ such that $\mathcal{T} = \langle e_1 \rangle \oplus \cdots \oplus \langle e_n \rangle$ and such that $t = \varpi_0^s e_1$ for some $s \geq 0$. The case $r = 0$ is trivial, so take $r \geq 1$. Then $r + s \leq e$. If $w' \in \varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}$ is such that $\varpi_0^s w' = w$ there is an isomorphism

$$\mathcal{T} \oplus (\varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}) \to A \tag{18.4.6}$$

sending $e_1 \mapsto e_1 + w'$, $e_i \mapsto e_i$ for $i \geq 2$, and $z \mapsto z$ (for any generator z of $\varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}$). This isomorphism takes t to $t + w$. $\square$

Given a finite torsion cyclic $\mathcal{O}_{F_0}$-module $N \cong \mathcal{O}_{F_0}/\varpi_0^a\mathcal{O}_{F_0}$, we set $\mathrm{ord}(N) \coloneqq a$.

Lemma 18.4.3. *Let $\mathcal{T}$ be a finite length $\mathcal{O}_{F_0}$-module, and assume $\mathcal{T}$ is ϖ^e-torsion for some $e \geq 0$. For any integer $b \geq 0$, form the $\mathcal{O}_{F_0}$-module $A_b \coloneqq \mathcal{T} \oplus (\varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0})$. The polynomial*

$$\alpha_b \coloneqq \sum_{\substack{\textit{cyclic submodules} \\ N \subseteq A_b}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_b/N), X) \tag{18.4.7}$$

stabilizes for $b > e$.

Proof. Applying $- \otimes_{\mathcal{O}_{F_0}} \mathbb{F}_q$ to the exact sequence

$$0 \to N \to A_b \to A_b/N \to 0 \tag{18.4.8}$$

shows

$$t_0(A_b/N) = \begin{cases} t_0(A_b) & \text{if } N \subseteq \varpi_0 A_b \\ t_0(A_b) - 1 & \text{if } N \not\subseteq \varpi_0 A_b \end{cases} \tag{18.4.9}$$

for any cyclic submodule $N \subseteq A_b$. We also have $t_0(A_b) = t_0(\mathcal{T}) + 1$ if $b > 1$.

There is a natural inclusion $A_b \to A_{b+1}$. For any cyclic submodule $N \subseteq A_b$, we have

$$t_0(A_b/N) = \begin{cases} t_0(A_{b+1}/N) - 1 & \text{if } N = \langle t + \varpi_0^{-b} \rangle \text{ with } t \in \varpi_0\mathcal{T} \\ t_0(A_{b+1}/N) & \text{otherwise} \end{cases} \tag{18.4.10}$$

where $\varpi_0^{-b} \in \varpi_0^{-b}\mathcal{O}_{F_0}/\mathcal{O}_{F_0}$. Assume $b > e$. Then, in the first case above, the element $t \in \mathcal{T}$ is uniquely determined by N (using $b > e$). The cyclic submodules $N \subseteq A_{b+1}$ with $N \not\subseteq A_b$ are of the form $N\langle t + \varpi_0^{-b-1} \rangle$ for a unique $t \in \mathcal{T}$.

We thus have

$$
\begin{aligned}
&\alpha_{b+1} - \alpha_b && (18.4.11)\\
&= \sum_{\substack{t \in \mathcal{T} \\ N = \langle t + \varpi_0^{-b-1} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_{b+1}/N), X) + \sum_{\substack{t \in \varpi_0 \mathcal{T} \\ N = \langle t + \varpi_0^{-b} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_{b+1}/N), X) && (18.4.12)\\
&\quad - \sum_{\substack{t \in \varpi_0 \mathcal{T} \\ N = \langle t + \varpi_0^{-b} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_b/N), X) && (18.4.13)
\end{aligned}
$$

where the sums run over $t \in \mathcal{T}$ or $t \in \varpi \mathcal{T}$, as indicated. We compute

$$
\sum_{\substack{t \in \mathcal{T} \\ N = \langle t + \varpi_0^{-b-1} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_{b+1}/N), X) = |\mathcal{T}| X^{b+1} \mathfrak{m}(t_0(\mathcal{T}), X) \tag{18.4.14}
$$

where $|\mathcal{T}|$ is the cardinality of $\mathcal{T}$. For any integer $a \geq 0$, we have the identity $\mathfrak{m}(a+1, X) - \mathfrak{m}(a, x) = -q^a X \mathfrak{m}(a, X)$, so we compute

$$
\begin{aligned}
&\sum_{\substack{t \in \varpi_0 \mathcal{T} \\ N = \langle t + \varpi_0^{-b} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_{b+1}/N), X) - \sum_{\substack{t \in \varpi_0 \mathcal{T} \\ N = \langle t + \varpi_0^{-b} \rangle}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A_b/N), X) && (18.4.15)\\
&= -|\varpi_0 \mathcal{T}| q^{t_0(\mathcal{T})} X^{b+1} \mathfrak{m}(t_0(\mathcal{T}), X). && (18.4.16)
\end{aligned}
$$

But the exact sequence

$$
0 \to \varpi_0 \mathcal{T} \to \mathcal{T} \to \mathcal{T}/\varpi_0 \mathcal{T} \to 0 \tag{18.4.17}
$$

shows that $|\mathcal{T}| = |\varpi_0 \mathcal{T}| q^{t_0(\mathcal{T})}$ since $t_0(\mathcal{T}) = \dim_{\mathbb{F}_q} \mathcal{T}/\varpi_0 \mathcal{T}$ by definition. Substituting into (18.4.11) shows $\alpha_{b+1} - \alpha_b = 0$. $\square$

Lemma 18.4.4. *The polynomial* $\mathrm{Den}_{L^\flat, x}(X)^\circ_n - X \mathrm{Den}_{L^\flat, x'}(X)^\circ_n$ *stabilizes for* $\mathrm{val}(x) > 2a_{\max}(L^\flat)$.

Proof. As the x' dependence of $\mathrm{Den}_{L^\flat, x'}(X)^\circ_n$ is only on $\mathrm{val}(x')$, we may assume $x' = \varpi_1^{-1} x$ without loss of generality. Assume $\mathrm{val}(x) > 2a_{\max}(L^\flat)$. The lemma is trivial if $L^\flat$ is not integral (the polynomial is 0), so assume $L^\flat$ is integral.

Inspecting (18.3.3) shows that $\mathrm{Den}_{L^\flat, x}(X)^\circ_n - X \mathrm{Den}_{L^\flat, x'}(X)^\circ_n$ is equal to

$$
\sum_{\substack{L \subseteq M \subseteq M^* \\ M \cap L^\flat_F = L^\flat \\ \varpi_1^{-1} x \notin M}} X^{\ell(M/L)} \mathfrak{m}(t(M), X) \tag{18.4.18}
$$

where the sum runs over lattices $M \subseteq V$ (similar reasoning was used at the beginning of the proof of Lemma 18.3.5). For each M in the above sum, note that M/L is cyclic (again, $L^\flat \subseteq M$ is a saturated sublattice, so there is a direct sum decomposition $M = L^\flat \oplus \langle \xi \rangle$ (not necessarily orthogonal) for some $\xi \in V$). By Lemma 18.4.1, it is enough to show that

$$\sum_{\substack{L \subseteq M \subseteq M^* \\ M_1 \cap L^\flat_F = L^\flat_1 \\ M/L \text{ is cyclic} \\ \varpi_1^{-1} x \notin M}} X^{\ell(M/L)} \mathfrak{m}(t(M), X) \tag{18.4.19}$$

stabilizes for $\mathrm{val}(x) > 2a_{\max}(L^\flat)$, where the sum runs over lattices $M \subseteq V$ (because the difference between (18.4.19) and (18.4.18) is (18.4.4)).

We find that (18.4.19) equals

$$\sum_{\substack{L_1 \subseteq M_1 \subseteq L_2^* \\ M_1 \cap L^\flat_F = L^\flat_1 \\ \varpi_1^{-1} x_1 \notin M_1}} \sum_{\substack{L_2 \subseteq M_2 \subseteq M_1^* \\ M_2/L_2 \text{ is cyclic}}} X^{\ell(M/L)} \mathfrak{m}(t(M), X) \tag{18.4.20}$$

where the outer sum runs over lattices $M_1 \subseteq V_1$, the right-most sum runs over lattices $M_2 \subseteq V_2$, and $M = M_1 \oplus M_2$. Note that the lattices M_1 always satisfy M_1/L_1 being cyclic, because $M_1 \cap L^\flat_F = L^\flat_1$ implies $M_1 = L^\flat_1 \oplus \langle y_1 + \varpi_1^{-e_1} x \rangle$ where $y_1 \in L^{\flat *}_2 / L^\flat_1$ is of exact $\varpi_1^{e_1}$-torsion.

To prove the lemma, it is enough to check that (18.4.20) does not change if x is replaced with $\varpi_2 x$. The set of lattices $M_1 \subseteq V_1$ appearing in the outer sum is indexed elements $y_1 \in L^{\flat *}_2 / L^\flat_1$ (since e_1 is determined by y_1, in the above notation), and hence does not change if x is replaced by $\varpi_2 x$ (here using $\mathrm{val}(x) > a_{\max}(L^\flat)$ to ensure $M_1 \subset L_2^*$ for any choice of y_1). Note also that $\ell_0(M_1/L_1) = e_1$ and hence does not change when x is replaced by $\varpi_2 x$.

For the rest of the proof, fix an M_1 as in the outer sum of (18.4.20). We will show that the inner sum of (18.4.20) does not change if x is replaced by $\varpi_2 x$.

Set $A = M_1^*/L_2$. The inner sum is

$$X^{\ell_0(M_1/L_1)} \sum_{\substack{\text{cyclic submodules} \\ N \subseteq A}} X^{\mathrm{ord}(N)} \mathfrak{m}(t_0(A/N), X). \tag{18.4.21}$$

We already discussed that the factor $X^{\ell_0(M_1/L_1)}$ does not change when x is replaced by $\varpi_2 x$. On the other hand, we have $A \cong \mathrm{Hom}_{\mathcal{O}_{F_0}}(L_2^*/M_1, F_0/\mathcal{O}_{F_0})$ so $A \cong L_2^*/M_1$ (non-canonically). If $b \coloneqq \mathrm{val}(x)$ and $\mathcal{T} \coloneqq L^{\flat *}_2 / L^\flat_1$, then Lemma 18.4.2 shows $A \cong (\mathcal{T}/\langle y_1 \rangle) \oplus (\varpi_0^{-b} \mathcal{O}_{F_0}/\mathcal{O}_{F_0})$, where y_1 is associated to M_1 as above (since the submodule $(M_1/L_1) \subseteq L_2^*/L_1$ is cyclic and generated by $y_1 + \varpi^{-e_1} x_1$ where y_1 is of exact $\varpi_1^{e_1}$-torsion).

Now Lemma 18.4.3 implies that the sum in (18.4.21) does not change if x is replaced by $\varpi_2 x$. $\square$

Proof of Proposition 18.3.2 in split case. Assume F/F_0 is split. As remarked at the beginning of Section 18.4, it is enough to show that $\mathrm{Den}(X, L) - X\mathrm{Den}(X, L')$ stabilizes for $\mathrm{val}(x) > 2a_{\max}(L^\flat)$. We have

$$\mathrm{Den}(X, L) - X\mathrm{Den}(X, L') = \sum_{L^\flat \subseteq M^\flat \subseteq M^{\flat *}} \mathrm{Den}_{L^\flat, x}(X)^\circ_n - X\mathrm{Den}_{L^\flat, x'}(X)^\circ_n \tag{18.4.22}$$

by definition (see (18.3.4)), so Lemma 18.4.4 proves the claimed stabilization. $\square$

18.5 Limits

We continue in the setup of Section 18.3 but now assume $\varepsilon(V) = -1$ if F/F_0 is nonsplit. Recall also the definitions in (18.3.3) and (18.3.4).

Let $M^\flat \subseteq L^\flat_F$ be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$. If F/F_0 is nonsplit, set

$$\partial\mathrm{Den}_{L^\flat}(x)_n := -[\breve{F} : \breve{F}_0] \frac{d}{dX}\bigg|_{X=1} \mathrm{Den}(X, L)_n \quad \partial\mathrm{Den}_{M^\flat, \mathscr{H}}(x)^\circ_n := -[\breve{F} : \breve{F}_0] \frac{d}{dX}\bigg|_{X=1} \mathrm{Den}_{M^\flat, x}(X)^\circ_n$$

$$\partial\mathrm{Den}_{L^\flat, \mathscr{H}}(x)_n := \sum_{\substack{L \subseteq N \subseteq N^* \\ N^\flat = N \cap L^\flat_F \\ t(N^\flat) \leq 1}} \partial\mathrm{Den}_{N^\flat, \mathscr{H}}(x)^\circ_n \qquad \partial\mathrm{Den}_{L^\flat, \mathscr{V}}(x)_n := \partial\mathrm{Den}_{L^\flat}(x)_n - \partial\mathrm{Den}_{L^\flat, \mathscr{H}}(x)_n.$$

If F/F_0 is split, set

$$\mathrm{Den}_{L^\flat}(x)_n := \mathrm{Den}(X, L)_n\bigg|_{X=1} \qquad \mathrm{Den}_{M^\flat, \mathscr{H}}(x)^\circ_n := \mathrm{Den}_{M^\flat, x}(X)^\circ_n\bigg|_{X=1}$$

$$\mathrm{Den}_{L^\flat, \mathscr{H}}(x)_n := \sum_{\substack{L \subseteq N \subseteq N^* \\ N^\flat = N \cap L^\flat_F \\ t(N^\flat) \leq 1}} \mathrm{Den}_{N^\flat, \mathscr{H}}(x)^\circ_n \qquad \mathrm{Den}_{L^\flat, \mathscr{V}}(x)_n := \mathrm{Den}_{L^\flat}(x)_n - \mathrm{Den}_{L^\flat, \mathscr{H}}(x)_n.$$

The above sums run over lattices $N \subseteq V$ (so $N^\flat$ varies). These definitions also apply for any $x \notin L^\flat_F$ (not necessarily perpendicular to $L^\flat_F$), as long as we take $L = L^\flat + \langle x \rangle$.

Lemma 18.5.1. *If F/F_0 is split, then* $\mathrm{Den}_{L^\flat, \mathscr{V}}(x)_n = 0$ *for all x.*

Proof. Inspecting (18.2.2) shows that $\mathrm{Den}(X, M)^\circ_n = 0$ unless $M = M^*$. Lemma 18.3.4 implies $\mathrm{Den}_{N^\flat, x}(X)^\circ_n|_{X=1} = 0$ unless $t(N^\flat) \leq 1$, i.e. $\mathrm{Den}_{L^\flat}(x)_n = \mathrm{Den}_{L^\flat, \mathscr{H}}(x)$. $\square$

Given $x \in V$ with $(x,x) \neq 0$, we set $\mathrm{val}''(x) \coloneqq \mathrm{val}'(x)$ if F/F_0 is not inert (resp. $\mathrm{val}''(x) \coloneqq (\mathrm{val}(x) - 1)/2$ if F/F_0 is inert) to save space. We say a limit *stabilizes* if the argument of the limit becomes constant.

Proposition 18.5.2. *If F/F_0 is nonsplit, we have*

$$\partial\mathrm{Den}^*(L^\flat)_n = 2[\breve{F} : \breve{F}_0]^{-1} \lim_{x \to 0} \left(\partial\mathrm{Den}_{L^\flat}(x')_n - \mathrm{val}''(x)\mathrm{Den}^*(L^\flat)_n\right). \tag{18.5.1}$$

If F/F_0 is split, we have

$$\partial\mathrm{Den}^*(L^\flat)_n = \lim_{x \to 0} \left(\mathrm{Den}_{L^\flat}(x')_n - \mathrm{val}(x)\mathrm{Den}^*(L^\flat)_n\right). \tag{18.5.2}$$

The expressions are 0 *if $L^\flat$ is not integral, and all limits stabilize for* $\mathrm{val}(x) \gg 0$. *If $L^\flat$ is integral and F/F_0 is nonsplit (resp. split), then the limits stabilize when* $\mathrm{val}(x) > a_{\max}(L^\flat)$ *(resp.* $\mathrm{val}(x) > 2a_{\max}(L^\flat)$*).*

Proof. We emphasize that we are following Setup 18.3.1; in particular, we have $x' \to 0$ as $x \to 0$. Assume $L^\flat$ is integral (as the lemma is otherwise clear) and assume $\mathrm{val}(x) > a_{\max}(L^\flat)$. The key input is the induction formula from Proposition 18.3.2.

Case F/F_0 is nonsplit: Multiply the induction formula from Proposition 18.3.2 by $X^{-\mathrm{val}'(L^\flat)/2}$, and call the resulting expression $(*)$ (temporary notation). Taking one derivative of $(*)$ at $X = 1$ yields

$$\mathrm{Den}^*(L^\flat)_n = \partial\mathrm{Den}_{L^\flat}(x)_n - \partial\mathrm{Den}_{L^\flat}(x')_n. \tag{18.5.3}$$

Here we used $\mathrm{Den}(1, L)_n = \mathrm{Den}(1, L')_n = 0$ because $\varepsilon(V) = -1$ causes a sign in the functional equation (16.1.5). Taking two derivatives of $(*)$ at $X = 1$ yields the identity

$$\begin{aligned} &\mathrm{val}'(L^\flat)\partial\mathrm{Den}_{L^\flat}(x)_n + [\breve{F} : \breve{F}_0] \frac{d^2}{dX^2}\bigg|_{X=1} \mathrm{Den}(X, L)_n \\ &= (\mathrm{val}'(L^\flat) - 4[\breve{F} : \breve{F}_0]^{-1})\partial\mathrm{Den}_{L^\flat}(x')_n + [\breve{F} : \breve{F}_0] \frac{d^2}{dX^2}\bigg|_{X=1} \mathrm{Den}(X, L')_n + \partial\mathrm{Den}^*(L^\flat)_n. \end{aligned} \tag{18.5.4}$$

Again using $\varepsilon(V) = -1$, we apply the functional equation for $\mathrm{Den}(X, L)$ (16.1.5) to find

$$\frac{d^2}{dX^2}\bigg|_{X=1} \mathrm{Den}(X, L)_n = (\mathrm{val}(L) - 1)\frac{d}{dX}\bigg|_{X=1} \mathrm{Den}(X, L)_n \tag{18.5.5}$$

$$= -(\mathrm{val}(L) - 1)[\breve{F} : \breve{F}_0]^{-1}\partial\mathrm{Den}_{L^\flat}(x)_n \tag{18.5.6}$$

(the second equality is by definition) and similarly for L'. We also have

$$\mathrm{val}'(L^\flat) = \mathrm{val}(L) - 2[\breve{F} : \breve{F}_0]^{-1}\mathrm{val}''(x) - 1 \qquad \mathrm{val}(L) = \mathrm{val}(L') + 2[\breve{F} : \breve{F}_0]^{-1}. \tag{18.5.7}$$

Substituting all displayed equations into (18.5.4) proves the claim.

Case F/F_0 is split: Evaluating the induction formula from Proposition 18.3.2 at $X = 1$ yields

$$\mathrm{Den}^*(L^\flat)_n = \mathrm{Den}(1, L)_n - \mathrm{Den}(1, L')_n. \tag{18.5.8}$$

Multiplying both sides of the induction formula by $X^{-\mathrm{val}(L^\flat)/2}$ and taking one derivative at $X = 1$, we find

$$\begin{aligned}
&\mathrm{val}(L^\flat)\mathrm{Den}(1, L)_n - 2\frac{d}{dX}\bigg|_{X=1} \mathrm{Den}(X, L)_n && (18.5.9)\\
&= (\mathrm{val}(L^\flat) - 2)\mathrm{Den}(1, L')_n - 2\frac{d}{dX}\bigg|_{X=1} \mathrm{Den}(X, L')_n + \partial\mathrm{Den}^*(L^\flat)_n.
\end{aligned}$$

The functional equation (16.1.5) implies

$$2\frac{d}{dX}\bigg|_{X=1} \mathrm{Den}(X, L)_n = \mathrm{val}(L)\mathrm{Den}(1, L)_n \tag{18.5.10}$$

and similarly for L'. We also have

$$\mathrm{val}(L^\flat) = \mathrm{val}(L) - \mathrm{val}(x) \qquad \mathrm{val}(L) = \mathrm{val}(L') + 1. \tag{18.5.11}$$

Substituting all displayed equations into (18.5.9) proves the claim. □

Corollary 18.5.3. *Let $M^\flat \subseteq L^\flat_F$ be any full rank integral lattice with $t(M^\flat) \leq 1$. If F/F_0 is nonsplit, the following formulas hold.*

(1) $\partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n = 2[\breve{F} : \breve{F}_0]^{-1} \lim_{x\to 0} \partial\mathrm{Den}_{L^\flat,\mathscr{V}}(x)_n$

(2) $\partial\mathrm{Den}^*_{\mathscr{H}}(L^\flat)_n = 2[\breve{F} : \breve{F}_0]^{-1} \lim_{x\to 0} \big(\partial\mathrm{Den}_{L^\flat,\mathscr{H}}(x')_n - \mathrm{val}''(x)\mathrm{Den}^*(L^\flat)_n\big)$

(3) $\partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n = 2[\breve{F} : \breve{F}_0]^{-1} \lim_{x\to 0} \big(\partial\mathrm{Den}_{M^\flat,\mathscr{H}}(x')^\circ_n - \mathrm{val}''(x)\mathrm{Den}^*(M^\flat)^\circ_n\big)$.

If F/F_0 is split, the following formulas hold.

(1) $\partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n = \lim_{x\to 0} \mathrm{Den}_{L^\flat,\mathscr{V}}(x)_n$

(2) $\partial\mathrm{Den}^*_{\mathscr{H}}(L^\flat)_n = \lim_{x\to 0} \big(\mathrm{Den}_{L^\flat,\mathscr{H}}(x')_n - \mathrm{val}(x)\mathrm{Den}^*(L^\flat)_n\big)$

(3) $\partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n = \lim_{x\to 0} \big(\mathrm{Den}_{M^\flat,\mathscr{H}}(x')^\circ_n - \mathrm{val}(x)\mathrm{Den}^*(M^\flat)^\circ_n\big)$.

All limits stabilize for $\mathrm{val}(x) \gg 0$. The expressions (1) and (2) are 0 if $L^\flat$ is not integral. If $L^\flat$ is integral and F/F_0 is nonsplit (resp. split), then the limits in (1) and (2) stabilize when $\mathrm{val}(x) > a_{\max}(L^\flat)$ (resp. $\mathrm{val}(x) > 2a_{\max}(M^\flat)$). If F/F_0 is nonsplit (resp. split), the limits in (3) stabilizes when $\mathrm{val}(x) > a_{\max}(M^\flat)$ (resp. $\mathrm{val}(x) > 2a_{\max}(M^\flat)$).

Proof. Denote the result of Proposition 18.5.2 as (0). We have (3) $\implies$ (2) (in all cases, nonsplit or split), by summing over $M^\flat$ containing $L^\flat$. We have (0) $\implies$ (3) by taking $L^\flat = M^\flat$ and inducting on $\mathrm{val}(M^\flat)$ (starting with the base cases of $M^\flat$ being maximal integral (still with $t(M^\flat) \leq 1$), in which case $\partial\mathrm{Den}^*(M^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n$, and similarly for $\mathrm{Den}^*(M^\flat)_n$, as well as $\partial\mathrm{Den}_{M^\flat}(x')_n$ (nonsplit) and $\mathrm{Den}_{M^\flat}(x')_n$ (split)). Since (0) = (1) + (2), we conclude that (0) $\implies$ (1) as well. $\square$

The following lemma is the geometric counterpart of Corollary 18.5.3(1) (in the special case when $\alpha = {}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}}$ for a non-degenerate Hermitian $\mathcal{O}_F$-lattice $L^\flat \subseteq \mathbf{W}$).

Lemma 18.5.4. *Take $F_0 = \mathbb{Q}$, and assume $p \neq 2$ if $F/\mathbb{Q}_p$ is ramified. Let $\mathcal{Z} \to \operatorname{Spec} \overline{k}$ be a proper scheme equipped with a closed immersion $\mathcal{Z} \hookrightarrow \mathcal{N}$. Given any $\alpha \in \mathrm{gr}_1 K_0'(\mathcal{Z})$, we have*

$$\deg_{\overline{k}}(\alpha \cdot \mathcal{E}^\vee) = \lim_{w \to 0} \deg_{\overline{k}}(\alpha \cdot {}^{\mathbb{L}}\mathcal{Z}(w)) \tag{18.5.12}$$

where the limit runs over $w \in \mathbf{W}$. The limit stabilizes for w satisfying $\mathcal{Z} \subseteq \mathcal{Z}(w)$.

Proof. We may assume $F/\mathbb{Q}_p$ is nonsplit, as otherwise $\mathrm{gr}_1 K_0'(\mathcal{Z}) = 0$ (Section 5.4) for dimension reasons and the lemma is trivial.

For any fixed $w \in \mathbf{W}$, there exists $e \gg 0$ such that $\mathcal{Z} \subseteq \mathcal{Z}(p^e w)$ (over a quasi-compact base scheme, p^e times any quasi-homomorphism is a homomorphism for $e \gg 0$). Hence $\mathcal{Z} \subseteq \mathcal{Z}(w)$ for all $w \in \mathbf{W}$ lying in a sufficiently small neighborhood of 0.

Assume $w \in \mathbf{W}$ is such that $\mathcal{Z} \subseteq \mathcal{Z}(w)$. Write $\mathcal{I}(w) \subseteq \mathcal{O}_{\mathcal{N}}$ for the ideal sheaf of $\mathcal{Z}(w)$ (recall that $\mathcal{Z}(w)$ is a Cartier divisor, see Section 5.5). The lemma now follows from the "linear invariance" argument in the proof of [LL22, Lemma 2.55(3)] (valid in the inert case as well, using [How19]). Alternatively, the proof of linear invariance (particularly [How19, Definition 4.2] (inert) [LL22, Lemma 2.39] (ramified)) exhibits a canonical isomorphism $\mathcal{E} \otimes \mathcal{O}_{Z(w)} \cong \mathcal{I}(w) \otimes \mathcal{O}_{Z(w)}$ via Grothendieck–Messing theory. $\square$

Proof of Theorem 18.1.2. The horizontal part of the theorem was already verified in Proposition 18.2.1, so it remains to show $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n$.

If $F/\mathbb{Q}_p$ is split, then ${}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} = 0$ and so $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n = 0$. Applying Corollary 18.5.3(1) with $V = \mathbf{V}$, we find $\partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n = 0$ since $\mathrm{Den}_{L^\flat,\mathscr{V}}(x)_n = 0$ for all x (Lemma 18.5.1).

Next assume $F/\mathbb{Q}_p$ is nonsplit. For any $w \in \mathbf{W}$ not in $L^\flat_F$, we have $\deg_{\overline{k}}({}^{\mathbb{L}}\mathcal{Z}(L^\flat)_{\mathscr{V}} \cdot {}^{\mathbb{L}}\mathcal{Z}(w)) = \partial\mathrm{Den}_{L^\flat,\mathscr{V}}(w)_n$ by [LZ22a, Theorem 8.2.1] (inert) and [LL22, Theorem 2.7] (the "vertical" parts of the main results of loc. cit..). Lemma 18.5.4 implies $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n = 2[\breve{F} : \breve{F}_0]^{-1} \lim_{w \to 0} \partial\mathrm{Den}_{L^\flat,\mathscr{V}}(w)_n$. Restricting to w perpendicular to $L^\flat_F$, the limiting formula in Corollary 18.5.3(1) now implies $\mathrm{Int}_{\mathscr{V}}(L^\flat)_n = \partial\mathrm{Den}^*_{\mathscr{V}}(L^\flat)_n$. $\square$

Remark 18.5.5. Suppose $F_0 = \mathbb{Q}_p$, suppose $F/\mathbb{Q}_p$ is nonsplit, and assume $p \neq 2$. Let $M^\flat \subseteq \mathbf{V}$ be a non-degenerate integral $\mathcal{O}_F$-lattice of rank $n-1$ with $t(M^\flat) \leq 1$. As above, let $\mathcal{Z}(M^\flat)^\circ \subseteq \mathcal{N}$ be the associated quasi-canonical lifting cycle.

For any nonzero $w \in \mathbf{V}$ not in $M^\flat_F$, we have $\deg_{\overline{k}}(\mathcal{Z}(M^\flat)^\circ \cap \mathcal{Z}(w)) = \partial\mathrm{Den}_{M^\flat,\mathscr{H}}(w)^\circ_n$ by [KR11, Proposition 8.4] (inert, see also [LZ22a, Corollary 5.4.6, Theorem 6.1.3]) and [LL22, Corollary 2.46] (ramified), i.e. the "horizontal" parts of the main results of loc. cit..

The "horizontal part" of our main theorem showed $\mathrm{Int}_{\mathscr{H}}(M^\flat)^\circ_n = \partial\mathrm{Den}^*_{\mathscr{H}}(M^\flat)^\circ_n$ (Proposition 18.2.1). Using also the special value formula in Lemma 18.1.3, our limiting result Corollary 18.5.3(3) is equivalent to the geometric statement

$$\begin{aligned} &2\deg \mathcal{Z}(M^\flat)^\circ \cdot \delta_{\mathrm{tau}}(\mathrm{val}'(M^\flat)) \\ &= 2[\breve{F} : \breve{\mathbb{Q}}_p]^{-1} \lim_{x\to 0} \left(\deg_{\overline{k}}(\mathcal{Z}(M^\flat)^\circ_n \cap \mathcal{Z}(x)) - \mathrm{val}''(\varpi x) \deg \mathcal{Z}(M^\flat)^\circ\right) \end{aligned} \tag{18.5.13}$$

(limiting over nonzero x perpendicular to $M^\flat$) where $\delta_{\mathrm{tau}}(\mathrm{val}'(M^\flat))$ is the "local change of tautological height", as in (7.2.7) (which is $-1/2$ times the "local change of Faltings height" $\delta_{\mathrm{Fal}}(\mathrm{val}'(M^\flat))$).

To prove our main theorem, we verified (18.5.13) indirectly by the computations in Section 18.2. Direct computations are also possible.

19 Archimedean local identity

Let V be the non-degenerate $\mathbb{C}/\mathbb{R}$ Hermitian space of rank n and signature $(n-1,1)$. We freely use notation for the Hermitian symmetric domain $\mathcal{D}$ and its special cycles (Section 8.1) as well as the Archimedean local Whittaker functions $W^*_{T,\infty}(s)^\circ_n$ for $T \in \mathrm{Herm}_m(\mathbb{R})$ (complex Hermitian matrices) with $\det T \neq 0$ (Section 15.2). Here $W^*_{T,\infty}(s)^\circ_n$ denotes the function $W^*_{T,v}(1,s)^\circ_n$ in the notation of loc. cit..

19.1 Statement of identity

Our main Archimedean local identity ("Archimedean local arithmetic Siegel–Weil") is the following. The notation $(-)^\perp$ denotes orthogonal complement.

Theorem 19.1.1. *Let $\underline{x} \in V^m$ be an m-tuple with nonsingular Gram matrix, and set $T = (\underline{x}, \underline{x})$. If $\mathrm{span}(\underline{x})^\perp$ is anisotropic, we have*

$$\int_{\mathcal{D}} [\xi(\underline{x})] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m} = \frac{d}{ds}\bigg|_{s=-s_0} W^*_{T,\infty}(s)^\circ_n \tag{19.1.1}$$

where $s_0 = (n-m)/2$.

In Theorem 19.1.1, integration of the current $[\xi(\underline{x})] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m}$ over $\mathcal{D}$ is understood in the sense described in Section 4.5. The displayed integral is convergent (combine Lemma 8.3.3 and Lemma 8.3.1). The local functional equation (Lemma 16.2.1) implies that the derivative of $W^*_{T,\infty}(s)^\circ_n$ at $s = s_0$ and $s = -s_0$ are the same up to a simple sign.

The case $m = n$ of Theorem 19.1.1 is the content of [Liu11, Theorem 4.17] (when translating to Liu's notation, recall also that $W^*_{T,\infty}(0)^\circ_n = 0$ when $m \leq n$ for non positive-definite T, as discussed in Section 15.2). We do not give a new proof of this case. Indeed, we reduce the other cases of Theorem 19.1.1 to the case $m = n$ by the following limiting result.

Proposition 19.1.2. *Let $T^\flat \in \mathrm{Herm}_m(\mathbb{R})$ be a matrix with $\det T^\flat \neq 0$, assume $m \leq n$, and set $s_0 = (n-m)/2$. Assume that either $m = n-1$ or that $T^\flat$ is not positive definite. Given $t = \mathrm{diag}(t_{m+1}, \ldots, t_n) \in \mathrm{Herm}_{n-m}(\mathbb{R})$, set $T = \mathrm{diag}(t, T^\flat)$.*

$$\frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n = \lim_{t \to 0^\pm} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n + (\log |t|_{F_v^+} + \log(4\pi) - \Gamma'(1)) W^*_{T^\flat,\infty}(-s_0)^\circ_n \right) \tag{19.1.2}$$

where $|t|_{F_v^+} \coloneqq |\det t|_{F_v^+}$, and where the sign on $0^\pm$ (meaning all t_j have this sign) is

$$\begin{cases} - & \text{if } T^\flat \text{ is positive definite} \\ + & \text{else.} \end{cases} \tag{19.1.3}$$

Remark 19.1.3. In the situation of Proposition 19.1.2, recall

$$W^*_{T^\flat,\infty}(-s_0)^\circ_n = \begin{cases} 1 & \text{if } T^\flat \text{ is positive definite} \\ 0 & \text{else} \end{cases} \tag{19.1.4}$$

(see Section 15.2). Due to this vanishing, the term $(\log |t|_{F_v^+} + \log(4\pi) - \Gamma'(1))$ from Proposition 19.1.2 should not be taken seriously outside the positive definite $T^\flat$ case (especially if $m \neq n-1$).

If $T^\flat$ has signature (p, q) for $q \geq 2$, we also have

$$\frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n = \frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n = 0 \tag{19.1.5}$$

for any $t \in \mathrm{Herm}_{n-m}(\mathbb{R})$ with $\det t \neq 0$ by [Shi82, Theorem 4.2, (4.34.K)]. Thus Proposition 19.1.2 holds for $T^\flat$ of signature (p, q) when $q \geq 2$ (both sides of the identity are 0).

The proof of the remaining cases of Proposition 19.1.2 will occupy most of the rest of Section 19. The case of $T^\flat$ having signature $(m-1, 1)$ is completed in Section 19.4, and the case of positive definite $T^\flat$ is completed in Section 19.5. We also obtain an explicit formula for both sides of (19.1.1) when T is positive definite, namely (19.5.4) (the formula is a polynomial in the eigenvalues of T^{-1}).

Once the proposition is proved, Theorem 19.1.1 follows (and is equivalent to the proposition for any given $T^\flat$, which is the T in Theorem 19.1.1) by the following argument.

Proof of equivalence of Theorem 19.1.1 and Proposition 19.1.2. Let $T^\flat$ be as in Proposition 19.1.2. We may assume $T^\flat$ has signature $(m,0)$ or $(m-1,1)$ by Remark 19.1.3. Suppose $\underline{x}^\flat = (x_1^\flat, \ldots, x_m^\flat) \in V^m$ satisfies $(\underline{x}^\flat, \underline{x}^\flat) = T^\flat$. Given an orthogonal basis $\underline{x}^\# = (x_{m+1}, \ldots, x_n)$ of $\operatorname{span}(\underline{x}^\flat)^\perp$, set $t_j = (x_j, x_j)$ for $j \geq m+1$, set $t = (t_{m+1}, \ldots, t_n)$, set $\underline{x} = (x_{m+1}, \ldots, x_n, x_1^\flat, \ldots, x_m^\flat) \in V^n$, and set $T = (\underline{x}, \underline{x})$. We have

$$\left.\frac{d}{ds}\right|_{s=0} W^*_{T,\infty}(s)^\circ_n = \int_{\mathcal{D}} [\xi(\underline{x})] = \int_{\mathcal{D}} [\xi(\underline{x}^\flat)] \wedge \omega(\underline{x}^\#) + \int_{\mathcal{D}(\underline{x}^\flat)} \xi(x) \tag{19.1.6}$$

where the first equality is the $m = n$ case of Theorem 19.1.1 (already proved by Liu as cited above) and the second identity is by definition.

Using the pointwise convergence $\lim_{a \to 0} \omega(ax) = c_1(\widehat{\mathcal{E}}^\vee)$ on $\mathcal{D}$ for each $x \in V$ (8.2.8), we have

$$\lim_{\underline{x}^\# \to 0} \int_{\mathcal{D}} [\xi(\underline{x}^\flat)] \wedge \omega(\underline{x}^\#) = \int_{\mathcal{D}} [\xi(\underline{x}^\flat)] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m} \tag{19.1.7}$$

(say, where the limit runs over $\underline{x}^\# = (a_{m+1}x_{m+1}, \ldots, a_n x_n)$ as $a_j \to 0$ for all j) by dominated convergence (applying estimate from the proof of Lemma 8.3.1 and convergence from Lemma 8.3.3, particularly convergence of (8.3.8)).

The closed submanifold $\mathcal{D}(\underline{x}^\flat) \subseteq \mathcal{D}$ is a single point if $T^\flat$ is positive definite (in which case we assumed $m = n-1$), and is empty if $T^\flat$ is not positive definite. We thus have

$$\int_{\mathcal{D}(\underline{x}^\flat)} \xi(x) = \begin{cases} -\operatorname{Ei}(4\pi(x_n, x_n)) & \text{if } T^\flat \text{ is positive definite} \\ 0 & \text{else.} \end{cases} \tag{19.1.8}$$

Recall asymptotics for the function Ei (8.2.2) and recall $\Gamma'(1) = -\gamma$. Recall the special value formulas from (19.1.4). We substitute into (19.1.6) to obtain

$$\int_{\mathcal{D}} [\xi(\underline{x}^\flat)] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m} = \lim_{t \to 0^\pm} \left(\left.\frac{d}{ds}\right|_{s=0} W^*_{T,\infty}(s)^\circ_n + (\log|t|_{F_v^+} + \log(4\pi) - \Gamma'(1)) W^*_{T^\flat,\infty}(-s_0)^\circ_n \right) \tag{19.1.9}$$

(where the sign on $0^\pm$ is the sign of t, determined by the signature of $T^\flat$) which proves the claimed equivalence. $\square$

Remark 19.1.4. For any $T \in \mathrm{Herm}_m(\mathbb{R})$ with $\det T \neq 0$, recall that the (normalized) Archimedean local Whittaker function $W^*_{T,\infty}(s)^\circ_n$ satisfies a certain "linear invariance" property, i.e. the local Whittaker function is unchanged if we replace T by ${}^t\overline{k}Tk$ for any $k \in U(m)$ where $U(m)$ is the usual positive definite unitary group in standard coordinates (see Section 15.2). It is thus enough to prove Proposition 19.1.2 when $T^\flat$ is diagonal.

Remark 19.1.5. Using the linear invariance property for Whittaker functions, the limiting relation in (19.1.9) implies that the quantity

$$\int_{\mathcal{D}} [\xi(\underline{x}^\flat)] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m} \tag{19.1.10}$$

from Theorem 19.1.1 is similarly linearly invariant (i.e. does not change if $\underline{x}^\flat$ is replaced by $\underline{x}^\flat \cdot k$ for any $k \in U(m)$, where $\underline{x}^\flat$ is viewed as a row vector of elements in V). Stated alternatively, we observe that the linear invariance result of Liu for $\int_{\mathcal{D}}[\xi(\underline{x}^\flat)] \wedge c_1(\widehat{\mathcal{E}}^\vee)^{n-m}$ when $m = n$ [Liu11, Proposition 4.10] can be used to prove the analogous linear invariance in our setting via limiting, even before we have proved Theorem 19.1.1 or Proposition 19.1.2.

19.2 Computation when $n = 2$

Before proving Theorem 19.1.1 via Proposition 19.1.2 in the later sections, we check the $n = 2$ case of Theorem 19.1.1 by direct computation (the case $n = 1$ and $m \neq n$ is trivial as both sides of the identity are trivially 0). The proof for general n (which proceeds differently, not relying on the $n = 2$ computation) begins in Section 19.3 below.

Take $n = 2$ throughout Section 19.2, and suppose $T \in \mathbb{R}$ is nonzero. By [Shi82, (1.29) and (3.3)] (translation via (13.2.13)) and some rearranging, we have the formula

$$W^*_{T,\infty}(s)^\circ_n = \Gamma(s-1/2)^{-1}|4\pi T|^{s-1/2} \int_0^\infty e^{-4\pi T u}(u+1)^{s+1/2}u^{s-3/2}\, du \tag{19.2.1}$$

$$= 1 + \Gamma(s-1/2)^{-1}|4\pi T|^{s-1/2} \int_0^\infty e^{-4\pi T u}((u+1)^{s+1/2}-1)u^{s-3/2}\, du \tag{19.2.2}$$

if $T > 0$, where the integrals in (19.2.1) and (19.2.2) are convergent for $\mathrm{Re}(s) > 1/2$ and $\mathrm{Re}(s) > -1/2$ respectively. We similarly have

$$W^*_{T,\infty}(s)^\circ_n = \Gamma(s-1/2)^{-1}|4\pi T|^{s-1/2} \int_1^\infty e^{4\pi T u}(u-1)^{s+1/2}u^{s-3/2}\, du \tag{19.2.3}$$

if $T < 0$, where displayed integral is convergent for $\mathrm{Re}(s) > -3/2$.

Proposition 19.2.1. *Given any nonzero $T \in \mathbb{R}$ and any $x \in V$ with $T = (x, x)$, we have*

$$\int_{\mathcal{D}} \xi(x) c_1(\widehat{\mathcal{E}}^\vee) = -\frac{d}{ds}\bigg|_{s=1/2} W^*_{T,\infty}(s)^\circ_n = \begin{cases} (-4\pi T)^{-1} & \textit{if } T > 0 \\ (4\pi T)^{-1} e^{4\pi T} - \mathrm{Ei}(4\pi T) & \textit{if } T < 0. \end{cases} \tag{19.2.4}$$

The preceding proposition (proved below) shows that Theorem 19.1.1 holds when $n = 2$ (the functional equation implies $-\frac{d}{ds}|_{s=1/2} W^*_{T,\infty}(s)^\circ_n = \frac{d}{ds}|_{s=-1/2} W^*_{T,\infty}(s)^\circ_n$).

Lemma 19.2.2. *For any nonzero $T \in \mathbb{R}$, we have*

$$-\frac{d}{ds}\bigg|_{s=1/2} W^*_{T,\infty}(s)^\circ_n = \begin{cases} (-4\pi T)^{-1} & \textit{if } T > 0 \\ (4\pi T)^{-1} e^{4\pi T} - \mathrm{Ei}(4\pi T) & \textit{if } T < 0. \end{cases} \tag{19.2.5}$$

Proof. Recall that $\Gamma(s)^{-1} = s + O(s^2)$ near $s = 0$. The integrals in (19.2.2) and (19.2.3) are convergent and holomorphic at $s = 1/2$. Directly evaluating the integrals at $s = 1/2$ gives the claimed formulas. □

Lemma 19.2.3. *With $x \in V$ and $T \in \mathbb{R}$ as in the statement of Proposition 19.2.1, we have*

$$\int_{\mathcal{D}} \xi(x) c_1(\widehat{\mathcal{E}}^\vee) = \begin{cases} (-4\pi T)^{-1} & \textit{if } T > 0 \\ (4\pi T)^{-1} e^{4\pi T} - \mathrm{Ei}(4\pi T) & \textit{if } T < 0. \end{cases} \tag{19.2.6}$$

Proof. By (8.3.2), we have

$$c_1(\widehat{\mathcal{E}}^\vee) = \frac{1}{2\pi i} \partial\overline{\partial}(\log R(x, z)) = \frac{1}{2\pi i} \frac{dz \wedge d\overline{z}}{(1 - z\overline{z})^2}. \tag{19.2.7}$$

If $T > 0$, we have

$$\begin{aligned} \int_{\mathcal{D}} \xi(x) c_1(\widehat{\mathcal{E}}^\vee) &= \int_{\mathcal{D}} \int_1^\infty e^{-4\pi T u z\overline{z}(1-z\overline{z})^{-1}} u^{-1} \, du \, \frac{1}{2\pi i} \frac{dz \wedge d\overline{z}}{(1 - z\overline{z})^2} && (19.2.8) \\ &= -2 \int_0^1 \int_1^\infty e^{-4\pi T u r^2 (1-r^2)^{-1}} u^{-1} (1 - r^2)^{-2} r \, du \, dr && (19.2.9) \\ &= -\int_1^\infty \int_0^\infty e^{-4\pi T v u} u^{-1} \, dv \, du && (19.2.10) \\ &= (-4\pi T)^{-1} && (19.2.11) \end{aligned}$$

via the change of variables $v = r^2(1 - r^2)^{-1}$.

If $T < 0$, we have

$$\int_{\mathcal{D}} \xi(x) c_1(\widehat{\mathcal{E}}^\vee) = \int_{\mathcal{D}} \int_1^\infty e^{4\pi T u (1 - z\overline{z})^{-1}} u^{-1} \, du \frac{1}{2\pi i} \frac{dz \wedge d\overline{z}}{(1 - z\overline{z})^2} \tag{19.2.12}$$

$$= -2 \int_0^1 \int_1^\infty e^{4\pi T u (1 - r^2)^{-1}} u^{-1} (1 - r^2)^{-2} r \, du \, dr \tag{19.2.13}$$

$$= - \int_1^\infty \int_1^\infty e^{4\pi T v u} u^{-1} \, dv \, du \tag{19.2.14}$$

$$= (4\pi T)^{-1} \int_1^\infty e^{4\pi T u} u^{-2} \, du \tag{19.2.15}$$

via the change of variables $v = (1 - r^2)^{-1}$. We also have

$$\int_1^\infty e^{4\pi T u} u^{-2} \, du = e^{4\pi T} - (4\pi T) \operatorname{Ei}(4\pi T) \tag{19.2.16}$$

via integration by parts. □

Proof of Proposition 19.2.1. Already proved by direct computation in Lemmas 19.2.2 and 19.2.3. □

19.3 More on Archimedean local Whittaker functions

We use some special functions studied by Shimura [Shi82] to describe the Archimedean Whittaker functions $W^*_{T,\infty}(s)^\circ_n$ from above. We allow arbitrary $n \in \mathbb{Z}$ for the moment.

We first recall Shimura's definitions. Given an integer $m \geq 0$, set

$$\Gamma_m(s) = \pi^{m(m-1)/2} \prod_{k=0}^{m-1} \Gamma(s - k) \tag{19.3.1}$$

as in [Shi82, (1.17.K)], where Γ is the usual gamma function. Given Hermitian matrices h, h', the notation $h > h'$ will mean that $h - h'$ is positive definite. For

$$\alpha, \beta \in \mathbb{C} \qquad g \in \operatorname{Herm}_m(\mathbb{R})_{>0} \qquad h \in \operatorname{Herm}_m(\mathbb{R})$$
$$z \in \mathcal{H}' \coloneqq \{z = x + iy \in M_{m,m}(\mathbb{C}) \text{ with } x, y \in \operatorname{Herm}_m(\mathbb{R}) \text{ and } x > 0\}$$

we set

$$\xi(g,h;\alpha,\beta) \coloneqq \int_{\mathrm{Herm}_m(\mathbb{R})} e^{-2\pi i \operatorname{tr}(hx)} \det(x+ig)^{-\alpha} \det(x-ig)^{-\beta}\, dx \tag{19.3.2}$$

$$\eta(g,h;\alpha,\beta) \coloneqq \int_{\substack{\mathrm{Herm}_m(\mathbb{R}) \\ x>h \\ x>-h}} e^{-\operatorname{tr}(gx)} \det(x+h)^{\alpha-m} \det(x-h)^{\beta-m}\, dx \tag{19.3.3}$$

$$\zeta_m(z;\alpha,\beta) \coloneqq \int_{\mathrm{Herm}_m(\mathbb{R})_{>0}} e^{-\operatorname{tr}(zx)} \det(x+1_m)^{\alpha-m} \det(x)^{\beta-m}\, dx \tag{19.3.4}$$

$$\omega_m(z;\alpha,\beta) \coloneqq \Gamma_m(\beta)^{-1} \det(z)^{\beta} \zeta_m(z;\alpha,\beta) \tag{19.3.5}$$

$$\zeta_{p,q}(g;\alpha,\beta) \coloneqq e^{-\operatorname{tr}(g)/2} \int_{\substack{\mathrm{Herm}_m(\mathbb{R}) \\ x+\mathrm{diag}(1_p,0)>0 \\ x+\mathrm{diag}(0,1_q)>0}} e^{-\operatorname{tr}(gx)} \det(x+\mathrm{diag}(1_p,0))^{\alpha-m} \det(x+\mathrm{diag}(0,1_q))^{\beta-m}\, dx \tag{19.3.6}$$

initially defined for $\mathrm{Re}(\alpha), \mathrm{Re}(\beta) \gg 0$, where $p, q \geq 0$ are integers with $p+q=m$. All implicit measures in the integrals are Euclidean. See Remark 13.2.1 for the log branch convention.

The special functions $\xi, \eta, \zeta_m, \omega_m, \zeta_{p,q}$ are copied from [Shi82, (1.25), (1.26), (3.2), (3.6), (4.16)], respectively. Formulas relating ξ and η, relating η and ζ_m, and relating η and $\zeta_{p,q}$ are given in [Shi82, (1.29), (3.3), (4.18)]. These will be used implicitly in our computations below.

Recall that $\omega_m(z;\alpha,\beta)$ admits holomorphic continuation to all $(z,\alpha,\beta) \in \mathcal{H}' \times \mathbb{C}^2$ (by [Shi82, Theorem 3.1]), and that $\Gamma_q(\alpha-p)^{-1}\Gamma_p(\beta-q)^{-1}\zeta_{p,q}(g;\alpha,\beta)$ admits holomorphic continuation to all (α,β), for any g (by [Shi82, Theorem 4.2]). We also recall the special value formula

$$\omega_m(z;m,\beta) = \omega_m(z;\alpha,0) = 1 \tag{19.3.7}$$

for all $\alpha, \beta \in \mathbb{C}$ [Shi82, (3.15)].

We will also use the differential operator $\Delta \coloneqq \det(\partial/\partial z_{j,k})$ on the space of matrices $z = (z_{j,k})_{j,k} \in M_{m,m}(\mathbb{C})$ as in [Shi82, (3.10.II)] (also [Liu11, (4-20)]). For any $u \in \mathrm{Herm}_m(\mathbb{R})_{>0}$, with $u^{1/2}$ denoting its unique positive definite Hermitian square-root, we have the relation

$$\begin{aligned}&(-1)^m \Delta(e^{-\operatorname{tr} uz} \det(uz)^{-\beta} \omega_m(u^{1/2} z u^{1/2};\alpha,\beta)) \\ &= e^{-\operatorname{tr} uz} \det(uz)^{-\beta} \det(u) \omega_m(u^{1/2} z u^{1/2};\alpha+1,\beta)\end{aligned} \tag{19.3.8}$$

where Δ is applied to the z variable, and where both sides are evaluated at $z \in \mathcal{H}'$. The preceding formula is a slight variant of [Shi82, (3.12)] and [Liu11, (4-21)] (and can be verified by similar reasoning). We will use this formula in its equivalent form

$$\begin{aligned}&(-1)^m e^{\operatorname{tr} uz} \Delta(e^{-\operatorname{tr} uz} \det(z)^{-\beta} \omega_m(u^{1/2} z u^{1/2};\alpha,\beta)) \\ &= \Gamma_m(\beta)^{-1} \det(u)^{\beta+1} \zeta_m(u^{1/2} z u^{1/2};\alpha+1,\beta).\end{aligned} \tag{19.3.9}$$

Remark 19.3.1.

(1) The special function ξ (which takes multiple arguments) should not be confused with the Green function from Section 8.2 (which takes one argument), as should be clear from context. The same applies to η the special function (which takes multiple arguments) and η the quadratic character (which takes one argument).

(2) The definition of $\zeta_{p,q}$ in [Shi82, (4.16)] has a running assumption that "g is diagonal", but we can make the same definition without this diagonal assumption.

(3) Liu also uses these functions but with slightly different normalizations [Liu11, §4A]. We follow Shimura's normalizations.

Given $T \in \mathrm{Herm}_m(\mathbb{R})$ with $\det T \neq 0$, we set (non-standard)

$$W_T^*(\alpha, \beta) \coloneqq e^{2\pi \operatorname{tr} T} \frac{2^{m(m-1)}\pi^{-m\beta}}{(-2\pi i)^{m(\alpha-\beta)}} \Gamma_m(\alpha) |\det T|^{-\alpha+m} \xi(1_m, T; \alpha, \beta) \tag{19.3.10}$$

for $\alpha, \beta \in \mathbb{C}$ initially defined for $\mathrm{Re}(\alpha), \mathrm{Re}(\beta) \gg 0$. We have

$$W_{T,\infty}^*(s)_n^\circ = W_T^*(\alpha, \beta) \qquad \text{when } \alpha = s - s_0 + n \text{ and } \beta = s - s_0 \tag{19.3.11}$$

where $s_0 = (n-m)/2$ (see (13.2.13)).

For any $c \in \mathrm{GL}_m(\mathbb{C})$ such that $c 1_{p,q} {}^t\overline{c} = T$ (where $1_{p,q} = \mathrm{diag}(1_p, -1_q) \in M_{m,m}(\mathbb{R})$), and with $g \coloneqq {}^t\overline{c}c$, we have

$$\begin{aligned} W_T^*(\alpha, \beta) &= e^{2\pi \operatorname{tr} T} (2\pi)^{2m\beta} \pi^{-m\beta} \Gamma_m(\beta)^{-1} |\det T|^{-\alpha+m} 2^{m(m-\alpha-\beta)} |\det T|^{\alpha+\beta-m} \eta(2\pi g, 1_{p,q}; \alpha, \beta) \\ &= e^{2\pi \operatorname{tr} T} \Gamma_m(\beta)^{-1} |\det 4\pi T|^{\beta} \zeta_{p,q}(4\pi g; \alpha, \beta). \end{aligned} \tag{19.3.12}$$

When T is positive definite, our conventions imply

$$W_T^*(\alpha, \beta) = \omega_m(4\pi g; \alpha, \beta). \tag{19.3.13}$$

Lemma 19.3.2. *Suppose $T \in \mathrm{Herm}_m(\mathbb{R})$ has $\det T \neq 0$. If T is positive definite (resp. not positive definite), the function $W_T^*(\alpha, \beta)$ admits holomorphic continuation to all $(\alpha, \beta) \in \mathbb{C}^2$ (resp. for $\mathrm{Re}(\alpha) > m - 1$ and all β). In this region, we have*

$$\frac{\partial}{\partial \alpha} W_T^*(\alpha, \beta) = 0 \tag{19.3.14}$$

for $\beta = 0$ (resp. $\beta \in \mathbb{Z}_{\leq 0}$).

Proof. Let T have signature (p,q) and let g be as above. The holomorphic continuation of $\Gamma_p(\beta-q)^{-1}\Gamma_q(\alpha-p)^{-1}\zeta_{p,q}(4\pi g;\alpha,\beta)$ to all $(\alpha,\beta)\in\mathbb{C}^2$ (as recalled above from [Shi82, Theorem 4.2]) implies that $W_T^*(\alpha,\beta)$ admits holomorphic continuation to the region claimed.

When T is positive definite, (19.3.7) implies $(\partial/\partial\alpha)W_T^*(\alpha,0)=0$. If T is not positive definite, the function $\Gamma_m(\beta)^{-1}\Gamma_p(\beta-q)$ has a zero at every $\beta\in\mathbb{Z}_{\leq 0}$, which implies $W_T^*(\alpha,\beta)=0$ for all $\beta\in\mathbb{Z}_{\leq 0}$. Thus $(\partial/\partial\alpha)W_T^*(\alpha,\beta)=0$ for all $b\in\mathbb{Z}_{\leq 0}$ in this case. $\square$

Suppose $m\geq 1$. For any $g=(-4\pi)^{-1}\operatorname{diag}(a,b)\in\operatorname{Herm}_m(\mathbb{R})_{>0}$ with $a\in\operatorname{Herm}_{m-1}(\mathbb{R})_{<0}$ and $b\in\mathbb{R}_{<0}$, we have (as in [Shi82, (4.25)] and also [Liu11, (4-15)])

$$
\begin{aligned}
&e^{2\pi\operatorname{tr}g}\zeta_{m-1,1}(4\pi g;\alpha,\beta) && (19.3.15)\\
&=\int_{\mathbb{C}^{m-1}} e^{\operatorname{tr}(aww^*)+bw^*w}\zeta_1(-b(1+w^*w);\beta,\alpha-m+1) && (19.3.16)\\
&\qquad\cdot e^{\operatorname{tr}(-au/2)}\eta(-a,u/2;\alpha,\beta-1)\ dw\\
&=\int_{\mathbb{C}^{m-1}} e^{\operatorname{tr}(aww^*)+bw^*w}\zeta_1(-b(1+w^*w);\beta,\alpha-m+1) && (19.3.17)\\
&\qquad\cdot\det(u)^{\alpha+\beta-m}\zeta_{m-1}(-u^{1/2}au^{1/2};\alpha,\beta-1)\ dw
\end{aligned}
$$

with $w\in\mathbb{C}^{m-1}$ viewed as column vectors, with $w^*\coloneqq {}^t\overline{w}$, with $u=1_{m-1}+ww^*$, with $u^{1/2}$ the unique positive definite Hermitian square-root of u, and with dw being the Euclidean measure.

We next specialize (19.3.17) to $\alpha=m$. We have

$$e^{b(1+w^*w)}\zeta_1(-b(1+w^*w);\beta,1)=\int_1^\infty e^{b(1+w^*w)x}x^{\beta-1}\ dx. \tag{19.3.18}$$

Combining (19.3.9) and (19.3.7), we also find

$$\det(u)^\beta\zeta_{m-1}(-u^{1/2}au^{1/2};m,\beta-1)=(-1)^{m-1}e^{-\operatorname{tr}au}\Delta|_{z=-a}(e^{-\operatorname{tr}uz}\det(z)^{-\beta+1}). \tag{19.3.19}$$

Hence, we have

$$
\begin{aligned}
&(-1)^{m-1}e^b\Gamma_{m-1}(\beta-1)^{-1}e^{2\pi\operatorname{tr}g}\zeta_{m-1,1}(4\pi g;m,\beta) && (19.3.20)\\
&=\int_{\mathbb{C}^{m-1}}\int_1^\infty e^{\operatorname{tr}(aww^*)}e^{b(1+w^*w)x}x^{\beta-1} && (19.3.21)\\
&\qquad\cdot e^{-\operatorname{tr}(1_{m-1}+ww^*)a}\Delta|_{z=-a}(e^{-\operatorname{tr}(1_{m-1}+ww^*)z}\det(z)^{-\beta+1})\ dx\ dw.
\end{aligned}
$$

These rearrangements are initially valid for $\operatorname{Re}(\beta)\gg 0$, but in fact hold for all $\beta\in\mathbb{C}$ by analytic continuation (see also [Shi82, (3.8)] for estimates on ζ_1 and ζ_{m-1} giving convergence).

The next lemma generalizes a calculation of Liu [Liu11, Lemma 4.7], and will be used to re-express (19.3.21) more explicitly. In the statement and proof below, we adopt the following notation from [Liu11, Lemma 4.7]: given a matrix $u \in M_{m,m}(\mathbb{C})$ and sets $I, J \subseteq \{1, \ldots m\}$ of the same cardinality, the symbol $|u^{I,J}|$ (resp. $|u_{I,J}|$) will mean the determinant of the matrix obtained from u by discarding (resp. keeping) the rows in indexed by I and the columns indexed by J.

Lemma 19.3.3. *Given any $u \in M_{m,m}(\mathbb{C})$ and $z_0 \in \mathrm{Herm}_m(\mathbb{R})_{>0}$ with z_0 diagonal, we have*

$$\Delta|_{z=z_0}(e^{\operatorname{tr} uz} \det(z)^s) = e^{\operatorname{tr} uz_0} \det(z_0)^s \sum_{t=0}^{m} \sum_{\substack{J=\{j_1<\cdots<j_t\} \\ J \subseteq \{1,\ldots,m\}}} \left(\prod_{k=1}^{t} (s+k-1) \right) |g_{0,J,J}|^{-1} |u^{J,J}| \tag{19.3.22}$$

for all $s \in \mathbb{C}$, where the inner sum runs over all subsets $J \subseteq \{1, \ldots, m\}$ of size t.

Proof. Observe that (upon fixing u and z_0) the expression

$$e^{-\operatorname{tr} uz_0} \det(z_0)^{-s} \Delta|_{z=z_0}(e^{\operatorname{tr} uz} \det(z)^s) \tag{19.3.23}$$

is a polynomial in s. Hence it is enough to prove the lemma holds for all $s \in \mathbb{Z}_{\geq 1}$. The case $s = 1$ is given by the proof of [Liu11, Lemma 4.7] via combinatorial calculation. For all $s \in \mathbb{Z}_{\geq 1}$, a similar calculation shows

$$e^{-\operatorname{tr} uz_0} \det(z_0)^{-s} \Delta|_{z=z_0}(e^{\operatorname{tr} uz} \det(z)^s) = \sum_{t=0}^{m} \sum_{\substack{J=\{j_1<\cdots<j_t\} \\ J \subseteq \{1,\ldots,m\}}} N_{s,t} \cdot |z_{0,J,J}|^{-1} |u^{J,J}| \tag{19.3.24}$$

for all $s \in \mathbb{Z}_{\geq 1}$, where $N_{s,t}$ is the number of tuples $((\sigma_1, J_1), \ldots, (\sigma_s, J_s))$ where $J_i \subseteq J$ are disjoint subsets (possibly empty) with $\bigcup J_i = J$ and each σ_i is a permutation of J_i. If $|J_i|$ denotes the cardinality of J_i, then there are $\binom{t+s-1}{s-1}$ possibilities for the tuple $(|J_1|, \ldots, |J_s|)$, and each such tuple admits $t!$ corresponding tuples $((\sigma_1, J_1), \ldots, (\sigma_s, J_s))$. Hence $N_{s,t} = \binom{t+s-1}{s-1} t! = \prod_{k=1}^{t} (s+k-1)$. $\square$

19.4 Limiting identity: non positive definite $T^\flat$

Take integers $m, n \geq 1$, assume $m \leq n$, and set $s_0 = (n-m)/2$. Given $a = \mathrm{diag}(a_1, \ldots, a_{n-1}) \in \mathrm{Herm}_{n-1}(\mathbb{R})_{<0}$ and $b \in \mathbb{R}_{<0}$, set $a^\flat = \mathrm{diag}(a_{n-m+1}, \ldots, a_{n-1}) \in \mathrm{Herm}_{n-m}(\mathbb{R})$ and

$$T = (-4\pi)^{-1} \operatorname{diag}(a, -b) \qquad T^\flat = (-4\pi)^{-1} \operatorname{diag}(a^\flat, -b)$$
$$g = (-4\pi)^{-1} \operatorname{diag}(a, b) \qquad g^\flat = (-4\pi)^{-1} \operatorname{diag}(a^\flat, b).$$

We have $T, g \in \mathrm{Herm}_n(\mathbb{R})$ and $T^\flat, g^\flat \in \mathrm{Herm}_m(\mathbb{R})$.

We have

$$\begin{aligned}W^*_{T^\flat}(\alpha, \beta) &= e^{2\pi \operatorname{tr} T^\flat} \Gamma_m(\beta)^{-1} |\det 4\pi T^\flat|^\beta \zeta_{m-1,1}(4\pi g^\flat; \alpha, \beta) \\ &= e^b |\det 4\pi T^\flat|^\beta \pi^{-m+1} \Gamma(\beta)^{-1} \Gamma_{m-1}(\beta - 1)^{-1} e^{2\pi \operatorname{tr} g^\flat} \zeta_{m-1,1}(4\pi g^\flat; \alpha, \beta)\end{aligned}$$

which implies

$$\begin{aligned}&\frac{\partial}{\partial \beta} W^*_{T^\flat}(m, \beta) && (19.4.1)\\ &= \left(\frac{d}{d\beta} \Gamma(\beta)^{-1} \right) |\det 4\pi T^\flat|^\beta \pi^{-m+1} e^b \Gamma_{m-1}(\beta - 1)^{-1} e^{2\pi \operatorname{tr} g^\flat} \zeta_{m-1,1}(4\pi g^\flat; m, \beta)\end{aligned}$$

whenever both sides are evaluated at $\beta \in \mathbb{Z}_{\leq 0}$.

Equation (19.3.11) and Lemma 19.3.2 imply

$$\frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n = \frac{\partial}{\partial \beta}\bigg|_{\beta = m-n} W^*_{T^\flat}(m, \beta). \tag{19.4.2}$$

Since $\Gamma(s)^{-1}$ has residue $(-1)^{n-m}(n-m)!$ at $s = m-n$, we use (19.4.1) and (19.3.21) to find

$$\begin{aligned}&\frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n && (19.4.3)\\ &= (-1)^{n-m}(n-m)! |\det 4\pi T^\flat|^{m-n} (-\pi)^{-m+1} && (19.4.4)\\ &\quad \cdot \int_{\mathbb{C}^{m-1}} \int_1^\infty e^{\operatorname{tr}(aww^*)} e^{b(1+w^*w)x} x^{m-n-1} \\ &\qquad \cdot e^{-\operatorname{tr}(1_{m-1}+ww^*)a^\flat} \Delta|_{z=-a^\flat} (e^{-\operatorname{tr}(1_{m-1}+ww^*)z} \det(z)^{n-m+1}) \, dx \, dw.\end{aligned}$$

Next, we write $w = (w_1, \ldots, w_m)$ and apply Lemma 19.3.3 to find (using $\det(1 + ww^*) =$

$1 + w^* w$ as in [Shi82, Lemma 2.2])

$$
\begin{aligned}
&\frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n \\
&= (-1)^{n+1}(n-m)!|\det 4\pi T^\flat|^{m-n}\pi^{-m+1}\det(-a^\flat)^{n-m} \\
&\quad \cdot \sum_{t=0}^{m-1} \sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m-1\}}} \Bigg(\prod_{k=1}^{m-1-t}(n-m+k)(a_{i_1}\cdots a_{i_t}) \\
&\quad \int_{\mathbb{C}^{m-1}}\int_1^\infty e^{\mathrm{tr}(aww^*)}e^{b(1+w^*w)x}x^{m-n-1}(1+w_{i_1}\overline{w}_{i_1}+\cdots+w_{i_t}\overline{w}_{i_t})\ dx\ dw\Bigg) \\
&= (-1)^{n+1}\pi^{-m+1}(-b)^{m-n} \\
&\quad \cdot \sum_{t=0}^{m-1} \sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m-1\}}} \Bigg((n-1-t)!(a_{i_1}\cdots a_{i_t}) \\
&\quad \int_{\mathbb{C}^{m-1}}\int_1^\infty e^{\mathrm{tr}(aww^*)}e^{b(1+w^*w)x}x^{m-n-1}(1+w_{i_1}\overline{w}_{i_1}+\cdots+w_{i_t}\overline{w}_{i_t})\ dx\ dw\Bigg).
\end{aligned}
\tag{19.4.5}
$$

We have used exponential decay of the function $\int_1^\infty e^{cx}x^{m-n-1}$ as $c \to -\infty$ for convergence estimates (to rearrange integrals). The previous formulas also hold when $T^\flat, m, g^\flat, a^\flat$ are replaced by T, n, g, a (the latter is just the special case $m = n$).

For the reader's convenience, we recall the formulas (which will be used below)

$$
\int_{\mathbb{R}^2} e^{c(x^2+y^2)}\ dx\ dy = -\pi c^{-1} \qquad \int_{\mathbb{R}^2}(x^2+y^2)e^{c(x^2+y^2)}\ dx\ dy = \pi c^{-2} \tag{19.4.6}
$$

valid for any $c \in \mathbb{R}_{<0}$.

Proof of Proposition 19.1.2 when $T^\flat$ is not positive definite. It is enough to check the case where $T^\flat$ is diagonal and signature $(m-1,1)$, by Remarks 19.1.3 and 19.1.4. Take notation as above. There is nothing to check when $m = n$. Otherwise, we may show

$$
\lim_{\substack{a_i\to 0\\ i=1,\ldots,n-m}} \frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n = \frac{d}{ds}\bigg|_{s=-s_0} W^*_{T^\flat,\infty}(s)^\circ_n \tag{19.4.7}
$$

via (19.4.5). Indeed, interchanging the limit and integrals (dominated convergence) and integrating out the variables $w_1, \ldots, w_{n-m}$ gives the claim (using the left identity in (19.4.6)). $\square$

19.5 Limiting identity: positive definite $T^\flat$

Take any integer $n \geq 1$ and set $m = n - 1$, so that $s_0 = (n-m)/2 = 1/2$. Given $a = \mathrm{diag}(a_1, \ldots, a_{n-1}) \in \mathrm{Herm}_{n-1}(\mathbb{R})_{<0}$ and $b \in \mathbb{R}_{<0}$, set

$$T^\flat = (-4\pi)^{-1} a \quad \text{and} \quad T = (-4\pi)^{-1} \operatorname{diag}(a, -b). \tag{19.5.1}$$

Equation (19.3.11) and Lemma 19.3.2 imply

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,\infty}(s)^\circ_n = -\frac{d}{ds}\bigg|_{s=1/2} W^*_{T^\flat,\infty}(s)^\circ_n = -\frac{\partial}{\partial \beta}\bigg|_{\beta=0} W^*_{T^\flat}(n,\beta) \tag{19.5.2}$$

where the first equality is via the functional equation from Lemma 16.2.1. We have

$$W^*_{T^\flat}(n,\beta) = \Gamma_m(\beta)^{-1}\det(-a)^\beta \zeta_m(-a;n,\beta) = (-1)^m e^{-\operatorname{tr} a}\det(-a)^{-1}\Delta|_{z=1_m}(e^{\operatorname{tr} az}\det(z)^{-\beta}) \tag{19.5.3}$$

where the first equality is by (19.3.5) and (19.3.13), and the second equality is by (19.3.9) and (19.3.7). Applying Lemma 19.3.3 then yields

$$-\frac{\partial}{\partial \beta}\bigg|_{\beta=0} W^*_{T^\flat}(n,\beta) = \det(a)^{-1}\sum_{t=0}^{m-1}\sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}}(m-1-t)!(a_{i_1}\cdots a_{i_t}). \tag{19.5.4}$$

Before proceeding, we define several functions which serve only to aid computation in Section 19.5. Set

$$d_m := \sum_{t=0}^{m}\sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}}(m-t)!(a_{i_1}\cdots a_{i_t}) \tag{19.5.5}$$

$$q_m(x) := (x+a_1)^{-1}\cdots(x+a_m)^{-1} \tag{19.5.6}$$

$$r_m(x) := 1-(x+a_1)^{-1}-\cdots-(x+a_m)^{-1} \tag{19.5.7}$$

$$h_m(x) := q(x)e^x\sum_{k=0}^{m-1}\sum_{t=0}^{k}\sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}}(m-1-k)!a_{i_1}\cdots a_{i_t}x^{k-t} \tag{19.5.8}$$

$$u_m(x) := \sum_{t=0}^{m}\sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}}(m-t)!a_{i_1}\cdots a_{i_t}(1-(x+a_{i_1})^{-1}-\cdots-(x+a_{i_t})^{-1}) \tag{19.5.9}$$

$$f_m(x) := q_m(x)u_m(x) \tag{19.5.10}$$

where dependence on $a_1, \ldots, a_m$ is suppressed from notation.

Next, we consider (19.4.5) for the matrix T. Changing variables $x \mapsto x/b$ and computing the dw integral (using (19.4.6)), we find

$$\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n = -\int_{-\infty}^{1/b} f_m(x)e^x x^{-1}\, dx = -\operatorname{Ei}(b) + \int_{-\infty}^{1/b} (1-f_m(x))e^x x^{-1}\, dx \quad (19.5.11)$$

with $m = n-1$ as above, and where Ei is the exponential integral function from Section 8.2.

Lemma 19.5.1. *We have $f_m(x) = 1 + O(x)$ near $x = 0$.*

Proof. In the lemma statement, the variables $a_1, \ldots, a_m$ are understood to be fixed (and negative). Since $f_m(x)$ is a rational function of x, it is enough to check $f_m(0) = 1$, i.e. that

$$\sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}} (m-t)!a_{j_1}^{-1}\cdots a_{j_{m-t}}^{-1}(1-a_{i_1}^{-1}-\cdots-a_{i_t}^{-1}) = 1$$

where $\{j_1, \ldots, j_{m-t}\} = \{1, \ldots, m\} \setminus \{i_1, \ldots, i_t\}$. This holds because the sum telescopes, i.e. for any given $t = 0, \ldots, m-1$, we have

$$\sum_{\substack{I=\{i_1<\cdots<i_t\}\\ I\subseteq\{1,\ldots,m\}}} (m-t)!a_{j_1}^{-1}\cdots a_{j_{m-t}}^{-1} = \sum_{\substack{I'=\{i'_1<\cdots<i'_{t+1}\}\\ I'\subseteq\{1,\ldots,m\}}} (m-t-1)!a_{j'_1}^{-1}\cdots a_{j'_{m-t-1}}^{-1}(a_{i'_1}^{-1}+\cdots+a_{i'_t}^{-1})$$

where $\{j'_1, \ldots, j'_{m-t-1}\} = \{1, \ldots, m\} \setminus \{i'_1, \ldots, i'_{t+1}\}$. $\square$

Lemma 19.5.2. *We have $\frac{d}{dx}h_m(x) = (1-f_m(x))e^x x^{-1}$.*

Proof. We prove this by induction on m. The case $m = 0$ is clear, as both sides of the identity are 0. Next, suppose the claim holds for some m. We write $f_{m+1}(x)$, $h_{m+1}(x)$, etc. for the corresponding functions formed with respect to the tuple $(a_1, \ldots, a_m, a_{m+1})$ for any given choice of $a_{m+1} \in \mathbb{R}_{<0}$. Observe that we have an inductive formula

$$h_{m+1}(x) = h_m(x) + q_{m+1}(x)d_m \quad (19.5.12)$$

which implies

$$\frac{d}{dx}h_{m+1}(x) - \frac{d}{dx}h_m(x) = q_{m+1}(x)r_{m+1}(x)e^x d_m. \quad (19.5.13)$$

So it is enough to check $f_m(x) - f_{m+1}(x) = xq_{m+1}(x)r_{m+1}(x)d_m$, which is equivalent to checking

$$(x + a_{m+1})u_m(x) - u_{m+1}(x) = xr_{m+1}(x)d_m. \quad (19.5.14)$$

To see that this holds, we first compute

$$\begin{aligned}
&u_{m+1} - (a_{m+1}u_m(x) - a_{m+1}(x+a_{m+1})^{-1}d_m) \\
&= \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m+1-t)!a_{i_1}\cdots a_{i_t}(1-(x+a_{i_1})^{-1} - \cdots - (x+a_{i_t})^{-1}).
\end{aligned}$$

Using the identity $a_{m+1}(x+a_{m+1})^{-1} = 1 - x(x+a_{m+1})^{-1}$, we see that (19.5.14) is equivalent to the identity

$$\begin{aligned}
&xu_m(x) - xr_{m+1}(x)d_m + (1-x(x+a_{m+1})^{-1})d_m \\
&= \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m+1-t)!a_{i_1}\cdots a_{i_t}(1-(x+a_{i_1})^{-1} - \cdots - (x+a_{i_t})^{-1}).
\end{aligned} \tag{19.5.15}$$

To see that the latter identity holds, we compute

$$\begin{aligned}
&xu_m(x) - xr_{m+1}(x)d_m + (1-x(x+a_{m+1})^{-1})d_m = xu_m(x) - xr_m(x) + d_m \\
&= \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m-t)!a_{i_1}\cdots a_{i_t}(m+1-t-a_{j_1}(x+a_{j_1})^{-1} - \cdots - a_{j_{m-t}}(x+a_{j_{m-t}})^{-1}) \\
&= \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m+1-t)!a_{i_1}\cdots a_{i_t} \\
&\quad - \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m-t)!a_{i_1}\cdots a_{i_t}(a_{j_1}(x+a_{j_1})^{-1} + \cdots + a_{j_{m-t}}(x+a_{j_{m-t}})^{-1})
\end{aligned}$$

where $\{j_1,\ldots,j_{m-t}\} = \{1,\ldots,m\} \setminus \{i_1,\ldots,i_t\}$.

We thus find that (19.5.15) is equivalent to the identity

$$\begin{aligned}
&\sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m+1-t)!a_{i_1}\cdots a_{i_t}((x+a_{i_1})^{-1} + \cdots + (x+a_{i_t})^{-1}) \\
&= \sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} (m-t)!a_{i_1}\cdots a_{i_t}(a_{j_1}(x+a_{j_1})^{-1} + \cdots + a_{j_{m-t}}(x+a_{j_{m-t}})^{-1})
\end{aligned}$$

with $\{j_1,\ldots,j_{m-t}\}$ as above, and this identity holds because both expressions are equal to

$$\sum_{t=0}^{m} \sum_{\substack{I=\{i_1<\cdots<i_t\} \\ I\subseteq\{1,\ldots,m\}}} \sum_{i=1}^{t} (m+1-t)!a_{i_1}\cdots a_{i_t}(x+a_{i_i})^{-1}. \qquad \square$$

Proof of Proposition 19.1.2 when $T^\flat$ is positive definite. We may assume $T^\flat$ is diagonal by Remark 19.1.4. With T and $T^\flat$ as above, we find

$$\lim_{b \to 0^-} \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(s)^\circ_n + \mathrm{Ei}(b) \right) = \int_{-\infty}^{0} (1 - f_m(x)) e^x x^{-1} \, dx \tag{19.5.16}$$

via (19.5.11) (and Lemma 19.5.1 for convergence of the integral). The asymptotics for $\mathrm{Ei}(b)$ as $b \to 0^-$ (8.2.2) show that it is enough to verify the identity

$$\frac{d}{ds}\bigg|_{s=-1/2} W^*_{T^\flat,\infty}(s)^\circ_n = \int_{-\infty}^{0} (1 - f_m(x)) e^x x^{-1} \, dx. \tag{19.5.17}$$

The left-hand side was computed in (19.5.4) (via (19.5.2)). The right-hand side is equal to $h_m(0)$ (in the notation above) via the explicit antiderivative result from Lemma 19.5.2. Inspecting the formula for $h_m(x)$ shows that the claimed identity holds. $\square$

Part VII

Siegel–Weil

Our main results (arithmetic Siegel–Weil) are in Section 22. We also give some explicit formulas for special values (local Siegel–Weil and geometric Siegel–Weil) in Sections 20 and 21. These special value formulas will be needed as ingredients in the proofs of our arithmetic Siegel–Weil results.

20 Local Siegel–Weil

We need precise information about special values of local Whittaker functions (i.e. local Siegel–Weil, with explicit constants as in Lemma 20.3.6). The application to uniformization of special cycles is Lemma 20.4.1. These results do not seem available in the literature in the generality or explicitness that we need. We omit some computations (but give statements) for arguments which are presumably routine or similar to arguments available in the literature.

20.1 Volume forms

Given a scheme X which is smooth and equidimensional over a field A, a *volume form* (or *gauge form*) on X will mean a nowhere vanishing (algebraic) differential form of top degree on X. When X is also affine and A is a local field, the set $X(A)$ has the natural structure of an A-analytic manifold (in the sense of [Ser06, Part II, Chapter III]). In this case, a volume form on X defines a Borel measure on $X(A)$ in a standard way (see [Wei82, §2.2]).

We use volume forms to normalize various Haar measures. Let B be a degree 2 étale algebra over a field A of characteristic $\neq 2$, and write $b \mapsto \overline{b}$ for the nontrivial involution on B. Let V be a B/A Hermitian space which is free of rank n, and set $G = U(V)$. Fix a nonnegative integer $m \leq n$, and choose translation invariant volume forms α and β on V^m and Herm_m respectively (viewed as group schemes over A). The forms α and β have degrees $2nm$ and m^2 respectively.

Consider the moment map

$$\begin{aligned} V^m &\xrightarrow{\mathcal{T}} \mathrm{Herm}_m \\ \underline{x} &\longmapsto (\underline{x}, \underline{x}). \end{aligned} \tag{20.1.1}$$

We assume $n \geq m$, and write $V^m_{\mathrm{reg}} \subseteq V^m$ for the open subscheme where $\det \mathcal{T}$ is invertible. A tangent space calculation shows that $\mathcal{T}$ is smooth when restricted to V^m_{reg}.

Given $T \in \mathrm{Herm}_m(A)$, we write $\Omega_T \subseteq V^m$ for the fiber of the moment map over T. If $\underline{x} \in V^m(A)$ has Gram matrix $T = (\underline{x}, \underline{x})$, then $g \mapsto g^{-1} \cdot \underline{x}$ defines a morphism $\iota_{\underline{x}} \colon G \to \Omega_T$. If $\det T$ is invertible, then a dimension count and tangent space calculation show that $\iota_{\underline{x}}$ is smooth. If $\det T$ is invertible, if A is a local field, and if $G_{\underline{x}} \subseteq G$ denotes the stabilizer of

$\underline{x}$, then the induced map $G_{\underline{x}}(A)\backslash G(A) \to \Omega_T(A)$ is a homeomorphism (surjectivity is from Witt's theorem, and openness is from the submersivity of $G(A) \to \Omega_T(A)$, which in turn comes from smoothness of $\iota_{\underline{x}}$).

Lemma 20.1.1. *There exists an algebraic differential form ν on V^m_{reg} of degree $m(2n-m)$ satisfying the following conditions.*

(1) We have $\alpha = \mathcal{T}^*(\beta) \wedge \nu$.

(2) For the $G \times \mathrm{Res}_{B/A} \mathrm{GL}_m$ action on V^m_{reg} given by $x \mapsto gxh^{-1}$ for $(g,h) \in G \times \mathrm{Res}_{F/F^+} \mathrm{GL}_m$, we have $(g,h)^*\nu = \det({}^t\overline{h}h)^{m-n}\nu$.

(3) For each $x \in V^m_{\mathrm{reg}}$, the restriction of $\nu \colon T_x(V^m) \to \mathbb{G}_a$ to $\ker d\mathcal{T}_x$ is nonzero.

(4) For any fixed non-degenerate subspace $V^\flat \subseteq V$ which is free of rank m, and for $\underline{x} \in V^{\flat m}_{\mathrm{reg}}(A)$, the differential form

$$\det(\underline{x}, \underline{x})^{m-n} \iota^*_{\underline{x}} \nu \tag{20.1.2}$$

on G is independent of the choice of $\underline{x}$. This form is right G-invariant.

Proof. The case $m = n$ is stated in [KR14, §10]. The analogue of that case for orthogonal groups is discussed in [KRY06, Lemmas 5.3.1, 5.3.2] (there stated and proved for three-dimensional quadratic spaces). The present lemma may be proved by a similar computation.

Part (4) follows from part (2) (where "non-degenerate subspace" means that the restriction of the Hermitian pairing is non-degenerate). In part (3), $x \in V^m_{\mathrm{reg}}$ means $x \in V^m_{\mathrm{reg}}(S)$ for some suppressed A-scheme S, and we similarly abused notation in part (2). In part (3), the symbol $\mathbb{G}_a$ denotes the additive group scheme. □

20.2 Special value formula

We retain notation from Section 20.1, and specialize to the case where B/A is the extension F_v/F^+_v where F^+_v is a local field of characteristic $\neq 2$. If F^+_v is Archimedean, we assume F_v/F^+_v is $\mathbb{C}/\mathbb{R}$. We often use subscripts v to emphasize F^+_v being a local field, e.g. we write $\underline{x}_v$ for elements of $V^m_{\mathrm{reg}}(F^+_v)$.

Fix a nontrivial additive character $\psi_v \colon F^+_v \to \mathbb{C}^\times$. We write db_v for the self-dual Haar measure on $\mathrm{Herm}_m(F^+_v)$ with respect to the trace pairing $(b,c) \mapsto \psi_v(\mathrm{tr}(bc))$. We also write $d\underline{x}_v$ for the self-dual Haar measure on $V^m(F^+_v)$ with respect to the pairing $\psi_v(\mathrm{tr}_{F_v/F^+_v}(\mathrm{tr}(-,-)))$.

Fix translation-invariant volume forms α and β as in Section 20.1. These determine Haar measures $d_\beta b_v$ and $d_\alpha \underline{x}_v$ on $\mathrm{Herm}_m(F_v^+)$ and $V^m(F_v^+)$ respectively. Define positive real constants $c_v(\alpha, \psi_v)$ and $c_v(\beta, \psi_v)$ such that

$$d_\alpha \underline{x}_v = c_v(\alpha, \psi_v) d\underline{x}_v \qquad d_\beta b_v = c_v(\beta, \psi_v) db_v. \tag{20.2.1}$$

Suppose $T \in \mathrm{Herm}_m(F_v^+)$ is a matrix with $\det T \neq 0$. For the rest of Section 20.2, fix a differential form ν as in Lemma 20.1.1. The restriction of $\det(T)^{m-n}\nu$ to Ω_T is a G-invariant volume form on Ω_T, and we write $d_{T,\nu}\underline{x}_v$ for the resulting measure on $\Omega_T(F_v^+)$.

It is known that there exists a constant $c_{T,v}$ (depending on T, the measure $d_{T,\nu}\underline{x}_v$, and the character ψ_v) such that

$$W_{T,v}(s_0, \Phi_{\varphi_v}) = c_{T,v} \int_{\Omega_T(F_v^+)} \varphi_v(\underline{x}_v) \, d_{T,\nu}\underline{x}_v \tag{20.2.2}$$

holds for any Schwartz function $\varphi_v \in \mathcal{S}(V^m(F_v^+))$ (see [Ich04, Lemma 5.1, Lemma 5.2]). Here we set $s_0 \coloneqq (n-m)/2$ as usual, and Φ_{φ_v} is the Siegel–Weil section associated with φ_v (Section 14.2). If $\Omega_T(F_v^+) = \emptyset$, we thus have $W_{T,v}(s_0, \Phi_{\varphi_v}) = 0$ for all φ_v.

We may compute the constant $c_{T,v}$ by evaluating (20.2.2) on any nonzero nonnegative Schwartz function φ_v. We may take φ_v to have support which is compact and contained in $V^m_{\mathrm{reg}}(F_v^+)$. The relation $\alpha = \mathcal{T}^*(\beta) \wedge \nu$ and an "integrate along the fibers of $\mathcal{T}$" computation (similar to the proof of [KRY06, Proposition 5.3.3]) gives

$$c_{T,v} = \frac{\gamma_{\psi_v}(V)^{-m} c_v(\beta, \psi_v)}{c_v(\alpha, \psi_v)} |\det T|^{n-m}_{F_v^+}. \tag{20.2.3}$$

Here $\gamma_{\psi_v}(V)$ is the Weil index, as appearing in the Weil representation (Section 14.2).

Lemma 20.2.1 (Local Siegel–Weil). *Let V be an F_v/F_v^+ Hermitian space of rank n, and let $\psi_v \colon F_v^+ \to \mathbb{C}^\times$ be a nontrivial additive character. Fix a non-degenerate subspace $V^\flat \subseteq V$ which is free of rank m, and fix a Haar measure on $U(V^{\flat\perp})(F_v^+)$.*

There exists a unique Haar measure dg_v on $G(F_v^+)$ such that, for any basis $\underline{x}_v \in V^{\flat m}$ of $V^\flat$ and any Schwartz function $\varphi_v \in \mathcal{S}(V^m(F_v^+))$, we have

$$W_{T,v}(s_0, \Phi_{\varphi_v}) = \gamma_{\psi_v}(V)^{-m} |\det T|^{n-m}_{F_v^+} \int_{G_{\underline{x}_v}(F_v^+) \backslash G(F_v^+)} \varphi_v(g_v^{-1}\underline{x}_v) \, dg_v \tag{20.2.4}$$

for the corresponding quotient measure, where $T = (\underline{x}_v, \underline{x}_v)$ is the Gram matrix of $\underline{x}_v$ (and where the Haar measure on $G_{\underline{x}_v}(F_v^+)$ is induced by the canonical identification $G_{\underline{x}_v} \cong U(V^{\flat\perp})$).

Proof. Select any basis $\underline{x}_v$ of $V^\flat$. Set $\omega_1 = \det(\underline{x}_v, \underline{x}_v)^{m-n} \iota_{\underline{x}}^* \nu$ (temporary notation). We know ω_1 does not depend on the choice of $\underline{x}_v$, by Lemma 20.1.1(4). Let ω_2 be a right G-invariant differential form of degree $(n-m)^2$ on G such what $\omega_1 \wedge \omega_2$ is a nowhere vanishing differential form of top degree n^2 (also right G-invariant). The volume form $\omega_1 \wedge \omega_2$ on G defines a Haar measure on $G(F_v^+)$. The restriction $\omega_2|_{G_{\underline{x}}}$ is a volume form on $G_{\underline{x}}$ (by smoothness of $\iota_{\underline{x}}$), and defines a Haar measure on $G_{\underline{x}_v}(F_v^+)$. The resulting quotient measure on $G_{\underline{x}}(F_v^+) \backslash G(F_v^+) \cong \Omega_T(F_v^+)$ is precisely the measure for the volume form $(\det T)^{m-n} \nu|_{\Omega_T}$ on Ω_T.

The lemma then follows from (20.2.2) and the constant calculated in (20.2.3). $\square$

Remark 20.2.2. Consider the situation of Lemma 20.2.1, and suppose $V^{\flat\prime} \subseteq V$ is a subspace which is isomorphic to $V^\flat$ as a Hermitian space. Suppose $f_v \in U(V)(F_v^+)$ satisfies $f_v(V^\flat) = V^{\flat\prime}$, and equip $U(V^{\flat\prime\perp})(F_v^+)$ with the Haar measure induced from $U(V^{\flat\perp})(F_v^+)$ via f_v. If dg_v and dg_v' are the induced Haar measures on $G(F_v^+)$ corresponding to $V^\flat$ and $V^{\flat\prime}$ respectively (Lemma 20.2.1), a change of variables shows $dg_v = dg_v'$.

20.3 Explicit Haar measures

For our application to uniformization of special cycles (Section 20.4), we need to explicitly compute the Haar measures from Lemma 20.2.1 in a few cases. The main result of this subsection is Lemma 20.3.6, and the other lemmas are auxiliary.

We retain notation from Section 20.2. In addition, we assume that F_v^+ is non-Archimedean and that ψ_v is unramified. Let ϖ_0 be a uniformizer of F_v^+. For F_v/F_v^+ ramified in Section 20.3, we assume that F_v^+ has residue characteristic $\neq 2$ and write $\varpi \in F_v$ for a uniformizer satisfying $\varpi^\sigma = -\varpi$, where ϖ^σ denotes the Galois conjugate of ϖ under the Galois involution σ of F_v over F_v^+.

Let M_2° be the rank 2 self-dual lattice described in Section 14.2, and write $U(M_2^\circ)$ for the group of (unitary) automorphisms of M_2°. Let q_v be the residue cardinality of F_v^+.

The next lemma should be compared with Witt's theorem for lattices with quadratic forms, as in [Mor79].

Lemma 20.3.1. *For any given $c \in \mathcal{O}_{F_v^+}^\times$, the group $U(M_2^\circ)$ acts transitively on the set*

$$\{x \in M_2^\circ : (x,x) = c\}. \tag{20.3.1}$$

If F_v/F_v^+ is inert, the same holds for any $c \in \varpi_0 \mathcal{O}_{F_v^+}^\times$.

Proof. Given $y \in M_2^\circ$, we write $\langle y \rangle \subseteq M_2^\circ$ for the submodule generated by y. If F_v/F_v^+ is ramified, we view ϖ as a generator of the different ideal $\mathfrak{d}$, and we otherwise view 1 as a generator of $\mathfrak{d}$. Choose a basis e_1, e_2 of M_2° with Gram matrix given by (14.2.2). In this basis, we also consider the elements

$$w' = \begin{pmatrix} 0 & 1 \\ \epsilon & 0 \end{pmatrix} \qquad m(a) = \begin{pmatrix} a & 0 \\ 0 & \overline{a}^{-1} \end{pmatrix} \qquad n(b) = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \qquad \epsilon = \begin{cases} -1 & \text{if } F_v/F_v^+ \text{ is ramified} \\ 1 & \text{else} \end{cases} \tag{20.3.2}$$

of $U(M_2^\circ)$ (acting on column vectors), where $a \in \mathcal{O}_{F_v}^\times$ and $b \in \mathcal{O}_{F_v}$ satisfies $\overline{b} = -\epsilon b$.

Case 1. Assume F_v/F_v^+ is unramified and $c \in \mathcal{O}_{F_v^+}^\times$. Given any $x \in M_2^\circ$ with $(x,x) = c$, there exists an orthogonal direct sum decomposition $M_2^\circ = \langle x \rangle \oplus \langle y \rangle$ for some $y \in M_2^\circ$ with $(y,y) = 1$ (by self-dualness). Via this decomposition, the lemma is clear in this case.

Case 2. Assume F_v/F_v^+ is ramified and $c \in \mathcal{O}_{F_v^+}^\times$. Suppose $x = a_1 e_1 + a_2 e_2 \in M_2^\circ$ with $(x,x) = c$. Without loss of generality, we may assume $a_2 \in \mathcal{O}_{F_v}^\times$ (replace x with $w'x$ if necessary), and we may further assume $a_2 = 1$ (replace x with $m(\overline{a}_2)x$). We then have $\mathrm{tr}_{F_v/F_v^+}(\varpi^{-1}a_1) = -c$. Given another $x' = a_1' e_1 + e_2 \in M_2^\circ$ with $(x', x') = c$, we take $b = a_1' - a_1$ and have $n(b)x = x'$.

Case 3. Assume F_v/F_v^+ is inert and $c \in \varpi_0 \mathcal{O}_{F_v^+}^\times$. Suppose $x = a_1 e_1 + a_2 e_2 \in M_2^\circ$ with $(x,x) = c$. Without loss of generality, we may assume $a_2 = 1$ and $\mathrm{tr}_{F_v/F_v^+}(a_1) = c$ (argue as in Case 2). Given another $x' = a_1' e_1 + e_2 \in M_2^\circ$ with $(x', x') = c$, we take $b = a_1' - a_1$ and have $n(b)x = x'$. $\square$

Remark 20.3.2. If F_v/F_v^+ is nonsplit, the proof of Lemma 20.3.1 shows that the $U(M_2^\circ)$ orbits in $(M_2^\circ \otimes_{\mathcal{O}_{F_v}} F_v) \setminus \{0\}$ are completely classified by pairs $(c, r) \in F_v^+ \times \mathbb{Z}$, i.e. that the map

$$\begin{aligned} U(M_2^\circ)\backslash((M_2^\circ \otimes_{\mathcal{O}_{F_v}} F_v) \setminus \{0\}) &\longrightarrow F_v^+ \times \mathbb{Z} \\ x &\longmapsto (x,x), \min\{r \in \mathbb{Z} : \varpi^r x \in M_2^\circ\} \end{aligned} \tag{20.3.3}$$

is a bijection (say, setting $\varpi = \varpi_0$ in the inert case).

Lemma 20.3.3. *Let L be a self-dual hermitian $\mathcal{O}_{F_v}$-lattice of rank n. Any isomorphism between self-dual sublattices of L extends to a (unitary) automorphism of L. The same holds for almost self-dual lattices of rank $n-1$.*

Proof. Any self-dual lattice $L^\flat \subseteq L$ admits a (unique) orthogonal direct sum decomposition $L = L^\flat \oplus L^\#$ where $L^\#$ is also self-dual. This immediately implies the claim for self-dual sublattices of L, as self-dual lattices are unique up to isomorphism (for a fixed rank).

Next, assume that F_v/F_v^+ is nonsplit and that $L^\flat \subseteq L$ is almost self-dual of rank $n-1$. There is an orthogonal direct sum decomposition $L^\flat = L^{\flat\flat} \oplus L^{\flat\#}$, where $L^{\flat\flat}$ is self-dual of rank $n-2$ and $L^{\flat\#}$ is almost self-dual of rank 1. We also have an orthogonal direct sum decomposition $L = L^{\flat\flat} \oplus L^\#$ where $L^\#$ is self-dual of rank 2.

Suppose $L^{\flat\prime} \subseteq L$ is another almost self-dual lattice of rank $n-1$, equipped with an isomorphism $L^\flat \to L^{\flat\prime}$. Applying the result just proved above (in the case of rank $n-2$ self-dual sublattices), we may assume there is an orthogonal decomposition $L^{\flat\prime} = L^{\flat\flat} \oplus L^{\flat\prime\#}$ where $L^{\flat\#} \cong L^{\flat\prime\#}$. We thus reduce to the case $n=2$ (the claim for $L^\#$), which was proved in Lemma 20.3.1. $\square$

Lemma 20.3.4. *Assume F_v/F_v^+ is nonsplit, and let V be an F_v/F_v^+ Hermitian space of rank n, and assume that V contains a full-rank self-dual lattice. Suppose $L^\flat \subseteq V$ is a non-degenerate lattice of rank $n-1$ satisfying $L^\flat \subseteq L^{\flat *}$ and $t(L^\flat) \leq 1$. Then $L^\flat$ is contained in a self-dual lattice of rank n.*

Proof. Recall that $t(L^\flat) \coloneqq \dim_k((L^{\flat *}/L^\flat) \otimes k)$ where k is the residue field of $\mathcal{O}_{F_v}$.

Let $L^\flat \subseteq V$ be as in the lemma statement. The existence of such $L^\flat$ implies $n \geq 2$. There exists an orthogonal decomposition $L^\flat = L^{\flat\flat} \oplus L^{\flat\#}$ where $L^{\flat\flat}$ is self-dual of rank $n-2$. Replacing V with the orthogonal complement of $L^{\flat\flat}$, we reduce immediately to the case $n=2$, which we now assume.

Let ϖ be a uniformizer for F_v (take $\varpi = \varpi_0$ if F_v/F_v^+ is inert). We may take $V = M_2^\circ \otimes F_v$, where M_2° is as in Lemma 20.3.1. We also choose a standard basis e_1, e_2 for M_2° and consider the elements $w', m(a), n(b) \in U(V)$ as in the proof of that lemma (now allowing $a \in F_v^\times$ and allowing $b \in F_v$ satisfying $\overline{b} = -\epsilon b$).

The rank one lattice $L^\flat$ is generated by an element $x = a_1 e_1 + a_2 e_2$ for some $a_1, a_2 \in F_v$ (such that (x,x) is nonzero and lies in $\mathcal{O}_{F_v^+}$). It is enough to check chat the orbit $U(V) \cdot x$ intersects M_2°. Acting on x by $m(a) \in U(V)$ for suitable a, we see that it is enough to check the case where $a_2 = 1$ and $a_1 \in F_v^\times$.

If F_v/F_v^+ is inert, there exists $a' \in \mathcal{O}_{F_v}$ such that $\mathrm{tr}_{F_v/F_v^+}(a') = (x,x)$ since $\mathcal{O}_{F_v}$ is self-dual with respect to the trace pairing. If F_v/F_v^+ is ramified, there exists $a' \in \mathcal{O}_{F_v}$ such that $\mathrm{tr}_{F_v/F_v^+}(-\varpi^{-1}a') = (x,x)$ since $\mathcal{O}_{F_v}$ and $\varpi^{-1}\mathcal{O}_{F_v}$ are dual. In either case, we can take $b = a' - a_1$, and have $n(b)x \in M_2^\circ$. $\square$

Lemma 20.3.5. *In the situations of Lemma 20.3.1, choose $x \in M_2^\circ$ with $(x,x) = c$ and form the orthogonal complement lattice $x^\perp \subseteq M_2^\circ$ (of rank one). We view both $U(M_2^\circ)$ and $U(x^\perp)$ as subgroups of $U(M_2^\circ \otimes F_v)$.*

Viewing $U(x^\perp)$ as the norm-one subgroup of $\mathcal{O}_{F_v}^\times$, we have

$$U(M_2^\circ) \cap U(x^\perp) = \{\alpha \in \mathcal{O}_{F_v}^\times : \alpha\overline{\alpha} = 1, \text{ and } \alpha \equiv 1 \pmod{c\mathfrak{d}\mathcal{O}_{F_v}}\} \subseteq U(x^\perp). \tag{20.3.4}$$

The subgroup $U(M_2^\circ) \cap U(x^\perp) \subseteq U(x^\perp)$ has index

$$\begin{cases} 1 & \text{if } c \in \mathcal{O}_{F_v^+}^\times \text{ and } F_v/F_v^+ \text{ is unramified} \\ 2 & \text{if } c \in \mathcal{O}_{F_v^+}^\times \text{ and } F_v/F_v^+ \text{ is ramified} \\ q_v + 1 & \text{if } c \in \varpi_0\mathcal{O}_{F_v^+}^\times \text{ and } F_v/F_v^+ \text{ is inert.} \end{cases} \tag{20.3.5}$$

Proof. We express elements of $U(M_2^\circ \otimes F_v)$ in a standard basis e_1, e_2 of M_2°, as in the proof of Lemma 20.3.1.

Case 1. Assume F_v/F_v^+ is unramified and $c \in \mathcal{O}_{F_v^+}^\times$. We then have $U(M_2^\circ) \cap U(x^\perp) = U(x^\perp)$, as follows immediately from an orthogonal direct sum decomposition $M_2^\circ = \langle x \rangle \oplus \langle y \rangle$ as in the proof of Lemma 20.3.1 Case 1.

Case 2. Assume F_v/F_v^+ is ramified and $c \in \mathcal{O}_{F_v^+}^\times$. By the proof of Lemma 20.3.1 Case 2, we may assume (after conjugating $U(M_2^\circ \otimes F_v)$ by an appropriate element of $U(M_2^\circ)$) that $x = a_1 e_1 + e_2$ for some $a_1 \in \mathcal{O}_{F_v}$, where $a_1 - \overline{a}_1 = -\varpi c$. Then $\overline{a}_1 e_1 + e_2$ is orthogonal to x. For every $\alpha \in \mathcal{O}_{F_v}^\times$, the matrix

$$\begin{pmatrix} a_1 & \overline{a}_1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} a_1 & \overline{a}_1 \\ 1 & 1 \end{pmatrix}^{-1} = (-\varpi c)^{-1} \begin{pmatrix} a_1 - \overline{a}_1\alpha & (-1+\alpha)a_1\overline{a}_1 \\ 1 - \alpha & -\overline{a}_1 + a_1\alpha \end{pmatrix} \tag{20.3.6}$$

lies in $U(M_2^\circ)$ if and only if $\alpha \equiv 1 \pmod{\varpi\mathcal{O}_{F_v}}$. The claim about index follows from surjectivity of the reduction modulo ϖ map

$$\{\alpha \in \mathcal{O}_{F_v}^\times : \alpha\overline{\alpha} = 1\} \to \{\alpha \in \mathbb{F}_{q_v}^\times : \alpha^2 = 1\} \tag{20.3.7}$$

(surjectivity is by smoothness of the corresponding unitary group over $\operatorname{Spec} \mathcal{O}_{F_v^+}$).

Case 3. Assume F_v/F_v^+ is inert and $c \in \varpi_0\mathcal{O}_{F_v^+}^\times$. By the proof of Lemma 20.3.1 Case 3, we may assume (after conjugating $U(M_2^\circ \otimes F_v)$ by an appropriate element of $U(M_2^\circ)$) that $x = a_1 e_1 + e_2$ for some $a_1 \in \mathcal{O}_{F_v}$, where $a_1 + \overline{a}_1 = c$. Then $-\overline{a}_1 e_1 + e_2$ is orthogonal to x. For every $\alpha \in \mathcal{O}_{F_v}^\times$, the matrix

$$\begin{pmatrix} a_1 & -\overline{a}_1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} a_1 & -\overline{a}_1 \\ 1 & 1 \end{pmatrix}^{-1} = c^{-1} \begin{pmatrix} a_1 + \overline{a}_1\alpha & (1-\alpha)a_1\overline{a}_1 \\ 1 - \alpha & \overline{a}_1 + a_1\alpha \end{pmatrix} \tag{20.3.8}$$

lies in $U(M_2^\circ)$ if and only if $\alpha \equiv 1 \pmod{\varpi_0\mathcal{O}_{F_v}}$. The claim about index follows from surjectivity of the reduction modulo ϖ_0 map

$$\{\alpha \in \mathcal{O}_{F_v}^\times : \alpha\overline{\alpha} = 1\} \to \{\alpha \in \mathbb{F}_{q_v^2}^\times : \alpha\overline{\alpha} = 1\} \tag{20.3.9}$$

(surjectivity is by smoothness of the corresponding unitary group over $\operatorname{Spec} \mathcal{O}_{F_v^+}$). $\square$

We take a particular choice of Schwartz function φ_v in the next lemma, which immediately determines the Haar measure for other choices of φ_v in Lemma 20.2.1. If mn is odd and F_v/F_v^+ is inert with F_v^+ of residue characteristic 2, we also require $F_v^+ = \mathbb{Q}_2$ (because of Lemma 14.2.1).

Lemma 20.3.6. *Take $m = n-1$ or $m = n$ and $s_0 := (n-m)/2$. Assume the rank n Hermitian space V contains a full-rank self-dual lattice L of full rank. Let $K_v \subseteq G = U(V)$ be the stabilizer of such a lattice L.*

Consider any $\underline{x}_v \in V^m(F_v^+)$ with nonsingular Gram matrix $T = (\underline{x}_v, \underline{x}_v) \in \mathrm{Herm}_m(F_v^+)$. Let $\mathbf{1}_L$ be the characteristic function of L, and set $\varphi_v = \mathbf{1}_L^{\otimes m} \in \mathcal{S}(V^m(F_v^+))$.

Give $G(F_v^+)$ the Haar measure which assigns volume 1 to K_v. Give $G_{\underline{x}_v}(F_v^+)$ the Haar measure which assigns volume 1 to the (unique) maximal open compact subgroup. We have

$$W^*_{T,v}(s_0)^\circ_n = \frac{1}{e}\int_{G_{\underline{x}_v}(F_v^+)\backslash G(F_v^+)} \varphi_v(g_v^{-1}\underline{x}_v)\, dg_v \qquad e := \begin{cases} 2 & \text{if } F_v/F_v^+ \text{ is ramified and } m = n-1 \\ 1 & \text{else} \end{cases} \tag{20.3.10}$$

with respect to the associated quotient measure.

Proof. Recall that $W^*_{T,v}(s)^\circ_n$ is our notation for a certain normalized spherical Whittaker function (Section 15.3), which is a rescaled version of $W_{T,v}(s, \Phi_{\varphi_v})$.

In the lemma statement, the stabilizer in $G(F_v^+)$ of any full-rank self-dual lattice in V has volume 1 (because any such stabilizer is conjugate to K_v). To verify (20.3.10), we can (and will) replace L by any full-rank self-dual lattice in V (by Lemma 20.2.1 again).

Let $V^\flat \subseteq V$ be the rank m subspace spanned by $\underline{x}_v$. Then $V^\flat$ is free of rank m. By Lemma 20.2.1, it is enough to show (20.3.10) holds for one choice of basis $\underline{x}_v$ for $V^\flat$. We choose $\underline{x}_v$ to be a basis for a full-rank lattice $L^\flat \subseteq V^\flat$ which is

$$\begin{cases} \text{self-dual} & \text{if } V^\flat \text{ contains a full-rank self-dual lattice} \\ \text{almost self-dual} & \text{else.} \end{cases} \tag{20.3.11}$$

Note that $V^\flat$ always contains a full-rank self-dual lattice if F_v/F_v^+ is split.

Case 1. Assume $L^\flat$ is self-dual. There exists a rank $n-m$ self-dual lattice $L^\# \subseteq V$ which is orthogonal to $L^\flat$. Form the rank n self-dual lattice $L = L^\flat \oplus L^\#$. Any isomorphism between self-dual sublattices of L lifts to an element of $K_v = U(L)$ (Lemma 20.3.3). This implies that $g_v \mapsto \varphi_v(g_v^{-1}\underline{x}_v)$ is the characteristic function of $G_{\underline{x}_v}(F_v^+)\backslash(G_{\underline{x}_v}(F_v^+)K_v)$.

We know that $K_v \cap G_{\underline{x}_v}(F_v^+)$ is the unique maximal open compact subgroup in $G_{\underline{x}_v}(F_v^+)$ (i.e. $U(L^{\#})$). We compute

$$\int_{G_{\underline{x}_v}(F_v^+)\backslash G(F_v^+)} \varphi_v(g_v^{-1}\underline{x}_v)\, dg_v = \mathrm{vol}(G_{\underline{x}_v}(F_v^+)\backslash(G_{\underline{x}_v}(F_v^+)K_v)) = \frac{\mathrm{vol}(K_v)}{\mathrm{vol}(K_v \cap G_{\underline{x}_v}(F_v^+))} = 1. \tag{20.3.12}$$

Since $T = (\underline{x}_v, \underline{x}_v)$ and $\underline{x}_v$ is a basis for the self-dual lattice $L^\flat$, we also know $W^*_{T,v}(s_0)_n = 1$ (see (15.5.7); note that $V^\flat$ containing self-dual lattice means that F_v/F_v^+ is unramified if m is odd).

Case 2 Assume that $L^\flat$ is almost self-dual and that F_v/F_v^+ is ramified. Then $n \geq 2$ and $m = n-1$. There is an orthogonal direct sum decomposition $L^\flat = L^{\flat\flat} \oplus L^{\flat\#}$, where $L^{\flat\flat}$ is self-dual of rank $m-1$ and $L^{\flat\#}$ is almost self-dual of rank 1. There exists a rank 2 self-dual lattice $L^{\#} \subseteq V$ which is orthogonal to $L^{\flat\flat}$. We can assume $L^{\flat\#} \subseteq L^{\#}$ (Lemma 20.3.4). Form the rank n self-dual lattice $L = L^{\flat\flat} \oplus L^{\#}$. Any isomorphism between rank $n-1$ almost self-dual sublattices in L lifts to an element of $K_v = U(L)$ (Lemma 20.3.3). This implies that $g_v \mapsto \varphi_v(g_v^{-1}\underline{x}_v)$ is the characteristic function of $G_{\underline{x}_v}(F_v^+)\backslash(G_{\underline{x}_v}(F_v^+)K_v)$.

We know that $K_v \cap G_{\underline{x}_v}(F_v^+) = U(L^{\#}) \cap G_{\underline{x}_v}(F_v^+)$ has index 2 inside the unique maximal open compact subgroup of $G_{\underline{x}_v}(F_v^+)$ (reduces immediately to the case $n = 2$, which is Lemma 20.3.5). We compute

$$\int_{G_{\underline{x}_v}(F_v^+)\backslash G(F_v^+)} \varphi_v(g_v^{-1}\underline{x}_v)\, dg_v = \mathrm{vol}(G_{\underline{x}_v}(F_v^+)\backslash(G_{\underline{x}_v}(F_v^+)K_v)) = \frac{\mathrm{vol}(K_v)}{\mathrm{vol}(K_v \cap G_{\underline{x}_v}(F_v^+))} = 2. \tag{20.3.13}$$

Since $T = (\underline{x}_v, \underline{x}_v)$ and since $\underline{x}_v$ is a basis for the almost self-dual lattice $L^\flat$, we also know $W^*_{T,v}(s_0)^\circ_n = 1$ (15.5.7).

Case 3 Assume that $L^\flat$ is almost self-dual and that F_v/F_v^+ is inert. This implies $n \geq 2$ and $m = n-1$. Arguing as in Case 2 (use the same notation; the first paragraph applies verbatim), again apply Lemma 20.3.3 and Lemma 20.3.5 to compute

$$\int_{G_{\underline{x}_v}(F_v^+)\backslash G(F_v^+)} \varphi_v(g_v^{-1}\underline{x}_v)\, dg_v = \mathrm{vol}(G_{\underline{x}_v}(F_v^+)\backslash(G_{\underline{x}_v}(F_v^+)K_v)) = \frac{\mathrm{vol}(K_v)}{\mathrm{vol}(K_v \cap G_{\underline{x}_v}(F_v^+))} = q_v + 1. \tag{20.3.14}$$

When $n = 2$, we have $\mathrm{Den}^*(X, L^\flat)_n = q_v X^{-1/2} + X^{1/2}$ (follows from the relevant Cho–Yamauchi type formula; see [LZ22a, Example 3.5.2] [FYZ24, Theorem 2.2]). The "cancellation" property for local densities and self-dual lattices (15.5.11) implies $\mathrm{Den}^*(X, L^\flat)_n = q_v X^{-1/2} + X^{1/2}$ for $n \geq 2$. We thus have $W^*_{T,v}(s_0)^\circ_n = \mathrm{Den}^*(1, L^\flat)_n = q_v + 1$. $\square$

20.4 Uniformization degrees for special cycles

The purpose of this section is to express the groupoid cardinality of (20.4.4) in terms of special values of local Whittaker functions, with explicit constants (Lemma 20.4.1). This groupoid has already appeared as a "uniformization degree" for special cycles (see (11.5.12), also Sections, 11.8, 11.9, and 12.4). This calculation will be needed for our main arithmetic Siegel–Weil results (Section 22.1).

Let F/F^+ be a CM extension of number fields, with respective adèle rings $\mathbb{A}_F$ and $\mathbb{A}$ and finite adèle rings $\mathbb{A}_{F,f}$ and $\mathbb{A}_f$, etc.. As in Part V, we write v for places of F^+ with completions F^+_v, and set $F_v \coloneqq F \otimes_{F^+} F^+_v$.

Let $T \in \mathrm{Herm}_m(F^+)$ be a Hermitian matrix (with F-coefficients) for any integer $m \geq 0$. Set $m^\flat \coloneqq \mathrm{rank}(T)$. For each place v, select any $a_v \in \mathrm{GL}_m(F_v)$ such that ${}^t\overline{a}_v^{-1} T a_v^{-1} = \mathrm{diag}(0, T^\flat_v)$ for some $T^\flat_v \in \mathrm{Herm}_{m^\flat}(F^+_v)$ with $\det T^\flat_v \neq 0$. For each v, choose any decomposition (Iwasawa decomposition)

$$a_v = \begin{pmatrix} 1_{m-m^\flat} & * \\ 0 & 1_{m^\flat} \end{pmatrix} \begin{pmatrix} a^\#_v & 0 \\ 0 & a^\flat_v \end{pmatrix} k_v \qquad k_v \in \begin{cases} \mathrm{GL}_m(\mathcal{O}_{F_v}) & \text{if } v \text{ is non-Archimedean} \\ U(m) & \text{if } v \text{ is Archimedean,} \end{cases} \tag{20.4.1}$$

where $a^\#_v \in \mathrm{GL}_{m-m^\flat}(F_v)$, $a^\flat_v \in \mathrm{GL}_{m^\flat}(F_v)$, and $U(m) \subseteq \mathrm{GL}_m(\mathbb{C})$ is the unitary group for the standard diagonal positive definite Hermitian pairing.

Let L be a non-degenerate Hermitian $\mathcal{O}_F$-lattice of any rank n, set $V \coloneqq L \otimes_{\mathcal{O}_F} F$, and let $G = U(V)$ be the associated unitary group. Set $s^\flat_0 \coloneqq (n - m^\flat)/2$. For any place v of F^+_v, we set $V_v \coloneqq V \otimes_{F^+} F^+_v$. Let $K_{L,f} = \prod K_{L,v} \subseteq U(V)(\mathbb{A}_f)$ be the adèlic stabilizer of L (i.e. $K_{L,v}$ is the stabilizer of $L_v \coloneqq L \otimes_{\mathcal{O}_{F^+}} \mathcal{O}_{F^+_v}$ for every place $v < \infty$ of F^+_v). Fix a place v_0 of F^+_v (Archimedean or non-Archimedean). Assume V_v is positive definite for every Archimedean $v \neq v_0$.

Given $\underline{x}^{v_0}_f \in (V \otimes_{F^+} \mathbb{A}^{v_0}_f)^m$, we define the "away from v_0 special cycle" (compare Sections 11.2 and 12.1)

$$\mathcal{Z}(\underline{x}^{v_0}_f) \coloneqq \{g_f \in G(\mathbb{A}^{v_0}_f)/K^{v_0}_{L,f} : g^{-1}_{f,v}\underline{x}_v \in L_v \text{ for all non-Archimedean } v \neq v_0\} \tag{20.4.2}$$

where $\underline{x}_v \in V^m_v$ is the v-component of $\underline{x}^{v_0}_f$.

Fix a nontrivial additive character ψ_v for each place v. Assume ψ_v is unramified if $v < \infty$, and assume $\psi_v(x) = e^{2\pi i x}$ if $F^+_v = \mathbb{R}$. For every non-Archimedean place $v \neq v_0$, set $\varphi_v \coloneqq \mathbf{1}^m_{L_v}$ (characteristic function of $L^m_v \subseteq V^m_v$) and set

$$\varphi^{v_0}_f = \otimes_{\substack{v<\infty \\ v \neq v_0}} \varphi_v \in \mathcal{S}(V(\mathbb{A}^{v_0}_f)^m). \tag{20.4.3}$$

Similarly set $\varphi_v^\flat := \mathbf{1}_{L_v}^{m^\flat} \in \mathcal{S}(V(F_v^+)^{m^\flat})$ for such v.

For every place v of F_v^+, let $\eta_v \colon F_v^{+\times} \to \{\pm 1\}$ be the quadratic character associated to F_v/F_v^+. Let $\chi_v \colon F_v^\times \to \mathbb{C}^\times$ be any character satisfying $\chi_v|_{F_v^{+\times}} = \eta_v^n$. Form the associated Siegel–Weil standard section $\Phi_{\varphi_v} \in I(s, \chi_v)$ (Section 14.2) for every place $v < \infty$ with $v \neq v_0$. To simplify slightly, we assume that 2-adic places of F^+ are unramified in F for the rest of Section 20.4.

For $v < \infty$ with $v \neq v_0$, the local Whittaker function variant $\tilde{W}^*_{T_v^\flat, v}(a_v^\flat, s, \Phi_{\varphi_v})_n$ does not depend on the choice of a_v or $a_v^\flat$. Indeed, the $\mathrm{GL}_m(\mathcal{O}_{F_v})$-equivalence class of the Hermitian matrix ${}^t\overline{a}_v^\flat T_v^\flat a_v^\flat$ does not depend on the choice of a_v (follows from the invariance properties in (15.3.4)). For $v \mid \infty$ with $v \neq v_0$, the local Whittaker function variant $\tilde{W}^*_{T_v^\flat, v}(a_v^\flat, s)_n^\circ$ similarly does not depend on the choice of a_v or $a_v^\flat$, as the $U(m)$-equivalence class of ${}^t\overline{a}_v^\flat T_v^\flat a_v^\flat$ is well-defined (then apply (15.2.4)).

Given any tuple $\underline{x} \in V^m$ which spans a non-degenerate Hermitian space, we write $G_{\underline{x}} \subseteq G$ for the stabilizer of $\underline{x}$ (i.e. the unitary group of the orthogonal complement $\mathrm{span}(\underline{x})^\perp \subseteq V$). We write $\underline{x}_f^{v_0}$ for the image of $\underline{x}$ in $(V \otimes_{F^+} \mathbb{A}_f^{v_0})^m$.

Suppose there exists $\underline{x} \in V^m$ with Gram matrix $(\underline{x}, \underline{x})$. Fix such an $\underline{x}$, and assume $\mathrm{span}(\underline{x})^\perp$ is definite at every Archimedean place. Let $K_{\underline{x}, v_0} \subseteq G_{\underline{x}}(F_{v_0}^+)$ be any open compact subgroup, and assume $K_{\underline{x}, v_0}(F_{v_0}^+) = G_{\underline{x}}(F_{v_0}^+)$ if v_0 is Archimedean.

We are mostly interested in applying Lemma 20.4.1 below when $m^\flat \geq n - 1$ and L_v is self-dual for all $v < \infty$ with $v \neq v_0$. The result and proof is simpler in that case, and the lemma may not be optimal otherwise.

Lemma 20.4.1. *Consider the groupoid quotient*

$$\left[G_{\underline{x}}(F^+) \backslash \left(G_{\underline{x}}(F_{v_0}^+)/K_{\underline{x}, v_0} \times \mathcal{Z}(\underline{x}_f^{v_0}) \right) \right]. \tag{20.4.4}$$

The displayed groupoid has finite automorphism groups and finitely many isomorphism classes. Its groupoid cardinality is

$$C \cdot \prod_{\substack{v \mid \infty \\ v \neq v_0}} \tilde{W}^*_{T_v^\flat, v}(a_v^\flat, s_0^\flat)_n^\circ \prod_{\substack{v < \infty \\ v \neq v_0}} \tilde{W}^*_{T_v^\flat, v}(a_v^\flat, s_0^\flat, \Phi_{\varphi_v^\flat})_n \tag{20.4.5}$$

for some volume constant $C \in \mathbb{Q}_{>0}$ which we describe in the following three situations.

(1) Suppose v_0 is Archimedean. Assume the local characters $(\psi_v)_v$ and $(\chi_v)_v$ arise from global characters $\psi \colon F^+ \backslash \mathbb{A} \to \mathbb{C}^\times$ and $F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$. The constant C may depend on V, n, $m^\flat$, F, and the isomorphism classes of the local Hermitian lattices $\{L_v\}_{v<\infty}$. The constant C does not otherwise depend on T or $V^\flat$ or $\underline{x}$.

(2) *Suppose $m^\flat = n$ (with v_0 not necessarily Archimedean). Then*

$$C = \prod_{\substack{v<\infty \\ v \neq v_0}} c_v \tag{20.4.6}$$

for some constants $c_v \in \mathbb{Q}_{>0}$, all but finitely many of which are 1. For any given $v < \infty$ with $v \neq v_0$, the constant c_v may depend on the local Hermitian lattice L_v and the quadratic extension F_v/F_v^+, but otherwise does not depend on T or V or $\underline{x}$ or v_0 or F/F^+.

If L_v is self-dual, then $c_v = 1$.

(3) *Suppose $m^\flat = n-1$ (with v_0 not necessarily Archimedean). Assume $K_{\underline{x},v_0} \subseteq G_{\underline{x}}(F_v^+)$ is the unique maximal open compact subgroup. Then there are constants $c'_v \in \mathbb{Q}_{>0}$ such that*

$$C = \frac{2^{1-o(\Delta)} h_F}{w_F h_{F^+} \cdot \#(\mathcal{O}_F^\times/(W\mathcal{O}_{F^+}^\times))} \prod_{\substack{v<\infty \\ v \neq v_0}} c'_v \tag{20.4.7}$$

where $o(\Delta)$ is the number of prime ideals of $\mathcal{O}_{F^+}$ which ramify in $\mathcal{O}_F$, where h_F (resp. h_{F^+}) is the class number of F (resp. F^+), where w_F (resp. W) is the number of (resp. group of) roots of unity in F. All but finitely many c'_v are equal to 1.

For each $v < \infty$ with $v \neq v_0$, the constant c'_v may depend on the local Hermitian lattice L_v, the quadratic extension F_v/F_v^+, and the local invariant $\varepsilon(V_v^\flat) \in \{\pm 1\}$. The constant c'_v does not otherwise depend on T or V or $V^\flat$ or $\underline{x}$ or v_0 or F/F^+.

If L_v is self-dual, then $c'_v = 1$ if F_v/F_v^+ is unramified (resp. $c'_v = 2$ if F_v/F_v^+ is ramified).

Proof. For the moment, we allow v_0 Archimedean or not. The groupoid in the lemma statement indeed has finite stabilizer groups, by discreteness of $G_{\underline{x}}(F^+)$. Take any factorizable open compact subgroup $K_{\underline{x}} = \prod_v K_{\underline{x},v} \subseteq G_{\underline{x}}(\mathbb{A})$. Assume $K_{\underline{x},v} = G_{\underline{x},v}(F_v^+)$ for every Archimedean v, and assume $K_{\underline{x},v} = K_{\underline{x},v_0}$ is the open compact subgroup fixed in the lemma statement when $v = v_0$. For each v, define $\underline{x}_v^\flat = [x_{1,v}^\flat, \ldots, x_{m^\flat,v}^\flat] \in V_v^{m^\flat}$ to be the tuple satisfying $\underline{x} \cdot a_v^{-1} = [0, \ldots, 0, x_{1,v}^\flat, \ldots, x_{m^\flat,v}^\flat]$ (so $T_v^\flat = (\underline{x}_v^\flat, \underline{x}_v^\flat)$).

We have $\tilde{W}^*_{T_v^\flat,v}(a_v^\flat, s_0)_n^\circ = 1$ for all Archimedean $v \neq v_0$ by positive definite-ness of $T_v^\flat$ (Section 15.2). For all but finitely many v, the Hermitian matrix ${}^t\overline{a}_v^\flat T_v^\flat a_v^\flat$ defines a self-dual $\mathcal{O}_{F_v}$-lattice (first check the case where the collection $(a_v)_v$ comes from a single element $a \in \mathrm{GL}_m(F)$; then recall that $\tilde{W}^*_{T_v^\flat,v}(a_v^\flat, s, \Phi_{\varphi_v})_n$ does not depend on the choice of a_v or $a_v^\flat$). For such non-Archimedean $v \neq v_0$, we have $\tilde{W}^*_{T_v^\flat,v}(a_v^\flat, s, \Phi_{\varphi_v^\flat})_n = \tilde{W}^*_{T_v^\flat,v}(a_v^\flat, s)_n^\circ = 1$ if L_v

is self-dual (see (15.5.7), Remark 15.5.1, and the invariance property in (15.3.4)). Hence $\tilde{W}^*_{T^\flat_v,v}(a^\flat_v, s, \Phi_{\varphi^\flat_v})_n = 1$ for all but finitely many v.

Choose Haar measures $dg_{x,v}$ on $G_{\underline{x}}(F^+_v)$ for each v. Assume that $\mathrm{vol}_{dg_{x,v}}(K_{\underline{x},v}) \in \mathbb{Q}$ for all v, that $\mathrm{vol}_{dg_{x,v}}(K_{\underline{x},v}) = 1$ for all but finitely many v, and that $\mathrm{vol}_{dg_{x,v}}(K_{\underline{x},v}) = 1$ if $v = v_0$ or if $v \mid \infty$.

For $v < \infty$ with $v \neq v_0$, we give $G(F^+_v)$ the unique Haar measure dg_v such that

$$W^*_{T^{\flat\prime}_v,v}(1, s^\flat_0, \Phi_{\varphi^\flat_v})_n = \int_{G_{\underline{x}_v}(F^+_v)\backslash G(F^+_v)} \varphi^\flat_v(g_v^{-1}\underline{x}'_v)\ dg_v \tag{20.4.8}$$

for any tuple $\underline{x}'_v \in V^m_v$ (temporary notation) with nonsingular Gram matrix $T^{\flat\prime}_v \coloneqq (\underline{x}'_v, \underline{x}'_v)$ (Lemma 20.2.1). The integral is taken with respect to the quotient measure induced by $dg_{x,v}$. This measure dg_v on $G(F^+_v)$ may depend on n, $m^\flat$, the isomorphism class of L_v (as the normalization defining $\tilde{W}^*_{T^\flat,v}$ depended on L_v) as well as the local invariant $\varepsilon(V^\flat_v)$ (Remark 20.2.2). The measure dg_v does not otherwise depend on $T^\flat_v$. Note $\mathrm{vol}_{dg_v}(K_{L,v}) \in \mathbb{Q}_{>0}$ for any $v < \infty$ with $v \neq v_0$, since the left-hand side of (20.4.8) lies in $\mathbb{Q}$ by Lemma 15.4.2. We have $\mathrm{vol}_{dg_v}(K_{L,v}) = 1$ for all but finitely many v (cf. the proof of Lemma 20.3.6; we have $W^*_{T^\flat_v,v}(s^\flat_0)^\circ_n = 1$ for all but finitely many v). We equip $G(\mathbb{A}^{v_0}_f)$ with the product measure $dg = \prod_{\substack{v<\infty \\ v\neq v_0}} dg_v$.

Using the Haar measures specified above, we may unfold the groupoid cardinality as

$$\deg\left[G_{\underline{x}}(F^+)\backslash\left(G_{\underline{x}}(F^+_{v_0})/K_{\underline{x},v_0} \times \mathcal{Z}(\underline{x}^{v_0}_f)\right)\right] \tag{20.4.9}$$

$$= \mathrm{vol}_{dg}(K^{v_0}_{L,f})^{-1} \int_{G_{\underline{x}}(F^+)\backslash((\prod_{\substack{v=v_0 \\ \text{or } v\mid\infty}} G_{\underline{x}}(F^+_v))\times G(\mathbb{A}^{v_0}_f))} \varphi^{v_0}_f(g^{-1}\underline{x})\ dg \tag{20.4.10}$$

$$= \mathrm{vol}(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A}))\mathrm{vol}_{dg}(K^{v_0}_{L,f})^{-1} \left(\int_{G_{\underline{x}}(\mathbb{A}^{v_0}_f)\backslash G(\mathbb{A}^{v_0}_f)} \varphi^{v_0}_f(g^{-1}\underline{x})\ dg\right) \tag{20.4.11}$$

$$= \mathrm{vol}(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A}))\mathrm{vol}_{dg}(K^{v_0}_{L,f})^{-1} \prod_{\substack{v<\infty \\ v\neq v_0}} \int_{G_{\underline{x}}(F^+_v)\backslash G(F^+_v)} \varphi^\flat_v(g_v^{-1}\underline{x}^\flat_v a^\flat_v)\ dg_v \tag{20.4.12}$$

$$= C \prod_{\substack{v<\infty \\ v\neq v_0}} \tilde{W}^*_{T^\flat_v,v}(a^\flat_v, s^\flat_0, \Phi_{\varphi^\flat_v})_n \tag{20.4.13}$$

with

$$C \coloneqq \mathrm{vol}(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A})) \prod_{\substack{v<\infty \\ v\neq v_0}} \mathrm{vol}_{dg_v}(K_{L,v})^{-1}. \tag{20.4.14}$$

Note that the integrals are absolutely convergent, since the integrands are continuous and compactly supported. This unfolding also shows that the groupoid in (20.4.4) has finitely many isomorphism classes.

(1) Suppose v_0 is Archimedean. Recall that the Tamagawa number of any nontrivial unitary group is 2 [Ich04, Section 4]. After scaling one of the non-Archimedean local measures $dg_{x,v}$ by an element of $\mathbb{Q}_{>0}$, we may assume $\prod_v dg_{x,v}$ is the Tamagawa measure on $G_{\underline{x}}(\mathbb{A})$. If $v \mid \infty$, let dg_v be the Haar measure on $G(F_v^+)$ given by Lemma 20.2.1 (induced by $dg_{x,v}$). For $v \mid \infty$, the local invariant $\varepsilon(V_v^\flat)$ is already determined by V and the requirement that $V_v^{\flat\perp}$ is definite. Hence the measures dg_v for $v \mid \infty$ do not depend on $V^\flat$ (apply Remark 20.2.2).

   By construction of the measures in Lemma 20.2.1 (via invariant differentials), we find that $\prod_v dg_v$ equals the Tamagawa measure on $G(\mathbb{A})$ up to scaling by a constant which may depend on the lattices $\{L_v\}_{v<\infty}$ as well as n and $m^\flat$ (coming from our normalization of local Whittaker functions $\widetilde{W}^*_{T^\flat,v}$, Section 15.3). We conclude that the measure dg on $G(\mathbb{A}_f)$ may depend on V, n, $m^\flat$, F, and the lattices $\{L_v\}_{v<\infty}$, but it does not otherwise depend on T or $V^\flat$ or $\underline{x}$.

(2) Suppose $m^\flat = n$. Then $G_{\underline{x}}$ is the trivial group. Take $\mathrm{vol}_{dg_{x,v}}(K_{\underline{x},v}) = 1$ for all v. Consider $v < \infty$ with $v \neq v_0$ and set $c_v = \mathrm{vol}_{dg_v}(K_{L,v})^{-1}$. If L_v is self-dual, then $c_v = 1$ by Lemma 20.3.6. In general, dg_v may depend on L_v (but not on T or $T_v^\flat$).

(3) Suppose $m^\flat = n - 1$. Then $G_{\underline{x}}$ is isomorphic to the norm-one torus inside $\mathrm{Res}_{F/F^+} \mathbb{G}_m$. Assume $K_{\underline{x},v} \subseteq G_{\underline{x}}(F_v^+)$ is the unique maximal open compact subgroup for every v. Take $\mathrm{vol}_{dg_{x,v}}(K_{\underline{x},v}) = 1$ for all v. Consider $v < \infty$ with $v \neq v_0$ and set $c'_v = \mathrm{vol}_{dg_v}(K_{L,v})^{-1}$. If L_v is self-dual, then $c'_v = 1$ if F_v/F_v^+ is unramified (resp. $c'_v = 2$ if F_v/F_v^+ is ramified) by Lemma 20.3.6. In general, dg_v may depend on L_v, $m^\flat$ and the local invariant $\varepsilon(V_v^\flat)$ (but not on T or $T_v^\flat$).

   We have

$$\mathrm{vol}(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A})) = \deg[G_{\underline{x}}(F^+)\backslash (G_{\underline{x}}(\mathbb{A})/K_{\underline{x}})] = \frac{\deg(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A})/K_{\underline{x}})}{w_F}$$

   where $\deg[-]$ denotes groupoid cardinality and $\deg(-)$ denotes set cardinality. We have

$$\deg(G_{\underline{x}}(F^+)\backslash G_{\underline{x}}(\mathbb{A})/K_{\underline{x}}) = 2^{u-t} h_F h_{F^+}^{-1}, \tag{20.4.15}$$

   where t is the number of prime ideals of F^+ which ramify in F, and where $u \in \mathbb{Z}$ is such that $H^1(\mathrm{Gal}(F/F^+), \mathcal{O}_F^\times) \cong (\mathbb{Z}/2\mathbb{Z})^u$ [Ono85, (9)]. A group cohomology computation (omitted) shows that $2^{-u} = \#(\mathcal{O}_F^\times/(W\mathcal{O}_{F^+}^\times))/2$ (where # also means cardinality). □

21 Geometric Siegel–Weil

For our main results (arithmetic Siegel–Weil), we will need a special value formula for degrees of 0-cycles in the generic fiber (Section 21.1). The result on complex volumes (Section 21.2) will not be needed, but may be of independent interest. Let $F/\mathbb{Q}$ be an imaginary quadratic field, with accompanying notation as in Part I. We also write h_F (resp. w_F) for the class number (resp. cardinality $|\mathcal{O}_F^\times|$).

21.1 Degrees of 0-cycles

Let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of signature $(n-1,1)$ (not assuming n is even). Let $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ be the associated moduli stack (Sections 3.1 and 3.2). Recall that $d_L \in \mathbb{Z}$ is a certain integer associated to L, with $d_L = 1$ if L is self-dual when $2 \nmid \Delta$. Let $V \coloneqq L \otimes_{\mathcal{O}_F} F$ be the associated $F/\mathbb{Q}$ Hermitian space.

Consider an integer m with $m = n$ or $m = n-1$. Pick any embedding $F \to \mathbb{C}$, and set $\mathcal{M}_{\mathbb{C}} \coloneqq \mathcal{M} \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$, etc.. Given $T \in \operatorname{Herm}_m(\mathbb{Q})$ with $\operatorname{rank} T = n-1$, recall that there is an associated special cycle $\mathcal{Z}(T) \to \mathcal{M}$. The base change $\mathcal{Z}(T)_{\mathbb{C}}$ is smooth, proper, and quasi-finite (and of dimension zero) over $\operatorname{Spec} \mathbb{C}$ (Lemma 3.5.5, also Lemma 4.7.4).

For each place v of $\mathbb{Q}$, select any $a_v \in \mathrm{GL}_m(F_v)$ such that ${}^t\overline{a}_v^{-1} T a_v^{-1} = \operatorname{diag}(0, T_v^\flat)$ for some $T_v^\flat \in \operatorname{Herm}_{n-1}(F_v^+)$ with $\det T_v^\flat \neq 0$. Choose any $a_v^\flat \in \mathrm{GL}_{n-1}(F_v)$ associated to a_v via the Iwasawa decomposition, as in (20.4.1) (if $m = n-1$, we can just take $a_v^\flat = a_v$).

For formation of local Whittaker functions, we use the standard additive character $\psi \colon \mathbb{Q} \backslash \mathbb{A} \to \mathbb{C}^\times$ with $\psi_\infty(x) = e^{2\pi i x}$. Suppose $\chi \coloneqq F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$ is a character satisfying $\chi|_{\mathbb{A}^\times} = \eta^n$, where η is the quadratic character associated to $F/\mathbb{Q}$. For each prime p, we let $\varphi_v^\flat = \mathbf{1}_{L_p}^{n-1} \in \mathcal{S}(V(\mathbb{Q}_p)^{n-1})$ where $\mathbf{1}_{L_p}$ is the characteristic function of the lattice $L_p \subseteq V(\mathbb{Q}_p)$.

Proposition 21.1.1. *Let $C \in \mathbb{Q}_{>0}$ be the volume constant from Lemma 20.4.1(3), for the Hermitian space V and with $v_0 = \infty$ in the notation of loc. cit.. In the situation above, we have*

$$\deg \mathcal{Z}(T)_{\mathbb{C}} = \frac{h_F}{w_F} C \cdot \tilde{W}^*_{T^\flat_\infty, \infty}(a^\flat_\infty, 1/2)^\circ_n \prod_p \tilde{W}^*_{T^\flat_p, p}(a^\flat_p, 1/2, \Phi_{\varphi^\flat_v})_n. \tag{21.1.1}$$

Proof. As in Section 20.1, we write $\Omega_T(R) \coloneqq \{\underline{x} \in (V \otimes_{\mathbb{Q}} R)^m : (\underline{x}, \underline{x}) = T\}$ for $\mathbb{Q}$-algebras R. Here $\deg \mathcal{Z}(T)_{\mathbb{C}}$ denotes the (stacky) degree of $\mathcal{Z}(T)_{\mathbb{C}}$ over $\operatorname{Spec} \mathbb{C}$, as explained at the end of Appendix A.1.

Suppose there is no tuple $\underline{x} \in V^m$ such that $(\underline{x}, \underline{x}) = T$. By the Hasse principle, we conclude $\Omega_T(\mathbb{Q}_{v_0}) = \emptyset$ for some place v_0 of $\mathbb{Q}$. Since $\operatorname{rank}(T) < n$, we must have $v_0 = \infty$ (i.e.

for $F_v < \infty$, any non-degenerate Hermitian F_v vector space of rank $n-1$ embeds into any non-degenerate Hermitian F_v vector space of rank n). We conclude that $T^{\flat}_{\infty}$ (and ${}^t\overline{a}^{\flat}_{\infty} T^{\flat}_{\infty} a^{\flat}_{\infty}$) has signature $(n-1-r, r)$ for some $r \geq 2$. The proposition holds in this case because $\tilde{W}^*_{T^{\flat}_{\infty},\infty}(a^{\flat}_{\infty}, 1/2)^{\circ}_n = 0$ (by (15.2.6) or (20.2.2)).

Suppose there exists $\underline{x} \in V^m$ such that $(\underline{x}, \underline{x}) = T$. For such $\underline{x}$, write $\underline{x}_{\infty} \in V^m_{\mathbb{R}}$ and $\underline{x}_f \in (V \otimes_{\mathbb{Q}} \mathbb{A}_f)^m$ for the respective images. By complex uniformization of special cycles (Section 11.5), we have

$$\deg \mathcal{Z}(T)_{\mathbb{C}} = \frac{h_F}{w_F} \cdot \deg \mathcal{D}(\underline{x}_{\infty}) \cdot \deg \left[U(V)(\mathbb{Q}) \backslash \coprod_{\substack{\underline{x} \in V^m \\ (\underline{x},\underline{x}) = T}} \mathcal{D}(\underline{x}_f) \right]. \tag{21.1.2}$$

Here $\deg \mathcal{D}(\underline{x}_{\infty})$ is the degree of the Archimedean local special cycle $\mathcal{D}(\underline{x}_{\infty}) \subseteq \mathcal{D}$ (Section 8.2) for any $\underline{x} \in V^m$ with $(\underline{x}, \underline{x}) = T$. We know $\mathcal{D}(\underline{x}_{\infty})$ is a single point if T is positive semidefinite, and empty otherwise. Hence we have $\deg \mathcal{D}(\underline{x}_{\infty}) = \tilde{W}^*_{T^{\flat}_{\infty},\infty}(a^{\flat}_{\infty}, 1/2)^{\circ}_n$ (by (15.2.6), the right-hand side is 1 if $T^{\flat}_{\infty}$ is positive definite and 0 otherwise).

We then use Lemma 20.4.1 to evaluate the groupoid cardinality in (21.1.2). $\square$

Remark 21.1.2. Suppose $2 \nmid \Delta$ and that L is self-dual (for the trace pairing, as is our running convention). We then have $C = 2h_F/w_F$ in Proposition 21.1.1. Take any $a \in \mathrm{GL}_m(F)$ such that ${}^t\overline{a}^{-1} T a^{-1} = \mathrm{diag}(0, T^{\flat})$ where $T^{\flat} \in \mathrm{Herm}_{n-1}(\mathbb{Q})$ with $\det T^{\flat} \neq 0$. For each place v of $\mathbb{Q}$, let $a_v \coloneqq a \in \mathrm{GL}_m(F_v)$. Set $a^{\flat} = (a^{\flat}_v)_v \in \mathrm{GL}_m(\mathbb{A}_F)$ (running over places v of $\mathbb{Q}$) in the notation above. The proposition then states

$$\deg \mathcal{Z}(T)_{\mathbb{C}} = 2\frac{h_F^2}{w_F^2} \cdot \tilde{E}^*_{T^{\flat}}(a^{\flat}, 1/2)^{\circ}_n. \tag{21.1.3}$$

Remark 21.1.3. As observed by Li and Zhang [LZ22a, Remark 4.6.2], Proposition 21.1.1 may be proved using Rapoport–Zink non-Archimedean uniformization in essentially the same way. Indeed, the horizontal local special cycle $\mathcal{Z}(T)_{\mathscr{H}} \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ is proper, quasi-finite, and flat (Lemma 4.7.4), so we may calculate its degree in the fiber over any geometric point of $\operatorname{Spec} \mathcal{O}_F[1/d_L]$. Fix a geometric point in characteristic $p > 0$. Assume $p \neq 2$ if 2 is nonsplit in $\mathcal{O}_F$, assume L_p is self-dual, and assume either $p \nmid \Delta$ or that L is self-dual and $2 \nmid \Delta$. Consider the n-dimensional positive definite non-degenerate Hermitian space $\mathbf{V}$ with $\varepsilon(\mathbf{V}_p) = -1$ and $\varepsilon(\mathbf{V}_{\ell}) = \varepsilon(V_{\ell})$ for any $\ell \neq p$.

Using non-Archimedean uniformization, we may then argue as in the proof of Proposition 21.1.1 (see (11.9.6)), using the special value formula for degrees of local special cycles (Lemma 18.1.3), and the formula for uniformization degrees (Lemma 20.4.1) for $\mathbf{V}$ and $v_0 = p$.

21.2 Complex volumes

Assume 2 is unramified in $\mathcal{O}_F$. For even integers $n \in \mathbb{Z}_{>0}$, we show that the global normalizing factors $\Lambda_n(s)_n^\circ$ (Section 17.1) encode complex volumes of certain unitary Shimura varieties (Propositions 21.2.1 and 21.2.3).

First consider $n \equiv 0 \pmod 4$. Let V be the unique $F/\mathbb{Q}$ Hermitian space of signature $(n, 0)$ which satisfies $\varepsilon(V_p) = 1$ for all primes p (with ε as in Section 2.2). Set $G := U(V)$, let $L \subseteq V$ be a full-rank self-dual lattice, and write $K_{L,f} \subseteq G(\mathbb{A}_f)$ for the adèlic stabilizer of L. The following proposition should be a special case of a unitary analogue of the classical Siegel mass formula. It is included for comparison with the analogous volume identity for a signature $(n-1, 1)$ unitary complex Shimura variety. The left-hand side counts self-dual positive definite $\mathcal{O}_F$-lattices of rank n, weighted by the inverses of the sizes of their automorphism groups.

Proposition 21.2.1. *We have*

$$\#[G(\mathbb{Q})\backslash(G(\mathbb{A}_f)/K_{L,f})] = 2\Lambda_n(0)_n^\circ \tag{21.2.1}$$

where the left-hand side denotes groupoid cardinality.

Proof. Let $\psi\colon \mathbb{Q}\backslash\mathbb{A} \to \mathbb{C}^\times$ be the standard additive character with $\psi_\infty(s) = e^{2\pi i x}$. Let $\chi\colon \mathbb{A}_F^\times \to \mathbb{C}^\times$ be the trivial character.

For $v = \infty$, let $\varphi_v(\underline{x}) = e^{2\pi i \operatorname{tr}(\underline{x}, \underline{x})} \in \mathcal{S}(V(\mathbb{R})^n)$ and let $T \in \operatorname{Herm}_n(\mathbb{R})$ be an arbitrary positive definite matrix. For $v < \infty$ corresponding to a prime p, let $\varphi_v = \mathbf{1}_{L_v}^n \in \mathcal{S}(V(\mathbb{Q}_p)^n)$ and let T be the Gram matrix for any basis of L_v. For such T, we have $W^*_{T,v}(s_0)_n^\circ = 1$ for all v (Sections 15.2 and 15.5.7). Recall $W^*_{T,v}(s)_n^\circ = \Lambda_{T,v}(s)_n^\circ W_{T,v}(s, \Phi_{\varphi_v})$ if $v < \infty$ (resp. $W^*_{T,v}(s)_n^\circ e^{-2\pi \operatorname{tr}(T)} = \Lambda_{T,v}(s)_n^\circ W_{T,v}(s, \Phi_{\varphi_v})$ if $v = \infty$); see Section 14.2.

Using these data, the local Siegel–Weil formula and its proof (Lemma 20.2.1) for each place v of $\mathbb{Q}$ shows that $\operatorname{vol}(G(\mathbb{R}) \times K_{L,f})^{-1} = \Lambda_n(0)_n^\circ$ for the Tamagawa measure on $G(\mathbb{A})$. Since G has Tamagawa number 2 [Ich04, §4], the proposition follows. $\square$

Next, consider $n \equiv 2 \pmod 4$. Let V be the unique n-dimensional $F/\mathbb{Q}$ Hermitian space of signature $(n-1, 1)$ which satisfies $\varepsilon(V_p) = 1$ for all primes p. Again, set $G := U(V)$, let $L \subseteq V$ be a full-rank self-dual lattice, and write $K_{L,f} \subseteq G(\mathbb{A}_f)$ for the adèlic stabilizer of L. For sufficiently small open compact $K_f \subseteq G(\mathbb{A}_f)$, there is a complex (analytic) Shimura variety

$$\operatorname{Sh}_{K_f,\mathbb{C}} = G(\mathbb{A})\backslash(\mathcal{D} \times G(\mathbb{A}_f)/K_f) \tag{21.2.2}$$

of dimension $n-1$, where $\mathcal{D}$ is the Hermitian symmetric domain from Section 8.1 (the V of loc. cit. is our $V_\mathbb{R}$, with $\mathbb{C} = F \otimes_\mathbb{Q} \mathbb{R}$-action). The metrized tautological bundle $\widehat{\mathcal{E}}^\vee$ of loc. cit.

descends to $\mathrm{Sh}_{K_f,\mathbb{C}}$. For any open compact $K'_f \subseteq G(\mathbb{A}_f)$ and any sufficiently small $K_f \subseteq K'_f$, we set

$$\mathrm{vol}(\mathrm{Sh}_{K_f,\mathbb{C}}) := \int_{\mathrm{Sh}_{K_f,\mathbb{C}}} c_1(\widehat{\mathcal{E}})^{n-1} \qquad \mathrm{vol}(\mathrm{Sh}_{K'_f,\mathbb{C}}) := \frac{1}{[K'_f : K_f]} \mathrm{vol}(\mathrm{Sh}_{K_f,\mathbb{C}}). \tag{21.2.3}$$

If $K_{L',f} \subseteq G(\mathbb{A}_f)$ is the adèlic stabilizer of a full-rank lattice $L' \subseteq V$ which is self-dual for the Hermitian pairing, the quantity $\mathrm{vol}(\mathrm{Sh}_{K'_{L,f},\mathbb{C}})$ was computed explicitly in [BH21, Theorem A]. We show that the level $K_{L,f}$ (self-dual for the *trace pairing*) removes the additional factors at ramified primes in loc. cit., and that the resulting complex volume agrees with $2\Lambda_n(0)^\circ_n$ exactly.

The volume identity should also follow from [LL21, Footnote 11] (or possibly other geometric Siegel–Weil results). We instead compute $\mathrm{vol}(\mathrm{Sh}_{K_{L,f},\mathbb{C}})$ using [BH21, Theorem A] by calculating the "change of level" via the following lemma.

Lemma 21.2.2. *Let E_v^+ be a non-Archimedean local field of odd residue cardinality q_v, and let E_v/E_v^+ be a ramified quadratic extension with involution $a \mapsto a^\sigma$.*

Let W be a rank $2d$ non-degenerate E_v/E_v^+ Hermitian space, and assume W contains a full-rank lattice $M \subseteq W$ which is self-dual (for the trace pairing). Let $M' \subseteq W$ be any full-rank lattice which is self-dual for the Hermitian pairing.

If $K, K' \subseteq U(W)$ are the stabilizers of M and M' respectively, we have

$$\frac{\mathrm{vol}(K)}{\mathrm{vol}(K')} = 2^{-1}(1 + q_v^d) \tag{21.2.4}$$

for any Haar measure on $U(W)$.

Proof. We know that any two full-rank lattices in W which are self-dual (resp. self-dual for the Hermitian form) are isomorphic [Jac62, Proposition 8.1] (false if E_v^+ is allowed to have residue characteristic 2). Hence $\mathrm{vol}(K)/\mathrm{vol}(K')$ does not depend on the choice of M and M' (nor the choice of Haar measure).

Let ϖ be a uniformizer of E_v, and assume $\varpi^\sigma = -\varpi$. The lattices M and M' admit bases with Gram matrices

$$\begin{pmatrix} 0 & \varpi^{-1} \\ -\varpi^{-1} & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{21.2.5}$$

respectively. Choose a basis $e_1, \ldots, e_{2d}$ for M with Gram matrix as above. We may assume that M' is the lattice with basis $e_1, \ldots, e_d, \varpi e_{d+1}, \ldots, \varpi e_{2d}$. Let $\overline{W}$ (resp. $\overline{W}'$) be the $2d$-dimensional vector space over $\mathbb{F}_{q_v}$ with symplectic pairing (resp. bilinear pairing) given by the block matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{resp.} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{21.2.6}$$

If $P_W \subseteq \mathrm{Sp}(\overline{W})$ and $P_{\overline{W}'} \subseteq \mathrm{O}(\overline{W})$ are the subgroups upper triangular matrices (in $d \times d$ blocks), we have

$$\#(K/(K \cap K')) = \#(\mathrm{Sp}(\overline{W})(\mathbb{F}_{q_v})/P_{\overline{W}}(\mathbb{F}_{q_v})) \tag{21.2.7}$$
$$\#(K'/(K \cap K')) = \#(\mathrm{O}(\overline{W}')(\mathbb{F}_{q_v})/P_{\overline{W}'}(\mathbb{F}_{q_v})). \tag{21.2.8}$$

The lemma now follows from the formulas

$$\#\mathrm{Sp}(\overline{W})(\mathbb{F}_{q_v}) = q_v^{d^2} \prod_{i=1}^{d} (q_v^{2i} - 1) \qquad \#\mathrm{O}(\overline{W}')(\mathbb{F}_{q_v}) = 2q_v^{d(d-1)}(q_v^d + 1)^{-1} \prod_{i=1}^{d} (q_v^{2i} - 1)$$
$$\#P_{\overline{W}}(\mathbb{F}_{q_v}) = q_v^{d(d+1)/2} \prod_{i=1}^{d} (q_v^d - q_v^{i-1}) \qquad \#P_{\overline{W}'}(\mathbb{F}_{q_v}) = q_v^{d(d-1)/2} \prod_{i=1}^{d} (q_v^d - q_v^{i-1}). \qquad \square$$

We return to the global situation with $F/\mathbb{Q}$ as above and $L \subseteq V$ a self-dual lattice.

Proposition 21.2.3. *We have*

$$\mathrm{vol}(\mathrm{Sh}_{K_{L,f},\mathbb{C}}) = 2\Lambda_n(0)_n^{\circ}. \tag{21.2.9}$$

Proof. If $K_{L',f} \subseteq G(\mathbb{A}_f)$ is the adèlic stabilizer of a full-rank lattice $L' \subseteq V$ which is self-dual for the Hermitian pairing, the result [BH21, Theorem A] (see also [BH21, Theorem 5.5.1] to compare $c_1(\widehat{\mathcal{E}})$ with the Chern form of the metrized Hodge bundle; note our $\widehat{\mathcal{E}}$ is $\widehat{\mathcal{L}}$ in loc. cit. (up to restricting)) gives

$$\mathrm{vol}(\mathrm{Sh}_{K_{L',f},\mathbb{C}}) = \left[2^{1-o(\Delta)} \prod_{\ell \mid \Delta} (1 + \varepsilon(V_\ell)\ell^{-n/2}) \prod_{j=1}^{n} \frac{\Delta^{j/2}\Gamma(s+j)L(2s+j,\eta^j)}{2^j \pi^{s+j}} \right]_{s=0} \tag{21.2.10}$$

where $o(\Delta)$ is the number primes dividing Δ. We assumed $\varepsilon(V_\ell) = 1$ for all ℓ, and a direct computation shows

$$\Lambda_n(s)_n^{\circ} = \Delta^{n/2(s-1)} \prod_{j=1}^{n} \frac{\Delta^{j/2}\Gamma(s+j)L(2s+j,\eta^j)}{2^j \pi^{s+j}} \tag{21.2.11}$$

(using even-ness of n). The claim now follows from the computation of $\mathrm{vol}(K_{L,f})/\mathrm{vol}(K_{L',f})$ (for any Haar measure on $G(\mathbb{A}_f)$) from Lemma 21.2.2. Note that the only discrepancy between $\mathrm{vol}(K_{L,f})$ and $\mathrm{vol}(K_{L',f})$ is at ramified primes, since self-dual lattices for the Hermitian pairing are the same as self-dual lattices at unramified primes. $\square$

22 Arithmetic Siegel–Weil

As above, we write $F/\mathbb{Q}$ for an imaginary quadratic field with discriminant Δ, nontrivial involution $a \mapsto a^\sigma$, associated quadratic character $\eta \colon \mathbb{Q}^\times \backslash \mathbb{A}^\times \to \mathbb{C}^\times$, class number h_F, and $w_F \coloneqq |\mathcal{O}_F^\times|$. We fix the standard nontrivial additive character $\psi \colon \mathbb{Q}\backslash\mathbb{A} \to \mathbb{C}^\times$ with $\psi_\infty(x) = e^{2\pi i x}$ for the Archimedean place ∞. We allow $2 \mid \Delta$ unless otherwise specified.

22.1 Main theorems

This section contains the statements and proofs of our main global results (Theorem 22.1.1 and the secondary Theorem 22.1.6). Theorem 22.1.1 relies on essentially all preceding results in this work (except for the computations in Section 19.2 and Section 21.2). In the proof, we explain how to combine our local main results (proved in Part VI) and a (new) diagonalization argument to deal with singular T (including those which are not-necessarily $\mathrm{GL}_n(\mathcal{O}_F)$-conjugate to a block diagonal matrix with nonsingular diagonal blocks).

Assume $2 \nmid \Delta$, and let L be any non-degenerate self-dual Hermitian $\mathcal{O}_F$-lattice of signature $(n-1,1)$. Set $n \coloneqq \operatorname{rank} L$, and note $n \equiv 2 \pmod 4$ (by the global product formula for local invariants of Hermitian spaces; note $\varepsilon(L_p) = 1$ for all primes p).

Form the associated (smooth) moduli stack $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F$ (Section 3.2). We are imposing "no level structure" on $\mathcal{M}$ (i.e. $K_{0,f} \times K_f = K_{L_0,f} \times K_{L,f}$ in the notation of Section 3.4).

For any m, given $T \in \mathrm{Herm}_m(\mathbb{Q})$ (with F-coefficients), and given $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$ (with $\mathbb{C}$-coefficients), recall that there is an arithmetic special cycle class $[\widehat{\mathcal{Z}}(T)] \in \widehat{\mathrm{Ch}}^m(\mathcal{M})_{\mathbb{Q}}$ (Section 4.4) and a normalized T-th Fourier coefficient $E^*_T(y,s)^\circ_n$ (Section 17.1) of a $U(m,m)$ Eisenstein series. If $\operatorname{rank}(T) \geq n-1$ or if T is nonsingular and not positive definite, we are using the current $g_{T,y}$ from Section 12.4. The class $[\widehat{\mathcal{Z}}(T)]$ thus implicitly depends on y.

For special cycles $\mathcal{Z}(T)$ which are proper over $\operatorname{Spec} \mathcal{O}_F$, recall that we have defined certain arithmetic degrees without boundary contributions (4.7.1). These are the arithmetic degrees appearing in our main theorem below.

For use below, we record the expression

$$\frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} = -\frac{1}{2} L(2s+1,\eta)\Gamma(s+1)|\Delta|^{s+1/2}\pi^{-s-1} \tag{22.1.1}$$

which follows from our formula for the normalizing factor $\Lambda_m(s)^\circ_n$ (17.1.2). We thus have

$$\frac{\Lambda_n(0)^\circ_n}{\Lambda_{n-1}(1/2)^\circ_n} = -\frac{h_F}{w_F} \qquad \frac{d}{ds}\bigg|_{s=0} \left(\frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n}\right) = 2\frac{h_F}{w_F} h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^\vee} \tag{22.1.2}$$

where the left expression follows from the analytic class number formula, and $h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^\vee}$ is the height constant from (4.3.6).

Theorem 22.1.1 (Co-rank 1 arithmetic Siegel–Weil)**.** *Assume the prime* 2 *splits in* $\mathcal{O}_F$*.*

(1) For any $T \in \mathrm{Herm}_n(\mathbb{Q})$ *with* $\mathrm{rank}(T) = n-1$ *and any* $y \in \mathrm{Herm}_n(\mathbb{R})_{>0}$*, we have*

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T)]) = \frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n. \tag{22.1.3}$$

(2) For any $T^\flat \in \mathrm{Herm}_{n-1}(\mathbb{Q})$ *with* $\det T^\flat \neq 0$ *and any* $y^\flat \in \mathrm{Herm}_{n-1}(\mathbb{R})_{>0}$*, we have*

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T^\flat)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)) = 2\frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} \left(\frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} E^*_{T^\flat}(y^\flat, s+1/2)^\circ_n \right). \tag{22.1.4}$$

Proof. In the theorem statement, $[\widehat{\mathcal{Z}}(T)]$ and $[\widehat{\mathcal{Z}}(T^\flat)]$ are implicitly formed with respect to y and $y^\flat$, respectively. Note that $E^*_T(y,s)^\circ_n$ is a normalized Fourier coefficient for a $U(n,n)$ Eisenstein series, while $E^*_{T^\flat}(y^\flat,s)^\circ_n$ is a normalized Fourier coefficient for a $U(n-1,n-1)$ Eisenstein series. In the theorem statement, note that $\mathcal{Z}(T) \to \operatorname{Spec}\mathcal{O}_F$ and $\mathcal{Z}(T^\flat) \to \operatorname{Spec}\mathcal{O}_F$ are both proper (Lemma 4.7.5), so we may use (4.7.1) to define arithmetic degrees without boundary contributions.

Note that part (2) is the special case of part (1) when $T = \mathrm{diag}(0, T^\flat)$ and $y = \mathrm{diag}(1, y^\flat)$. This follows from the unfolding of Fourier coefficients in Corollary 17.2.2 (also the functional equation in Lemma 17.1.1) and from the definition of arithmetic degrees in (4.7.1).

Fix T and y as in the statement of part (1) (not necessarily block diagonal). Fix any prime p. It is enough to show that (22.1.3) holds modulo $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$ (i.e. as elements of the additive quotient $\mathbb{R}/(\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell)$), where the sum runs over primes $\ell \neq p$. Varying the prime p removes this discrepancy (giving an equality as elements of $\mathbb{R}$) because the real numbers $\log \ell$ (ranging over all primes ℓ in $\mathbb{Z}$) form a $\mathbb{Q}$-linearly independent set.

(*Step 1: Diagonalize*) For convenience, we fix an embedding $F \to \mathbb{C}$. Pick any $b \in \mathrm{GL}_m(F)$ such that ${}^t\overline{b}^{-1} T b^{-1} = \mathrm{diag}(0, T^\flat)$ for some $T^\flat \in \mathrm{Herm}_{n-1}(\mathbb{Q})$ with $\det T^\flat \neq 0$. We may (and do) assume $b \in \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)})$ as well. The proof below will show that the theorem holds modulo $\mathbb{Q} \cdot \log \ell$ for primes ℓ such that $b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})$.

For each place v of $\mathbb{Q}$, select any $b^\#_v \in \mathrm{GL}_1(F_v)$ and $b^\flat_v \in \mathrm{GL}_{n-1}(F_v)$ associated to an Iwasawa decomposition of $b_v \in \mathrm{GL}_n(F_v)$, as in (20.4.1) (where b_v denotes the image of b). Also consider the (unique) decomposition

$$b y {}^t\overline{b} = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^\# & 0 \\ 0 & y^\flat \end{pmatrix} \begin{pmatrix} 1 & 0 \\ {}^t\overline{c} & 1 \end{pmatrix} \tag{22.1.5}$$

as in (12.4.4), where $c \in M_{1,n-1}(\mathbb{C})$, $y^{\#} \in \mathbb{R}_{>0}$, and $y^{\flat} \in \mathrm{Herm}_{n-1}(\mathbb{R})_{>0}$. Pick any $a^{\#}_{\infty} \in \mathrm{GL}_1(\mathbb{C})$ and $a^{\flat}_{\infty} \in \mathrm{GL}_{n-1}(\mathbb{C})$ such that $a^{\#}_{\infty} {}^t\overline{a}^{\#}_{\infty} = y^{\#}$ and $a^{\flat}_{\infty} {}^t\overline{a}^{\flat}_{\infty} = y^{\flat}$.

Let $a^{\#} \in \mathrm{GL}_1(\mathbb{A}_F)$ be the element with component $a^{\#}_v \coloneqq b^{\#}_v$ for places $v < \infty$ and $a^{\#}_v \coloneqq a^{\#}_{\infty}$ for the place $v = \infty$. Similarly define $a^{\flat} \in \mathrm{GL}_{n-1}(\mathbb{A}_F)$, and set $a \coloneqq \mathrm{diag}(a^{\#}, a^{\flat}) \in \mathrm{GL}_n(\mathbb{A}_F)$.

By unfolding for co-rank 1 Fourier coefficients (Corollary 17.2.2) and Fourier coefficient invariance properties (see (13.3.3), (13.3.4), (15.2.4), and (15.3.4) for $U(m)$ invariance when $v \mid \infty$ and $\mathrm{GL}_m(\mathcal{O}_{F^+_v})$ invariance when $v < \infty$), we find

$$\begin{aligned}
E^*_T(y,s)^\circ_n = \chi_\infty(d)^{-1} \det(y)^{-n/2} E^*_{T'}(m(a),s)^\circ_n = \tilde{E}^*_{T'}(a,s)^\circ_n \\
= |a^{\#}|^s_F \frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} \tilde{E}^*_{T^\flat}(a^\flat, s+1/2)^\circ_n \\
- |a^{\#}|^{-s}_F \frac{\Lambda_n(-s)^\circ_n}{\Lambda_{n-1}(-s+1/2)^\circ_n} \tilde{E}^*_{T^\flat}(a^\flat, s-1/2)^\circ_n
\end{aligned}$$

where $T' \coloneqq \mathrm{diag}(0, T^\flat)$. We remind the reader that the notation $E^*_T(-,s)^\circ_n$ is overloaded (Section 17.1, also end of Section 13.2) and has slightly different meaning when "$-$" is $y \in \mathrm{Herm}_m(\mathbb{R})_{>0}$ versus $h \in U(m,m)(\mathbb{A})$ (e.g. $h = m(a)$).

(*Step 2: Leibniz rule*) Since $n \equiv 2 \pmod 4$, the functional equation for $\tilde{E}^*_{T^\flat}(a^\flat, s)^\circ_n$ (Lemma 17.1.1) implies

$$\begin{aligned}
&\frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n \\
&= 2 \frac{d}{ds}\bigg|_{s=0} \left(|a^{\#}|^s_F \frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} \tilde{E}^*_{T^\flat}(a^\flat, s+1/2)^\circ_n \right).
\end{aligned} \tag{22.1.6}$$

Since $\det T^\flat \neq 0$, we may factorize $\tilde{E}^*_{T^\flat}(a^\flat, s+1/2)^\circ_n$ into a product of (variants of) normalized local Whittaker functions (17.1.8). Also recall the formulas in (22.1.2). We have $|a^{\#}_\ell|_\ell = 1$ (ℓ-adic norm of $a^{\#}_\ell$) for any prime ℓ such that $b \in \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})$ (by construction, this includes $\ell = p$). By the Leibniz rule, we thus find

$$\left(\frac{2h_F}{w_F}\right)^{-1} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n \tag{22.1.7}$$

$$= 2h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^\vee} \tilde{E}^*_{T^\flat}(a^\flat, 1/2)^\circ_n \tag{22.1.8}$$

$$- \left(\frac{d}{ds}\bigg|_{s=1/2} \left(|a^\#_\infty|^s_\infty \tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty, s)^\circ_n\right)\right) \prod_\ell \tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, 1/2)^\circ_n \tag{22.1.9}$$

$$- \left(\frac{d}{ds}\bigg|_{s=1/2} \tilde{W}^*_{T^\flat,p}(a^\flat_p, s)^\circ_n\right) \prod_{v \neq p} \tilde{W}^*_{T^\flat,v}(a^\flat_v, 1/2)^\circ_n \tag{22.1.10}$$

$$- \sum_{\ell \neq p} \left(\frac{d}{ds}\bigg|_{s=1/2} |a^\#_\ell|^s_\ell \tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, s)^\circ_n\right) \prod_{v \neq \ell} \tilde{W}^*_{T^\flat,v}(a^\flat_v, 1/2)^\circ_n. \tag{22.1.11}$$

The product in (22.1.9) runs over all primes ℓ (not including the Archimedean place ∞). The products in (22.1.10) and (22.1.11) run over all places v of $\mathbb{Q}$ (with $v \neq p$ or $v \neq \ell$ as indicated), including $v = \infty$. The sum in (22.1.11) runs over all primes $\ell \neq p$. We remind the reader that $|a^\#_\infty|_\infty = \overline{a}^\#_\infty a^\#_\infty \in \mathbb{R}_{>0}$, by definition.

For all but finitely many primes ℓ, the Hermitian matrix ${}^t\overline{a}^\flat_\ell T^\flat a^\flat_\ell \in \mathrm{Herm}_{n-1}(\mathbb{Q}_\ell)$ defines a (non-degenerate) self-dual Hermitian $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$-lattice. For such ℓ, we have $\tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, s)^\circ_n$ identically equal to 1 (as a function in the s-variable). This follows from (15.5.7) and an invariance property for local Whittaker functions (15.3.4). In particular, the sums and products are finite in the right-hand side of (22.1.7).

For every prime ℓ, we have $\tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, s+1/2)^\circ_n \in \mathbb{Z}[\ell^{-1}, \ell^{-s}, \ell^s]$ (see (15.5.6), and again the invariance property in (15.3.4)). We also have $\tilde{W}^*_{T^\flat,v}(a^\flat_v, 1/2)^\circ_n \in \mathbb{Q}$ for all place v of $\mathbb{Q}$ (if $v \mid \infty$, this quantity is 1 if $T^\flat$ is positive definite and 0 otherwise by (15.2.6)). The quantity in (22.1.10) thus lies in $\mathbb{Q} \cdot \log p$, and the quantity in (22.1.11) thus lies in $\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell$.

As we explain below, every quantity on the right-hand side of (22.1.7) has geometric meaning via our main local results, at least modulo $\mathbb{Q} \cdot \log \ell$ for primes ℓ such that $b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})$.

(*Step 3a: Local geometric interpretation: complex degree*) Set $\mathcal{Z}(T)_{\mathbb{C}} = (\mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C})$ for the embedding $F \to \mathbb{C}$ fixed above. We have $\deg \mathcal{Z}(T)_{\mathbb{C}} = (\deg_F \mathcal{Z}(T) \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} F) = 2 \deg_{\mathbb{Q}}(\mathcal{Z}(T) \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} \mathbb{Q}) =: \deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}$. Here $\deg_F$ and $\deg_{\mathbb{Q}}$ denote stacky degrees over $\operatorname{Spec} F$ and $\operatorname{Spec} \mathbb{Q}$, respectively, as defined at the end of Appendix A.1.

By the geometric Siegel–Weil formula for Kudla–Rapoport 0-cycles over $\mathbb{C}$ (Proposition 21.1.1, also Remark 21.1.2), we conclude

$$\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}} = 2 \deg \mathcal{Z}(T)_{\mathbb{C}} = \frac{4h_F^2}{w_F^2} \tilde{E}^*_{T^\flat}(a^\flat, 1/2)^\circ_n. \tag{22.1.12}$$

This gives a geometric interpretation of (22.1.8).

(*Step 3b: Local geometric interpretation: at* ∞) We claim that

$$\mathrm{Int}_\infty(T,y) = -\frac{d}{ds}\bigg|_{s=1/2} \left(|a^{\#}_\infty|^s_\infty \tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty, s)^\circ_n\right) \quad \mathrm{mod} \sum_{\substack{\ell \text{ such that} \\ b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})}} \mathbb{Q}\cdot \log \ell \qquad (22.1.13)$$

where $\mathrm{Int}_\infty(T,y)$ is the geometric quantity defined in (12.4.15).

Indeed, (12.4.8) implies

$$\mathrm{Int}_\infty(T,y) = \qquad (22.1.14)$$

$$\begin{cases} \mathrm{Int}_\infty(T^\flat, a^\flat_\infty {}^t\overline{a}^\flat_\infty) - \log(|a^{\#}_\infty|_\infty) \quad \mathrm{mod} \sum_{\substack{\ell \text{ such that} \\ b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})}} \mathbb{Q}\cdot\log\ell & \text{if } T^\flat > 0 \\ \mathrm{Int}_\infty(T^\flat, a^\flat_\infty {}^t\overline{a}^\flat_\infty) \quad \mathrm{mod} \sum_{\substack{\ell \text{ such that} \\ b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_{(\ell)})}} \mathbb{Q}\cdot\log\ell & \text{if } T^\flat \not> 0. \end{cases} \qquad (22.1.15)$$

The notation $T^\flat > 0$ (resp. $T^\flat \not> 0$) means that $T^\flat$ is positive definite (resp. not positive definite). We have $\mathrm{Int}_\infty(T^\flat, a^\flat_\infty {}^t\overline{a}^\flat_\infty) = \mathrm{Int}_\infty({}^t\overline{a}^\flat_\infty T^\flat a^\flat_\infty, 1)$ (12.4.3). By our main Archimedean local identity (Theorem 19.1.1), we have $\mathrm{Int}_\infty({}^t\overline{a}^\flat_\infty T^\flat a^\flat_\infty, 1) = \frac{d}{ds}\big|_{s=-1/2} W^*_{{}^t\overline{a}^\flat_\infty T a^\flat_\infty,\infty}(s)^\circ_n$.

The Whittaker function invariance property (15.2.4) implies $W^*_{{}^t\overline{a}^\flat_\infty T^\flat a^\flat_\infty}(s)^\circ_n = \tilde{W}^*_{T^\flat}(a^\flat_\infty, s)^\circ_n$. By the Archimedean local functional equation (16.2.1) we have $\frac{d}{ds}\big|_{s=-1/2}\tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty,s)^\circ_n = -\frac{d}{ds}\big|_{s=1/2}\tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty,s)^\circ_n$. This is still true when $T^\flat$ has signature $(n-1-r,r)$ for $r \geq 2$, as both sides are zero in this case (by definition for the geometric side, and by (19.1.5) for the local Whittaker function). As already mentioned, recall that $\tilde{W}^*_{T^\flat}(a^\flat, 1/2)^\circ_n$ is 1 if $T^\flat$ is positive definite, and is 0 is $T^\flat$ is not positive definite (15.2.6). Now (22.1.13) follows from what we have just discussed.

Next, recall the global Archimedean intersection number $\mathrm{Int}_{\infty,\mathrm{global}}(T,y) = \int_{\mathcal{M}_{\mathbb{C}}} g_{T,y}$ (where $g_{T,y}$ is a current associated with T and y) as in (12.4.13). Recall the relation (12.4.14)

$$\mathrm{Int}_{\infty,\mathrm{global}}(T,y) = \frac{h_F}{w_F}\mathrm{Int}_\infty(T,y)\cdot \deg\left[U(V)(\mathbb{Q})\backslash \coprod_{\substack{\underline{x}\in V^n \\ (\underline{x},\underline{x})=T}} \mathcal{D}(\underline{x}_f)\right] \qquad (22.1.16)$$

where $V := L \otimes_{\mathcal{O}_F} F$ and $\mathcal{D}(\underline{x}_f)$ is a certain "away-from-∞" local special cycle (it is a discrete set), defined in Section 12.1. The displayed groupoid cardinality $\deg[\cdots]$ describes certain "complex uniformization degrees" (Section 12.4.13). If there exists $\underline{x} \in V^n$ with $(\underline{x},\underline{x}) = T$, the groupoid cardinality is

$$\deg\left[U(V)(\mathbb{Q})\backslash \coprod_{\substack{\underline{x}\in V^n \\ (\underline{x},\underline{x})=T}} \mathcal{D}(\underline{x}_f)\right] = \frac{2h_F}{w_F}\prod_{\ell}\tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, 1/2)^\circ_n \qquad (22.1.17)$$

by local Siegel–Weil as in Lemma 20.4.1 (with $v_0 = \infty$ in the notation of loc. cit.). If there does not exist such $\underline{x}$, then the Hasse principle implies that $T^\flat$ has signature $(n-1-r, r)$ for some $r \geq 2$ (compare the proof of Proposition 21.1.1). In this case, we have $\frac{d}{ds}\big|_{s=1/2} \tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty, s)^\circ_n = 0$ (19.1.5). In all cases, we thus have

$$\mathrm{Int}_{\infty,\mathrm{global}}(T, y) = -\frac{2h_F^2}{w_F^2} \left(\frac{d}{ds}\bigg|_{s=1/2} \left(|a^\#_\infty|^s_\infty \tilde{W}^*_{T^\flat,\infty}(a^\flat_\infty, s)^\circ_n \right) \right) \prod_\ell \tilde{W}^*_{T^\flat,\ell}(a^\flat_\ell, 1/2)^\circ_n \quad (22.1.18)$$

modulo $\sum_\ell \mathbb{Q} \cdot \log \ell$ for primes ℓ such that $b \notin \mathrm{GL}_n(\mathcal{O}_F \otimes_\mathbb{Z} \mathbb{Z}_{(\ell)})$. This give a geometric interpretation of (22.1.9).

(*Step 3c: Local geometric interpretation: at p*) Recall $\mathrm{Int}_p(T) \coloneqq \mathrm{Int}_{\mathscr{H},p}(T) + \mathrm{Int}_{\mathscr{V},p}(T)$ (11.9.9), where $\mathrm{Int}_{\mathscr{H},p}(T)$ is a "horizontal local intersection number" (11.9.1) and $\mathrm{Int}_{\mathscr{V},p}(T)$ is a "vertical local intersection number" (11.8.1) associated with T. The former describes "local change of tautological (or Faltings) height" and the latter describes degrees for "components in positive characteristic" in terms of local special cycles on Rapoport–Zink spaces.

We claim that

$$\mathrm{Int}_p(T) = -e_p \frac{d}{ds}\bigg|_{s=1/2} \tilde{W}^*_{T^\flat,p}(a^\flat_p, s)^\circ_n \quad (22.1.19)$$

where $e_p = 1$ if p is unramified (resp. $e_p = 2$ if p is ramified).

First note that the functional equation (16.1.4) implies $-\frac{d}{ds}\big|_{s=1/2} \tilde{W}^*_{T^\flat,p}(a^\flat_p, s)^\circ_n = \frac{d}{ds}\big|_{s=-1/2} \tilde{W}^*_{T^\flat,p}(a^\flat_p, s)^\circ_n$. The invariance property for Whittaker functions (15.3.4) implies $\tilde{W}^*_{T^\flat,p}(a^\flat_p, s)^\circ_n = \tilde{W}^*_{{}^t\overline{a}^\flat_p T^\flat a^\flat_p, p}(s)^\circ_n$.

Form the positive definite $F/\mathbb{Q}$ Hermitian spaces $\mathbf{W} \subseteq \mathbf{V}$ as in Section 11 (recall $\varepsilon(\mathbf{V}_p) = -1$ and $\varepsilon(\mathbf{V}_\ell) = \varepsilon(V_\ell)$ for all $\ell \neq p$). Set $\mathcal{O}_{F,p} \coloneqq \mathcal{O}_F \otimes_\mathbb{Z} \mathbb{Z}_p$. For any $\underline{\mathbf{x}}_p \in \mathbf{W}_p^n$ with Gram matrix T (such $\underline{\mathbf{x}}_p$ exists because $\mathrm{rank}(T) \leq n-1$; recall $\mathbf{W}$ has rank n if p is nonsplit and rank $n-1$ if p is split), there exists a basis of $L^\flat_p \coloneqq \mathrm{span}_{\mathcal{O}_{F,p}}(\underline{x}_p)$ with Gram matrix ${}^t\overline{a}^\flat_p T^\flat a^\flat_p$. Indeed, we have $a_p \in \mathrm{GL}_n(\mathcal{O}_{F_p})$ and $a^\flat_p \in \mathrm{GL}_{n-1}(\mathcal{O}_{F_p})$ by construction (and recall ${}^t\overline{a}_p^{-1} T a_p^{-1} = \mathrm{diag}(0, T^\flat)$ by definition). We remind the reader that (15.5.6) may be used to pass between (normalized) local densities and local Whittaker functions. We also pass between the notation $\mathrm{Den}^*(X, L^\flat_p)_n = \mathrm{Den}^*(X, {}^t\overline{a}^\flat_p T^\flat a^\flat_p)_n$ as explained in Section 15.5. Now (22.1.19) follows from our main non-Archimedean local identity (Theorem 18.1.2).

Next, recall the horizontal and vertical global intersection numbers $\mathrm{Int}_{\mathscr{H},p,\mathrm{global}}(T)$ and $\mathrm{Int}_{\mathscr{V},p,\mathrm{global}}(T)$ at p, associated with T (see (11.9.7) and (11.8.3)). These are elements of $\mathbb{Q} \cdot \log p$. Recall the $F/\mathbb{Q}$ Hermitian space $\mathbf{W}^\perp$ defined in Section 11.3, which satisfies $\mathbf{V} = \mathbf{W} \oplus \mathbf{W}^\perp$ (orthogonal direct sum). In particular, $\mathbf{W}^\perp = 0$ if p is nonsplit and $\dim_F \mathbf{W}^\perp = 1$ if p is split.

By (11.9.7) and (11.8.3) (and in the notation of loc. cit.), we have

$$\mathrm{Int}_{p,\mathrm{global}}(T) = \frac{h_F}{w_F}\mathrm{Int}_p(T) \cdot \deg\left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^n \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^{\perp}) / K_{1,\mathbf{L}_p^{\perp}} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right]. \tag{22.1.20}$$

The notation $\mathcal{Z}(\underline{\mathbf{x}}^p)$ means a certain "away-from-p" local special cycle (a discrete set), defined in Section 11.2. Recall that $K_{1,\mathbf{L}_p^{\perp}} \subseteq U(\mathbf{W}_p^{\perp})$ is the unique maximal open compact subgroup and $I_1 = U(\mathbf{W}) \times U(\mathbf{W}^{\perp})$ as algebraic groups over $\mathbb{Q}$ (Section 11.5). The displayed groupoid cardinality $\deg[\cdots]$ encodes certain "Rapoport–Zink non-Archimedean uniformization degrees".

If there exists $\underline{\mathbf{x}} \in \mathbf{W}^n$ with Gram matrix T, then local Siegel–Weil (Lemma 20.4.1) implies

$$\deg\left[I_1(\mathbb{Q}) \backslash \left(\coprod_{\substack{\underline{\mathbf{x}} \in \mathbf{W}^n \\ (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T}} U(\mathbf{W}_p^{\perp}) / K_{1,\mathbf{L}_p^{\perp}} \times \mathcal{Z}(\underline{\mathbf{x}}^p) \right) \right] = \frac{2h_F}{e_p w_F} \prod_{v \neq p} \tilde{W}^*_{T^{\flat},v}(a_v^{\flat}, 1/2)^{\circ}_n. \tag{22.1.21}$$

(in the notation of Lemma 20.4.1, take $v_0 = p$ and use the hermitian space $\mathbf{V}$ for the V in loc. cit.).

Set $\Omega_T(R) \coloneqq \{\underline{\mathbf{x}} \in (\mathbf{W} \otimes_{\mathbb{Q}} R)^n : (\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T\}$ for $\mathbb{Q}$-algebras R. If $\Omega_T(\mathbb{Q}) = \emptyset$, then the Hasse principle implies $\Omega_T(\mathbb{Q}_v) = \emptyset$ for some place v of $\mathbb{Q}$. We have $\Omega_T(\mathbb{Q}_p) \neq \emptyset$ (either p is nonsplit and $\mathbf{W} = \mathbf{V}$ and the claim follows because $\operatorname{rank} T < \operatorname{rank} \mathbf{W}$ (compare the proof of Proposition 21.1.1), or p is split and $\Omega_T(\mathbb{Q}_p) \neq \emptyset$ automatically). For all places v, we have $\Omega_T(\mathbb{Q}_v) = \emptyset$ if and only if $\Omega_{{}^t\overline{a}_v^{\flat} T^{\flat} a_v^{\flat}}(\mathbb{Q}_v) = \emptyset$ (where $\Omega_{{}^t\overline{a}_v^{\flat} T^{\flat} a_v^{\flat}}$ is defined like Ω_T but for $(n-1)$-tuples); this follows from our diagonalization of T (e.g. ${}^t\overline{a}_v^{-1} T a_v^{-1} = \operatorname{diag}(0, T^{\flat})$ for all $v < \infty$).

If $\Omega_T(\mathbb{Q}_v) = \emptyset$, we thus conclude $\tilde{W}^*_{T^{\flat},v}(a_v^{\flat}, 1/2)^{\circ}_n = \tilde{W}^*_{{}^t\overline{a}_v^{\flat} T^{\flat} a_v^{\flat},v}(1/2)^{\circ}_n = 0$ by the invariance property for local Whittaker functions (see (15.2.4) and (15.3.4)) and by local Siegel–Weil (20.2.2). Hence (22.1.21) holds even if there is no $\underline{\mathbf{x}} \in \mathbf{W}^n$ such that $(\underline{\mathbf{x}}, \underline{\mathbf{x}}) = T$ (both sides are 0 in this case).

We have shown

$$\mathrm{Int}_{p,\mathrm{global}}(T) = -\frac{2h_F^2}{w_F^2} \left(\frac{d}{ds}\bigg|_{s=1/2} \tilde{W}^*_{T^{\flat},p}(a_p^{\flat}, s)^{\circ}_n \right) \prod_{v \neq p} \tilde{W}^*_{T^{\flat},v}(a_v^{\flat}, 1/2)^{\circ}_n. \tag{22.1.22}$$

This gives a geometric interpretation for (22.1.10).

(*Step 4: Finish*) Recall the definition of arithmetic degree without boundary contributions

$\widehat{\deg}([\widehat{\mathcal{Z}}(T)])$ (4.7.1). In our current situation, this is

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T)]) \coloneqq \left(\int_{\mathcal{M}_{\mathbb{C}}} g_{T,y}\right) + \widehat{\deg}(\widehat{\mathcal{E}}^{\vee}|_{\mathcal{Z}(T)_{\mathscr{H}}}) + \sum_{\ell} \deg_{\mathbb{F}_\ell}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},\ell}) \log \ell$$

where the sum runs over all primes ℓ. By definition, we have

$$\begin{gathered} \int_{\mathcal{M}_{\mathbb{C}}} g_{T,y} = \mathrm{Int}_{\infty,\mathrm{global}}(T,y) \qquad \deg_{\mathbb{F}_\ell}({}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V},\ell}) \log \ell = \mathrm{Int}_{\mathscr{V},\ell,\mathrm{global}}(T) \\ \widehat{\deg}(\widehat{\mathcal{E}}^{\vee}|_{\mathcal{Z}(T)_{\mathscr{H}}}) = (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^{\vee}} + \sum_{\ell} \mathrm{Int}_{\mathscr{H},\ell,\mathrm{global}}(T) \end{gathered} \tag{22.1.23}$$

where $h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^{\vee}}$ is the height constant from (4.3.6). See (12.4.13) (Archimedean), (11.8.3) (vertical), and (11.9.8) (horizontal). For all primes ℓ, we have $\mathrm{Int}_{\mathscr{V},\ell,\mathrm{global}}(T) \in \mathbb{Q} \cdot \log \ell$ and $\mathrm{Int}_{\mathscr{H},\ell,\mathrm{global}}(T) \in \mathbb{Q} \cdot \log \ell$. These quantities are 0 for all but finitely many ℓ.

After multiplying both sides of (22.1.7) by $2(h_F/w_F)^2$, we apply the results of Steps 3a, 3b, and 3c above (see (22.1.12), (22.1.18), and (22.1.22)) to find

$$\frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^{\circ}_n = \widehat{\deg}([\widehat{\mathcal{Z}}(T)]) \tag{22.1.24}$$

as elements of $\mathbb{R}/(\sum_{\ell \neq p} \mathbb{Q} \cdot \log \ell)$. As we already discussed, varying p shows that this identity holds as an equality of real numbers. □

Remark 22.1.2 (Nonsingular arithmetic Siegel–Weil)**.** In the setup above (in particular, $n \equiv 2 \pmod 4$), consider any $T \in \mathrm{Herm}_n(\mathbb{Q})$ with $\det T \neq 0$ and any $y \in \mathrm{Herm}_n(\mathbb{R})_{>0}$. Assuming the prime 2 is split in $\mathcal{O}_F$, we still have

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T)]) = \frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^{\circ}_n \tag{22.1.25}$$

where the Green current for $[\widehat{\mathcal{Z}}(T)]$ is formed with respect to y, and where $\widehat{\deg}([\widehat{\mathcal{Z}}(T)])$ again denotes the arithmetic degree without boundary contributions as in (4.7.1). This should be compared with our preceding main theorem for singular T of co-rank 1 (Theorem 22.1.1).

Using the local theorems of Liu, Li–Zhang, and Li–Liu (cited below), one can prove (22.1.25) by a local decomposition as in the proof of Theorem 22.1.1 (no diagonalization procedure is necessary here) using the volume constant calculated in Lemma 21.1.1. This is possibly considered known to experts up to a volume constant by the cited local theorems. Nevertheless, the global statement is not available in the literature, so we have stated it. A sketch is provided below.

Decomposing $E^*_T(y,s)^\circ_n$ into a product of local Whittaker functions (Section 17.1), we find

$$\frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n = \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(y,s)^\circ_n\right)\prod_\ell W^*_{T,\ell}(0)^\circ_n \tag{22.1.26}$$

$$+\sum_p \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_n\right) W^*_{T,\infty}(y,0)^\circ_n \prod_{\ell\neq p} W^*_{T,\ell}(0)^\circ_n \tag{22.1.27}$$

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T)]) = \mathrm{Int}_{\infty,\mathrm{global}}(T,y) + \sum_p \mathrm{Int}_{p,\mathrm{global}}(T). \tag{22.1.28}$$

At most one of the summands is nonzero (see below), and all but finitely many $W^*_{T,\ell}(s)^\circ_n$ are identically equal to 1 as functions of s. In contrast with our main theorem, these intersection numbers $\mathrm{Int}_{p,\mathrm{global}}(T)$ are "purely vertical", without a mixed characteristic contribution.

In this setup, the local Archimedean theorem [Liu11, Theorem 4.1.7] (restated in our notation in Theorem 19.1.1) and the local Kudla–Rapoport theorems [LZ22a, Theorem 1.2.1] (inert) and [LL22, Theorem 2.7] (ramified, exotic smooth, even n) take the place of our main local identities (which were for co-rank 1 singular T). In combination with local Siegel–Weil with explicit constants (Lemma 20.4.1(1)), the cited local theorems imply

$$\mathrm{Int}_{\infty,\mathrm{global}}(T,y) = \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,\infty}(y,s)^\circ_n\right)\prod_p W^*_{T,p}(s)^\circ_n \tag{22.1.29}$$

$$\mathrm{Int}_{p,\mathrm{global}}(T) = \left(\frac{d}{ds}\bigg|_{s=0} W^*_{T,p}(s)^\circ_n\right) W^*_{T,\infty}(y,0)^\circ_n \prod_{\ell\neq p} W^*_{T,\ell}(0)^\circ_n \tag{22.1.30}$$

in our notation (end of Sections 12.4 and 11.8 respectively).

To apply local Siegel–Weil in the preceding discussion, we have in mind a (presumably routine) Hasse principle argument (compare [KR14, §9]). We briefly sketch this argument in our setup. For any prime p, set $\varepsilon_p(T) \coloneqq \eta_p((-1)^{n(n-1)/2}\det T)$ (the usual local invariant from Section 2.2), where $\eta_p \colon \mathbb{Q}_p^\times \to \{\pm 1\}$ is the local quadratic character associated to $F/\mathbb{Q}$.

We have $\mathrm{Int}_{\infty,\mathrm{global}}(T,y) = 0$ unless T has signature $(n-1,1)$ and $\varepsilon_p(T) = 1$ for all p. For such T, the special cycle $\mathcal{Z}(T)$ is empty (but may have a nontrivial Green current). We have $\mathrm{Int}_{p,\mathrm{global}}(T) = 0$ unless T is positive definite, $\varepsilon_p(T) = -1$, and $\varepsilon_\ell(T) = 1$ for all primes $\ell \neq p$. For such T, the special cycle $\mathcal{Z}(T)$ is supported in characteristic p (or empty). For all other T, the special cycle $\mathcal{Z}(T)$ is empty with Green current 0. These claims follow from e.g. uniformization of special cycles (e.g. Sections 12.4 (Archimedean) and 11.8 (non-Archimedean)) and the Hasse principle (e.g. applied to $\mathbf{V}$ from loc. cit. in the non-Archimedean case). In particular, $\mathrm{Int}_{p,\mathrm{global}}(T) = 0$ if p is split in $\mathcal{O}_F$, and $\mathcal{Z}(T)$ is empty over any split p.

On the analytic side, we have $W^*_{T,p}(0)^\circ_n = 0$ if $\varepsilon_p(T) = -1$ (by local Siegel–Weil (20.2.2), or the functional equation (16.1.4)) and $W^*_{T,\infty}(y,0)^\circ_n = 0$ if T is not positive definite (local Siegel–Weil again, or (15.2.6)). If T has signature $(n-r, r)$ for $r \geq 2$, we have $\frac{d}{ds}\big|_{s=0} W^*_{T,\infty}(y,s)^\circ_n = 0$ (19.1.5).

For the analogous global result (still $\det T \neq 0$ and $T \in \mathrm{Herm}_n$, central derivative) for an unramified CM extension of number fields F/F^+ where all 2-adic places are split (forcing $F^+ \neq \mathbb{Q}$) and a lattice L which is self-dual for the Hermitian pairing, see [LZ22a, Theorem 15.5.1] (at least up to a volume constant). For the analogous global result (still $\det T \neq 0$ and $T \in \mathrm{Herm}_n$, central derivative) for possibly ramified F/F^+ where all 2-adic places are split, on Krämer integral models (semistable reduction at ramified primes), and again L self-dual for the Hermitian pairing, see [HLSY23, Theorem 10.1] (at least up to a volume constant). For the result on Krämer models, one needs to correct the Eisenstein series derivative by special values of other Eisenstein series.

Remark 22.1.3. When $n \equiv 0 \pmod 4$, there is no non-degenerate self-dual signature $(n-1,1)$ Hermitian $\mathcal{O}_F$-lattice. In this case, Theorem 22.1.1(1) still holds in the sense that $\frac{d}{ds}\big|_{s=0} E^*_T(y,s)^\circ_n = 0$ (by the functional equation, Lemma 17.1.1).

Remark 22.1.4. We explain how Theorem 22.1.1 may be reformulated in terms of Faltings heights. Assume 2 is split in $\mathcal{O}_F$. Let $\widehat{\omega}$ be the metrized Hodge bundle on $\mathcal{M}$ as defined in Section 4.3. Take $T \in \mathrm{Herm}_n(\mathbb{Q})$ with $\mathrm{rank}(T) = n-1$. By (11.9.10), we have

$$\widehat{\deg}(\widehat{\omega}|_{\mathcal{Z}(T)_{\mathscr{H}}}) = (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot n \cdot h^{\mathrm{CM}}_{\mathrm{Fal}} - 2 \sum_p \mathrm{Int}_{\mathscr{H},p,\mathrm{global}}(T) \tag{22.1.31}$$

where $h^{\mathrm{CM}}_{\mathrm{Fal}}$ is the Faltings height of any elliptic curve with CM by $\mathcal{O}_F$ (as in (4.3.5)). By definition of Faltings height, we have

$$\widehat{\deg}(\widehat{\omega}|_{\mathcal{Z}(T)_{\mathscr{H}}}) = 2 \sum_{\alpha' \in \mathcal{Z}(T)(\mathbb{C})} |\operatorname{Aut}(\alpha')|^{-1} h_{\mathrm{Fal}}(A) \tag{22.1.32}$$

where $\alpha' = (A_0, \iota_0, \lambda_0, A, \iota, \lambda) \in \mathcal{Z}(T)(\mathbb{C})$ (choose $F \to \mathbb{C}$), and where $h_{\mathrm{Fal}}(A)$ is the Faltings height of A (as in Section 9.1) after descent to any number field, with metric normalized as in (4.3.1). Alternatively, we could consider morphisms $\operatorname{Spec} \mathbb{C} \to \mathcal{M}$ over $\operatorname{Spec} \mathbb{Z}$, which would remove the factor of 2 in the previous formula.

Our main theorem (Theorem 22.1.1) admits the equivalent formulation

$$\begin{aligned}&\mathrm{Int}_{\infty,\mathrm{global}}(T,y) - \frac{1}{2}\widehat{\deg}(\widehat{\omega}|_{\mathcal{Z}(T)_{\mathscr{H}}}) + (\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}) \cdot (h^{\mathrm{CM}}_{\widehat{\mathcal{E}}^\vee} + \frac{n}{2} \cdot h^{\mathrm{CM}}_{\mathrm{Fal}}) + \sum_p \mathrm{Int}_{\mathscr{V},p,\mathrm{global}}(T) \\ &= \frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=0} E^*_T(y,s)^\circ_n \end{aligned} \tag{22.1.33}$$

via the decomposition in (22.1.23). We remind the reader that $\deg_{\mathbb{Z}} \mathcal{Z}(T)_{\mathscr{H}}$ is essentially a special value of a $U(n-1, n-1)$ Eisenstein series (22.1.12). For further discussion of the special case $n = 2$, see Section 22.2.

In the rest of Section 22.1, we discuss some results which are applicable even if L is not self-dual.

Allow possibly $2 \mid \Delta$, and let L be any non-degenerate Hermitian $\mathcal{O}_F$-lattice of signature $(n-1, 1)$ (with n not necessarily even). Select any character $\chi \colon F^\times \backslash \mathbb{A}_F^\times \to \mathbb{C}^\times$ such that $\chi|_{\mathbb{A}^\times} = \eta^n$, where η is the quadratic character associated with $F/\mathbb{Q}$. Set $V = L \otimes_{\mathcal{O}_F} F$, with associated local Hermitian space V_v for each place v of $\mathbb{Q}$. Suppose $m^\flat \geq 0$ is an integer. For each prime p, let $\varphi_p^\flat = \mathbf{1}_{L_p}^{m^\flat} \in \mathcal{S}(V_p^{m^\flat})$, form the local Siegel–Weil standard section $\Phi_{\varphi_v^\flat} \in I(\chi_v, s)$, and set

$$\Phi_L \coloneqq \Phi_\infty^{(n)} \bigotimes_p \Phi_{\varphi_p^\flat} \in I(\chi, s) \tag{22.1.34}$$

where the Archimedean component $\Phi_\infty^{(n)}$ is the standard (normalized) scalar weight section from Section 13.2. Form the associated classical $U(m^\flat, m^\flat)$ Eisenstein series $E(z^\flat, s, \Phi_L)_n$ for $z^\flat \in \mathcal{H}_{m^\flat}$, and consider the normalized Eisenstein series Fourier coefficients

$$E^*_{T^\flat}(y^\flat, s, \Phi_L)_n \coloneqq \left(\prod_p \gamma_{\psi_p}(V_p)^{m^\flat} \mathrm{vol}(L_p)^{-m^\flat} \right) \Lambda_{m^\flat}(s)^\circ_n E_{T^\flat}(y^\flat, s, \Phi_L)_n \tag{22.1.35}$$

for $T^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{Q})$. We are not sure whether this is a "good" normalization if L is not self-dual, so the preceding notation appears nowhere else in this work. As in Section 15.3, $\gamma_{\psi_p}(V_p)$ is a Weil index and $\mathrm{vol}(L_p)$ is the volume of L_p with respect to a certain self-dual Haar measure on V_p (these factors are 1 for all but finitely many p).

Form the moduli stack $\mathcal{M} \to \operatorname{Spec} \mathcal{O}_F[1/d_L]$ associated with L as in Section 3.1 (also Section 3.2).

Remark 22.1.5. Since the proof of Theorem 22.1.1 is local in nature, it is possible to use our local main theorems to prove variants for non self-dual L, up to discarding finitely many primes.

Set $m^\flat = n - 1$. Consider $T^\flat \in \mathrm{Herm}_{n-1}(\mathbb{Q})$ with $\det T^\flat \neq 0$. Let $C \in \mathbb{Q}_{>0}$ be the volume constant from Lemma 20.4.1(3), for the Hermitian space V and with $v_0 = \infty$ etc. in the notation of loc. cit.. Consider $y^\flat \in \mathrm{Herm}_{n-1}(\mathbb{R})_{>0}$. Form $[\widehat{\mathcal{Z}}(T^\flat)]$ with Green current with respect to $y^\flat$. Arguing as in the proof of our main theorem (Theorem 22.1.1) gives

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T^\flat)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)) = C \cdot \frac{d}{ds}\bigg|_{s=0} \left(\frac{\Lambda_n(s)^\circ_n}{\Lambda_{n-1}(s+1/2)^\circ_n} E^*_{T^\flat}(y^\flat, s+1/2, \Phi_L)_n \right) \mod \sum_{p \mid 2d_L} \mathbb{Q} \cdot \log p. \tag{22.1.36}$$

For proving (22.1.36), the diagonalization argument (Step 1) in the proof of Theorem 22.1.1 can be skipped. If 2 is split in $\mathcal{O}_F$, the expression "$2d_L$" in (22.1.36) may be replaced by "d_L".

In the case $n = 1$, recall that $\mathcal{M}$ extends smoothly (and nontrivially) over all of $\operatorname{Spec} \mathcal{O}_F$ (Remark 3.1.4). In this case, we need not discard any primes in (22.1.36). As $m^\flat = 0$, the normalized $U(m^\flat, m^\flat)$ Eisenstein series E^* is the constant function 1 in this case.

Recall that our main Archimedean local result was valid in arbitrary "codimension" for empty local special cycles with possibly nontrivial Green current ("purely Archimedean intersection number"). This has the following global consequence.

Theorem 22.1.6. *Let $m^\flat$ be any integer with $1 \leq m^\flat \leq n$. Consider $T^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{Q})$ which is nonsingular and not positive definite. Let $C \in \mathbb{Q}_{>0}$ be the volume constant from Lemma 20.4.1(1), for the Hermitian space V, the lattice L, and $v_0 = \infty$ in the notation of loc. cit..*

For any $y^\flat \in \mathrm{Herm}_{m^\flat}(\mathbb{R})_{>0}$, we have an equality of real numbers

$$\widehat{\deg}([\widehat{\mathcal{Z}}(T^\flat)] \cdot \widehat{c}_1(\widehat{\mathcal{E}}^\vee)^{n-m^\flat}) \coloneqq \int_{\mathcal{M}_\mathbb{C}} g_{T^\flat, y^\flat} \wedge c_1(\widehat{\mathcal{E}}_\mathbb{C}^\vee)^{n-m^\flat} = (-1)^{n-m^\flat} C \cdot \frac{h_F}{w_F} \frac{d}{ds}\bigg|_{s=s_0^\flat} E^*_{T^\flat}(y^\flat, s, \Phi_L)_n \tag{22.1.37}$$

where $s_0^\flat \coloneqq (n - m^\flat)/2$.

Proof. In the theorem statement, we set $\mathcal{M}_\mathbb{C} \coloneqq \mathcal{M} \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C}$ for either choice of embedding $F \to \mathbb{C}$. Recall that the special cycle $\mathcal{Z}(T^\flat)$ is empty by the non-positive definiteness (Section 3.3). The current $g_{T^\flat, y^\flat}$ associated with $[\widehat{\mathcal{Z}}(T^\flat)]$ is formed with respect to $y^\flat$, as usual.

Using our main Archimedean result (Theorem 19.1.1) and local Siegel–Weil (Lemma 20.4.1) for uniformization degrees, the theorem follows as in the proof of Theorem 22.1.1, Step (3a). Since $\det T^\flat \neq 0$, the proof is simpler here as the diagonalization argument of loc. cit. plays no role. Recall $W^*_{T^\flat, \infty}(y^\flat, s_0^\flat)^\circ_n = 0$ (15.2.6), so the derivatives of non-Archimedean Whittaker functions play no role. If $T^\flat$ has signature $(m^\flat - r, r)$ for $r \geq 2$, then both sides of (22.1.37) are zero. The sign $(-1)^{n-m^\flat}$ comes from the Archimedean local functional equation (Lemma 16.2.1), since Theorem 19.1.1 was stated at $s = -s_0^\flat$. $\square$

When $m^\flat = n$, the preceding result is due to Liu (see [Liu11, Theorem 4.17, Proof of Theorem 4.20] and also [LZ22a, Theorem 15.3.1]). We do not have a new proof of this case (we deduced our local result for arbitrary $m^\flat$ from Liu's result using our local limiting method).

22.2 Faltings heights of Hecke translates of CM elliptic curves

Using the Serre tensor construction, we restate part of the simplest case ($n = 2$) of our main theorem (Theorem 22.1.1) in more elementary terms, via Faltings heights of Hecke translates of CM elliptic curves (Corollary 22.2.2).

We assume $2 \nmid \Delta$, but allow 2 inert or split in $\mathcal{O}_F$ for the moment. When $n = 2$ and L is a self-dual Hermitian $\mathcal{O}_F$-lattice of signature $(1,1)$, recall

$$\mathcal{M} = \mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \mathscr{M}(1,1)^\circ \tag{22.2.1}$$

in the notation of Section 3.2. Recall that $\mathscr{M}_0$ is the moduli stack parameterizing $(A_0, \iota_0, \lambda_0)$ where A_0 is an elliptic curve with signature $(1,0)$ action ι_0 by $\mathcal{O}_F$, and λ_0 the unique principal polarization. Recall that $\mathscr{M}(1,1)^\circ$ is the closure of the generic fiber in the moduli stack of signature $(1,1)$ Hermitian abelian schemes (A, ι, λ) where $|\Delta| \cdot \lambda$ is a polarization with $\ker(|\Delta| \cdot \lambda) = A[\sqrt{\Delta}]$.

For integers $j > 0$, we first recall how to relate the special cycles $\mathcal{Z}(j) \to \mathcal{M}$ to Hecke translates of CM elliptic curves, as explained in [KR14, §14]. Our $|\Delta| \cdot \lambda$ is their λ.

Write $\mathscr{M}_{\mathrm{ell}}$ for the moduli stack of elliptic curves base-changed to $\operatorname{Spec} \mathcal{O}_F$. If $\mathcal{O}_F^* := \operatorname{Hom}_{\mathbb{Z}}(\mathcal{O}_F, \mathbb{Z})$, we write $\lambda_{\mathrm{tr}} \colon \mathcal{O}_F \to \mathcal{O}_F^*$ for the σ-linear map corresponding to the symmetric $\mathbb{Z}$-bilinear pairing $\operatorname{tr}_{F/\mathbb{Q}}(a^\sigma b)$ on $\mathcal{O}_F$. As in [KR14, §14], there is a *Serre tensor* morphism

$$\begin{aligned} \mathscr{M}_{\mathrm{ell}} &\xrightarrow{i_{\mathrm{Serre}}} \mathscr{M}(1,1)^\circ \\ E &\longmapsto E \otimes_{\mathbb{Z}} \mathcal{O}_F \end{aligned} \tag{22.2.2}$$

where $E \otimes_{\mathbb{Z}} \mathcal{O}_F$ is given the polarization $|\Delta|^{-1}(\lambda_E \otimes \lambda_{\mathrm{tr}}) \colon E \otimes_{\mathbb{Z}} \mathcal{O}_F \to E^\vee \otimes_{\mathbb{Z}} \mathcal{O}_F^*$. As we have seen previously, $E \otimes_{\mathbb{Z}} \mathcal{O}_F$ is (by definition) the functor given by $(E \otimes_{\mathbb{Z}} \mathcal{O}_F)(S') = E(S') \otimes_{\mathbb{Z}} \mathcal{O}_F$ for schemes S' (over the understood base for E).

For the rest of Section 22.2, we now assume $\mathcal{O}_F^\times = \{\pm 1\}$. In this case, the Serre tensor morphism is an open and closed immersion.[45] Indeed, i_{Serre} is proper (valuative criterion) and a monomorphism of algebraic stacks, hence a closed immersion of algebraic stacks. Since the source and target are Deligne–Mumford, smooth, finite type, and separated over $\operatorname{Spec} \mathcal{O}_F$ of the same relative dimension, this implies that i_{Serre} is also an open immersion.

[45]The hypothesis $\mathcal{O}_F^\times = \{\pm 1\}$ should be added in [KR14, Proposition 14.4], as otherwise $\operatorname{Aut}(E) \neq \operatorname{Aut}(E \otimes_{\mathbb{Z}} \mathcal{O}_F)$ (right-hand side means $\mathcal{O}_F$-linear automorphisms preserving the polarization) so $i_{\mathrm{Serre}} \colon \mathscr{M}_{\mathrm{ell}} \to \mathscr{M}(1,1)^\circ$ is not a monomorphism and hence cannot be a closed immersion in the sense of [SProject, Section 04YK]. The remaining arguments are the same at least if $2 \nmid \Delta$.

The class group $\mathrm{Cl}(\mathcal{O}_F)$ acts $\mathscr{M}(1,1)^\circ$ as follows. Given any fractional ideal $\mathfrak{a} \subseteq F$, set $\mathfrak{a}^\vee := \mathrm{Hom}_{\mathcal{O}_F}(\mathfrak{a}, \mathcal{O}_F)$, and consider the σ-linear map $\lambda_{\mathfrak{a}} \colon \mathfrak{a} \xrightarrow{\sim} \mathfrak{a}^\vee$ given by the perfect positive-definite Hermitian pairing $a, b \mapsto N(\mathfrak{a})^{-1} a^\sigma b$ on $\mathfrak{a}$. There is an induced automorphism of $\mathscr{M}(1,1)^\circ$ sending

$$(A, \iota, \lambda) \to (A \otimes_{\mathcal{O}_F} \mathfrak{a}, \iota, \lambda \otimes \lambda_{\mathfrak{a}}). \tag{22.2.3}$$

The action of $\mathrm{Cl}(\mathcal{O}_F)$ on $\mathscr{M}(1,1)^\circ$ is simply transitive on the set of connected components (see the proof of [KR14, Proposition 14.4]). There is a similar action of $\mathrm{Cl}(\mathcal{O}_F)$ on $\mathscr{M}_0$ which sends $(A_0, \iota_0, \lambda_0) \mapsto (A_0 \otimes_{\mathcal{O}_F} \mathfrak{a}, \iota_0, \lambda_0 \otimes \lambda_{\mathfrak{a}})$. Given a fractional ideal $\mathfrak{a} \subseteq F$, we write $f_{\mathfrak{a}} \colon \mathcal{M} \to \mathcal{M}$ for the induced automorphism just described.

Given any integer $j > 0$, the action of $\mathrm{Cl}(\mathcal{O}_F)$ preserves $\mathcal{Z}(j)$, in the sense that there is a 2-Cartesian diagram

$$\begin{tikzcd} \mathcal{Z}(j) \arrow[r, "\tilde{f}_{\mathfrak{a}}"] \arrow[d] & \mathcal{Z}(j) \arrow[d] \\ \mathcal{M} \arrow[r, "f_{\mathfrak{a}}"] & \mathcal{M} \end{tikzcd} \tag{22.2.4}$$

for any fractional ideal $\mathfrak{a}$, where $\tilde{f}_{\mathfrak{a}}$ sends

$$(A_0, \iota_0, \lambda_0, A, \iota, \lambda, x) \mapsto (A_0 \otimes_{\mathcal{O}_F} \mathfrak{a}, \iota_0, \lambda_0 \otimes \lambda_{\mathfrak{a}}, A \otimes_{\mathcal{O}_F} \mathfrak{a}, \iota, \lambda \otimes \lambda_{\mathfrak{a}}, x \otimes 1) \tag{22.2.5}$$

for $x \in \mathrm{Hom}_{\mathcal{O}_F}(A_0, A)$ satisfying $x^\dagger x = j$.

Consider the *j-th Hecke correspondence* $\mathcal{T}_j \to \mathscr{M}_0 \times_{\mathrm{Spec}\, \mathcal{O}_F} \mathscr{M}_{\mathrm{ell}}$, where $\mathcal{T}_j$ is the stack parameterizing tuples $(E_0, \iota_0, \lambda_0, E, w)$ for $(E_0, \iota_0, \lambda_0) \in \mathscr{M}_0$, for $E \in \mathscr{M}_{\mathrm{ell}}$, and $w \colon E \to E_0$ an isogeny of degree j.

Consider the map $\mathscr{M}_0 \times \mathscr{M}_{\mathrm{ell}} \to \mathcal{M}$ induced by i_{Serre} (and the identity on $\mathscr{M}_0$). The Kudla–Rapoport cycle $\mathcal{Z}(j)$ pulls back to the Hecke correspondence $\mathcal{T}_j$, i.e. there is a 2-Cartesian diagram

$$\begin{tikzcd} \mathcal{T}_j \arrow[r] \arrow[d] & \mathcal{Z}(j) \arrow[d] \\ \mathscr{M}_0 \times_{\mathrm{Spec}\, \mathcal{O}_F} \mathscr{M}_{\mathrm{ell}} \arrow[r] & \mathcal{M} \end{tikzcd} \tag{22.2.6}$$

where $\mathcal{T}_j \to \mathcal{Z}(j)$ sends

$$(E_0, \iota_0, \lambda_0, E, w) \mapsto (E_0, \iota_0, \lambda_0, E \otimes_{\mathbb{Z}} \mathcal{O}_F, \iota, \lambda_E \otimes \lambda_{\mathrm{tr}}, x_w) \tag{22.2.7}$$

(with λ_E denoting the unique principal polarization of E) and where $x_w \colon E_0 \to E \otimes_{\mathbb{Z}} \mathcal{O}_F$ is the $\mathcal{O}_F$-linear map such that $\sqrt{\Delta} x_w^\dagger \in \mathrm{Hom}_{\mathcal{O}_F}(E \otimes_{\mathbb{Z}} \mathcal{O}_F, E_0)$ corresponds to w via the adjunction

$$\mathrm{Hom}_{\mathcal{O}_F}(E \otimes_{\mathbb{Z}} \mathcal{O}_F, E_0) = \mathrm{Hom}(E, E_0). \tag{22.2.8}$$

Here, we are implicitly claiming $\deg(w) = x_w^\dagger x_w$. The fact that (22.2.6) is well-defined and 2-Cartesian is proved in [KR14, Proposition 14.5].

We next discuss the Eisenstein series of Theorem 22.1.1(2) in more elementary terms when $n = 2$. In this case, the $U(1,1)$ Eisenstein series $E^*(z,s)_2^\circ$ (with $m = 1$ in our usual notation, and normalized as in Section 17.1) admits the classical expression

$$E^*(z,s)_2^\circ = -\frac{\pi^{-s+1/2}}{8\pi^2}\Gamma(s+3/2)\zeta(2s+1)\sum_{\substack{c,d\in\mathbb{Z}\\(c,d)=1}}\frac{y^{s-1/2}}{(cz+d)^2|cz+d|^{2(s-1/2)}} \tag{22.2.9}$$

for $z = x + iy \in \mathcal{H}$, where $\mathcal{H} \subseteq \mathbb{C}$ is the usual upper-half space (here z corresponds to $z^\flat$ in Theorem 22.1.1(2)).

For nonzero $j \in \mathbb{Z}$, the (normalized) j-th Fourier coefficient of $E^*(z,s)_2^\circ$ factorizes into (normalized) local Whittaker functions

$$E^*_j(y,s)_2^\circ = W^*_{j,\infty}(y,s)_2^\circ \prod_p W^*_{j,p}(s)_2^\circ \tag{22.2.10}$$

as in Section 17.1. We have the formulas

$$W^*_{j,p}(s)_2^\circ = p^{v_p(j)(s+1/2)}\sigma_{-2s}(p^{v_p(j)}) \qquad \prod_p W^*_{j,p}(s)_2^\circ = |j|^{s+1/2}\sigma_{-2s}(|j|) \tag{22.2.11}$$

where $v_p(-)$ means p-adic valuation and

$$\sigma_s(|j|) \coloneqq \sum_{d \mid |j|} d^s \tag{22.2.12}$$

is the classical divisor function. These formulas for local Whittaker functions are likely classical, but they also follow from (18.2.8) on local densities (translation to local Whittaker functions via (15.5.6)). An integral expression for $W^*_{j,\infty}(y,s)_2^\circ$ may be found in Section 19.2. For $j > 0$, recall $W^*_{j,\infty}(y,1/2)_2^\circ = 1$ (15.2.6).

We require $j > 0$ for the rest of Section 22.2. Fix an embedding $F \to \mathbb{C}$. Given a CM elliptic curve $(E_0, \iota_0, \lambda_0) \in \mathscr{M}_0(\mathbb{C})$, we consider the set of *j-th Hecke translates* of E_0 given by

$$\mathcal{T}_j(E_0) \coloneqq \{(E_0, \iota_0, \lambda_0, E, w) \in \mathcal{T}_j(\mathbb{C})\}. \tag{22.2.13}$$

Phrased alternatively, the fiber of $\mathcal{T}_j \to \mathscr{M}_0$ over the point $\operatorname{Spec}\mathbb{C} \to \mathscr{M}_0$ corresponding to E_0 is a finite scheme over $\operatorname{Spec}\mathbb{C}$, and $\mathcal{T}_j(E_0)$ is its set of $\mathbb{C}$-points. We set

$$\deg \mathcal{T}_j(E_0) \coloneqq |\mathcal{T}_j(E_0)| \qquad h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) \coloneqq \sum_{E \in \mathcal{T}_j(E_0)} h_{\mathrm{Fal}}(E) \tag{22.2.14}$$

where $|-|$ denotes set cardinality, the sum runs over $(E_0, \iota_0, \lambda_0, E, w) \in \mathcal{T}_j(E_0)$, and $h_{\mathrm{Fal}}(E)$ denotes the Faltings height of E (with metric normalized as in (4.3.1), see also Section 9.1) after descending from $\mathbb{C}$ to any number field.

The following lemma states that the (total) Faltings height of j-th Hecke translates of a chosen elliptic curve with CM by $\mathcal{O}_F$ does not depend on the choice of CM elliptic curve. It should admit a general formulation in terms of Hecke correspondences over $\mathscr{M}_0$. We give a more elementary treatment in the spirit of this section.

Lemma 22.2.1. *Fix $j \in \mathbb{Z}_{>0}$. For any $(E_0, \iota_0, \lambda_0) \in \mathscr{M}_0(\mathbb{C})$ and $(E_0', \iota_0', \lambda_0') \in \mathscr{M}_0(\mathbb{C})$, we have*

$$\deg \mathcal{T}_j(E_0) = \deg \mathcal{T}_j(E_0') \qquad h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) = h_{\mathrm{Fal}}(\mathcal{T}_j(E_0')). \tag{22.2.15}$$

Proof. Given any $d \in \mathbb{Z}$, we claim that there exists an isogeny $\phi \colon E_0' \to E_0$ of degree prime to d. Consider

$$E_0(\mathbb{C}) = \mathbb{C}/\Lambda_0 \qquad E_0'(\mathbb{C}) = \mathbb{C}/\Lambda_0' \tag{22.2.16}$$

for lattices Λ_0 and Λ_0'. Without loss of generality, we may assume $\Lambda_0 = \mathcal{O}_F \subseteq \mathbb{C}$ and that $\Lambda_0' = \mathfrak{a}_0'$ for some fractional ideal $\mathfrak{a}_0' \subseteq \mathbb{C}$. By the Chinese remainder theorem, we can assume $\mathfrak{a}_0' \subseteq \mathcal{O}_F$ and that $\mathfrak{a}_0'$ has norm prime to d (without changing the ideal class of $\mathfrak{a}_0'$). The inclusion $\mathfrak{a}_0' \subseteq \mathcal{O}_F$ gives an isogeny $E_0' \to E_0$ of degree prime to d.

Let p be any prime. Let $\phi \colon E_0' \to E_0$ be an isogeny of degree prime to pj. As above, we view $\phi \colon E_0(\mathbb{C}) \to E_0'(\mathbb{C})$ as an inclusion of lattices $\Lambda_0' \to \Lambda_0$ of index prime to pj. There is an induced bijection

$$\begin{aligned} \mathcal{T}_j(E_0) &\longrightarrow \mathcal{T}_j(E_0') \\ \Lambda &\longmapsto \Lambda \cap \Lambda_0'. \end{aligned} \tag{22.2.17}$$

We are viewing Λ as the element $\mathbb{C}/\Lambda \to \mathbb{C}/\Lambda_0$ of $\mathcal{T}_j(E_0)$, and similarly for $\Lambda \cap \Lambda_0'$.

The isogeny $\mathbb{C}/(\Lambda \cap \Lambda_0') \to \mathbb{C}/\Lambda$ has degree $\deg \phi$, which is prime to p. As these elliptic curves are defined over $\overline{\mathbb{Q}}$, this isogeny also descends to $\overline{\mathbb{Q}}$. By the formula for change for Faltings height along an isogeny (9.2.4), we conclude $h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) - h_{\mathrm{Fal}}(\mathcal{T}_j(E_0')) \in \sum_{\ell | \deg \phi} \mathbb{Q} \cdot \log \ell$. Varying p shows $h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) = h_{\mathrm{Fal}}(\mathcal{T}_j(E_0'))$, as the real numbers $\log p$ are $\mathbb{Q}$-linearly independent for varying p. $\square$

Consider any $(E_0, \iota_0, \lambda_0) \in \mathscr{M}_0(\mathbb{C})$. Using (22.2.6) (Kudla–Rapoport cycle pulls back to Hecke correspondence), the geometric Siegel–Weil statement in Remark 21.1.2 implies

$$\frac{h_F^2}{w_F} \deg \mathcal{T}_j(E_0) = 2 \frac{h_F^2}{w_F^2} E_j^*(y, 1/2)_2^\circ \tag{22.2.18}$$

for any $y \in \mathbb{R}_{>0}$. On the left, one factor of h_F appears because the Serre tensor morphism $i_{\mathrm{Serre}} \colon \mathscr{M}_{\mathrm{ell}} \to \mathscr{M}(1,1)^\circ$ is the inclusion of one connected component (and $\mathscr{M}(1,1)^\circ$ has h_F connected components, by the action of $\mathrm{Cl}(\mathcal{O}_F)$ discussed above; we discussed that this action is compatible with Kudla–Rapoport cycles). On the left, the additional factor h_F/w_F appears via Lemma 22.2.1 (instead of summing over $\mathscr{M}_0(\mathbb{C})$, it is enough to consider a fixed E_0 and multiply by $h_F/w_F = \deg_{\mathbb{C}}(\mathscr{M}_0 \times_{\operatorname{Spec} \mathcal{O}_F} \operatorname{Spec} \mathbb{C})$).

By the formulas in (22.2.11) and surrounding discussion, this recovers the well-known identity $\deg \mathcal{T}_j(E_0) = \sigma_1(j)$ for degrees of Hecke correspondences (recall our running assumption $|\mathcal{O}_F^\times| = \{\pm 1\}$ for most of Section 22.2, i.e. $w_F = 2$).

In the next lemma, $h_{\mathrm{Fal}}^{\mathrm{CM}} = h_{\mathrm{Fal}}(E_0)$ is the Faltings height of any elliptic curve with CM by $\mathcal{O}_F$ (4.3.5). It is well known that this does not depend on the choice of CM elliptic curve (also follows from Lemma 22.2.1).

Corollary 22.2.2. *Suppose 2 is split in $\mathcal{O}_F$. For any integer $j > 0$ and any CM elliptic curve $(E_0, \iota_0, \lambda_0) \in \mathscr{M}_0(\mathbb{C})$, we have*

$$h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) - \sigma_1(j) \cdot h_{\mathrm{Fal}}^{\mathrm{CM}} = \frac{1}{2} \frac{d}{ds}\bigg|_{s=1/2} \left(j^{s+1/2} \sigma_{-2s}(j) \right). \tag{22.2.19}$$

Proof. Set $n = 2$ and consider the 2×2 matrix $T = \operatorname{diag}(0, j)$. Again using (22.2.6) to pull back Kudla–Rapoport cycles to Hecke correspondences, we have

$$2 \frac{h_F^2}{w_F} \left(2 h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) - 2 (\deg \mathcal{T}_j(E_0)) \cdot h_{\mathrm{Fal}}^{\mathrm{CM}} \right) = -2 \sum_p \mathrm{Int}_{\mathscr{H}, p, \mathrm{global}}(T) \tag{22.2.20}$$

in our previous notation (Remark 22.1.4). On the left, the outer factor of 2 has the same explanation as in (22.1.32) (see following discussion). The factor h_F^2/w_F has the same explanation as in (22.2.18), via Lemma 22.2.1 on Faltings height. The factor of 2 in $2 h_{\mathrm{Fal}}(\mathcal{T}_j(E_0))$ appears because $h_{\mathrm{Fal}}(E \otimes_{\mathbb{Z}} \mathcal{O}_F) = h_{\mathrm{Fal}}(E \times E) = 2 h_{\mathrm{Fal}}(E)$. The factor of 2 in $2 (\deg \mathcal{T}_j(E_0)) \cdot h_{\mathrm{Fal}}^{\mathrm{CM}}$ is the n in Remark 22.1.4.

In our previous notation, we have $\mathrm{Int}_{\mathscr{V}, p, \mathrm{global}}(T) = 0$ for all primes p as the vertical special cycle class ${}^{\mathbb{L}}\mathcal{Z}(T)_{\mathscr{V}, p}$ is 0 when $n = 2$ (Lemma 11.7.6). Hence we have $\mathrm{Int}_{p, \mathrm{global}}(T) = \mathrm{Int}_{\mathscr{H}, p, \mathrm{global}}(T) + \mathrm{Int}_{\mathscr{V}, \ell, \mathrm{global}}(T) = \mathrm{Int}_{\mathscr{H}, p, \mathrm{global}}(T)$.

Then (22.1.22) ("horizontal local part" of our main result) implies

$$\mathrm{Int}_{p, \mathrm{global}}(T) = -\frac{2 h_F^2}{w_F^2} \left(\frac{d}{ds}\bigg|_{s=1/2} W_{j,p}^*(s)_2^\circ \right) \prod_{\ell \neq p} W_{j,\ell}^*(1/2)_2^\circ \tag{22.2.21}$$

for all p (in the notation of loc. cit., take $T^\flat = j$, $a_v^\flat = 1$ for all $v < \infty$, and recall our notation $\tilde{W}^*_{T^\flat,v}(1,s)^\circ_n = W^*_{T^\flat,v}(s)^\circ_n$). Since $j > 0$, we have used $W^*_{j,\infty}(1/2)^\circ_2 = 1$ (15.2.6) as recalled above.

Combining (22.2.21) and (22.2.20) along with the formula $\deg \mathcal{T}_j(E_0) = \sigma_1(j)$, we obtain

$$h_{\mathrm{Fal}}(\mathcal{T}_j(E_0)) - \sigma_1(j) \cdot h^{\mathrm{CM}}_{\mathrm{Fal}} = \frac{1}{2} \frac{d}{ds}\bigg|_{s=1/2} \left(\prod_p W^*_{j,p}(1/2)^\circ_2 \right) \tag{22.2.22}$$

where the product runs over all primes (not including the Archimedean place). The corollary now follows from the formulas in (22.2.11). $\square$

Appendices

A K_0 groups

A.1 K_0 groups for Deligne–Mumford stacks

Suppose $\mathcal{X}$ is a Noetherian Deligne–Mumford stack. There are at least two different ways one might define K_0 groups for $\mathcal{X}$. One way is to define a K-theory spectrum for $\mathcal{X}$ using the K-theory spectra of schemes in the small étale site of $\mathcal{X}$, as in [Gil09, §2]. This is the approach used in [HM22]. Another way is to simply mimic a definition of K_0 for schemes and consider perfect complexes on the small étale site of $\mathcal{X}$. These two approaches will in general result in different K_0 groups [HM22, Remark A.2.4]. At least if $\mathcal{X}$ is regular (and, say, with the additional running hypotheses of [HM22, Appendix A]), there is a map from the latter K_0 group to the former K_0 group [HM22, (A.7),(A.8)].

In this paper, we take the latter approach and mimic constructions for schemes to define $K_0(\mathcal{X})$. Our definitions and notation will be analogous to those for schemes in [SProject, Section 0FDE]. When defining dimension/codimension filtrations on $K_0'(\mathcal{X})$ (with notation and hypotheses as below), we will require existence of a finite flat cover by a scheme (enough for our intended application). A similar approach appears in [YZ17, Appendix A] (at least for K_0'), but there the stacks are over a base field. We need a slightly more general setup which allows base schemes such as $\operatorname{Spec} R$ for Dedekind domains R.

Suppose $\mathcal{X}$ is a Deligne–Mumford stack. By an *$\mathcal{O}_{\mathcal{X}}$-module*[46], we mean a sheaf of modules on the small étale site[47] of $\mathcal{X}$. Similarly, *quasi-coherent $\mathcal{O}_{\mathcal{X}}$-modules* will mean quasi-coherent sheaves of modules on the small étale site. When $\mathcal{X}$ is locally Noetherian, we will also speak of *coherent $\mathcal{O}_{\mathcal{X}}$-modules* on the small étale site, which are the same as finitely presented quasi-coherent $\mathcal{O}_{\mathcal{X}}$-modules in this situation.

Suppose $\mathcal{X}$ is a locally Noetherian Deligne–Mumford stack. The category *Coh*$(\mathcal{O}_{\mathcal{X}})$ of coherent $\mathcal{O}_{\mathcal{X}}$-modules forms a weak Serre subcategory of the abelian category *Mod*$(\mathcal{O}_{\mathcal{X}})$ of $\mathcal{O}_{\mathcal{X}}$-modules (reduce to the case of small étale sites of schemes and apply [SProject, Lemma

[46] This is one of the only places where our conventions differ from the Stacks project [SProject, Chapter 06TF], which mostly works with sheaves on big sites (say, fppf and étale) for general algebraic stacks. Restriction from these two big sites to the small étale site (for Deligne–Mumford stacks) induces equivalences on categories of quasi-coherent sheaves. But the equivalences are not compatible with pushforward, and are also not compatible with exactness for $\mathcal{O}_{\mathcal{X}}$-modules on big sites versus small sites.

[47] The small étale site is as defined in [DM69, Definition 4.10]: the underlying category has objects which are pairs (U, f) for $f \colon U \to \mathcal{X}$ an étale morphism (definition as in [SProject, Definition 0CIL]) from a scheme U, and morphisms are pairs $(g, \xi) \colon (U, f) \to (U', f')$ where $g \colon U \to U'$ is a 1-morphism and $\xi \colon f \to f' \circ g$ is a 2-isomorphism.

05VG, Lemma 0GNB]). We may form derived categories such as

$$D(\mathcal{O}_{\mathcal{X}}) \quad D^b(\mathcal{O}_{\mathcal{X}}) \quad D_{perf}(\mathcal{O}_{\mathcal{X}}) \quad D^b_{Coh}(\mathcal{O}_{\mathcal{X}}) \quad D^b(Coh(\mathcal{O}_{\mathcal{X}})) \tag{A.1.1}$$

which denote the derived category of $\mathcal{O}_{\mathcal{X}}$-modules, bounded derived category of $\mathcal{O}_{\mathcal{X}}$-modules, derived category of perfect objects (definition as in [SProject, Section 08G4]) in $D(\mathcal{O}_{\mathcal{X}})$, bounded derived category of $\mathcal{O}_{\mathcal{X}}$-modules with coherent cohomology, and the bounded derived category of coherent $\mathcal{O}_{\mathcal{X}}$-modules, respectively.

If $\mathcal{X}$ happens to be a scheme, then $D_{perf}(\mathcal{O}_{\mathcal{X}})$ and $D^b_{Coh}(\mathcal{O}_{\mathcal{X}})$ and $D^b(Coh(\mathcal{O}_{\mathcal{X}}))$ will agree with the usual constructions using the small Zariski site instead of the small étale site, via comparison results such as [SProject, Lemma 08HG, Lemma 071Q, Lemma 05VG].

Definition A.1.1. Let $\mathcal{X}$ be a locally Noetherian Deligne–Mumford stack. We set

$$K_0(\mathcal{X}) \coloneqq K_0(D_{perf}(\mathcal{O}_{\mathcal{X}})) \qquad K'_0(\mathcal{X}) \coloneqq K_0(Coh(\mathcal{O}_{\mathcal{X}})). \tag{A.1.2}$$

Above, the left expression means K_0 of a triangulated category and the right expression means K_0 of an abelian category. If $\mathcal{X}$ is a locally Noetherian Deligne–Mumford stack, we have canonical identifications

$$K_0(Coh(\mathcal{O}_{\mathcal{X}})) = K_0(D^b(Coh(\mathcal{O}_{\mathcal{X}}))) = K_0(D^b_{Coh}(\mathcal{O}_{\mathcal{X}})) \tag{A.1.3}$$

as in [SProject, Lemma 0FDF] (the case of schemes) by general facts about derived categories (see also [SProject, Lemma 0FCS]).

If $\mathcal{X}$ is a quasi-compact locally Noetherian Deligne–Mumford stack, there is an inclusion $D_{perf}(\mathcal{O}_{\mathcal{X}}) \to D^b_{Coh}(\mathcal{O}_{\mathcal{X}})$ and a corresponding group homomorphism $K_0(\mathcal{X}) \to K'_0(\mathcal{X})$. If $\mathcal{X}$ is a regular locally Noetherian Deligne–Mumford stack (not necessarily quasi-compact), there is an inclusion $D^b_{Coh}(\mathcal{O}_{\mathcal{X}}) \to D_{perf}(\mathcal{O}_{\mathcal{X}})$ and a corresponding group homomorphism $K'_0(\mathcal{X}) \to K_0(\mathcal{X})$. If $\mathcal{X}$ is a locally Noetherian Deligne–Mumford stack which is both quasi-compact and regular, we have $D_{perf}(\mathcal{O}_{\mathcal{X}}) = D^b_{Coh}(\mathcal{O}_{\mathcal{X}})$ and a corresponding isomorphism

$$K_0(\mathcal{X}) \xrightarrow{\sim} K'_0(\mathcal{X}). \tag{A.1.4}$$

These claims follow from the corresponding facts for schemes [SProject, Lemma 0FDC] and comparison results mentioned previously.

The derived tensor product $\otimes^{\mathbb{L}}$ on $D(\mathcal{O}_{\mathcal{X}})$ gives $K_0(\mathcal{X})$ the structure of a commutative ring. Compatibility of $\otimes^{\mathbb{L}}$ with the case when $\mathcal{X}$ is also a scheme follows from the displayed equation in the proof of [SProject, Lemma 08HF] (comparison between the small Zariski and small étale sites).

We next describe dimension and codimension filtrations. Our setup for dimension theory is as in [SProject, Section 02QK]. That is, we work over a locally Noetherian and universally catenary base scheme S with a dimension function $\delta \colon |S| \to \mathbb{Z}$ (which we typically suppress). Typical setups will be $S = \operatorname{Spec} R$ for R a field or Dedekind domain, where δ is the dimension function sending closed points to 0. Any Deligne–Mumford stack $\mathcal{X}$ which is quasi-separated and locally of finite type over S inherits a dimension function $\delta_{\mathcal{X}} \colon |\mathcal{X}| \to \mathbb{Z}$ (work étale locally to pass to the case of schemes; the case of algebraic spaces is [SProject, Section 0EDS]). If $\mathcal{X}$ is equidimensional of dimension n, then $n - \delta_{\mathcal{X}}$ is also the codimension function (given by dimensions of local rings on étale covers by schemes).

For a scheme X which is locally of finite type over S, consider the full subcategory $\mathit{Coh}_{\leq d}(\mathcal{O}_X) \subseteq \mathit{Coh}(\mathcal{O}_X)$ consisting of coherent $\mathcal{O}_X$-modules $\mathcal{F}$ with $\dim(\operatorname{Supp}(\mathcal{F})) \leq d$. Then there is an increasing *dimension filtration* on $K'_0(X) = K_0(\mathit{Coh}(\mathcal{O}_X))$ given by the image

$$F_d K'_0(X) \coloneqq \operatorname{im}(K_0(\mathit{Coh}_{\leq d}(\mathcal{O}_X)) \to K_0(\mathit{Coh}(\mathcal{O}_X))) \tag{A.1.5}$$

as in [SProject, Section 0FEV]. We similarly consider the full subcategory $\mathit{Coh}^{\geq m}(\mathcal{O}_X) \subseteq \mathit{Coh}(\mathcal{O}_X)$ of coherent sheaves supported in codimension $\geq m$, and form the decreasing *codimension filtration*

$$F^m K'_0(X) \coloneqq \operatorname{im}(K_0(\mathit{Coh}^{\geq m}(\mathcal{O}_X)) \to K_0(\mathit{Coh}(\mathcal{O}_X))). \tag{A.1.6}$$

When X is equidimensional of dimension n, we have $F^m K'_0(X) = F_{n-m} K'_0(X)$.

For the case of Deligne–Mumford stacks, one could consider naive dimension/codimension filtrations on $K'_0(\mathcal{X})$ by mimicking the definition for schemes. This may not be a well-behaved notation, and we instead take the filtration defined in [YZ17, A.2.3] (with $\mathbb{Q}$-cofficients).

Definition A.1.2. For S as above, let $\mathcal{X}$ be a Deligne–Mumford stack which is quasi-separated and locally of finite type over S. Suppose there exists a finite flat surjection $\pi \colon U \to \mathcal{X}$ from a scheme U. Pick such a morphism π.

The *dimension filtration* on $K'_0(\mathcal{X})_{\mathbb{Q}}$ is the increasing filtration given by

$$F_d K'_0(\mathcal{X})_{\mathbb{Q}} \coloneqq \{\beta \in K'_0(\mathcal{X})_{\mathbb{Q}} : \pi^*\beta \in F_d K'_0(U)_{\mathbb{Q}}\} \subseteq K'_0(\mathcal{X})_{\mathbb{Q}} \tag{A.1.7}$$

for $d \in \mathbb{Z}$. If $\mathcal{X}$ is equidimensional, we also consider the decreasing *codimension filtration* on $K'_0(\mathcal{X})_{\mathbb{Q}}$ given by

$$F^m K'_0(\mathcal{X})_{\mathbb{Q}} \coloneqq \{\beta \in K'_0(\mathcal{X})_{\mathbb{Q}} : \pi^*\beta \in F^m K'_0(U)_{\mathbb{Q}}\} \subseteq K'_0(\mathcal{X})_{\mathbb{Q}} \tag{A.1.8}$$

for $m \in \mathbb{Z}$.

For $\mathcal{X}$ as in the preceding definition, the filtrations just defined give rise to graded pieces $\mathrm{gr}_d K_0'(\mathcal{X})_{\mathbb{Q}} := F_d K_0'(\mathcal{X})_{\mathbb{Q}} / F_{d-1} K_0'(\mathcal{X})_{\mathbb{Q}}$ and $\mathrm{gr}^m K_0'(\mathcal{X})_{\mathbb{Q}} := F^m K_0'(\mathcal{X})_{\mathbb{Q}} / F^{m+1} K_0'(\mathcal{X})_{\mathbb{Q}}$. If $\mathcal{X}$ as above is equidimensional of dimension n, we have $F^m K_0'(\mathcal{X})_{\mathbb{Q}} = F_{n-m} K_0'(\mathcal{X})_{\mathbb{Q}}$ for all $m \in \mathbb{Z}$.

Lemma A.1.3. *With notation as in Definition A.1.2, the filtrations $F_d K_0'(\mathcal{X})_{\mathbb{Q}}$ and $F^m K_0'(\mathcal{X})_{\mathbb{Q}}$ do not depend on the choice of finite flat surjection $\pi \colon U \to \mathcal{X}$. If $\mathcal{X}$ is a scheme, these filtrations recovers the usual filtrations.*

Proof. Suppose X is a scheme which is locally of finite type over S. If $Z_d(X)$ is the group of d-cycles on X, recall that there is an identification

$$K_0(Coh_{\leq d}(\mathcal{O}_X)/Coh_{\leq d-1}(\mathcal{O}_X)) \xrightarrow{\sim} Z_d(X) \tag{A.1.9}$$

which is compatible with flat pullback of constant relative dimension and finite pushforward [SProject, Lemma 02S9, Lemma 0FDR] (see also [SProject, Lemma 02MX]). For any finite flat surjection $\pi \colon U \to X$ of constant degree a, the map $\pi_* \pi^* \colon Z_d(X) \to Z_d(X)$ is multiplication by a. It follows that $\pi_* \pi^* \colon F_d K_0'(X)/F_{d-1} K_0'(X) \to F_d K_0'(X)/F_{d-1} K_0'(X)$ is multiplication by a. This is an isomorphism after tensoring by $\mathbb{Q}$. When X is equidimensional, this gives the corresponding statement for the codimension filtration as well. This verifies the lemma when $\mathcal{X}$ is a scheme.

Let $\mathcal{X}$ be a Deligne–Mumford stack as in the lemma statement. Let $\pi \colon U \to \mathcal{X}$ and $\pi' \colon U' \to \mathcal{X}$ be two finite flat surjections, for schemes U and U'. Consider the fiber product $U \times_{\mathcal{X}} U'$ with its finite flat projections to U and U'. We then apply the preceding discussion to see that the filtrations do not depend on the choice of finite flat surjection.

These arguments are essentially the same as in [YZ17, A.2.3] (the arguments of loc. cit. are over a base field, so we have used different references). □

Remark A.1.4. As in [YZ17, A.2.3], it is possible to have $F_d K_0'(\mathcal{X})_{\mathbb{Q}} \neq 0$ for $d < 0$ in the situation of Definition A.1.2.

We are mainly interested in K_0 groups for the purpose of intersection theory, so we next discuss degree theory over a field. Suppose $S = \operatorname{Spec} k$ for a field k, and suppose $\mathcal{X}$ is a Deligne–Mumford stack which is proper over S. Again assuming that $\mathcal{X}$ admits a finite flat surjection from a scheme, there is a graded group homomorphism $\mathrm{gr}_* K_0'(\mathcal{X})_{\mathbb{Q}} \to \mathrm{Ch}_*(\mathcal{X})_{\mathbb{Q}}$ as defined in [YZ17, A.2.6] (pass to a finite flat surjection to reduce to the case of schemes).

There is a degree map $\deg \colon \mathrm{Ch}_0(\mathcal{X})_{\mathbb{Q}} \to \mathbb{Q}$ on 0-cycles which may be described as follows. Suppose $\mathcal{Z}$ is a quasi-separated finite type Deligne–Mumford stack over $\operatorname{Spec} k$ with separated diagonal, and assume the underlying topological space $|\mathcal{Z}|$ is a single point. If $V \to \mathcal{Z}$ is any finite flat surjection from a scheme V (which exists as in Remark 4.1.1), one can check that V is finite over $\operatorname{Spec} k$, and we take

$$\deg(\mathcal{Z}) \coloneqq \deg_k(\mathcal{Z}) \coloneqq \frac{\deg_k(V)}{\deg_{\mathcal{Z}}(V)} \tag{A.1.10}$$

where $\deg_k(V)$ (resp. $\deg_{\mathcal{Z}}(V)$) is the degree of the finite flat morphism $V \to \operatorname{Spec} k$ (resp. $V \to \mathcal{Z}$). It is straightforward to see that $\deg_k(\mathcal{Z})$ does not depend on the choice of $V \to \mathcal{Z}$ (compare [Vis89, Definition 1.15]). This generalizes immediately to the case where $|\mathcal{Z}|$ is instead a discrete finite set (add the degrees of its components). When $\mathcal{Z} = \emptyset$, we take $\deg(\mathcal{Z}) \coloneqq 0$.

There is an induced degree map

$$\deg \colon \mathrm{gr}_0 K_0'(\mathcal{X})_{\mathbb{Q}} \to \mathbb{Q}. \tag{A.1.11}$$

Consider a class $\beta = \sum_i b_i [\mathcal{F}_i] \in F_0 K_0'(\mathcal{X})_{\mathbb{Q}}$ where each $\mathcal{F}_i$ is a coherent sheaf on $\mathcal{X}$ (we do not assume $[\mathcal{F}_i] \in F_0 K_0'(\mathcal{X})_{\mathbb{Q}}$ for any given i). Select any finite flat surjection $\pi \colon U \to \mathcal{X}$ with U a scheme. If π has constant degree a, we have

$$\deg(\beta) = \frac{1}{a} \deg(\pi^* \beta) = \frac{1}{a} \sum_i b_i \cdot \chi(\pi^* \mathcal{F}_i) \tag{A.1.12}$$

where χ denotes Euler characteristic. We can give a similar description for general finite flat surjections π by decomposing $\mathcal{X}$ into its connected components. On account of (A.1.12), we may write $\chi(\beta) \coloneqq \deg(\beta)$ and think of $\chi \colon \mathrm{gr}_0 K_0'(\mathcal{X})_{\mathbb{Q}} \to \mathbb{Q}$ as a "stacky Euler characteristic" (compare usage in [KR14, Definition 11.4]). We caution, however, that we have only defined χ on $\mathrm{gr}_0 K_0'(\mathcal{X})_{\mathbb{Q}}$ and have not defined $\chi(\mathcal{F})$ for a general coherent sheaf $\mathcal{F}$ on $\mathcal{X}$.

We conclude this subsection with a lemma which we will use to decompose $K_0'(\mathcal{X})$ in terms of irreducible components of $\mathcal{X}$. A similar lemma for formal schemes is [Zha21, Lemma B.1].

Lemma A.1.5. *Let $\mathcal{X}$ be a locally Noetherian Deligne–Mumford stack. Let $\pi_1 \colon \mathcal{Z}_1 \to \mathcal{X}$ and $\pi_2 \colon \mathcal{Z}_2 \to \mathcal{X}$ be closed immersions of Deligne–Mumford stacks with corresponding ideal sheaves $\mathcal{I}_1$ and $\mathcal{I}_2$. Assume that the diagonals of $\mathcal{X}$, $\mathcal{Z}_1$, and $\mathcal{Z}_2$ are representable by schemes (e.g. if $\mathcal{X}$ is separated).*

Assume that $\mathcal{X} = \mathcal{Z}_1 \cup \mathcal{Z}_2$ scheme-theoretically (meaning $\mathcal{I}_1 \cap \mathcal{I}_2 = 0$). There are mutually *inverse isomorphisms*

$$\begin{aligned} \tfrac{K_0'(\mathcal{X})}{K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2)} &\longleftrightarrow \tfrac{K_0'(\mathcal{Z}_1)}{K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2)} \oplus \tfrac{K_0'(\mathcal{Z}_2)}{K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2)} \\ [\mathcal{F}] &\longmapsto ([\pi_1^* \mathcal{F}_1], [\pi_2^* \mathcal{F}_2]) \\ [\pi_{1,*}\mathcal{F}_1] + [\pi_{2,*}\mathcal{F}_2] &\longleftarrow\!\shortmid ([\mathcal{F}_1], [\mathcal{F}_2])\ . \end{aligned} \tag{A.1.13}$$

Here, $\mathcal{F}$, $\mathcal{F}_1$, and $\mathcal{F}_2$ stand for coherent sheaves on $\mathcal{X}$, $\mathcal{Z}_1$, and $\mathcal{Z}_2$ respectively.

Proof. The condition about diagonals is included for technical convenience. Some additional explanation on notation in the lemma statement: the symbol $\mathcal{Z}_1 \cap \mathcal{Z}_2$ denotes the closed substack $\mathcal{Z}_1 \times_{\mathcal{X}} \mathcal{Z}_2$ of $\mathcal{X}$, with associated ideal sheaf $\mathcal{I}_1 + \mathcal{I}_2$, and we have also written $K_0'(\mathcal{X})/K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2) \coloneqq \operatorname{coker}(K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2) \to K_0'(\mathcal{X}))$ etc. (the latter map may not be injective).

Consider the short exact sequence

$$0 \to \mathcal{O}_{\mathcal{X}}/(\mathcal{I}_1 \cap \mathcal{I}_2) \to \mathcal{O}_{\mathcal{X}}/\mathcal{I}_1 \oplus \mathcal{O}_{\mathcal{X}}/\mathcal{I}_2 \to \mathcal{O}_{\mathcal{X}}/(\mathcal{I}_1 + \mathcal{I}_2) \to 0. \tag{A.1.14}$$

Tensoring by any coherent sheaf $\mathcal{F}$ on $\mathcal{X}$, we find that $\operatorname{Tor}_1^{\mathcal{O}_{\mathcal{X}}}(\mathcal{F}, \mathcal{O}_{\mathcal{X}}/\mathcal{I}_1)$ is an $\mathcal{O}_{\mathcal{X}}/(\mathcal{I}_1 + \mathcal{I}_2)$-module, and similarly with $\mathcal{I}_2$ instead of $\mathcal{I}_1$. This shows that the displayed projection maps $\mathcal{F} \mapsto \pi_1^* \mathcal{F}$ and $\mathcal{F} \mapsto \pi_2^* \mathcal{F}$ are well-defined (i.e. that they are additive in short exact sequences and hence descend to the given quotients of K_0'-groups). Since $\operatorname{Tor}_1^{\mathcal{O}_{\mathcal{X}}}(\mathcal{F}, \mathcal{O}_{\mathcal{X}}/(\mathcal{I}_1 + \mathcal{I}_2))$ is an $\mathcal{O}_{\mathcal{X}}/(\mathcal{I}_1 + \mathcal{I}_2)$-module, the Tor long exact sequence of the displayed short exact sequence also shows that $[\mathcal{F}] = [\mathcal{F} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{O}_{\mathcal{X}}/\mathcal{I}_1] + [\mathcal{F} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{O}_{\mathcal{X}}/\mathcal{I}_2]$ in $K_0'(\mathcal{X})/K_0'(\mathcal{Z}_1 \cap \mathcal{Z}_2)$. □

A.2 K_0 groups with supports along finite morphisms

Suppose X is a separated regular Noetherian scheme. There is an established intersection theory for K_0 groups with supports along closed subsets of X, and the intersection pairing is multiplicative with respect to codimension filtrations (after tensoring by $\mathbb{Q}$) [GS87]. However, we will need a slightly more general setup which allows for "supports along finite morphisms". This is needed because the special cycles $\mathcal{Z}(T) \to \mathcal{M}$ (Section 3.3) are not literally cycles but are instead finite unramified morphisms.

Intersection theory with supports along finite morphisms is also discussed in [HM22, Appendix A.4] for a similar purpose. They are not able to show the codimension filtration is multiplicative in general [HM22, Remark A.4.2], but can show multiplicativity for intersections against classes of codimension 1 (in the case of supports along finite unramified morphisms) [HM22, Proposition A.4.4].

We have two main objectives in this section (besides fixing notation). Our first objective is to comment on another situation where the codimension filtration is multiplicative (namely, when the finite supports become disjoint unions of closed immersions after finite flat base change to a regular scheme). The (short) proof reduces to the case of supports along closed immersions. This is relevant for us because of Lemma 3.4.5, which says that each special cycle $\mathcal{Z}(T) \to \mathcal{M}$ becomes a disjoint union of closed immersions after finite étale base change, at least after inverting the prime p in the cited lemma. For $\mathcal{M}$ associated to a Hermitian lattice of signature $(n-r, r)$, intersecting special cycles over $\mathcal{M}$ involves multiplicativity for classes of codimension r (not covered by [HM22, Proposition A.4.4] when $r > 1$).

Our second objective is to explain intersection theory with supports along finite morphisms of Deligne–Mumford stacks in terms of the K_0 groups of Appendix A.1. A stacky theory is also considered in [HM22, Appendix A.4], but the K_0 groups we use are slightly different (as discussed at the beginning of Appendix A.1). The setup we consider agrees with [HM22, Appendix A.4] for schemes.

Lemma A.2.1. *Consider a 2-commutative diagram of algebraic stacks*

$$\begin{tikzcd} \mathcal{Z} \times_{\mathcal{X}} \mathcal{W} \arrow[r] \arrow[d] \arrow[dr, "h"] & \mathcal{W} \arrow[d, "g"] \\ \mathcal{Z} \arrow[r, "f"] & \mathcal{X} \end{tikzcd} \tag{A.2.1}$$

with outer square 2-Cartesian, where $\mathcal{X}$ is a separated regular Noetherian Deligne–Mumford stack and the morphisms f and g (and hence h) are finite.

There is a bilinear pairing

$$\begin{aligned} K_0'(\mathcal{Z}) \times K_0'(\mathcal{W}) &\longrightarrow K_0'(\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}) \\ (\mathcal{F}, \mathcal{G}) &\longmapsto \textstyle\sum_i (-1)^i \mathrm{Tor}_i^{\mathcal{O}_{\mathcal{X}}}(f_* \mathcal{F}, g_* \mathcal{G}) \end{aligned} \tag{A.2.2}$$

where $\mathcal{F}$ and $\mathcal{G}$ stand for coherent $\mathcal{O}_{\mathcal{Z}}$-modules and coherent $\mathcal{O}_{\mathcal{W}}$-modules, respectively. We have a commutative diagram

$$\begin{tikzcd} K_0'(\mathcal{Z}) \times K_0'(\mathcal{W}) \arrow[r] \arrow[d, "f_* \times g_*"] & K_0'(\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}) \arrow[d, "h_*"] \\ K_0'(\mathcal{X}) \times K_0'(\mathcal{X}) \arrow[r] & K_0'(\mathcal{X}) \end{tikzcd} \tag{A.2.3}$$

where vertical arrows are pushforward and the lower horizontal arrow is the bilinear pairing from the ring structure on $K_0'(\mathcal{X}) \cong K_0(\mathcal{X})$.

Proof. If $\mathcal{F}$ is a coherent $\mathcal{O}_{\mathcal{Z}}$-module and $\mathcal{G}$ is a coherent $\mathcal{O}_{\mathcal{W}}$-module, we may form the object $(f_*\mathcal{F} \otimes^{\mathbb{L}} g_*\mathcal{G})$ in $D_{perf}(\mathcal{O}_{\mathcal{X}})$. For each object $U \to \mathcal{X}$ in the small étale site of $\mathcal{X}$ (i.e. U is a scheme with an étale morphism to $\mathcal{X}$), the restriction $(f_*\mathcal{F} \otimes^{\mathbb{L}} g_*\mathcal{G})|_U \in D_{perf}(\mathcal{O}_U)$ carries natural $\mathcal{O}_U$-linear actions of $(f_*\mathcal{O}_{\mathcal{Z}})|_U$ and $(f_*\mathcal{O}_{\mathcal{W}})|_U$. The resulting cohomology sheaves $\mathrm{Tor}_i^{\mathcal{O}_{\mathcal{X}}}(f_*\mathcal{F}, g_*\mathcal{G}) = H^{-i}(f_*\mathcal{F} \otimes^{\mathbb{L}} g_*\mathcal{G})$ (a priori coherent $\mathcal{O}_{\mathcal{X}}$-modules) are thus sheaves of $(f_*\mathcal{O}_{\mathcal{Z}}) \otimes_{\mathcal{O}_{\mathcal{X}}} (g_*\mathcal{O}_{\mathcal{W}})$-algebras. There is a canonical isomorphism $(f_*\mathcal{O}_{\mathcal{Z}}) \otimes_{\mathcal{O}_{\mathcal{X}}} (g_*\mathcal{O}_{\mathcal{W}}) \to h_*\mathcal{O}_{\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}}$ of $\mathcal{O}_{\mathcal{X}}$-algebras. Since h is affine, we obtain a lift (up to canonical isomorphism) of each $\mathrm{Tor}_i^{\mathcal{O}_{\mathcal{X}}}(f_*\mathcal{F}, g_*\mathcal{G})$ to a coherent sheaf of $\mathcal{O}_{\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}}$ modules (to pass between quasi-coherent $h_*\mathcal{O}_{\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}}$-modules and quasi-coherent $\mathcal{O}_{\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}}$-modules, we may take an étale surjection of $\mathcal{X}$ from a scheme, use the corresponding result for the small étale site of schemes which is [SProject, Lemma 08AI], and reduce to a statement about glueing data on the small étale sites of $\mathcal{Z} \times_{\mathcal{X}} \mathcal{W}$ and $\mathcal{X}$).

The procedure just described descends to K_0' groups and gives the pairing in the lemma statement. □

We think of the map $K_0'(\mathcal{Z}) \times K_0'(\mathcal{W}) \to K_0'(\mathcal{Z} \times_{\mathcal{X}} \mathcal{W})$ from the preceding lemma as an *intersection pairing* "with supports along finite morphisms".

Next, fix a base scheme S with dimension function δ as in Appendix A.1. Suppose $\mathcal{X}$ is a Deligne–Mumford stack which is quasi-separated and locally of finite type over S. We assume that $\mathcal{X}$ is equidimensional of dimension n, and we also assume that $\mathcal{X}$ admits a finite flat surjection from a scheme in order to define dimension and codimension filtrations as in Definition A.1.2.

Consider a finite morphism $f\colon \mathcal{Z} \to \mathcal{X}$ from a Deligne–Mumford stack $\mathcal{Z}$. We define a "relative codimension" filtration on $K_0'(\mathcal{Z})_{\mathbb{Q}}$ by setting

$$F^m_{\mathcal{X}} K_0'(\mathcal{Z})_{\mathbb{Q}} \coloneqq F_{n-m} K_0'(\mathcal{Z})_{\mathbb{Q}}. \tag{A.2.4}$$

We similarly set $\mathrm{gr}^m_{\mathcal{X}} K_0'(\mathcal{Z})_{\mathbb{Q}} \coloneqq F^m_{\mathcal{X}} K_0'(\mathcal{Z})_{\mathbb{Q}} / F^{m+1}_{\mathcal{X}} K_0'(\mathcal{Z})_{\mathbb{Q}}$. The subscript $\mathcal{X}$ is meant to remind of the dependence on $\mathcal{X}$.

Lemma A.2.2. *Let $\mathcal{X}$ be a regular Noetherian Deligne–Mumford stack which is separated and finite type over S. Assume that $\mathcal{X}$ is equidimensional. Let $f\colon \mathcal{Z} \to \mathcal{X}$ and $g\colon \mathcal{W} \to \mathcal{X}$ be finite morphisms from Deligne–Mumford stacks $\mathcal{Z}$ and $\mathcal{W}$.*

Assume that there exists a finite flat surjection $\pi\colon U \to \mathcal{X}$ with U a regular Noetherian scheme, such that $\mathcal{Z} \times_{\mathcal{X}} U \to U$ and $\mathcal{W} \times_{\mathcal{X}} U \to U$ are both disjoint unions of closed immersions. Then the intersection pairing of Lemma A.2.1 restricts to a pairing

$$F^s_{\mathcal{X}} K_0'(\mathcal{Z})_{\mathbb{Q}} \times F^t_{\mathcal{X}} K_0'(\mathcal{W})_{\mathbb{Q}} \to F^{s+t}_{\mathcal{X}} K_0'(\mathcal{Z} \times_{\mathcal{X}} \mathcal{W})_{\mathbb{Q}} \tag{A.2.5}$$

for any $s, t \in \mathbb{Z}$.

Proof. We use the shorthand $\mathcal{Z}_U \coloneqq \mathcal{Z} \times_{\mathcal{X}} U$ and $\mathcal{W}_U \coloneqq \mathcal{W} \times_{\mathcal{X}} U$. If we abuse notation and also write π for the natural projections $\mathcal{Z}_U \to \mathcal{Z}$ and $\mathcal{W}_U \to \mathcal{W}$ and $\mathcal{Z}_U \times_U \mathcal{W}_U \to \mathcal{Z} \times_{\mathcal{X}} \mathcal{W}$, we have $(\pi^* \alpha) \cdot (\pi^* \beta) = \pi^*(a \cdot \beta)$ for any $\alpha \in K_0'(\mathcal{Z})_{\mathbb{Q}}$ and $\beta \in K_0'(\mathcal{W})_{\mathbb{Q}}$. By definition of the dimension filtration (Definition A.1.2), it is enough to check that the intersection pairing over U restricts to

$$F_U^s K_0'(\mathcal{Z}_U)_{\mathbb{Q}} \times F_U^t K_0'(\mathcal{W}_U)_{\mathbb{Q}} \to F_U^{s+t} K_0'(\mathcal{Z}_U \times_U \mathcal{W}_U)_{\mathbb{Q}} \tag{A.2.6}$$

(i.e. respects filtrations). We have thus reduced to the case where $\mathcal{X}$ is a scheme and $\mathcal{Z} \to \mathcal{X}$ and $\mathcal{W} \to \mathcal{X}$ are disjoint unions of closed immersions, and we assume these conditions hold for the rest of the proof. Write $\mathcal{Z} = \coprod_i \mathcal{Z}_i$ where each $\mathcal{Z}_i \to \mathcal{X}$ is a closed immersion of schemes, and similarly write $\mathcal{W} = \coprod_j \mathcal{W}_j$. By a result of Gillet–Soulé [GS87, Proposition 5.5], the pairing $F_{\mathcal{X}}^s K_0'(\mathcal{Z}_i)_{\mathbb{Q}} \times F_{\mathcal{X}}^t K_0'(\mathcal{W}_j)_{\mathbb{Q}} \to K_0'(\mathcal{Z}_i \times_{\mathcal{X}} \mathcal{W}_j)_{\mathbb{Q}}$ factors through $F_{\mathcal{X}}^{s+t} K_0'(\mathcal{Z}_i \times_{\mathcal{X}} \mathcal{W}_j)_{\mathbb{Q}}$. We may decompose $F_{\mathcal{X}}^s K_0'(\mathcal{Z})_{\mathbb{Q}} = \bigoplus_i F_{\mathcal{X}}^s K_0'(\mathcal{Z}_i)_{\mathbb{Q}}$ and $F_{\mathcal{X}}^t K_0'(\mathcal{W})_{\mathbb{Q}} = \bigoplus_j F_{\mathcal{X}}^t K_0'(\mathcal{W}_j)_{\mathbb{Q}}$. Commutativity of the diagram

$$\begin{tikzcd}
F_{\mathcal{X}}^s K_0'(\mathcal{Z}_i)_{\mathbb{Q}} \times F_{\mathcal{X}}^t K_0'(\mathcal{W}_j)_{\mathbb{Q}} \arrow[r] \arrow[d] & F_{\mathcal{X}}^{s+t} K_0'(\mathcal{Z}_i \times_{\mathcal{X}} \mathcal{W}_j)_{\mathbb{Q}} \arrow[d] \\
(\bigoplus_i F_{\mathcal{X}}^s K_0'(\mathcal{Z}_i)_{\mathbb{Q}}) \times (\bigoplus_j F_{\mathcal{X}}^t K_0'(\mathcal{W}_j)_{\mathbb{Q}}) \arrow[r] & K_0'(\mathcal{Z} \times_{\mathcal{X}} \mathcal{W})_{\mathbb{Q}}
\end{tikzcd} \tag{A.2.7}$$

for each i, j gives the claim. ☐

B Miscellany on p-divisible groups

We collect some terminology/notation and miscellaneous facts about p-divisible groups, which we use freely.

B.1 Terminology

Suppose S is a formal scheme[48] and suppose $\mathcal{P}$ is a property of morphisms of schemes which is fppf local on the target and stable under arbitrary base-change. A sheaf X on $(Sch/S)_{fppf}$ is *represented by a relative scheme with property* $\mathcal{P}$ *over* S if, for every scheme T over S, the restriction sheaf $X|_T$ is represented by a scheme with property $\mathcal{P}$ over T.

[48]The formal schemes we use are the "préschémas formels" of [EGAI, §10]. Given a formal scheme, the notation $(Sch/S)_{fppf}$ means the site whose objects are morphisms $T \to S$ for schemes T, where coverings are fppf.

Fix a prime p. A *p-divisible group* over a formal scheme S is a sheaf X of abelian groups on $(Sch/S)_{fppf}$ which satisfies the following conditions.

(1) (p-divisibility) The multiplication by p map $[p]\colon X \to X$ is a surjection of sheaves.

(2) (p^∞-torsion) The natural map $X[p^\infty] := \varinjlim_n X[p^n] \to X$ is an isomorphism, where $X[p^n] \subseteq X$ are the p^n-torsion subsheaves.

(3) (representable p-power-torsion) The sheaves $X[p^n]$ are represented by finite locally free relative schemes over S for all $n \geq 1$.

If S is an adic (e.g. locally Noetherian) formal scheme and $\mathscr{I}$ is an ideal sheaf of definition on S, giving a p-divisible group over S is the same as giving p-divisible groups X_n over each scheme $S_n := (S, \mathcal{O}_S/\mathscr{I}^n)$ with isomorphisms $X_{n+1}|_{S_n} \xrightarrow{\sim} X_n$.

For a general formal scheme S, we say a p-divisible group X over S has *height* h if $X[p]$ is finite locally free relative scheme over S of degree p^h. In general, h is understood as a locally constant function on S.

If p is locally topologically nilpotent on S (equivalently, S is a formal scheme over $\operatorname{Spf} \mathbb{Z}_p$) and if X is a p-divisible group over S, there is an associated sheaf $\operatorname{Lie} X$ on $(Sch/S)_{fppf}$ (constructed as in [SGA3II, Definition 3.2]). By work of Messing [Mes72, Theorem 3.3.18], it is known that $\operatorname{Lie} X$ is a finite locally free sheaf of modules on $(Sch/S)_{fppf}$. We refer to the dual $\Omega_X := (\operatorname{Lie} X)^\vee$ as a *Hodge bundle*. If r is the rank of $\operatorname{Lie} X$, we say that X has *dimension* r (in general, r is a locally constant $\mathbb{Z}_{\geq 0}$-valued function). In this case, we write $\omega_X := \bigwedge^r \Omega_X$ for the top exterior power and also call ω_X a *Hodge bundle.*

If p is locally topologically nilpotent on the formal scheme S, a *formal p-divisible group* X over S is a p-divisible group over S such that, fppf (equivalently, Zariski) locally on any $T \in \operatorname{Obj}(Sch/S)_{fppf}$, the pointed fppf sheaf X is isomorphic to $\operatorname{Spf} \mathcal{O}_T[[x_1, \ldots, x_r]]$ for some r (possibly varying). See [Mes72, Proposition II.4.4] for equivalent characterizations.

Given p-divisible groups X and Y over a general formal scheme S, a *quasi-homomorphism* is a global section of the sheaf $\underline{\operatorname{Hom}}(X, Y) \otimes_{\mathbb{Z}} \mathbb{Q}$ on $(Sch/S)_{fppf}$. We write $\operatorname{Hom}^0(X, Y)$ for the space of quasi-homomorphisms $X \to Y$, and similarly $\operatorname{End}^0(X) = \operatorname{Hom}^0(X, X)$. Given a quasi-compact scheme T with a map $T \to S$, we have $\operatorname{Hom}^0(X_T, Y_T) = \operatorname{Hom}(X_T, Y_T) \otimes_{\mathbb{Z}} \mathbb{Q}$. If X and Y are equipped with an action by a ring R, then $\operatorname{Hom}^0_R(X, Y)$ will denote the R-linear quasi-homomorphisms.

A morphism $f\colon X \to Y$ of p-divisible groups over S is an *isogeny* if f is a surjection of fppf sheaves and $\ker f$ is represented by a finite locally free relative scheme over S. If $\ker f$ is finite locally free of rank p^r, we say that f has *degree* p^r and *height* r. A *quasi-isogeny*

$f\colon X \to Y$ is a quasi-homomorphism which, locally on $(Sch/S)_{fppf}$, is of the form $f = p^n g$ for $n \in \mathbb{Z}$ and an isogeny g. If the p-divisible group X has height h, such a quasi-isogeny $f = p^n g$ is said to have *degree* $p^{nh} \deg(g)$ and *height* $nh + \operatorname{height}(g)$. We write $\operatorname{Isog}(X, Y)$ (resp. $\operatorname{Isog}^0(X, Y)$) for the isogenies (resp. quasi-isogenies) $X \to Y$. We write $\operatorname{Isog}(X)$ (resp. $\operatorname{Isog}^0(X)$) for self-isogenies (resp. self quasi-isogenies) of X.

A p divisible group X over S is *étale* if $X[p]$ is an étale relative scheme. This implies that each $X[p^n]$ is an étale relative scheme. If R is a Noetherian Henselian local ring, we say that a p-divisible group X over $\operatorname{Spec} R$ is *connected* if $X[p]$ is connected. This implies that each $X[p^n]$ is connected.

Given any p-divisible group X over a general formal scheme S, there is a *dual* p-divisible group $X^\vee$. A *polarization* of X is an isogeny $\lambda\colon X \to X^\vee$ satisfying $\lambda^\vee = -\lambda$. The polarization is *principal* if λ is an isomorphism. A *quasi-polarization* is a quasi-isogeny $f\colon X \to X^\vee$ such that mf is a polarization for some $m \in \mathbb{Q}_p^\times$. Suppose X and Y are p-divisible groups over S with quasi-polarizations $\lambda_X\colon X \to X^\vee$ and $\lambda_Y\colon Y \to Y^\vee$. Given any $x \in \operatorname{Hom}^0(Y, X)$ with dual $x^\vee \in \operatorname{Hom}^0(X^\vee, Y^\vee)$, we set $x^\dagger \coloneqq \lambda_Y^{-1} \circ x^\vee \circ \lambda_X \in \operatorname{Hom}^0(X, Y)$, and call the resulting map $\dagger\colon \operatorname{Hom}^0(Y, X) \to \operatorname{Hom}^0(X, Y)$ the *Rosati involution*.

Over an algebraically closed field, we say that a p-divisible group is supersingular if all slopes of its isocrystal are equal to $1/2$, and we say that it is ordinary if all slopes of its isocrystal are either 0 or 1. A p-divisible group over an arbitrary formal scheme is *supersingular* (resp. *ordinary*) if it is supersingular (resp. ordinary) for every geometric fiber.

Over any algebraically closed field, there is a unique étale p-divisible group of height r (namely the constant sheaf $(\underline{\mathbb{Q}_p/\mathbb{Z}_p})^r$). Over any algebraically closed field of characteristic p, there is also a unique p-divisible group of height r with all slopes of its isocrystal being 1 (namely $\boldsymbol{\mu}_{p^\infty}^r \coloneqq (\varinjlim_e \boldsymbol{\mu}_{p^e})^r \cong (\underline{\mathbb{Q}_p/\mathbb{Z}_p}^\vee)^r$, given by p-th power roots of unity). Since the connected étale exact sequence of any p-divisible group over a perfect field is (canonically) split, we conclude that $\boldsymbol{\mu}_{p^\infty}^{n-r} \times (\underline{\mathbb{Q}_p/\mathbb{Z}_p})^r$ is the unique ordinary p-divisible group of height n and dimension $n - r$ over any algebraically closed field.

By *Drinfeld rigidity* we mean the following phenomenon: if $S_0 \to S$ is a finite order thickening of schemes over $\operatorname{Spf} \mathbb{Z}_p$, and X, Y are p-divisible groups over S, any quasi-homomorphism of $X \to Y$ over S_0 lifts uniquely to a quasi-homomorphism over S [And03, Theorem 2.2.3] (alternative proof: Grothendieck–Messing theory).

If A is a relative abelian scheme over a general formal scheme S, there is an associated p-divisible group $A[p^\infty] \coloneqq \varinjlim_n A[p^n]$, where $A[p^n]$ is the p^n-torsion subfunctor of A. If p is locally topologically nilpotent on S, there is a canonical identification $\operatorname{Lie} A \cong \operatorname{Lie} A[p^\infty]$.

Given a p-divisible group X over a formal scheme S and given a finite free $\mathbb{Z}_p$-module M

of some rank $d \geq 0$, there is the *Serre tensor construction* p-divisible group $X \otimes_{\mathbb{Z}_p} M$ given by the functor

$$(X \otimes_{\mathbb{Z}_p} M)(T) := X(T) \otimes_{\mathbb{Z}_p} M \tag{B.1.1}$$

for schemes T over S. Any choice of $\mathbb{Z}_p$-basis for M gives an isomorphism $X \otimes_{\mathbb{Z}_p} M \cong X^d$ as p-divisible groups. This construction is functorial in M: in particular, any $\mathbb{Z}_p$-algebra R acting on M also acts on $X \otimes_{\mathbb{Z}_p} M$. The resulting R-action on $X \otimes_{\mathbb{Z}_p} M$ is the *Serre tensor R-action*. There is a canonical identification $(X \otimes_{\mathbb{Z}_p} M)^\vee \cong X^\vee \otimes_{\mathbb{Z}_p} M^\vee$ where $M^\vee := \mathrm{Hom}_{\mathbb{Z}_p}(M, \mathbb{Z}_p)$. More generally, see [Con04, §7].

B.2 Isogeny criterion

We explain a criterion for a morphism of p-divisible groups to be an isogeny (Lemma B.2.2). This should be well-known.[49]

Lemma B.2.1. *Let S be a scheme, and let H, G, and Q be commutative group schemes over S which are locally of finite presentation. Suppose*

$$0 \to H \to G \xrightarrow{f} Q \to 0$$

is an exact sequence of fppf sheaves of abelian groups. If $G \to S$ is finite locally free and $Q \to S$ is separated, then

(1) The map $f \colon G \to Q$ is finite locally free.

(2) The group schemes Q and H are finite locally free over S.

Proof. Since f is a surjection of fppf sheaves, it is a surjection on underlying topological spaces. We also know that f is locally of finite presentation because both G and Q are locally of finite presentation over S [SProject, Lemma 00F4]. Since $G \to S$ is flat, the fibral flatness criterion [EGAIV3, 11.3.11] implies that flatness of f may be checked fiberwise over S, i.e. it is enough to check flatness of the base-change $G_{k(s)} \to Q_{k(s)}$ for each $s \in S$. The exact sequence

$$0 \to H_{k(s)} \to G_{k(s)} \to Q_{k(s)} \to 0$$

shows that $G_{k(s)} \to Q_{k(s)}$ is an $H_{k(s)}$-torsor in the fppf topology, hence flat. This shows that f is fppf. Since $Q \to S$ is separated and $G \to S$ is finite, we know that f is also finite, hence finite locally free. Moreover, the fibral flatness criterion also implies that Q is flat over S. We also conclude that $Q \to S$ is proper via [SProject, Lemma 03GN].

[49]The only reference I know is the sketch in [Far05, Lemme 9]. We spell out the argument for completeness.

Since $H = \ker(f)$ and f is an fppf morphism, we know $H \to S$ is fppf as well. Since $Q \to S$ is separated, the identity section $S \to Q$ is a closed immersion, hence $H = \ker(f)$ is a closed subscheme of G. Since $G \to S$ is finite, we conclude that $H \to S$ is also finite, hence finite locally free.

We have already seen that $Q \to S$ is flat, proper, and locally of finite presentation. To check that $Q \to S$ is finite, it is enough to check that it has finite fibers, which follows because $G \to Q$ is surjective and $G \to S$ is finite. $\square$

Lemma B.2.2. *Let S be a formal scheme. Let $f\colon X \to Y$ be a homomorphism of p-divisible groups over S. Then f is an isogeny if and only if, locally on $(Sch/S)_{fppf}$, there exists a homomorphism $g\colon Y \to X$ such that*

$$g \circ f = [p^N] \qquad f \circ g = [p^N]$$

for some integer $N \geq 0$, where $[p^N]$ denotes multiplication by p^N.

Moreover, given an isogeny f, such g, N will exist globally on S if S is quasi-compact or has finitely many connected components. If f is an isogeny of constant degree p^n, we may take $N = n$.

Proof. If $f\colon X \to Y$ is an isogeny, then Y is the fppf sheaf quotient of X by $\ker(f)$. If S is a quasi-compact formal scheme or if S has finitely many connected components, we have $\ker f \subseteq X[p^N]$ for N large, so $g \circ f = [p^N]$ for some homomorphism $g\colon Y \to X$. We also have $f \circ g \circ f = [p^N] \circ f$. Since f is an epimorphism of fppf sheaves, we conclude that $f \circ g = [p^N]$.

Conversely, suppose that locally on $(Sch/S)_{fppf}$ there exists a homomorphism $g\colon Y \to X$ and an integer $N \geq 0$ as in the lemma statement. Since the property of being an isogeny may be checked locally on $(Sch/S)_{fppf}$, we may assume that S is a scheme and that g, N exist globally on S. Since $f \circ g = [p^N]$, we see that f is a surjection of fppf sheaves. It remains only to check that $\ker f$ is representable by a finite locally free group scheme over S.

We know that $\ker(f) \subseteq X[p^N]$ and $\ker(g) \subseteq Y[p^N]$. We have $\ker(f) = \ker(X[p^N] \to Y[p^N])$ and $\ker(g) = \ker(Y[p^N] \to X[p^n])$. Since $X[p^N]$ and $Y[p^N]$ represented by finite locally free group schemes over S, we see that $\ker(f)$ and $\ker(g)$ are represented by schemes which are finite and locally of finite presentation over S.

We have short exact sequences

$$0 \to \ker(f) \to X[p^N] \xrightarrow{f} \ker(g) \to 0$$
$$0 \to \ker(g) \to Y[p^N] \xrightarrow{g} \ker(f) \to 0$$

of fppf sheaves of abelian groups. By Lemma B.2.1, we conclude that $\ker(f)$ and $\ker(g)$ are finite locally free group schemes over S. $\square$

Lemma B.2.3. *Let S be a formal scheme. Let X and Y be p-divisible groups over S. Then $f \in \operatorname{Hom}^0(X, Y)$ is a quasi-isogeny if and only if it is invertible, meaning there exists $g \in \operatorname{Hom}^0(Y, X)$ (necessarily unique) with $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.*

Proof. Invertibility and the property of being a quasi-isogeny can both be checked locally on $(Sch/S)_{fppf}$, so the lemma follows from Lemma B.2.2. ☐

B.3 p-divisible groups over Spec A and Spf A

The following facts are implicitly used, e.g. throughout Parts II and III.

Lemma B.3.1. *Let A be an adic Noetherian ring. There are equivalences of categories*

$$\{\textit{finite schemes over } \operatorname{Spec} A\} \to \{\textit{finite relative schemes over } \operatorname{Spf} A\}$$
$$\{\textit{finite locally free schemes over } \operatorname{Spec} A\} \to \{\textit{finite locally free relative schemes over } \operatorname{Spf} A\}$$
$$\{p\textit{-divisible groups over } \operatorname{Spec} A\} \to \{p\textit{-divisible groups over } \operatorname{Spf} A\}$$

given by base change, i.e. restriction of fppf sheaves along the inclusion $(Sch/\operatorname{Spf} A)_{fppf} \to (Sch/\operatorname{Spec} A)_{fppf}$.

Proof. For the statements about finite relative schemes, the quasi-inverse functor is given by $\operatorname{Spf} R \mapsto \operatorname{Spec} R$ for finite A-algebras R (topologized so that R is an adic ring and the map $A \to R$ is adic). This also gives the quasi-inverse functor for finite locally free relative schemes (check using the local criterion for flatness). For the statement about p-divisible groups (which follows from the other statements), see [Mes72, 4.15, Lemma II.4.16] or [dJo95, Lemma 2.4.4]. ☐

Lemma B.3.2. *Let A be an adic Noetherian ring, and let $\phi \colon X \to Y$ be a homomorphism of p-divisible groups over $\operatorname{Spec} A$. Then ϕ is an isogeny if and only if $\phi_{\operatorname{Spf} A} \colon X_{\operatorname{Spf} A} \to Y_{\operatorname{Spf} A}$ is an isogeny.*

Proof. Follows from Lemma B.3.1 and the isogeny criterion from Lemma B.2.2. ☐

For adic Noetherian rings A, we may thus pass between p-divisible groups over $\operatorname{Spec} A$ and $\operatorname{Spf} A$ without loss of information, and similarly for finite locally free relative schemes. We abuse notation in this way: for example, if A is a domain, the *generic fiber* of a p-divisible group over $\operatorname{Spf} A$ will refer to its generic fiber as a p-divisible group over $\operatorname{Spec} A$.

To avoid potential confusion, we remark on three situations where p-divisible groups may have different properties when considered over $\operatorname{Spec} A$ versus over $\operatorname{Spf} A$.

Remark B.3.3. Let A be an adic Noetherian ring, and suppose p is topologically nilpotent in A. Let X be a p-divisible group over $\operatorname{Spec} A$. By work of Messing, [Mes72, §II], the sheaf $\operatorname{Lie}(X_{\operatorname{Spf} A})$ (in the sense of [SGA3II]) is locally free of finite rank on $(Sch/\operatorname{Spf} A)_{fppf}$. However, $\operatorname{Lie} X$ (viewed as a sheaf on $(Sch/\operatorname{Spec} A)_{fppf}$) is *not* necessarily locally free.

For example, consider $A = \mathbb{Z}_p$ and $X = \boldsymbol{\mu}_{p^\infty} := \varinjlim \boldsymbol{\mu}_{p^n}$, where $\boldsymbol{\mu}_{p^n}$ is the group scheme of p^n-th roots of unity. Then the p-divisible group X over $\operatorname{Spec} \mathbb{Z}_p$ is étale in the generic fiber, but connected of dimension 1 in the special fiber. We find that $\operatorname{Lie} X|_{\operatorname{Spec} \mathbb{Q}_p} = 0$ but $\operatorname{Lie} X|_{\operatorname{Spec} \mathbb{F}_p}$ is free of rank 1, so $\operatorname{Lie} X$ cannot be a locally free sheaf of modules on $(Sch/\operatorname{Spec} A)_{fppf}$.

Thus, when writing $\operatorname{Lie} X$ in this situation, we always mean (by abuse of notation) to view X as a p-divisible group over $\operatorname{Spf} A$, so that $\operatorname{Lie} X$ will be a finite locally free sheaf on $(Sch/\operatorname{Spf} A)_{fppf}$. Similarly, if we say X has dimension r, we mean that the finite locally free sheaf $\operatorname{Lie} X$ on $(Sch/\operatorname{Spf} A)_{fppf}$ has rank r.

Remark B.3.4. Let A be an adic Noetherian ring, and let X be a p-divisible group over $\operatorname{Spec} A$. In general, there are sections of $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ which do not arise as sections of $X \to \operatorname{Spec} A$. Indeed, sections of $X \to \operatorname{Spec} A$ correspond precisely to torsion sections of $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ (use quasi-compactness of $\operatorname{Spec} A$). But $X_{\operatorname{Spf} A} \to \operatorname{Spf} A$ may have many non-torsion sections, e.g. when $A = \mathbb{Z}_p$ and $X_{\operatorname{Spf} A}$ is a formal p-divisible group, hence $X_{\operatorname{Spf} A} \cong \operatorname{Spf} \mathbb{Z}_p[[X_1, \ldots, X_r]]$ as pointed fppf sheaves on $(Sch/\operatorname{Spf} \mathbb{Z}_p)_{fppf}$. There will be uncountably many non-torsion sections in this situation. This makes a difference in Section 6.1, for example, where some statements are correct over $\operatorname{Spf} R$ (which is the written version) but incorrect over $\operatorname{Spec} R$.

Remark B.3.5. Let A be an adic Noetherian ring. By our conventions, it is *not* true that any quasi-homomorphism of p-divisible groups over $\operatorname{Spf} A$ necessarily lifts to a quasi-homomorphism of p-divisible groups over $\operatorname{Spec} A$. See Example 7.1.1. On the other hand, homomorphisms and isogenies will lift (uniquely) by the preceding lemmas.

C Quasi-compactness of special cycles

Besides fixing notation, the purpose of this appendix is to prove a quasi-compactness statement for special cycles (explicit proofs of other properties, e.g. having finite fibers, are more readily available in the literature, e.g. [KR14, Proposition 2.9]).[50] A similar proof of quasi-

[50]It is possible to have a locally of finite type morphism of locally Noetherian schemes which has finite fibers and satisfies the valuative criterion for properness, but which is not quasi-compact.

compactness (in the context of special divisors on some orthogonal Shimura varieties) is [AGHMP17, Proposition 2.7.2].

C.1 Terminology

Suppose A and B are abelian schemes over a base scheme S. We write $\underline{\mathrm{Hom}}(A, B)$ for the fppf sheaf (on S) of homomorphisms of abelian schemes. Then the sheaf of *quasi-homomorphisms* is $\underline{\mathrm{Hom}}^0(A, B) := \underline{\mathrm{Hom}}(A, B) \otimes_{\mathbb{Z}} \mathbb{Q}$. We write $\mathrm{Hom}^0(A, B)$ for the space of global sections and call elements $x \in \mathrm{Hom}^0(A, B)$ *quasi-homomorphisms*, sometimes writing $x \colon A \to B$. If S is quasi-compact, we have $\mathrm{Hom}^0(A, B) = \mathrm{Hom}(A, B) \otimes_{\mathbb{Z}} \mathbb{Q}$. When $A = B$, we often use the notation $\underline{\mathrm{End}}(A)$, $\underline{\mathrm{End}}^0(A)$, and $\mathrm{End}^0(A)$ instead, and often use the term *quasi-endomorphism*. We write $\mathrm{Isog}(A, B)$ for the set of isogenies $A \to B$. We write $\underline{\mathrm{Isog}}(A, B) \subseteq \underline{\mathrm{Hom}}(A, B)$ for the subsheaf of sets consisting of isogenies, and $\underline{\mathrm{Isog}}^0(A, B) \subseteq \underline{\mathrm{Hom}}^0(A, B)$ for the subsheaf of *quasi-isogenies*, meaning those quasi-homomorphisms which are locally of the form mf for some isogeny f and some nonzero integer $m \in \mathbb{Z}$. We write $\mathrm{Isog}(A, B)$ (resp. $\mathrm{Isog}^0(A, B)$) for the set of *isogenies* (resp. *quasi-isogenies*), consisting of global sections of $\underline{\mathrm{Isog}}(A, B)$ (resp. $\underline{\mathrm{Isog}}^0(A, B)$). We write $\mathrm{Isog}(A)$ and $\mathrm{Isog}^0(A)$ for the self-isogenies and self quasi-isogenies of A. A *quasi-polarization* of A is a quasi-isogeny $A \to A^\vee$ which is locally of the form $m\lambda$ for some polarization λ and some positive integer $m \in \mathbb{Z}_{>0}$.

Suppose the abelian schemes A and B are equipped with quasi-polarizations $\lambda_A \colon A \to A^\vee$ and $\lambda_B \colon B \to B^\vee$. Given any $x \in \mathrm{Hom}^0(B, A)$ with dual $x^\vee \in \mathrm{Hom}^0(A^\vee, B^\vee)$, we set $x^\dagger := \lambda_B^{-1} \circ x^\vee \circ \lambda_A \in \mathrm{Hom}^0(A, B)$, and call the resulting map $\dagger \colon \mathrm{Hom}^0(B, A) \to \mathrm{Hom}^0(A, B)$ the *Rosati involution*. Given m-tuples $\underline{x}, \underline{y} \in \mathrm{Hom}^0(B, A)^m$ with $\underline{x} = [x_1, \ldots, x_m]$ and $\underline{y} = [y_1, \ldots, y_m]$, we write $(\underline{x}, \underline{y})$ for the $m \times m$ matrix whose i, j-th entry is $x_i^\dagger y_j$. We say that $(\underline{x}, \underline{x})$ is the *Gram matrix* of $\underline{x}$. If $S = \operatorname{Spec} k$ for a field k, the $\mathbb{Q}$-bilinear pairing

$$\begin{aligned} \mathrm{Hom}^0(B, A) \times \mathrm{Hom}^0(B, A) &\longrightarrow \mathbb{Q} \\ x, y &\longmapsto \mathrm{tr}(x^\dagger y) \end{aligned} \tag{C.1.1}$$

is symmetric and positive definite ("positivity of the Rosati involution"), where $\mathrm{tr} \colon \mathrm{End}^0(A) \to \mathbb{Q}$ is the trace for $\mathrm{End}^0(A)$ acting on the $\mathbb{Q}$-vector space $\mathrm{End}^0(A)$ by left multiplication.

C.2 Proof

We continue in the setup of Appendix C.1.

Given any $y \in \mathrm{End}^0(B)$, define a functor $\mathcal{Z}(y)\colon (Sch/S)^{\mathrm{op}} \to Set$ as

$$\mathcal{Z}(y) \coloneqq \{x \in \underline{\mathrm{Hom}}(B, A) : x^\dagger x = y\}. \tag{C.2.1}$$

We will check that $\mathcal{Z}(y)$ is representable by a scheme which is finite, unramified, and of finite presentation over S.

Lemma C.2.1. *The functor $\mathcal{Z}(y)$ is represented by a scheme over S. The structure morphism $\mathcal{Z}(y) \to S$ is separated and locally of finite presentation.*

Proof. By a standard limit argument (e.g. using [SProject, Lemma 01ZM]) we may reduce to the case where S is Noetherian, affine, and connected. It is also enough to check the case where λ_A and λ_B are polarizations, not just quasi-polarizations.

Existence of the product polarization $\lambda_B \times \lambda_A$ on $B \times A$ implies that $B \times A$ admits a relatively ample line bundle over S. Thus the Hilbert functor $\mathrm{Hilb}_{B \times A}$ is represented by a scheme, each of whose connected components is locally projective over S (in the sense of [SProject, Definition 01W8]), see [Nit05, Theorem 5.15] and [SProject, Lemma 0DPF]. By [SProject, Lemma 0D1B], we know there is a locally closed immersion

$$\mathcal{Z}(y) \to \mathrm{Hilb}_{B \times A}$$

which sends $x\colon B \to A$ to its graph $(1 \times x)\colon B \to B \times A$. In particular, $\mathcal{Z}(y)$ is represented by a scheme which is separated and locally of finite presentation over S. $\square$

Lemma C.2.2. *The structure morphism $\mathcal{Z}(y) \to S$ is quasi-compact.*

Proof. Again, we may reduce to the case where S is affine, Noetherian, and connected by a standard limit argument. It is also enough to check the case where λ_A and λ_B are polarizations, not just quasi-polarizations.

Consider the graph morphisms

$$B \xrightarrow{1 \times \lambda_B} B \times B^\vee$$
$$A \xrightarrow{1 \times \lambda_A} A \times A^\vee \; .$$

If $\mathcal{P}_B$ and $\mathcal{P}_A$ denote the Poincaré bundles on $B \times B^\vee$ and $A \times A^\vee$ respectively, we know that $\mathcal{L}_B \coloneqq (1 \times \lambda_B)^* \mathcal{P}_B$ and $\mathcal{L}_A \coloneqq (1 \times \lambda_A)^* \mathcal{P}_A$ are relatively ample line bundles on B and A, respectively, over S. If $\pi_B\colon B \times A \to B$ and $\pi_A\colon B \times A \to A$ are the natural projections, we know $\mathcal{E} \coloneqq \pi_B^* \mathcal{L}_B \otimes \pi_A^* \mathcal{L}_A$ is a relatively ample line bundle on $B \times A$. Moreover, $\mathcal{E}$ is isomorphic to the pullback of the Poincaré bundle $\mathcal{P}_{B \times A}$ along the graph of the polarization

$\lambda_B \times \lambda_A$ of $B \times A$. Let $m \in \mathbb{Z}_{\geq 1}$ be any integer such that $m \cdot \lambda_B$ and $m^2 \cdot y$ are both honest homomorphisms (rather than quasi-homomorphisms).

As above, write $\mathrm{Hilb}_{B \times A}$ for the Hilbert scheme associated with $B \times A$. Given a numerical polynomial $P \colon \mathbb{Z} \to \mathbb{Z}$, we write $\mathrm{Hilb}^P_{B \times A} \subseteq \mathrm{Hilb}_{B \times A}$ for the open and closed subscheme corresponding to the Hilbert polynomial P with respect to the line bundle $\mathcal{E}^{\otimes m^2}$ on $B \times A$. That is, for a S-scheme T, we have

$$\mathrm{Hilb}^P_{B \times A}(T) \coloneqq \{ Z \in \mathrm{Hilb}_{B \times A}(T) : \chi(Z_t, \mathcal{E}^{\otimes m^2 n}|_{Z_t}) = P(n) \text{ for all } n \in \mathbb{Z} \text{ and } t \in T \}$$

(where Z_t is the fiber of $Z \to T$ over $t \in T$ and χ denotes Euler characteristic). We know that $\mathrm{Hilb}^P_{B \times A}(T)$ is locally projective over S [Nit05, Theorem 5.15], hence quasi-compact over S.

As in the proof of Lemma C.2.1, there is a locally closed immersion $\mathcal{Z}(y) \to \mathrm{Hilb}_{B \times A}$ which sends $x \in \mathcal{Z}(y)$ to its graph $1 \times x \colon B \to B \times A$. To show that $\mathcal{Z}(y)$ is quasi-compact, it is enough to check that $\mathcal{Z}(y) \to \mathrm{Hilb}_{B \times A}$ factors through $\mathrm{Hilb}^P_{B \times A}$ for some fixed numerical polynomial P (possibly depending on y).

Consider the line bundle $\mathcal{F} \coloneqq \mathcal{L}_B^{\otimes m^2} \otimes ((1 \times \lambda_B)^*(m^2 y \times 1)^* \mathcal{P}_B)$ on B. For any point $s \in S$, there is a numerical polynomial $P \colon \mathbb{Z} \to \mathbb{Z}$ such that

$$P(n) = \chi(B_s, \mathcal{F}^{\otimes n}|_{B_s}) \quad \text{for all } n \in \mathbb{Z} \tag{C.2.2}$$

as in [SProject, Lemma 0BEM]. The polynomial P does not depend on s because S is connected and the Euler characteristics are locally constant as a function of s (using flatness and properness and the standard facts [SProject, Lemma 0BDJ] and [SProject, Section 07VJ]).

Let T be a scheme over S, and suppose $x \in \mathcal{Z}(y)(T)$. View x as an element of $\mathrm{Hilb}_{B \times A}(T)$ as above. We claim that $x \in \mathrm{Hilb}^P_{B \times A}(T)$. By taking a base-change to T, we may assume $T = S$ without loss of generality (to lighten notation). It is enough to check $\mathcal{F} \cong (1 \times x)^* \mathcal{E}^{\otimes m^2}$.

First observe $(1 \times x)^* \mathcal{E}^{\otimes m^2} \cong \mathcal{L}_B^{\otimes m^2} \otimes x^* \mathcal{L}_A^{\otimes m^2}$. It is thus enough to verify the identity $x^* \mathcal{L}_A^{\otimes m^2} \cong (1 \times \lambda_B)^*(m^2 y \times 1)^* \mathcal{P}_B$. Consider the commutative diagram

$$\begin{tikzcd}
A \arrow[r, "1\times\lambda_A"] & A \times A^\vee & \\
B \arrow[u, "mx"] \arrow[r, "1\times\lambda_B"] & B \times B^\vee \arrow[u, "mx\times mx^{\dagger\vee}"'] \arrow[r, "mx\times 1"] \arrow[d, "m^2y\times 1"'] & A \times B^\vee \arrow[ul, "1\times mx^{\dagger\vee}"'] \arrow[dl, "mx^\dagger\times 1"] \\
 & B \times B^\vee &
\end{tikzcd}.$$

There exists an isomorphism $(mx^\dagger \times 1)^* \mathcal{P}_B \cong (1 \times mx^{\dagger\vee})^* \mathcal{P}_A$ (this characterizes $mx^{\dagger\vee}$ as the dual of $mx^\dagger$). Recall also that $m^* \mathcal{L}_A \cong \mathcal{L}_A^{\otimes m^2}$ (consider a similar diagram as above, with

$A = B$ and $x = y = 1$, and recall that the pullback of $\mathcal{P}_A$ along $(m \times 1)\colon A \times A^\vee \to A \times A^\vee$ is isomorphic to $\mathcal{P}_A^{\otimes m^2}$ because $m = m^\vee$). These facts prove the claimed identity $x^*\mathcal{L}_A^{\otimes m^2} \cong (1 \times \lambda_B)^*(m^2 y \times 1)^*\mathcal{P}_B$. □

Lemma C.2.3. *The functor $\mathcal{Z}(y)$ is represented by a scheme over S, and the structure morphism $\mathcal{Z}(y) \to S$ is finite, unramified, and of finite presentation.*

Proof. Again, we may reduce to the case where S is Noetherian by a standard limit argument. By Lemmas C.2.1 and C.2.2, we already know that $\mathcal{Z}(y)$ is represented by a scheme which is separated and of finite presentation over S.

To see that $\mathcal{Z}(y) \to S$ is proper, we can use the valuative criterion for discrete valuation rings [SProject, Lemma 0207] because S is Noetherian. This valuative criterion holds by the Néron mapping property for abelian schemes over discrete valuation rings.

For unramifiedness, it is enough to check that $\mathcal{Z}(y) \to S$ is formally unramified (i.e. satisfies the infinitesimal lifting criterion of [SProject, Lemma 02HE]). Formal unramifiedness holds because of rigidity for morphisms of abelian schemes as in [MFK94, Corollary 6.2].

Since unramified morphisms of schemes are locally quasi-finite, and since proper locally quasi-finite morphisms of schemes are finite, the lemma is proved. □

Recall that if $\mathcal{X}$ and $\mathcal{Y}$ are categories fibered in groupoids over the fppf site of some base scheme with $\mathcal{Y}$ being a Deligne–Mumford stack, and if $f\colon \mathcal{X} \to \mathcal{Y}$ is a morphism which is representable by algebraic spaces, then $\mathcal{X}$ is also a Deligne–Mumford stack [SProject, Comment 2142]. This can be used in combination with Lemma C.2.3 to verify that various stacks in this work are Deligne–Mumford.

References

[AGHMP17] Fabrizio Andreatta, Eyal Z. Goren, Benjamin Howard, and Keerthi Madapusi Pera. “Height pairings on orthogonal Shimura varieties”. In: *Compos. Math.* 153.3 (2017), pp. 474–534.

[AGHMP18] Fabrizio Andreatta, Eyal Z. Goren, Benjamin Howard, and Keerthi Madapusi Pera. “Faltings heights of abelian varieties with complex multiplication”. In: *Ann. of Math. (2)* 187.2 (2018), pp. 391–531.

[And03] Yves André. *Period mappings and differential equations. From $\mathbb{C}$ to $\mathbb{C}_p$.* Vol. 12. MSJ Memoirs. Tôhoku-Hokkaidô lectures in arithmetic geometry, With appendices by F. Kato and N. Tsuzuki. Mathematical Society of Japan, Tokyo, 2003, pp. viii+246.

[Aut03] Pascal Autissier. “Hauteur des correspondances de Hecke”. In: *Bull. Soc. Math. France* 131.3 (2003), pp. 421–433.

[BBK07] Jan H. Bruinier, José I. Burgos Gil, and Ulf Kühn. “Borcherds products and arithmetic intersection theory on Hilbert modular surfaces”. In: *Duke Math. J.* 139.1 (2007), pp. 1–88.

[BBM82] Pierre Berthelot, Lawrence Breen, and William Messing. *Théorie de Dieudonné cristalline. II.* Vol. 930. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1982, pp. x+261.

[BF06] Jan Hendrik Bruinier and Jens Funke. “Traces of CM values of modular functions”. In: *J. Reine Angew. Math.* 594 (2006), pp. 1–33.

[BGS94] Jean-Benoît Bost, Henri Gillet, and Christophe Soulé. “Heights of projective varieties and positive Green forms”. In: *J. Amer. Math. Soc.* 7.4 (1994), pp. 903–1027.

[BH21] Jan Hendrik Bruinier and Benjamin Howard. *Arithmetic volumes of unitary Shimura varieties.* 2021. arXiv: 2105.11274v2 [math.NT].

[BHKRY20] Jan H. Bruinier, Benjamin Howard, Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. “Modularity of generating series of divisors on unitary Shimura varieties”. In: *Astérisque* 421, Diviseurs arithmétiques sur les variétés orthogonales et unitaires de Shimura (2020), pp. 7–125.

[BHKRY20II] Jan H. Bruinier, Benjamin Howard, Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. “Modularity of generating series of divisors on unitary Shimura varieties II: Arithmetic applications”. In: *Astérisque* 421, Diviseurs arithmétiques sur les variétés orthogonales et unitaires de Shimura (2020), pp. 127–186.

[BY21] Jan Hendrik Bruinier and Tonghai Yang. “Arithmetic degrees of special cycles and derivatives of Siegel Eisenstein series”. In: *J. Eur. Math. Soc. (JEMS)* 23.5 (2021), pp. 1613–1674.

[CLZ] Ryan Chen, Weixiao Lu, and Wei Zhang. *Faltings heights and the subleading terms of adjoint L-functions*. In preparation.

[Con04] Brian Conrad. “Gross-Zagier revisited”. In: *Heegner points and Rankin L-series*. Vol. 49. Math. Sci. Res. Inst. Publ. With an appendix by W. R. Mann. Cambridge Univ. Press, Cambridge, 2004, pp. 67–163.

[Deu41] Max Deuring. “Die Typen der Multiplikatorenringe elliptischer Funktionenkörper”. In: *Abh. Math. Sem. Hansischen Univ.* 14 (1941), pp. 197–272.

[dJo95] Aise Johan de Jong. “Crystalline Dieudonné module theory via formal and rigid geometry”. In: *Inst. Hautes Études Sci. Publ. Math.* 82 (1995), 5–96 (1996).

[DM69] Pierre Deligne and David Mumford. “The irreducibility of the space of curves of given genus”. In: *Inst. Hautes Études Sci. Publ. Math.* 36 (1969), pp. 75–109.

[DY19] Tuoping Du and Tonghai Yang. “Arithmetic Siegel-Weil formula on $X_0(N)$”. In: *Adv. Math.* 345 (2019), pp. 702–755.

[EGAI] Alexander Grothendieck. “Éléments de géométrie algébrique. I. Le langage des schémas”. In: *Inst. Hautes Études Sci. Publ. Math.* 4 (1960), p. 228.

[EGAIII1] Alexander Grothendieck. “Éléments de géométrie algébrique. III. Étude cohomologique des faisceaux cohérents. I”. In: *Inst. Hautes Études Sci. Publ. Math.* 11 (1961), p. 167.

[EGAIV3] Alexander Grothendieck. “Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. III”. In: *Inst. Hautes Études Sci. Publ. Math.* 28 (1966), p. 255.

[Eme20] Matthew Emerton. *Formal algebraic stacks.* http://web.archive.org/web/20220709032216/http://www.math.uchicago.edu/~emerton/pdffiles/formal-stacks.pdf (archived July 2022). 2020.

[Fal86] Gerd Faltings. “Finiteness theorems for abelian varieties over number fields”. In: *Arithmetic geometry (Storrs, Conn., 1984).* Translated from the German original [Invent. Math. **7**3 (1983), no. 3, 349–366; MR0718935; ibid. **7**5 (1984), no. 2, 381; MR0732554] by Edward Shipz. Springer, New York, 1986, pp. 9–27.

[Far05] Laurent Fargues. *A Course on Lubin-Tate spaces and p-divisible groups. 2nd Course: p-divisible groups, formal Lie groups, Dieudonné theory.* http://web.archive.org/web/20230113023653/https://webusers.imj-prg.fr/~laurent.fargues/Notes_Cours.html (archived January 2023). Course notes. 2005.

[FC90] Gerd Faltings and Ching-Li Chai. *Degeneration of abelian varieties.* Vol. 22. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. With an appendix by David Mumford. Springer-Verlag, Berlin, 1990, pp. xii+316.

[FYZ21] Tony Feng, Zhiwei Yun, and Wei Zhang. *Higher theta series for unitary groups over function fields.* To appear in *Ann. Sci. Éc. Norm. Supér.* 2021. arXiv: 2110.07001v2 `[math.NT]`.

[FYZ24] Tony Feng, Zhiwei Yun, and Wei Zhang. “Higher Siegel–Weil formula for unitary groups: the non-singular terms”. In: *Invent. Math.* 235.2 (2024), pp. 569–668.

[Gil09] Henri Gillet. “Arithmetic intersection theory on Deligne-Mumford stacks”. In: *Motives and algebraic cycles.* Vol. 56. Fields Inst. Commun. Amer. Math. Soc., Providence, RI, 2009, pp. 93–109.

[GK93] Benedict H. Gross and Kevin Keating. “On the intersection of modular correspondences”. In: *Invent. Math.* 112.2 (1993), pp. 225–245.

[Gro78] Benedict H. Gross. “On the periods of abelian integrals and a formula of Chowla and Selberg”. In: *Invent. Math.* 45.2 (1978). With an appendix by David E. Rohrlich, pp. 193–211.

[Gro86] Benedict H. Gross. "On canonical and quasicanonical liftings". In: *Invent. Math.* 84.2 (1986), pp. 321–326.

[Gro97] Benedict H. Gross. "On the motive of a reductive group". In: *Invent. Math.* 130.2 (1997), pp. 287–313.

[GS19] Luis E. Garcia and Siddarth Sankaran. "Green forms and the arithmetic Siegel-Weil formula". In: *Invent. Math.* 215.3 (2019), pp. 863–975.

[GS87] Henri Gillet and Christophe Soulé. "Intersection theory using Adams operations". In: *Invent. Math.* 90.2 (1987), pp. 243–277.

[GS90] Henri Gillet and Christophe Soulé. "Arithmetic intersection theory". In: *Inst. Hautes Études Sci. Publ. Math.* 72 (1990), 93–174 (1991).

[GZ86] Benedict H. Gross and Don B. Zagier. "Heegner points and derivatives of L-series". In: *Invent. Math.* 84.2 (1986), pp. 225–320.

[HKS96] Michael Harris, Stephen S. Kudla, and William J. Sweet. "Theta dichotomy for unitary groups". In: *J. Amer. Math. Soc.* 9.4 (1996), pp. 941–1004.

[HLSY23] Qiao He, Chao Li, Yousheng Shi, and Tonghai Yang. "A proof of the Kudla-Rapoport conjecture for Krämer models". In: *Invent. Math.* 234.2 (2023), pp. 721–817.

[HM22] Benjamin Howard and Keerthi Madapusi. *Kudla's modularity conjecture on integral models of orthogonal Shimura varieties*. 2022. arXiv: 2211.05108v1 [math.NT].

[Hör14] Fritz Hörmann. *The geometric and arithmetic volume of Shimura varieties of orthogonal type*. Vol. 35. CRM Monograph Series. American Mathematical Society, Providence, RI, 2014, pp. vi+152.

[How12] Benjamin Howard. "Complex multiplication cycles and Kudla-Rapoport divisors". In: *Ann. of Math. (2)* 176.2 (2012), pp. 1097–1171.

[How15] Benjamin Howard. "Complex multiplication cycles and Kudla-Rapoport divisors, II". In: *Amer. J. Math.* 137.3 (2015), pp. 639–698.

[How19] Benjamin Howard. "Linear invariance of intersections on unitary Rapoport-Zink spaces". In: *Forum Math.* 31.5 (2019), pp. 1265–1281.

[HSY21] Qiao He, Yousheng Shi, and Tonghai Yang. "Kudla program for unitary Shimura varieties (in Chinese)". In: *Sci. Sin. Math.* 51.10 (2021). https://doi.org/10.1360/SSM-2021-0002 or https://web.archive.org/web/20211208041220/https://people.math.wisc.edu/~qhe36/ASW.pdf, pp. 1595–1626.

[HSY23] Qiao He, Yousheng Shi, and Tonghai Yang. "Kudla-Rapoport conjecture for Krämer models". In: *Compos. Math.* 159.8 (2023), pp. 1673–1740.

[Ich04] Atsushi Ichino. "A regularized Siegel-Weil formula for unitary groups". In: *Math. Z.* 247.2 (2004), pp. 241–277.

[Ich07] Atsushi Ichino. "On the Siegel-Weil formula for unitary groups". In: *Math. Z.* 255.4 (2007), pp. 721–729.

[Ike08] Tamotsu Ikeda. "On the lifting of Hermitian modular forms". In: *Compos. Math.* 144.5 (2008), pp. 1107–1154.

[Jac62] Ronald Jacobowitz. "Hermitian forms over local fields". In: *Amer. J. Math.* 84 (1962), pp. 441–465.

[JL70] Hervé Jacquet and Robert P. Langlands. *Automorphic forms on* GL(2). Lecture Notes in Mathematics, Vol. 114. Springer-Verlag, Berlin-New York, 1970, pp. vii+548.

[Kar79] Martin L. Karel. "Functional equations of Whittaker functions on p-adic groups". In: *Amer. J. Math.* 101.6 (1979), pp. 1303–1325.

[Kat79] Nicholas M. Katz. "Slope filtration of F-crystals". In: *Journées de Géométrie Algébrique de Rennes (Rennes, 1978), Vol. I.* Vol. 63. Astérisque. Soc. Math. France, Paris, 1979, pp. 113–163.

[Kat99] Hidenori Katsurada. "An explicit formula for Siegel series". In: *Amer. J. Math.* 121.2 (1999), pp. 415–452.

[Kot92] Robert E. Kottwitz. "Points on some Shimura varieties over finite fields". In: *J. Amer. Math. Soc.* 5.2 (1992), pp. 373–444.

[KR00] Stephen S. Kudla and Michael Rapoport. "Cycles on Siegel threefolds and derivatives of Eisenstein series". In: *Ann. Sci. École Norm. Sup. (4)* 33.5 (2000), pp. 695–756.

[KR11] Stephen S. Kudla and Michael Rapoport. "Special cycles on unitary Shimura varieties I. Unramified local theory". In: *Invent. Math.* 184.3 (2011), pp. 629–682.

[KR14] Stephen S. Kudla and Michael Rapoport. "Special cycles on unitary Shimura varieties II: Global theory". In: *J. Reine Angew. Math.* 697 (2014), pp. 91–157.

[KR88] Stephen S. Kudla and Stephen Rallis. "On the Weil-Siegel formula". In: *J. Reine Angew. Math.* 387 (1988), pp. 1–68.

[KR99] Stephen S. Kudla and Michael Rapoport. "Arithmetic Hirzebruch-Zagier cycles". In: *J. Reine Angew. Math.* 515 (1999), pp. 155–244.

[Krä03] Nicole Krämer. "Local models for ramified unitary groups". In: *Abh. Math. Sem. Univ. Hamburg* 73 (2003), pp. 67–80.

[KRY04] Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. "Derivatives of Eisenstein series and Faltings heights". In: *Compos. Math.* 140.4 (2004), pp. 887–951.

[KRY06] Stephen S. Kudla, Michael Rapoport, and Tonghai Yang. *Modular forms and special cycles on Shimura curves*. Vol. 161. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2006, pp. x+373.

[KS97] Stephen S. Kudla and W. Jay Sweet Jr. "Degenerate principal series representations for $\mathrm{U}(n, n)$". In: *Israel J. Math.* 98 (1997), pp. 253–306.

[Kud04] Stephen S. Kudla. "Special cycles and derivatives of Eisenstein series". In: *Heegner points and Rankin L-series*. Vol. 49. Math. Sci. Res. Inst. Publ. Cambridge Univ. Press, Cambridge, 2004, pp. 243–270.

[Kud94] Stephen S. Kudla. "Splitting metaplectic covers of dual reductive pairs". In: *Israel J. Math.* 87.1-3 (1994), pp. 361–401.

[Kud97a] Stephen S. Kudla. "Algebraic cycles on Shimura varieties of orthogonal type". In: *Duke Math. J.* 86.1 (1997), pp. 39–78.

[Kud97b] Stephen S. Kudla. "Central derivatives of Eisenstein series and height pairings". In: *Ann. of Math. (2)* 146.3 (1997), pp. 545–646.

[Li23] Chao Li. "From sum of two squares to arithmetic Siegel-Weil formulas". In: *Bull. Amer. Math. Soc. (N.S.)* 60.3 (2023), pp. 327–370.

[Li24] Chao Li. *Geometric and arithmetic theta correspondences*. To appear in the proceedings of the IHES 2022 Summer School on the Langlands Program. 2024. arXiv: 2402.12159v1 [math.NT].

[Liu11] Yifeng Liu. "Arithmetic theta lifting and L-derivatives for unitary groups, I". In: *Algebra Number Theory* 5.7 (2011), pp. 849–921.

[LL21] Chao Li and Yifeng Liu. "Chow groups and L-derivatives of automorphic motives for unitary groups". In: *Ann. of Math. (2)* 194.3 (2021), pp. 817–901.

[LL22] Chao Li and Yifeng Liu. "Chow groups and L-derivatives of automorphic motives for unitary groups, II". In: *Forum Math. Pi* 10 (2022), Paper No. e5, 71.

[LMB00] Gérard Laumon and Laurent Moret-Bailly. *Champs algébriques*. Vol. 39. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2000, pp. xii+208.

[LZ22a] Chao Li and Wei Zhang. "Kudla–Rapoport cycles and derivatives of local densities". In: *J. Amer. Math. Soc.* 35.3 (2022), pp. 705–797.

[LZ22b] Chao Li and Wei Zhang. "On the arithmetic Siegel-Weil formula for GSpin Shimura varieties". In: *Invent. Math.* 228.3 (2022), pp. 1353–1460.

[Mad23] Keerthi Madapusi. *Derived special cycles on Shimura varieties*. 2023. arXiv: 2212.12849v3 [math.NT].

[Mes72] William Messing. *The crystals associated to Barsotti-Tate groups: with applications to abelian schemes*. Lecture Notes in Mathematics, Vol. 264. Springer-Verlag, Berlin-New York, 1972, pp. iii+190.

[Meu07] Volker Meusers. "Canonical and quasi-canonical liftings in the split case". In: *Astérisque* 312 (2007), pp. 87–98.

[MFK94] David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*. Third. Vol. 34. Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)]. Springer-Verlag, Berlin, 1994, pp. xiv+292.

[Mih22] Andreas Mihatsch. "Relative unitary RZ-spaces and the arithmetic fundamental lemma". In: *J. Inst. Math. Jussieu* 21.1 (2022), pp. 241–301.

[Mor79] Karl A. Morin-Strom. "Witt's theorem for modular lattices". In: *Amer. J. Math.* 101.6 (1979), pp. 1181–1192.

[Mum85] David Mumford. *Abelian varieties.* Tata Institute of Fundamental Research Studies in Mathematics. With appendices by C. P. Ramanujam and Yuri Manin, Reprint of the second (1974) edition. 1985, pp. x+279.

[Nit05] Nitin Nitsure. "Construction of Hilbert and Quot schemes". In: *Fundamental algebraic geometry.* Vol. 123. Math. Surveys Monogr. Amer. Math. Soc., Providence, RI, 2005, pp. 105–137.

[NT91] Yukiyoshi Nakkajima and Yuichiro Taguchi. "A generalization of the Chowla-Selberg formula". In: *J. Reine Angew. Math.* 419 (1991), pp. 119–124.

[Ono85] Takashi Ono. "A generalization of Gauss's theorem on the genera of quadratic forms". In: *Proc. Japan Acad. Ser. A Math. Sci.* 61.4 (1985), pp. 109–111.

[Oor74] Frans Oort. "Subvarieties of moduli spaces". In: *Invent. Math.* 24 (1974), pp. 95–119.

[OZ02] Frans Oort and Thomas Zink. "Families of p-divisible groups with constant Newton polygon". In: *Doc. Math.* 7 (2002), pp. 183–201.

[Pap00] Georgios Pappas. "On the arithmetic moduli schemes of PEL Shimura varieties". In: *J. Algebraic Geom.* 9.3 (2000), pp. 577–605.

[Qiu22] Congling Qiu. *Modularity of arithmetic special divisors for unitary Shimura varieties (with an appendix by Yujie Xu).* 2022. arXiv: 2204.13457v2 [math.NT].

[Rao93] Ramaswamy Ranga Rao. "On some explicit formulas in the theory of Weil representation". In: *Pacific J. Math.* 157.2 (1993), pp. 335–371.

[RSZ17] Michael Rapoport, Brian Smithling, and Wei Zhang. "On the arithmetic transfer conjecture for exotic smooth formal moduli spaces". In: *Duke Math. J.* 166.12 (2017), pp. 2183–2336.

[RSZ18] Michael Rapoport, Brian Smithling, and Wei Zhang. "Regular formal moduli spaces and arithmetic transfer conjectures". In: *Math. Ann.* 370.3-4 (2018), pp. 1079–1175.

[RSZ20] Michael Rapoport, Brian Smithling, and Wei Zhang. "Arithmetic diagonal cycles on unitary Shimura varieties". In: *Compos. Math.* 156.9 (2020), pp. 1745–1824.

[RSZ21] Michael Rapoport, Brian Smithling, and Wei Zhang. “On Shimura varieties for unitary groups”. In: *Pure Appl. Math. Q.* 17.2 (2021), pp. 773–837.

[RZ96] Michael Rapoport and Thomas Zink. *Period spaces for p-divisible groups*. Vol. 141. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 1996, pp. xxii+324.

[Ser06] Jean-Pierre Serre. *Lie algebras and Lie groups*. Vol. 1500. Lecture Notes in Mathematics. 1964 lectures given at Harvard University, Corrected fifth printing of the second (1992) edition. Springer-Verlag, Berlin, 2006, pp. viii+168.

[SGA3II] Michel Demazure. “Fibrés tangents, algèbres de Lie”. In: *Schémas en Groupes (Sém. Géométrie Algébrique, Inst. Hautes Études Sci., 1963), Fasc. 1, Exposé 2*. Inst. Hautes Études Sci., Paris, 1963, p. 40.

[Shi22] Yousheng Shi. “Special cycles on unitary Shimura curves at ramified primes”. In: *Manuscripta Mathematica* (2022).

[Shi82] Goro Shimura. “Confluent hypergeometric functions on tube domains”. In: *Math. Ann.* 260.3 (1982), pp. 269–302.

[Shi97] Goro Shimura. *Euler products and Eisenstein series*. Vol. 93. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997, pp. xx+259.

[SProject] The Stacks project authors. *The Stacks project*. https://stacks.math.columbia.edu. 2024.

[SSTT22] Ananth N. Shankar, Arul Shankar, Yunqing Tang, and Salim Tayou. “Exceptional jumps of Picard ranks of reductions of K3 surfaces over number fields”. In: *Forum Math. Pi* 10 (2022), Paper No. e21, 49.

[SSY23] Siddarth Sankaran, Yousheng Shi, and Tonghai Yang. “A genus two arithmetic Siegel-Weil formula on $X_0(N)$”. In: *Trans. Amer. Math. Soc.* 376.6 (2023), pp. 3995–4041.

[ST68] Jean-Pierre Serre and John Tate. “Good reduction of abelian varieties”. In: *Ann. of Math. (2)* 88 (1968), pp. 492–517.

[SU99] Lucien Szpiro and Emmanuel Ullmo. “Variation de la hauteur de Faltings dans une classe de $\overline{\mathbf{Q}}$-isogénie de courbe elliptique”. In: *Duke Math. J.* 97.1 (1999), pp. 81–97.

[Tan99] Victor Tan. “Poles of Siegel Eisenstein series on $\mathrm{U}(n,n)$”. In: *Canad. J. Math.* 51.1 (1999), pp. 164–175.

[Tat67a] John T. Tate. “p-divisible groups”. In: *Proc. Conf. Local Fields (Driebergen, 1966)*. Springer, Berlin, 1967, pp. 158–183.

[Tat67b] John T. Tate. “Fourier analysis in number fields, and Hecke’s zeta-functions”. In: *Algebraic Number Theory (Proc. Instructional Conf., Brighton, 1965)*. Thompson, Washington, D.C., 1967, pp. 305–347.

[Tat79] John T. Tate. “Number theoretic background”. In: *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*. Proc. Sympos. Pure Math., XXXIII. Amer. Math. Soc., Providence, R.I., 1979, pp. 3–26.

[Ter11] Ulrich Terstiege. “Intersections of arithmetic Hirzebruch-Zagier cycles”. In: *Math. Ann.* 349.1 (2011), pp. 161–213.

[Ter13] Ulrich Terstiege. “Intersections of special cycles on the Shimura variety for $\mathrm{GU}(1,2)$”. In: *J. Reine Angew. Math.* 684 (2013), pp. 113–164.

[Vis89] Angelo Vistoli. “Intersection theory on algebraic stacks and on their moduli spaces”. In: *Invent. Math.* 97.3 (1989), pp. 613–670.

[Vol10] Inken Vollaard. “The supersingular locus of the Shimura variety for $\mathrm{GU}(1,s)$”. In: *Canad. J. Math.* 62.3 (2010), pp. 668–720.

[VW11] Inken Vollaard and Torsten Wedhorn. “The supersingular locus of the Shimura variety of $\mathrm{GU}(1,n-1)$ II”. In: *Invent. Math.* 184.3 (2011), pp. 591–627.

[Wei64] André Weil. “Sur certains groupes d’opérateurs unitaires”. In: *Acta Math.* 111 (1964), pp. 143–211.

[Wei82] André Weil. *Adeles and algebraic groups*. Vol. 23. Progress in Mathematics. With appendices by M. Demazure and Takashi Ono. Birkhäuser, Boston, Mass., 1982, pp. iii+126.

[Wew07] Stefan Wewers. “Canonical and quasi-canonical liftings”. In: *Astérisque* 312 (2007), pp. 67–86.

[Wu16] Haifeng Wu. *The supersingular locus of unitary Shimura varieties with exotic good reduction*. 2016. arXiv: 1609.08775v2 `[math.AG]`.

[WW73] Edmund T. Whittaker and George N. Watson. *A course of modern analysis. An introduction to the general theory of infinite processes and of analytic functions: with an account of the principal transcendental functions*. Fourth edition. Reprinted. Cambridge University Press, New York, 1973, p. 608.

[Yam11] Shunsuke Yamana. "On the Siegel-Weil formula: the case of singular forms". In: *Compos. Math.* 147.4 (2011), pp. 1003–1021.

[Yan04] Tonghai Yang. "Faltings heights and the derivative of Zagier's Eisenstein series". In: *Heegner points and Rankin L-series*. Vol. 49. Math. Sci. Res. Inst. Publ. Cambridge Univ. Press, Cambridge, 2004, pp. 271–284.

[YZ17] Zhiwei Yun and Wei Zhang. "Shtukas and the Taylor expansion of L-functions". In: *Ann. of Math. (2)* 186.3 (2017), pp. 767–911.

[YZ18] Xinyi Yuan and Shou-Wu Zhang. "On the averaged Colmez conjecture". In: *Ann. of Math. (2)* 187.2 (2018), pp. 533–638.

[YZZ13] Xinyi Yuan, Shou-Wu Zhang, and Wei Zhang. *The Gross-Zagier formula on Shimura curves*. Vol. 184. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2013, pp. x+256.

[Zha21] Wei Zhang. "Weil representation and arithmetic fundamental lemma". In: *Ann. of Math. (2)* 193.3 (2021), pp. 863–978.

[Zha95] Shouwu Zhang. "Positive line bundles on arithmetic varieties". In: *J. Amer. Math. Soc.* 8.1 (1995), pp. 187–221.

[Zhu23a] Baiqing Zhu. *Arithmetic Siegel-Weil formula on $\mathcal{X}_0(N)$*. 2023. arXiv: 2304.10696v2 `[math.NT]`.

[Zhu23b] Baiqing Zhu. *Arithmetic Siegel-Weil formula on $\mathcal{X}_0(N)$: singular terms*. 2023. arXiv: 2401.06368v1 `[math.NT]`.