Algebra & Number Theory

Volume 14
2020
No. 9

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2020 is US \$415/year for the electronic version, and \$620/year (+\$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2295

The Brauer group of the moduli stack of elliptic curves

Benjamin Antieau and Lennart Meier

We compute the Brauer group of $\mathcal{M}_{1,1}$, the moduli stack of elliptic curves, over $\operatorname{Spec}\mathbb{Z}$, its localizations, finite fields of odd characteristic, and algebraically closed fields of characteristic not 2. The methods involved include the use of the parameter space of Legendre curves and the moduli stack $\mathcal{M}(2)$ of curves with full (naive) level 2 structure, the study of the Leray–Serre spectral sequence in étale cohomology and the Leray spectral sequence in fppf cohomology, the computation of the group cohomology of S_3 in a certain integral representation, the classification of cubic Galois extensions of $\mathbb{Q}$, the computation of Hilbert symbols in the ramified case for the primes 2 and 3, and finding p-adic elliptic curves with specified properties.

1. Introduction

Brauer groups of fields have been considered since the times of Brauer and Noether; later Grothendieck generalized Brauer groups to the case of arbitrary schemes. Although both the definition via Azumaya algebras and the cohomological definition generalize to arbitrary Deligne–Mumford stacks, Brauer groups of stacks have so far mostly been neglected especially for stacks containing arithmetic information. Some exceptions are the use of Brauer groups of root stacks, as in the work of Chan and Ingalls [2005] on the minimal model program for orders on surfaces, the work of Auel, Bernardara and Bolognesi [Auel et al.

Benjamin Antieau was supported by NSF Grants DMS-1461847 and DMS-1552766. Lennart Meier was supported by DFG SPP 1786.
MSC2010: primary 14F22; secondary 14H52, 14K10.
Keywords: Brauer groups, moduli of elliptic curves, level structures, Hilbert symbols.

2014] on derived categories of families of quadrics, and the work of Lieblich [2011], who computed the Brauer group $\mathrm{Br}(\mathrm{B}\mu_n)$ over a field and applied it to the period-index problem. In this paper, we study the Brauer group $\mathrm{Br}(\mathcal{M})$ of the moduli stack of elliptic curves $\mathcal{M} = \mathcal{M}_{1,1}$.

The case of the Picard group has been considered before. Mumford [1965] showed that $\mathrm{Pic}(\mathcal{M}_k) = \mathrm{H}^1_{\text{ét}}(\mathcal{M}_k, \mathbb{G}_m) \cong \mathbb{Z}/12$ if k is a field of characteristic not dividing 6, and that the Picard group is generated by the Hodge bundle λ. The bundle λ is characterized by the property that $u^*\lambda \cong p_*\Omega^1_{E/T}$ when $u : T \to \mathcal{M}$ classifies a family $p : E \to T$ of elliptic curves. This calculation was extended by Fulton and Olsson who showed that $\mathrm{Pic}(\mathcal{M}_S) \cong \mathrm{Pic}(\mathbb{A}^1_S) \oplus \mathbb{Z}/12$ whenever S is a reduced scheme [Fulton and Olsson 2010].

In contrast, an equally uniform description of $\mathrm{Br}(\mathcal{M}_S)$ does not seem possible (even if we assume that S is regular noetherian); both the result and the proofs depend much more concretely on the arithmetic on S. The following is a sample of our results in ascending order of difficulty. We view (5) as the main result of this paper.

Theorem 1.1. (1) $\mathrm{Br}(\mathcal{M}_k) = 0$ *if k is an algebraically closed field of characteristic not* 2,[1]

(2) $\mathrm{Br}(\mathcal{M}_k) \cong \mathbb{Z}/12$ *if k is a finite field of characteristic not* 2,

(3) $\mathrm{Br}(\mathcal{M}_{\mathbb{Q}}) \cong \mathrm{Br}(\mathbb{Q}) \oplus \bigoplus_{p \not\equiv 3 \bmod 4} \mathbb{Z}/4 \oplus \bigoplus_{p \equiv 3 \bmod 4} \mathbb{Z}/2 \oplus \mathrm{H}^1(\mathbb{Q}, C_3)$, *where p runs over all primes and* -1,

(4) $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]}) \cong \mathrm{Br}\big(\mathbb{Z}\big[\tfrac{1}{2}\big]\big) \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/4$, *and*

(5) $\mathrm{Br}(\mathcal{M}) = 0$.

In all cases the nontrivial classes can be explicitly described via cyclic algebras (see Lemma 7.2, Proposition 9.6 and Remark 9.8). In general, the p-primary torsion for $p \geq 5$ is often easy to control via the following theorem.

Theorem 1.2. *Let S be a regular noetherian scheme and $p \geq 5$ prime. Assume that $S[1/(2p)] = S_{\mathbb{Z}[1/(2p)]}$ is dense in S and that $\mathcal{M}_S \to S$ has a section. Then, the natural map ${}_p\mathrm{Br}'(S) \to {}_p\mathrm{Br}'(\mathcal{M}_S)$ is an isomorphism.*

Here, $\mathrm{Br}'(\mathcal{M}_S)$ denotes the cohomological Brauer group, which agrees with $\mathrm{Br}(\mathcal{M}_S)$ whenever S is affine and at least one prime is invertible on S.

Let us now explain how to compute the 2- and 3-primary torsion in the example of $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$. We use the S_3-Galois cover $\mathcal{M}(2) \to \mathcal{M}_{\mathbb{Z}[1/2]}$, where $\mathcal{M}(2)$ is the moduli stack of elliptic curves with full (naive) level 2-structure. The Leray–Serre spectral sequence reduces the problem of computing $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ to understanding the low-degree $\mathbb{G}_m$-cohomology of $\mathcal{M}(2)$ *together* with the action of S_3 on these cohomology groups and to the computation of differentials.

To understand the groups themselves, it is sufficient to use the fact that $\mathcal{M}(2) \cong \mathrm{B}C_{2,X}$, the classifying stack of cyclic C_2-covers over X, where X is the parameter space of Legendre curves. Explicitly, X is

[1] Minseon Shin [2019] has proved that when k is algebraically closed of characteristic 2 one has $\mathrm{Br}(\mathcal{M}_k) \cong \mathbb{Z}/2$.

the arithmetic surface

$$X = \mathbb{P}^1_{\mathbb{Z}[1/2]} - \{0, 1, \infty\} = \operatorname{Spec} \mathbb{Z}\left[\tfrac{1}{2}, t^{\pm 1}, (t-1)^{-1}\right],$$

and the universal Legendre curve over X is defined by the equation

$$y^2 = x(x-1)(x-t).$$

The Brauer group of $\mathrm{B}C_{2,X}$ can be described using a Leray–Serre spectral sequence as well, but this description is not S_3-equivariant, which causes some complications, and we have to use the S_3-equivariant map from $\mathcal{M}(2)$ to X, its coarse moduli space, to get full control. Knowledge about the Brauer group of X leads for the 3-primary torsion to the following conclusion.

Theorem 1.3. *Let S be a regular noetherian scheme. If 6 is a unit on S, then there is an exact sequence*

$$0 \to {}_3\operatorname{Br}'(S) \to {}_3\operatorname{Br}'(\mathcal{M}_S) \to \mathrm{H}^1(S, C_3) \to 0,$$

which is noncanonically split.

There is a unique cubic Galois extension of $\mathbb{Q}$ which is ramified at most at (2) and (3), namely $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$. This shows that the cokernel of $\operatorname{Br}\left(\mathbb{Z}\left[\frac{1}{6}\right]\right) \hookrightarrow \operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/6]})$ is $\mathbb{Z}/3$. The proof that this extra class does not extend to $\mathcal{M}_{\mathbb{Z}[1/2]}$, which is similar to the strategy discussed below for the 2-torsion, uses the computation of cubic Hilbert symbols at the prime 3. Putting these ingredients together, we conclude that $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ is a 2-group, and further computations first over $\mathbb{Z}\left[\frac{1}{2}, i\right]$ and then over $\mathbb{Z}\left[\frac{1}{2}\right]$ let us deduce the structure in Theorem 1.1(4). The corresponding general results on 2-torsion are contained in Proposition 9.3 and Theorem 9.1. They are somewhat more complicated to state so we omit them from this introduction.

To show that $\operatorname{Br}(\mathcal{M}) = 0$, we need an extra argument since all our arguments using the Leray–Serre spectral sequence presuppose at least that 2 is inverted. Note first that the map $\operatorname{Br}(\mathcal{M}) \to \operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ is an injection. Thus, we have only to show that the nonzero classes in $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ do not extend to $\mathcal{M}$. Our general method is the following. For each nonzero class α in the Brauer group of $\mathcal{M}_{\mathbb{Z}[1/2]}$, we exhibit an elliptic curve over $\operatorname{Spec} \mathbb{Z}_2$ such that the restriction of α to $\operatorname{Spec} \mathbb{Q}_2$ is nonzero. Such an elliptic curve defines a morphism $\operatorname{Spec} \mathbb{Z}_2 \to \mathcal{M}$ and we obtain a commutative diagram

$$\begin{tikzcd}
\operatorname{Br}(\mathbb{Q}_2) & \operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]}) \arrow[l] \\
\operatorname{Br}(\mathbb{Z}_2) \arrow[u] & \operatorname{Br}(\mathcal{M}) \arrow[l] \arrow[u]
\end{tikzcd} \tag{1.4}$$

Together with the fact that $\operatorname{Br}(\mathbb{Z}_2) = 0$, this diagram implies that the class α cannot come from $\operatorname{Br}(\mathcal{M})$. This argument requires us to understand explicit generators for $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ and the computation of Hilbert symbols again.

We remark that the computation of $\operatorname{Br}(\mathcal{M})$ — while important in algebraic geometry and in the arithmetic of elliptic curves — was nevertheless originally motivated by considerations in chromatic homotopy theory, especially in the possibility in constructing twisted forms of the spectrum TMF of topological modular forms. The Picard group computations of Mumford and Fulton and Olsson are primary inputs into the computation of the Picard group of TMF due to Mathew and Stojanoska [2016].

Conventions. We will have occasion to use Zariski, étale, and fppf cohomology of schemes and Deligne–Mumford stacks. We will denote these by $\mathrm{H}^i_{\mathrm{Zar}}(X, F)$, $\mathrm{H}^i(X, F)$, and $\mathrm{H}^i_{\mathrm{pl}}(X, F)$ when F is an appropriate sheaf on X. Note in particular that without other adornment, $\mathrm{H}^i(X, F)$ or $\mathrm{H}^i(R, F)$ (when $X = \operatorname{Spec} R$) always denotes étale cohomology. If G is a group and F is a G-module, we let $\mathrm{H}^i(G, F)$ denote the group cohomology.

For all stacks $\mathcal{X}$ appearing in this paper, we will have $\mathrm{Br}'(\mathcal{X}) \cong \mathrm{H}^2(\mathcal{X}, \mathbb{G}_m)$. Thus, we will use the two groups interchangeably. When working over a general base S, we will typically state our results in terms of $\mathrm{Br}'(S)$ or $\mathrm{Br}'(\mathcal{M}_S)$. However, when working over an affine scheme, such as $S = \operatorname{Spec} \mathbb{Z}\left[\frac{1}{2}\right]$, we will write $\mathrm{Br}(S)$ or $\mathrm{Br}(\mathcal{M}_S)$. There should be no confusion as in all of these cases we will have $\mathrm{Br}(S) = \mathrm{Br}'(S)$ and $\mathrm{Br}(\mathcal{M}_S) = \mathrm{Br}'(\mathcal{M}_S)$ and so on by [de Jong 2005].

For an abelian group A and an integer n, we will denote by ${}_nA$ the subgroup of n-primary torsion elements: ${}_nA = \{x \in A : n^k x = 0 \text{ for some } k \geq 1\}$.

2. Brauer groups, cyclic algebras, and ramification

We review here some basic facts about the Brauer group, with special attention to providing references for those facts in the generality of Deligne–Mumford stacks. For more details about the Brauer group in general; see [Grothendieck 1968a].

Any Deligne–Mumford stack has an associated étale topos, and we can therefore consider étale sheaves and étale cohomology [Laumon and Moret-Bailly 2000].

Definition 2.1. If X is a quasicompact and quasiseparated Deligne–Mumford stack, the *cohomological Brauer group* of X is defined to be $\mathrm{Br}'(X) = \mathrm{H}^2(X, \mathbb{G}_m)_{\mathrm{tors}}$, the torsion subgroup of $\mathrm{H}^2(X, \mathbb{G}_m)$.

Because of its definition as the torsion in a cohomology group, the cohomological Brauer group is amenable to computation via Leray–Serre spectral sequences, long exact sequences, and so on, as we will see in the next sections. However, our main interest is in the *Brauer group* of Deligne–Mumford stacks.

Definition 2.2. An *Azumaya algebra* over a Deligne–Mumford stack X is a sheaf of quasicoherent $\mathcal{O}_X$-algebras $\mathcal{A}$ such that $\mathcal{A}$ is étale-locally on X isomorphic to $\mathcal{M}_n(\mathcal{O}_X)$, the sheaf of $n \times n$-matrices over $\mathcal{O}_X$, for some $n \geq 1$.

In particular, an Azumaya algebra $\mathcal{A}$ is a locally free $\mathcal{O}_X$-module, and the degree n appearing in the definition is a locally constant function. If the degree n is in fact constant, then $\mathcal{A}$ corresponds to a unique PGL_n-torsor on X because the group of k-algebra automorphisms of $\mathrm{M}_n(k)$ is isomorphic to $\mathrm{PGL}_n(k)$ for fields k. The exact sequence $1 \to \mathbb{G}_m \to \mathrm{GL}_n \to \mathrm{PGL}_n \to 1$ gives a boundary map

$$\delta : \mathrm{H}^1(X, \mathrm{PGL}_n) \to \mathrm{H}^2(X, \mathbb{G}_m).$$

For an Azumaya algebra $\mathcal{A}$ of degree n, we write $[\mathcal{A}]$ for the class $\delta(\mathcal{A})$ in $\mathrm{H}^2(X, \mathbb{G}_m)$. In general, when X has multiple connected components, its invariant $[\mathcal{A}] \in \mathrm{H}^2(X, \mathbb{G}_m)$ is computed on each component.

Example 2.3. (1) If E is a vector bundle on X of rank $n > 0$, then $\mathcal{A} = \mathcal{E}\mathrm{nd}(E)$, the sheaf of endomorphisms of E, is an Azumaya algebra on X. Indeed, in this case, $\mathcal{A}$ is even Zariski-locally equivalent to $\mathcal{M}_n(\mathcal{O}_X)$. The class of $\mathcal{A}$ in $\mathrm{H}^1(X, \mathrm{PGL}_n)$ is the image of E via $\mathrm{H}^1(X, \mathrm{GL}_n) \to \mathrm{H}^1(X, \mathrm{PGL}_n)$, so the long exact sequence in nonabelian cohomology implies that $[\mathcal{A}] = 0$ in $\mathrm{H}^2(X, \mathbb{G}_m)$.

(2) If $\mathcal{A}$ and $\mathcal{B}$ are Azumaya algebras on X, then $\mathcal{A} \otimes_{\mathcal{O}_X} \mathcal{B}$ is an Azumaya algebra.

Definition 2.4. Two Azumaya algebras $\mathcal{A}$ and $\mathcal{B}$ are *Brauer equivalent* if there are vector bundles E and F on X such that $\mathcal{A} \otimes_{\mathcal{O}_X} \mathcal{E}\mathrm{nd}(E) \cong \mathcal{B} \otimes_{\mathcal{O}_X} \mathcal{E}\mathrm{nd}(F)$. The *Brauer group* $\mathrm{Br}(X)$ of a Deligne–Mumford stack X is the multiplicative monoid of isomorphism classes of Azumaya algebras under tensor product modulo Brauer equivalence.

In terms of Azumaya algebras, addition is given by the tensor product $[\mathcal{A}] + [\mathcal{B}] = [\mathcal{A} \otimes_{\mathcal{O}_X} \mathcal{B}]$, and $-[\mathcal{A}] = [\mathcal{A}^{\mathrm{op}}]$. Here are the basic structural facts we will use about the Brauer group.

Proposition 2.5. (i) *The Brauer group of a quasicompact and quasiseparated Deligne–Mumford stack X is the subgroup of* $\mathrm{Br}'(X)$ *generated by* $[\mathcal{A}]$ *for $\mathcal{A}$ an Azumaya algebra on X.*

(ii) *The Brauer group* $\mathrm{Br}(X)$ *is a torsion group for any quasicompact and quasiseparated Deligne–Mumford stack X.*

(iii) *If X is a regular and noetherian Deligne–Mumford stack, then* $\mathrm{H}^2(X, \mathbb{G}_m)$ *is torsion, so in particular* $\mathrm{Br}'(X) \cong \mathrm{H}^2(X, \mathbb{G}_m)$.

(iv) *If X is a regular and noetherian Deligne–Mumford stack, and if $U \subseteq X$ is a dense open subset, then the restriction map* $\mathrm{H}^2(X, \mathbb{G}_m) \to \mathrm{H}^2(U, \mathbb{G}_m)$ *is injective.*

(v) *If X is a scheme with an ample line bundle, then* $\mathrm{Br}(X) = \mathrm{Br}'(X)$.

(vi) *If X is a regular and noetherian scheme with p invertible on X, then the morphism ${}_p\mathrm{H}^i(X, \mathbb{G}_m) \to {}_p\mathrm{H}^i(\mathbb{A}^1_X, \mathbb{G}_m)$ on p-primary torsion is an isomorphism for all $i \geq 0$.*

Proof. See [Grothendieck 1968a, Section 2] for points (i) and (ii). The proof of (iv) is analogous to that of [Lieblich 2008, Lemma 3.1.3.3], using an analogue of [Laumon and Moret-Bailly 2000, Proposition 15.4] to generalize [Lieblich 2008, Lemma 3.1.1.9] to algebraic stacks. See also [Auel et al. 2014, Proposition 1.26]. For (v), see [de Jong 2005].

For (iii), see for instance [Grothendieck 1968b, Proposition 1.4] in the case of schemes. We must generalize it to the case of a regular noetherian Deligne–Mumford stack X. We can assume that X is connected and hence irreducible as X is normal. Pick $U \subseteq X$ a dense open such that U admits a *finite* étale map $V \to U$ of degree n where V is a scheme. The composition $\mathrm{H}^2(U, \mathbb{G}_m) \to \mathrm{H}^2(V, \mathbb{G}_m) \to \mathrm{H}^2(U, \mathbb{G}_m)$ of restriction and transfer is multiplication by n. Since $\mathrm{H}^2(V, \mathbb{G}_m)$ is torsion, this implies that $\mathrm{H}^2(U, \mathbb{G}_m)$ is torsion. By (iv), $\mathrm{H}^2(X, \mathbb{G}_m)$ is torsion as well.

Finally, we have to prove (vi). For $i \leq 1$, the $\mathbb{A}^1$-invariance is even true before taking p-power torsion (see e.g., [Hartshorne 1977, II.6.6 and II.6.15]). Because p is invertible on S, the maps $\mathrm{H}^i(X, \mu_{p^m}) \to \mathrm{H}^i(\mathbb{A}^1_X, \mu_{p^m})$ are isomorphisms for all i and m. This is the $\mathbb{A}^1$-invariance of étale cohomology and a proof can be found in [Milne 1980, Corollary VI.4.20]. The short exact sequence

$$1 \to \mu_{p^m} \to \mathbb{G}_m \xrightarrow{p^m} \mathbb{G}_m \to 1$$

induces short exact sequences

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{H}^{i-1}(X, \mathbb{G}_m)/p^m \arrow[r] \arrow[d] & \mathrm{H}^i(X, \mu_{p^m}) \arrow[r] \arrow[d, "\cong"] & \mathrm{H}^i(X, \mathbb{G}_m)[p^m] \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathrm{H}^{i-1}(\mathbb{A}^1_X, \mathbb{G}_m)/p^m \arrow[r] & \mathrm{H}^i(\mathbb{A}^1_X, \mu_{p^m}) \arrow[r] & \mathrm{H}^i(\mathbb{A}^1_X, \mathbb{G}_m)[p^m] \arrow[r] & 0
\end{tikzcd}$$

Inductively, using the $\mathbb{A}^1$-invariance of $\mathrm{H}^{i-1}(X, \mathbb{G}_m)$ if $i \leq 2$ or of ${}_p\mathrm{H}^{i-1}(X, \mathbb{G}_m)$ if $i > 2$ as well as the fact that $\mathrm{H}^{i-1}(X, \mathbb{G}_m)$ is torsion for $i \geq 3$ since X is regular and noetherian [Grothendieck 1968b, Proposition 1.4], we see by the five lemma that $\mathrm{H}^i(X, \mathbb{G}_m)[p^m] \to \mathrm{H}^i(\mathbb{A}^1_X, \mathbb{G}_m)[p^m]$ is an isomorphism for all i and m and hence we also get an isomorphism ${}_p\mathrm{H}^i(X, \mathbb{G}_m) \to {}_p\mathrm{H}^i(\mathbb{A}^1_X, \mathbb{G}_m)$. □

Remark 2.6. At a couple points, we use another important fact, due to Gabber, which says that if $p : Y \to X$ is a surjective finite locally free map, if $\alpha \in \mathrm{Br}'(X)$, and if $p^*\alpha \in \mathrm{Br}(Y)$, then $\alpha \in \mathrm{Br}(X)$. This is already proved in [Gabber 1981, Chapter II, Lemma 4] for locally ringed topoi with strict hensel local rings, so we need to add nothing further in our setting.

By far the most important class of Azumaya algebras arising in arithmetic applications is the class of cyclic algebras. For a treatment over fields, see [Gille and Szamuely 2006, Section 2.5]. These algebras give a concrete realization of the cup product in fppf cohomology

$$\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, C_n) \times \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{H}^2_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{H}^2_{\mathrm{pl}}(\mathcal{X}, \mathbb{G}_m)$$

for an algebraic stack $\mathcal{X}$. Given that C_n and $\mathbb{G}_m$ are smooth and that the image of such a cup product is torsion, we can rewrite this as $\mathrm{H}^1(\mathcal{X}, C_n) \times \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{H}^2_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{Br}'(\mathcal{X})$. Given $\chi \in \mathrm{H}^1(\mathcal{X}, C_n)$ and $u \in \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$, we write $[(\chi, u)_n]$ or $[(\chi, u)]$ for the image of the cup product in $\mathrm{Br}'(\mathcal{X})$.

Fix an algebraic stack $\mathcal{X}$. Let $p(\chi) : \mathcal{Y} \to \mathcal{X}$ be the cyclic Galois cover defined by $\chi \in \mathrm{H}^1(\mathcal{X}, C_n)$. Then, $p(\chi)_*\mathcal{O}_{\mathcal{Y}}$ is a locally free $\mathcal{O}_{\mathcal{X}}$-algebra of finite rank which comes equipped with a canonical C_n-action. There is a natural isomorphism $\mathbf{Spec}_{\mathcal{X}}(p(\chi)_*\mathcal{O}_{\mathcal{Y}}) \cong \mathcal{Y} \to \mathcal{X}$.

The group $\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$ fits into a short exact sequence

$$0 \to \mathbb{G}_m(\mathcal{X})/n \to \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{Pic}(\mathcal{X})[n] \to 0. \tag{2.7}$$

It is helpful to have a more concrete description of $\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$, which will also show that the exact sequence (2.7) is noncanonically split. Let $\mathrm{H}(\mathcal{X}, n)$ be the abelian group of equivalence classes of pairs $(\mathcal{L}, s)$ where $\mathcal{L} \in \mathrm{Pic}(\mathcal{X})[n]$ and s is a choice of trivialization $s : \mathcal{O}_{\mathcal{X}} \to \mathcal{L}^{\otimes n}$. Two pairs $(\mathcal{L}, s)$ and $(\mathcal{M}, t)$

are equivalent if there is an isomorphism $g : \mathcal{L} \to \mathcal{M}$ and a unit $v \in \mathbb{G}_m(\mathcal{X})$ such that $g(s) = v^n t$. The group structure is given by tensor product of line bundles and of trivializations. The following construction is part of Kummer theory and is well-known in the scheme case (see e.g., [Milne 1980, page 125], which also shows part of Proposition 2.9 below).

Construction 2.8. Let $\mathcal{X}$ be an algebraic stack and fix a class $[u] \in \mathrm{H}(\mathcal{X}, n)$ with $u = (\mathcal{L}, s)$ for a line bundle $\mathcal{L}$ on $\mathcal{X}$ with a trivialization $s\colon \mathcal{O}_{\mathcal{X}} \to \mathcal{L}^{\otimes n}$. Define $\mathcal{O}_{\mathcal{X}}(\sqrt[n]{u})$ as $\bigoplus_{i\in\mathbb{Z}} \mathcal{L}^{\otimes i}/(s-1)$ and $\mathcal{X}(\sqrt[n]{u}) \to \mathcal{X}$ as the affine morphism $\mathbf{Spec}_{\mathcal{X}}\, \mathcal{O}_{\mathcal{X}}(\sqrt[n]{u}) \to \mathcal{X}$. It is easy to see that $\mathcal{X}(\sqrt[n]{u}) \to \mathcal{X}$ is an fppf μ_n-torsor and that this construction defines a group homomorphism $\mathrm{H}(\mathcal{X}, n) \to \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$. Indeed, if $\mathcal{X} = \operatorname{Spec} A$ and $\mathcal{L}$ is trivial, then $\mathcal{X}(\sqrt[n]{u}) \cong \operatorname{Spec} A(\sqrt[n]{s^{-1}}) \cong \operatorname{Spec} A(\sqrt[n]{s})$.

Proposition 2.9. *The map* $\mathrm{H}(\mathcal{X}, n) \to \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$ *is an isomorphism. In particular, there is a noncanonical splitting*

$$\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \cong \mathbb{G}_m(\mathcal{X})/n \oplus \operatorname{Pic}(\mathcal{X})[n].$$

Proof. We claim that the line bundle associated with the μ_n-torsor $\mathcal{X}(\sqrt[n]{u}) \to \mathcal{X}$ (via the map $\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \mathrm{H}^1(\mathcal{X}, \mathbb{G}_m)$ induced by the inclusion $\mu_n \to \mathbb{G}_m$) is exactly $\mathcal{L}$. Indeed, the obvious map from $\mathcal{X}(\sqrt[n]{u})$ to $\mathbf{Spec}_{\mathcal{X}}\big(\bigoplus_{i\in\mathbb{Z}} \mathcal{L}^{\otimes i}\big)$ is equivariant along the inclusion $\mu_n \to \mathbb{G}_m$ and the target is the $\mathbb{G}_m$-torsor associated with $\mathcal{L}$. Thus, the composition $\mathrm{H}(\mathcal{X}, n) \to \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n) \to \operatorname{Pic}(\mathcal{X})[n]$ is surjective. By (2.7), to prove that $\mathrm{H}(\mathcal{X}, n) \to \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$ is an isomorphism, it suffices to prove that the induced map from the kernel of this map to $\mathbb{G}_m(\mathcal{X})/n$ is an isomorphism. But, this follows immediately from the definition of $\mathrm{H}(\mathcal{X}, n)$.

It remains to construct the splitting. By Prüfer's theorem [Fuchs 1970, Theorem 17.2], $\operatorname{Pic}(\mathcal{X})[n]$ is a direct sum of cyclic groups. Thus, we have only to show that for every divisor k of n and each k-torsion element $[\mathcal{L}]$ in $\operatorname{Pic}(\mathcal{X})$, there exists a preimage in $\mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$ that is k-torsion. This preimage can be constructed as follows: choose a trivialization $s\colon \mathcal{O}_{\mathcal{X}} \to \mathcal{L}^{\otimes k}$ and take the μ_n-torsor associated with the μ_k-torsor $\mathcal{X}(\sqrt[k]{v}) \to \mathcal{X}$ with $v = (\mathcal{L}, s)$. □

Now, given χ and u as above, we let $\tilde{\mathcal{A}}_{\chi,u}$ be the coproduct

$$\Big(p(\chi)_*\mathcal{O}_{\mathcal{Y}} \coprod_{\mathcal{O}_{\mathcal{X}}} \mathcal{O}_{\mathcal{X}}(\sqrt[n]{u})\Big)$$

in the category of sheaves of quasicoherent (associative and unital) $\mathcal{O}_{\mathcal{X}}$-algebras. Finally, we let

$$\mathcal{A}_{\chi,u} = \tilde{\mathcal{A}}_{\chi,u}/(ab - ba^{g(\chi)})$$

be the quotient of $\tilde{\mathcal{A}}_{\chi,u}$ by the two-sided ideal generated by terms $ab - ba^{g(\chi)}$ where a is a local section of $p(\chi)_*\mathcal{O}_{\mathcal{Y}}$, b is a local section of $\mathcal{O}_{\mathcal{X}}(\sqrt[n]{u})$, and $a^{g(\chi)}$ denotes the action of $g(\chi)$ on $p(\chi)_*\mathcal{O}_{\mathcal{Y}}$ for $g(\chi)$ a fixed generator of the action of C_n on $p(\chi)_*\mathcal{O}_{\mathcal{Y}}$.

Lemma 2.10. *Given an algebraic stack* $\mathcal{X}$ *and classes* $\chi \in \mathrm{H}^1(\mathcal{X}, C_n)$ *and* $u \in \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$, *the algebra* $\mathcal{A}_{\chi,u}$ *is an Azumaya algebra on* $\mathcal{X}$.

Proof. It suffices to check this fppf-locally, and in particular we can assume that in fact $\mathcal{X}$ is a scheme X, that $\mathcal{L} \cong \mathcal{O}_{\mathcal{X}}$, and that χ classifies a C_n-Galois cover $p(\chi)\colon Y \to X$. In this case, we can write the $\mathcal{O}_X$-algebra $p(\chi)_*\mathcal{O}_Y$ Zariski locally as a quotient $\mathcal{O}_X[x]/f(x)$ for some monic polynomial $f(x)$ of degree n. Then, locally, we have that

$$\mathcal{A}_{\chi,u} \cong \mathcal{O}_X\langle x, y\rangle/(f(x), y^n - u, xy - yx^{g(\chi)}),$$

a quotient of the free algebra over $\mathcal{O}_X$ on generators x and y. Note that the sections $x^i y^j$ for $0 \le i, j \le n-1$ form a basis of $\mathcal{A}_{\chi,u}$ as an $\mathcal{O}_X$-module and in particular that $\mathcal{A}_{\chi,u}$ is (locally) a free $\mathcal{O}_X$-module. Examining the fibers of $\mathcal{A}_{\chi,u}$ over X, we obtain the usual definition of a cyclic algebra given in [Gille and Szamuely 2006, Proposition 2.5.2]. So, $\mathcal{A}_{\chi,u}$ is locally free with central simple fibers and [Grothendieck 1968a, Théorème 5.1] implies that $\mathcal{A}_{\chi,u}$ is Azumaya. □

The following proposition is well-known, but we do not know an exact reference. However, in the case of quaternion algebras, it is given in [Parimala and Srinivas 1992, Lemma 8].

Proposition 2.11. *Let X be a regular noetherian scheme and suppose that $\chi \in \mathrm{H}^1(X, C_n)$ and $u \in \mathrm{H}^1_{\mathrm{pl}}(X, \mu_n)$ are fixed classes. In the notation above, we have $[(\chi, u)_n] = [\mathcal{A}_{\chi,u}]$ in $\mathrm{Br}'(X)$. In particular $[(\chi, u)_n] \in \mathrm{Br}(X)$.*

Proof. We can assume that X is connected. As $\mathrm{Br}'(X) \to \mathrm{Br}'(K)$ is injective (see [Milne 1980, IV.2.6] or Proposition 2.5(iv)), it is enough to check this on a generic point $\operatorname{Spec} K$ of X. By definition and the previous example, $\mathcal{A}_{\chi,u}$ is a standard cyclic algebra over K as defined in [Gille and Szamuely 2006, Chapter 2]. They check in [loc. cit., Proposition 4.7.3] that $\mathcal{A}_{\chi,u}$ does indeed have Brauer class given by the cup product. See also the remark at the beginning of the proof of [loc. cit., Proposition 4.7.1]. □

Remark 2.12. The reader may notice that $\mathcal{A}_{\chi,u}$ is defined in complete generality, but that we only prove the equality $[(\chi, u)_n] = [\mathcal{A}_{\chi,u}]$ for regular noetherian schemes. In fact, this equality extends to arbitrary algebraic stacks, but a different argument is necessary. It is given at the end of Section 3.

We will abuse notation and write $(\chi, u)_n$ or even just (χ, u) for $\mathcal{A}_{\chi,u}$. This is called a *cyclic algebra*. If there is a primitive n-th root of unity $\omega \in \mu_n(\mathcal{X})$ and the cyclic Galois cover $\mathcal{Y} \to \mathcal{X}$ is obtained by adjoining an n-th root of an element $a \in \mathbb{G}_m(\mathcal{X})$, we write $(a, u) = (a, u)_\omega$ for the corresponding cyclic algebra, where we need the choice of ω to fix an isomorphism $\mathrm{H}^1(\mathcal{X}, C_n) \cong \mathrm{H}^1(\mathcal{X}, \mu_n)$. For $n = 2$, we obtain the classical notion of a quaternion algebra.

For us, the key point about the cyclic algebra is that it allows us to compute the ramification of a Brauer class explicitly. Before explaining this, we mention that by the Gabber–Česnavičius purity theorem the Brauer group of a regular noetherian scheme X is insensitive to throwing away high codimension subschemes.

Proposition 2.13 (purity [Gabber 1981, Chapter I; Česnavičius 2019]). *Let X be a regular noetherian scheme. If $U \subseteq X$ is a dense open subscheme with complement of codimension at least 2, then the restriction map $\mathrm{Br}'(X) \to \mathrm{Br}'(U)$ is an isomorphism.*

Let X be a regular noetherian scheme and let η be the scheme of generic points in X. The purity theorem reduces the problem of computing $\mathrm{Br}(X)$ from $\mathrm{Br}(\eta)$ to the problem of extending Brauer classes $\alpha \in \mathrm{Br}(\eta)$ over divisors in X. This is controlled by ramification theory. The following proposition is basically well-known, but we include a proof for the reader's convenience.

Proposition 2.14. *Let X be a regular noetherian scheme, $i\colon D \subseteq X$ a Cartier divisor that is regular with complement U, and n an integer. There is an exact sequence*

$$0 \to {}_n\mathrm{Br}'(X) \to {}_n\mathrm{Br}'(U) \xrightarrow{\mathrm{ram}_D} {}_n\mathrm{H}^3_D(X, \mathbb{G}_m) \to {}_n\mathrm{H}^3(X, \mathbb{G}_m) \to {}_n\mathrm{H}^3(U, \mathbb{G}_m). \qquad (2.15)$$

If n is prime to the residue characteristics of X, we have ${}_n\mathrm{H}^3_D(X, \mathbb{G}_m) \cong \mathrm{H}^1(D, {}_n\mathbb{Q}/\mathbb{Z})$.

Proof. By [Grothendieck 1968c, 6.1] or [Milne 1980, III.1.25] there is a long exact sequence

$$\mathrm{H}^2(X, \mathbb{G}_m) \to \mathrm{H}^2(U, \mathbb{G}_m) \to \mathrm{H}^3_D(X, \mathbb{G}_m) \to \mathrm{H}^3(X, \mathbb{G}_m) \to \mathrm{H}^3(U, \mathbb{G}_m).$$

By [Grothendieck 1968b, Propsition 1.4] all occurring groups are torsion so that the sequence is still exact after taking n-primary torsion. Furthermore, by Proposition 2.5 the first map is an injection.

We may assume that $n = p$ is a prime, in which case ${}_p\mathbb{Q}/\mathbb{Z} \cong \mathbb{Q}_p/\mathbb{Z}_p$. We have to show that ${}_p\mathrm{H}^3_D(X, \mathbb{G}_m) \cong \mathrm{H}^1(D, \mathbb{Q}_p/\mathbb{Z}_p)$. By either the relative cohomological purity theorem of Artin [SGA 4_3 1973, Théorème XVI.3.7 and 3.8] (when both X and D are smooth over some common base scheme S) or the absolute cohomological purity theorem of Gabber [Fujiwara 2002, Theorem 2.1], we have the following identifications of local cohomology sheaves: $\mathcal{H}^t_D(\mu_{p^\nu}) = 0$ for $t \neq 2$ and $\mathcal{H}^2_D(\mu_{p^\nu}) \cong i_*\mathbb{Z}/p^\nu(-1)$. It follows from the long exact sequence of local cohomology sheaves associated to the exact sequence $1 \to \mu_{p^\nu} \to \mathbb{G}_m \xrightarrow{p^\nu} \mathbb{G}_m \to 1$ that p acts invertibly on $\mathcal{H}^t_D(\mathbb{G}_m)$ for $t \neq 1, 2$. Moreover, since X is regular and noetherian, for every open $V \subset X$, the map $\mathrm{Pic}(V) \to \mathrm{Pic}(U \cap V)$ is surjective with kernel $(i_*\mathbb{Z})(V)$ by [Hartshorne 1977, II.6.5] and $\mathrm{Br}'(V) \to \mathrm{Br}'(U \cap V)$ is injective; thus $\mathcal{H}^2_D(\mathbb{G}_m) = 0$ and $\mathcal{H}^1_D(\mathbb{G}_m) \cong i_*\mathbb{Z}$. Therefore, the only contribution to p-primary torsion in $\mathrm{H}^3_D(X, \mathbb{G}_m)$ in the local to global spectral sequence

$$\mathrm{H}^s(X, \mathcal{H}^t_D(\mathbb{G}_m)) \Rightarrow \mathrm{H}^{s+t}_D(X, \mathbb{G}_m)$$

is ${}_p\mathrm{H}^2(X, i_*\mathbb{Z})$. We obtain

$${}_p\mathrm{H}^3_D(X, \mathbb{G}_m) \cong {}_p\mathrm{H}^2(X, \mathcal{H}^1_D(\mathbb{G}_m)) \cong {}_p\mathrm{H}^2(D, \mathbb{Z}) \cong \mathrm{H}^1(D, \mathbb{Q}_p/\mathbb{Z}_p)$$

as desired, where the last isomorphism holds because $\mathrm{H}^i(D, \mathbb{Q}) = 0$ for $i > 0$ since D is normal (see for example [Deninger 1988, 2.1]). □

Note that in all cases where we use Proposition 2.14, the easier relative cohomological purity theorem of Artin is applicable, so that in the end our paper does not rely on the more difficult results of Gabber and Česnavičius.

We will need to know a special case of the ramification map $\mathrm{ram}_D : \mathrm{Br}(U) \to \mathrm{H}^1(D, \mathbb{Q}/\mathbb{Z})$.

Proposition 2.16. *Let R be a discrete valuation ring with fraction field K and residue field k. Set $X = \operatorname{Spec} R$, $U = \operatorname{Spec} K$, and $x = \operatorname{Spec} k$. Let $(\chi, \pi)_n$ be a cyclic algebra over K, where χ is a degree n cyclic character of K, π is a uniformizing parameter of R (viewed as an element of $\mathbb{G}_m(K)/n$), and n is prime to the characteristic of k. Finally, let L/K be the cyclic Galois extension defined by χ. If the integral closure S of R in L is a discrete valuation ring with uniformizing parameter π_S, then $\operatorname{ram}_{(\pi)}(\chi, \pi)$ is the class of the cyclic extension $S/(\pi_S)$ over k.*

Proof. See [Saltman 1999, Lemma 10.2]. □

Finally, we discuss cyclic algebras over local fields and some implications for global calculations. Let K be a local field containing a primitive n-th root of unity ω. Then there is a pairing

$$\left(\frac{-,-}{\mathfrak{p}}\right) : \mathbb{G}_m(K)/n \times \mathbb{G}_m(K)/n \to \mu_n(K),$$

called the *Hilbert symbol* (where $\mathfrak{p}$ stands for the maximal ideal of the ring of integers of K). Our standard reference for this pairing is [Neukirch 1999, Section V.3]. If $\mathfrak{p}$ is generated by an element π, we will also write $\left(\frac{a,b}{\pi}\right)$. We will use Hilbert symbols to check whether explicitly defined cyclic algebras are zero in the Brauer group.

Proposition 2.17. *For $a, b \in K^\times$, the cyclic algebra $(a, b)_\omega$ is trivial in $\operatorname{Br}(K)$ if and only if $\left(\frac{a,b}{\mathfrak{p}}\right) = 1$.*

Proof. By Proposition V.3.2 of [Neukirch 1999] the Hilbert symbol $\left(\frac{a,b}{\mathfrak{p}}\right)$ equals 1 if and only if a is a norm from the extension $K(\sqrt[n]{b})|K$. By [Gille and Szamuely 2006, Corollary 4.7.7], this happens if and only if $(a, b)_\omega$ splits, i.e., defines the trivial class in $\operatorname{Br}(K)$. □

More generally, local class field theory calculates $\operatorname{Br}(\eta)$ when $\eta = \operatorname{Spec} K$ where K is a (nonarchimedean) local field. Let $X = \operatorname{Spec} R$ and $x = \operatorname{Spec} k$, where R is the ring of integers in K and k is the residue field of R. As $\mathrm{H}^3(X, \mathbb{G}_m) \cong \mathrm{H}^3(x, \mathbb{G}_m) = 0$ (for instance by [Grothendieck 1968c, Théorème 1.1]), we find from [loc. cit., Corollaire 2.2] that there is an exact sequence

$$0 \to \operatorname{Br}(X) \to \operatorname{Br}(\eta) \to \mathrm{H}^1(x, \mathbb{Q}/\mathbb{Z}) \to 0.$$

The idea is similar to that of Proposition 2.14, but here the proof is easier as $0 \to \mathbb{G}_m \to j_*\mathbb{G}_m \to i_*\mathbb{Z} \to 0$ is exact where $j : \eta \to X$ and $i : x \to X$. Since K is local, k is finite, so that $\mathrm{H}^1(x, \mathbb{Q}/\mathbb{Z}) \cong \mathbb{Q}/\mathbb{Z}$. However, since R is Henselian, $\operatorname{Br}(X) = \operatorname{Br}(x)$ (see [Grothendieck 1968a, Corollaire 6.2]), and $\operatorname{Br}(x) = 0$ by a theorem of Wedderburn (see [loc. cit., Proposition 1.5]).

Now, let K be a number field, and let R be a localization of the ring of integers of K. Set $\eta = \operatorname{Spec} K$ and $X = \operatorname{Spec} R$. In this case, by [loc. cit., Proposition 2.1], there is an exact sequence

$$0 \to \operatorname{Br}(X) \to \operatorname{Br}(\eta) \to \bigoplus_{\mathfrak{p} \in X^{(1)}} \operatorname{Br}(\operatorname{Spec} K_{\mathfrak{p}}),$$

where $X^{(1)}$ denotes the set of codimension 1 points of X. This exact sequence is compatible with (2.15) and with the exact sequence

$$0 \to \mathrm{Br}(\eta) \to \bigoplus_{\mathfrak{p}} \mathrm{Br}(\mathrm{Spec}\, K_{\mathfrak{p}}) \to \mathbb{Q}/\mathbb{Z} \to 0 \tag{2.18}$$

of class field theory (see [Neukirch et al. 2000, Theorem 8.1.17]). The sum ranges over the finite and the infinite places of K, and the map $\mathrm{Br}(\mathrm{Spec}\, K_{\mathfrak{p}}) \to \mathbb{Q}/\mathbb{Z}$ is the isomorphism described above when $\mathfrak{p}$ is a finite place, the natural inclusion $\mathbb{Z}/2 \to \mathbb{Q}/\mathbb{Z}$ when $K_{\mathfrak{p}} \cong \mathbb{R}$, and the natural map $0 \to \mathbb{Q}/\mathbb{Z}$ when $K_{\mathfrak{p}} \cong \mathbb{C}$. Using these sequences, we can compute the Brauer group of X.

The two fundamental observations we need about (2.18) are that a class $\alpha \in \mathrm{Br}(\eta)$ is ramified at no fewer than 2 places and that if K is purely imaginary, then $\alpha \in \mathrm{Br}(\eta)$ is ramified at no fewer than 2 *finite* places. The reader can easily verify the following examples.

Example 2.19. (1) $\mathrm{Br}(\mathbb{Z}) = 0$.

(2) $\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{p}\big]\big) \cong \mathbb{Z}/2$.

(3) $\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{pq}\big]\big) \cong \mathbb{Z}/2 \oplus \mathbb{Q}/\mathbb{Z}$.

(4) $\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{p}, \zeta_p\big]\big) = 0$.

We will use these computations and those like them throughout the paper, often without comment.

3. The low-dimensional $\mathbb{G}_m$-cohomology of $\mathrm{B}C_m$

Let S be a scheme. Write $C_{n,S}$ for the constant étale group scheme on the cyclic group C_n of order $n \geq 2$ over S. We will often suppress the base in the notation and simply write C_n when the base is clear from context. The purpose of this section is to make a basic computation of the $\mathbb{G}_m$-cohomology of the Deligne–Mumford stack $\mathrm{B}C_n = \mathrm{B}C_{n,S}$. In fact, we are only interested in the cases $n = 2$ and $n = 4$, but the general case is no more difficult.

The first tool for our computations of the étale cohomology of an étale sheaf $\mathcal{F}$ is the convergent Leray–Serre spectral sequence. If $\pi : Y \to X$ is a G-Galois cover where X and Y are Deligne–Mumford stacks, then this spectral sequence has the form

$$\mathrm{E}_2^{p,q} = \mathrm{H}^p(G, \mathrm{H}^q(Y, \pi^*\mathcal{F})) \Rightarrow \mathrm{H}^{p+q}(X, \mathcal{F}),$$

with differentials d_r of bidegree $(r, 1-r)$.

We will use the spectral sequence in this section for the C_n-Galois cover $\pi : S \to \mathrm{B}C_n$, where C_n acts trivially on S. In this case it is of the form

$$\mathrm{E}_2^{p,q} = \mathrm{H}^p(C_n, \mathrm{H}^q(S, \mathbb{G}_m)) \Rightarrow \mathrm{H}^{p+q}(\mathrm{B}C_{n,S}, \mathbb{G}_m)$$

and Figure 1 displays the low-degree part of its E_2-page. The fact that C_n acts trivially on the cohomology of S implies that the left-most column is simply the $\mathbb{G}_m$-cohomology of S. For F a constant C_n-module,

$$\begin{tikzcd}
\mathrm{H}^2(S, \mathbb{G}_m) & & & \\
\mathrm{Pic}(S) & \mathrm{Pic}(S)[n] \arrow[rrd] & \mathrm{Pic}(S)/n & \\
\mathbb{G}_m(S) & \mu_n(S) & \mathbb{G}_m(S)/n & \mu_n(S)
\end{tikzcd}$$

Figure 1. The E_2-page of the Leray–Serre spectral sequence computing $\mathrm{H}^i(\mathrm{B}C_n, \mathbb{G}_m)$.

we use the standard isomorphisms $\mathrm{H}^i(C_n, F) \cong F[n]$ when $i > 0$ odd, and $\mathrm{H}^i(C_n, F) \cong F/n$ when $i > 0$ is even. We are only interested in $\mathrm{H}^i(\mathrm{B}C_n, \mathbb{G}_m)$ for $0 \le i \le 2$.

Recall Grothendieck's theorem that the natural morphism $\mathrm{H}^i(X, G) \to \mathrm{H}^i_{\mathrm{pl}}(X, G)$ is an isomorphism for $i \ge 0$ when G is a smooth group scheme on X (such as C_n or $\mathbb{G}_m$). See [Grothendieck 1968c, Section 5, Thèoréme 11.7]. Note that this implies the agreement of étale and fppf cohomology on a Deligne–Mumford stack. For example, Grothendieck's theorem implies that the morphism from the Leray–Serre spectral sequence above to the analogous Leray–Serre spectral sequence

$$\mathrm{E}_2^{p,q} = \mathrm{H}^p(C_n, \mathrm{H}^q_{\mathrm{pl}}(S, \mathbb{G}_m)) \Rightarrow \mathrm{H}^{p+q}_{\mathrm{pl}}(\mathrm{B}C_n, \mathbb{G}_m)$$

for the fppf cohomology is an isomorphism; thus the comparison map

$$\mathrm{H}^i(\mathrm{B}C_n, \mathbb{G}_m) \to \mathrm{H}^i_{\mathrm{pl}}(\mathrm{B}C_n, \mathbb{G}_m)$$

is also an isomorphism. For $\mathbb{G}_m$-coefficients, we will thus not distinguish between étale and fppf cohomology in what follows.

We use these observations to compute the Picard and Brauer groups of $\mathrm{B}C_{n,S}$ via a Leray spectral sequence. The idea is borrowed from [Lieblich 2011, Section 4.1]. Consider the map to the coarse moduli space $c\colon \mathrm{B}C_{n,S} \to S$. We claim that

$$\begin{aligned}
\mathrm{R}^0_{\mathrm{pl}} c_* \mathbb{G}_m &= \mathbb{G}_m \\
\mathrm{R}^1_{\mathrm{pl}} c_* \mathbb{G}_m &= \mu_n \\
\mathrm{R}^2_{\mathrm{pl}} c_* \mathbb{G}_m &= 0.
\end{aligned}$$

Indeed, $\mathrm{R}^i_{\mathrm{pl}} c_* \mathbb{G}_m$ is the fppf-sheafification of $U \mapsto \mathrm{H}^i(\mathrm{B}C_{n,U}, \mathbb{G}_m)$ and in the Leray–Serre spectral sequences all classes in $\mathrm{H}^p(C_n, \mathrm{H}^q(U, \mathbb{G}_m))$ for $q > 0$ are killed by some fppf cover of U. Furthermore every unit has an n-th root fppf-locally so that the fppf-sheafification of the presheaf $\mathbb{G}_m/n$ vanishes. This implies the claim.

It follows that the fppf-Leray spectral sequence

$$\mathrm{E}_2^{p,q} = \mathrm{H}^p_{\mathrm{pl}}(S, \mathrm{R}^q_{\mathrm{pl}} c_* \mathbb{G}_m) \Rightarrow \mathrm{H}^{p+q}_{\mathrm{pl}}(\mathrm{B}C_{n,S}, \mathbb{G}_m) \tag{3.1}$$

for c takes the form given in Figure 2 in low degrees. As $\pi^* c^* = \mathrm{id}$, we see that the edge homomorphisms $\mathrm{H}^i(S, \mathbb{G}_m) \to \mathrm{H}^i(\mathrm{B}C_{n,S}, \mathbb{G}_m)$ from the bottom line in the Leray spectral sequence are all split injections. In particular, the displayed differentials $d_2^{0,1}$ and $d_2^{1,1}$ are zero.

$$\begin{tikzcd}
0 \\
\mu_n(S) \arrow[drr] & \mathrm{H}^1_{\mathrm{pl}}(S, \mu_n) \arrow[drr] \\
\mathbb{G}_m(S) & \mathrm{Pic}(S) & \mathrm{H}^2(S, \mathbb{G}_m) & \mathrm{H}^3(S, \mathbb{G}_m)
\end{tikzcd}$$

Figure 2. The E_2-page of the Leray spectral sequence (3.1) computing $\mathrm{H}^i(\mathrm{B}C_n, \mathbb{G}_m)$.

Proposition 3.2. *There is an isomorphism*

$$\mathbb{G}_m(\mathrm{B}C_{n,S}) \cong \mathbb{G}_m(S)$$

and short exact sequences

$$0 \to \mathrm{Pic}(S) \xrightarrow{c^*} \mathrm{Pic}(\mathrm{B}C_{n,S}) \to \mu_n(S) \to 0$$

and

$$\begin{gathered}
0 \to \mathrm{H}^2(S, \mathbb{G}_m) \xrightarrow{c^*} \mathrm{H}^2(\mathrm{B}C_{n,S}, \mathbb{G}_m) \xrightarrow{r} \mathrm{H}^1_{\mathrm{pl}}(S, \mu_n) \to 0, \\
0 \to \mathrm{Br}'(S) \xrightarrow{c^*} \mathrm{Br}'(\mathrm{B}C_{n,S}) \xrightarrow{r} \mathrm{H}^1_{\mathrm{pl}}(S, \mu_n) \to 0, \\
0 \to \mathrm{Br}(S) \xrightarrow{c^*} \mathrm{Br}(\mathrm{B}C_{n,S}) \xrightarrow{r} \mathrm{H}^1_{\mathrm{pl}}(S, \mu_n) \to 0,
\end{gathered}$$

which are split. The isomorphism and the short exact sequences are functorial in $\mathrm{B}C_{n,S}$ (*i.e., endomorphisms of* $\mathrm{B}C_{n,S}$ *induces endomorphisms of exact sequences in a functorial manner*), *but the splittings are only functorial in S.*

Proof. By the discussion above, the Leray spectral sequence proves everything except for the split exactness of the last two sequences. For the sequence involving the cohomological Brauer group, we just apply the torsion subgroup functor to the split exact sequence involving $\mathrm{H}^2(-, \mathbb{G}_m)$. By Remark 2.6 we furthermore see that $a = \pi^* c^* a \in \mathrm{H}^2(S, \mathbb{G}_m)$ is in $\mathrm{Br}(S)$ if and only if $c^* a \in \mathrm{Br}(\mathrm{B}C_{n,S})$, implying split exactness for the last exact sequence. □

Later on we will need not only the computation of the Brauer group of $\mathrm{B}C_n$, but also a description of the classes coming from the inclusion $\mathbb{G}_m(S)/n \hookrightarrow \mathrm{Br}(\mathrm{B}C_n)$, which is either defined via the Leray–Serre spectral sequence or using the splitting in Proposition 3.2 (the proof of the following lemma will, in particular, show that these two maps differ at most by a unit). These classes are described via the classical cyclic algebra construction from the previous section.

Lemma 3.3. *Let X be an algebraic stack and n a positive integer. Let* $\sigma \in \mathrm{H}^1(\mathrm{B}C_{n,X}, C_n)$ *be the class of the universal* C_n*-torsor* $X \to \mathrm{B}C_{n,X}$. *Then there is an integer k prime to n* (*that only depends on n*) *such that the map*

$$s : \mathrm{H}^1_{\mathrm{pl}}(X, \mu_n) \to \mathrm{H}^2(\mathrm{B}C_{n,X}, \mathbb{G}_m)$$

defined by $s(u) = k[(\sigma, u)_n]$ *is a section to the map r from Proposition 3.2.*

Proof. It suffices to consider the universal case of $X = \mathrm{B}_{\mathrm{pl}}\mu_n$ over $\operatorname{Spec}\mathbb{Z}$, the stack classifying fppf μ_n torsors. Note that $\mathrm{B}_{\mathrm{pl}}\mu_n$ is indeed a stack by [Stacks 2018, Tag 04UR] and is an algebraic stack by [Stacks 2018, Tag 06DC] with fppf atlas $\operatorname{Spec}\mathbb{Z} \to \mathrm{B}_{\mathrm{pl}}\mu_n$. Let $d : \mathrm{B}_{\mathrm{pl}}\mu_n \to \operatorname{Spec}\mathbb{Z}$ denote the structure map, and let $\mathrm{R}^q_{\mathrm{pl}}d_*\mu_n$ denote the derived functors of the push-forward in the fppf topos. Then, it is easy to see that $\mathrm{R}^0_{\mathrm{pl}}d_*\mu_n \cong \mu_n$ and we claim that $\mathrm{R}^1_{\mathrm{pl}}d_*\mu_n \cong C_n$. To see the latter isomorphism, consider the natural transformations

$$\operatorname{Hom}^{\mathrm{gp}}_{\operatorname{Spec} R}(\mu_{n,R}, \mathbb{G}_{m,R}) \to \mathrm{H}^1(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mathbb{G}_m)[n] \leftarrow \mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mu_n) \tag{3.4}$$

of presheaves on affine schemes over $\mathbb{Z}$; the first map sends a homomorphism $f\colon \mu_{n,R} \to \mathbb{G}_{m,R}$ to the image of the canonical class $\mathrm{H}^1(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mu_n)$ under f_* and the second map is part of the Kummer sequence. The leftmost term is a sheaf and it is a standard fact that it is represented by the constant étale group scheme C_n. See [Cornell et al. 1997, Section V.2.10] for example. The fppf-sheafification of the rightmost term is $\mathrm{R}^1_{\mathrm{pl}}d_*\mu_n$. To see that the induced map of sheaves are isomorphisms, it is sufficient to check on stalks in the fppf topology [Gabber and Kelly 2015, Remark 1.8, Theorem 2.3] and in particular if R is a Henselian local ring with algebraically closed residue field [loc. cit., Lemma 3.3]. If R is such a local ring, then $\mathbb{G}_m(R)/n = 0$ so that $\mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mu_n) \cong \mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mathbb{G}_m)[n]$. Using that $\operatorname{Pic}(R) = 0$, the Leray–Serre spectral sequence for the cover $\operatorname{Spec} R \to \mathrm{B}_{\mathrm{pl}}\mu_{n,R}$ shows that

$$\mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_{n,R}, \mathbb{G}_m) \cong \mathrm{H}^1_{\mathrm{group}}(\mu_{n,R}, \mathbb{G}_{m,R}) \cong \operatorname{Hom}^{\mathrm{gp}}_{\operatorname{Spec} R}(\mu_{n,R}, \mathbb{G}_{m,R}),$$

where $\mathrm{H}^1_{\mathrm{group}}(\mu_{n,R}, \mathbb{G}_{m,R})$ is the first cohomology of the cobar complex

$$\mathbb{G}_m(S) \to \mathbb{G}_m(\mu_{n,R}) \to \mathbb{G}_m(\mu_{n,R} \times_{\operatorname{Spec} R} \mu_{n,R}) \to \cdots$$

with differentials as in the usual definition of group cohomology. This shows that the morphisms in (3.4) are isomorphisms on fppf-stalks and thus that $\mathrm{R}^1_{\mathrm{pl}}d_*\mu_n \cong C_n$.

Now, the fppf-Leray spectral sequence for $d : \mathrm{B}_{\mathrm{pl}}\mu_n \to \operatorname{Spec}\mathbb{Z}$ yields an exact sequence

$$0 \to \mathrm{H}^1_{\mathrm{pl}}(\operatorname{Spec}\mathbb{Z}, \mu_n) \to \mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_n, \mu_n) \to \mathrm{H}^0(\operatorname{Spec}\mathbb{Z}, C_n) \to 0. \tag{3.5}$$

The right-hand term is isomorphic to $\mathbb{Z}/n$, and the sequence is split by applying the pullback map along $\operatorname{Spec}\mathbb{Z} \to \mathrm{B}_{\mathrm{pl}}\mu_n$. We denote by $\tau \in \mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_n, \mu_n)$ the class of the universal μ_n-torsor over $\mathrm{B}_{\mathrm{pl}}\mu_n$. This is (exactly) of order n and pulls back to zero on $\operatorname{Spec}\mathbb{Z}$.

Consider $c\colon \mathrm{B}C_{n,\mathrm{B}_{\mathrm{pl}}\mu_n} \to \mathrm{B}_{\mathrm{pl}}\mu_n$ and the class $\alpha = [(\sigma, c^*\tau)_n]$. The class of α has order (exactly) n as there are cyclic algebras of order n over fields, for example by [Gille and Szamuely 2006, Lemma 5.5.3]. As $\pi^*\alpha = 0$ for $\pi\colon \mathrm{B}_{\mathrm{pl}}\mu_n \to \mathrm{B}C_{n,\mathrm{B}_{\mathrm{pl}}\mu_n}$ the projection, it follows from Proposition 3.2 that $r(\alpha)$ in $\mathrm{H}^1_{\mathrm{pl}}(\mathrm{B}_{\mathrm{pl}}\mu_n, \mu_n)$ has order n as well. On the other hand, $r(\alpha)$ pulls back to zero over $\operatorname{Spec}\mathbb{Z}$ so it is a nonzero multiple of τ (using the split-exact sequence (3.5)). Thus, $r(\alpha) = m\tau$ for some m prime to n. This completes the proof if we set k to be a number such that $km \equiv 1 \bmod n$. □

Corollary 3.6. *Suppose that $\chi : X \to Y$ is a C_n-torsor for some positive integer n. Let $u \in \mathbb{G}_m(Y)/n$ be the class of a unit, and write α_u for the corresponding class in* $\mathrm{Br}'(Y)$ (*defined via the Leray–Serre* *spectral sequence). Then we have* $\alpha_u = k[(\chi, u)]$ *in* $\mathrm{Br}'(Y)$, *where k is some number prime to n which only depends on n.*

We do not know the value of k in the corollary. Perhaps it is always ± 1, as is the case in similar computations, such as the result of Lichtenbaum (see [Gille and Szamuely 2006, Theorem 5.4.10]), which computes the exact value of the map $\mathrm{Pic}(X_{\bar{k}})^G \cong \mathbb{Z} \to \mathrm{Br}(k)$ when X is a Severi–Brauer variety of a field with Galois group G, or the computation of [Gille and Quéguiner-Mathieu 2011] of the sign of the Rost invariant.

Proposition 3.7. *Let* $\mathcal{X}$ *be an algebraic stack and suppose that* $\chi \in \mathrm{H}^1(\mathcal{X}, C_n)$ *and* $u \in \mathrm{H}^1_{\mathrm{pl}}(\mathcal{X}, \mu_n)$ *are fixed classes. In the notation above, we have* $[(\chi, u)_n] = [\mathcal{A}_{\chi,u}]$ *in* $\mathrm{Br}'(\mathcal{X})$.

Proof. Both $[\mathcal{A}_{\chi,u}]$ and $[(\chi, u)_n]$ define classes in $\mathrm{H}^2_{\mathrm{pl}}(\mathrm{B}C_n \times \mathrm{B}_{\mathrm{pl}}\mu_n, \mathbb{G}_m)$. As at the end of the proof of Proposition 3.3, we see that $[\mathcal{A}_{\chi,u}] = k[(\chi, u)_n]$ for some k prime to n. We saw in Proposition 2.11 that they agree when pulled back to regular noetherian schemes. The result follows. □

4. A presentation of the moduli stack of elliptic curves

We will compute $\mathrm{Br}(\mathcal{M})$ using that it injects into $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ by Proposition 2.5(iv) and using a specific presentation of $\mathcal{M}_{\mathbb{Z}[1/2]}$, which we now describe. This presentation is standard and we claim no originality in our presentation of it. For references, see [Deligne and Rapoport 1973] or [Katz and Mazur 1985],

Definition 4.1. A *full level* 2 *structure* on an elliptic curve E over a base scheme S is a fixed isomorphism $(\mathbb{Z}/2)^2_S \to E[2]$, where $(\mathbb{Z}/2)^2_S$ denotes the constant group scheme on $(\mathbb{Z}/2)^2$ over S and $E[2]$ is the subgroupscheme of order 2 points in E. If there exists an isomorphism $(\mathbb{Z}/2)^2_S \cong E[2]$, an equivalent way of specifying a level 2 structure is to order the points of exact order 2 in $E(S)$ (over each connected component of S).

Remark 4.2. These full level 2 structures are sometimes called *naive* to distinguish them from the level structures considered by Drinfeld, which allow one to extend $\mathcal{M}(2)$ to a stack supported over all of $\mathrm{Spec}\,\mathbb{Z}$. We will not need this generalization in this paper. It is the subject of [Katz and Mazur 1985].

The moduli stack $\mathcal{M}(2)$ of elliptic curves with fixed level 2 structures is a regular noetherian Deligne–Mumford stack. Moreover, since the existence of a full level 2 structure implies that 2 is invertible in S (by [Katz and Mazur 1985, Corollary 2.3.2] for example), the functor $\mathcal{M}(2) \to \mathcal{M}$ which forgets the level structure factors through $\mathcal{M}_{\mathbb{Z}[1/2]}$. This map is clearly equivariant for the right S_3-action on $\mathcal{M}(2)$ that permutes the nonzero 2-torsion points and the trivial S_3-action on $\mathcal{M}$. Note that in general, being G-equivariant for a map $f : \mathcal{X} \to \mathcal{Y}$ of stacks with G-action is extra structure: for every $g \in G$ one has to provide compatible 2-morphism $\sigma_g : gf \to fg$ (see [Romagny 2005] for details). In the case of $\mathcal{M}(2) \to \mathcal{M}_{\mathbb{Z}[1/2]}$ though the equivariance is strict in the sense that all σ_g are the identity 2-morphisms of $\mathcal{M}(2) \to \mathcal{M}_{\mathbb{Z}[1/2]}$.

We have the following well-known statement; to fix ideas, we will provide a proof.

Lemma 4.3. *The map* $\mathcal{M}(2) \to \mathcal{M}_{\mathbb{Z}[1/2]}$ *is an* S_3*-Galois cover.*

Proof. It is enough to show that for every affine scheme Spec R over Spec $\mathbb{Z}\big[\frac{1}{2}\big]$ and every elliptic curve E over Spec R, we can find a full level 2 structure étale locally. Indeed, if there is one full level 2 structure on E, the map

$$(S_3)_{\operatorname{Spec} R} = S_3 \times \operatorname{Spec} R \to \operatorname{Spec} R \times_{\mathcal{M}_{\mathbb{Z}[1/2]}} \mathcal{M}(2)$$

from the constant group scheme on S_3 is an isomorphism since we get every other full level 2 structure on E by permuting the nonzero 2-torsion points.

The elliptic curve E defines an R-point $E : \operatorname{Spec} R \to \mathcal{M}_{\mathbb{Z}[1/2]}$ of the moduli stack of elliptic curves. Zariski locally we can assume the pullback $E^*\lambda$ of the Hodge bundle to be trivial, in which case there exists a nowhere vanishing invariant differential ω. By [Katz and Mazur 1985, Section 2.2], we can then write E in Weierstrass form over Spec R, which after a coordinate change takes the form

$$y^2 = x^3 + b_2x^2 + b_4x + b_6.$$

As a point (x, y) on E is 2-torsion if and only if $y = 0$, we have a full level 2 structure after adjoining the three roots e_1, e_2 and e_3 of $x^3 + b_2x^2 + b_4x + b_6$ to R. This defines an étale extension as the discriminant of this cubic polynomial does not vanish (because E is smooth). $\square$

Definition 4.4. A *Legendre curve* with parameter t over S is an elliptic curve E_t with Weierstrass equation

$$y^2 = x(x-1)(x-t).$$

As the discriminant of this equation is $16t^2(t-1)^2$, such an equation defines an elliptic (and hence Legendre) curve if and only if 2, t and $t-1$ are invertible on S.

The points $(0, 0)$, $(1, 0)$, and $(t, 0)$ define three nonzero 2-torsion points on E_t. Taking them in this order fixes a full level 2 structure on E. This defines a morphism

$$\pi : X \to \mathcal{M}(2),$$

where X is the parameter space of Legendre curves. In fact, X is an affine scheme, given as

$$X = \operatorname{Spec} \mathbb{Z}\big[\tfrac{1}{2}, t^{\pm 1}, (t-1)^{-1}\big] = \mathbb{A}^1_{\mathbb{Z}[1/2]} - \{0, 1\}.$$

We will use X throughout this paper to refer specifically to this moduli space of Legendre curves. In particular, X is naturally defined over $\mathbb{Z}\big[\frac{1}{2}\big]$. In general, given a scheme S, we let $X_S = \mathbb{A}^1_S - \{0, 1\}$. Note that this is a slight abuse of notation as we do not assume that 2 is invertible on S.

We equip the map $\pi : X \to \mathcal{M}(2)$ with the structure of a C_2-equivariant map with the trivial C_2-action on X and $\mathcal{M}(2)$ by choosing $\sigma_g : g\pi \to \pi g$ to be $[-1]$ (i.e., multiplication by -1 on the universal elliptic curve) for $g \in C_2$ the nontrivial element. Note that $[-1]$ fixes the level 2 structure and so indeed defines a natural automorphism of $\mathrm{id}_{\mathcal{M}(2)}$. The structure of a C_2-equivariant map on π induces a map $[X/C_2] = \mathrm{B}C_{2,X} \to \mathcal{M}(2)$.

Proposition 4.5. *The C_2-equivariant map $\pi : X \to \mathcal{M}(2)$ is a C_2-torsor. Thus, the map $\mathrm{B}C_{2,X} \to \mathcal{M}(2)$ is an equivalence.*

Proof. First we will show that an elliptic curve $E \to \operatorname{Spec} R$ with full level 2 structure can étale locally be brought into Legendre form. Our proof will be along the lines of [Silverman 2009, Proposition III.1.7], but we have to take a little bit more care.

As in the proof of Lemma 4.3, Zariski locally over $\operatorname{Spec} R$, we can write E in the form

$$y^2 = x^2 + b_2x^2 + b_4x + b_6$$

and the full level 2 structure allows us to factor the right-hand side as

$$(x - e_1)(x - e_2)(x - e_3),$$

where $(e_1, 0)$, $(e_2, 0)$ and $(e_3, 0)$ are the nonzero 2-torsion points. We set $p = e_2 - e_1$ and $q = e_3 - e_1$. By a linear coordinate change, we get $y^2 = x(x - p)(x - q)$.

Since the equation $y^2 = x(x - p)(x - q)$ defines an elliptic curve, p, q and $p - q$ are nowhere vanishing. Thus, the extension $R \to R[\sqrt{p}]$ is étale so that we can (and will) assume étale locally to have a (chosen) square root $\sqrt{p}$. Now, E is isomorphic to $y^2 = x(x - 1)(x - t)$ for $t = q/p$, where the isomorphism is given by $x = px'$ and $y = p^{3/2}y'$. Thus our original E is indeed étale locally (on the base) isomorphic to a Legendre curve as an elliptic curve with level 2 structure. It is moreover an elementary check with coordinate transformations that there is at most one choice of $t \in R$ such that the Legendre curve E_t with parameter t is isomorphic to E in $\mathcal{M}(2)$.

Now assume that our elliptic curve E over R is in Legendre form and assume further that $\operatorname{Spec} R$ is connected. By definition, for a commutative R-algebra R', an element of $(X \times_{\mathcal{M}(2)} \operatorname{Spec} R)(R')$ consists of a Legendre curve E_t together with an isomorphism of E_t to $E_{R'}$ in $\mathcal{M}(2)$. By assumption this set is nonempty and it is indeed a torsor under the group of automorphisms of $E_{R'}$ in $\mathcal{M}(2)$. By [Katz and Mazur 1985, Corollary 2.7.2], the only nontrivial automorphism of $E_{R'}$ with level 2 structure is $[-1]$. Thus, the C_2-action exactly interchanges the two elements of $(X \times_{\mathcal{M}(2)} \operatorname{Spec} R)(R')$ and we obtain a C_2-equivariant equivalence

$$C_2 \times \operatorname{Spec} R \simeq X \times_{\mathcal{M}(2)} \operatorname{Spec} R.$$

As every E with full level 2 structure satisfies étale locally our assumptions, this implies that $X \to \mathcal{M}(2)$ is a C_2-torsor.

By the general fact that for a G-torsor $\mathcal{X} \to \mathcal{Y}$, the induced map $[\mathcal{X}/G] \to \mathcal{Y}$ is an equivalence, we obtain in our case the equivalence $\mathrm{B}C_{2,X} \simeq \mathcal{M}(2)$. □

Corollary 4.6. *The map $c\colon \mathcal{M}(2) \to X$ sending $y^2 = (x - e_1)(x - e_2)(x - e_3)$ to $y^2 = x(x - 1)(x - (e_3 - e_1)/(e_2 - e_1))$ exhibits X as the coarse moduli space of $\mathcal{M}(2)$.*

Proof. The set of maps from $\mathcal{M}(2) \simeq \mathrm{B}C_{2,X}$ to X is in bijection with C_2-equivariant maps $X \to X$. Thus, a map $\mathcal{M}(2) \to X$ exhibits X as the coarse moduli space if and only if the precomposition with π is the identity. This is clearly the case for c. □

It follows that the right S_3-action on $\mathcal{M}(2)$ induces a right S_3 action on X. We can describe this explicitly as follows. Consider the generators $\sigma = (1\,3\,2)$ and $\tau = (2\,3)$ of $\mathrm{GL}_2(\mathbb{Z}/2) \cong S_3$, of orders 3 and 2, respectively. Then,

$$\sigma(t) = \frac{t-1}{t}, \quad \text{and} \quad \tau(t) = \frac{1}{t}.$$

By a simple computation, the map $c\colon \mathcal{M}(2) \to X$ defined above is strictly S_3-equivariant.

In contrast, the map $\pi : X \to \mathcal{M}(2)$ described above is *not* S_3-equivariant, as one notes for example by checking that the elliptic curves $y^2 = x(x-1)(x-t)$ and $y^2 = x(x-1)\bigl(x - \frac{1}{t}\bigr)$ are generally not isomorphic. To actually explain the correct S_3-action on $\mathrm{B}C_{2,X}$, we have to fix some notation.

Consider again an elliptic curve E given by $y^2 = (x-e_1)(x-e_2)(x-e_3)$. Set again $p = e_2 - e_1$ and $q = e_3 - e_1$ so that we can write E as

$$y^2 = x(x-p)(x-q).$$

The only possible coordinate changes fixing the form of this equation are the transformations $y \mapsto u^3 y$ and $x \mapsto u^2 x$; such a coordinate change results in multiplying the standard invariant differential $\omega = -dx/2y$ by u^{-1} and sending p to $u^2 p$ and q to $u^2 q$. Thus, $p\omega^{\otimes 2}$ and $q\omega^{\otimes 2}$ define canonical sections of $\lambda^{\otimes 2}$ on $\mathcal{M}(2)$, not dependent on any choice of Weierstrass form. Note that these sections are nowhere vanishing. We can consider the C_2-torsor $\mathcal{M}(2)(\sqrt{p}) \to \mathcal{M}(2)$ defined as the cyclic cover $\mathbf{Spec}_{\mathcal{M}(2)}\bigl(\bigoplus_{i\in\mathbb{Z}} \lambda^{\otimes i}/(1-p)\bigr)$. Étale locally on some $\operatorname{Spec} R$, we can trivialize λ so that p becomes an element of R and the C_2-torsor becomes $\operatorname{Spec} R[\sqrt{p}] \to \operatorname{Spec} R$. The C_2-torsor $\mathcal{M}(2)(\sqrt{p}) \to \mathcal{M}(2)$ is equivalent to $X \to \mathcal{M}(2)$. Indeed, we have shown in the proof of Proposition 4.5 that the latter has a section as soon as we have a chosen square root of p.

As $g^*\lambda$ for $g \in S_3$ on $\mathcal{M}(2)$ is canonically isomorphic to λ (as this is pulled back from $\mathcal{M}$), we have an action of S_3 on $\mathrm{H}^0(\mathcal{M}(2), \lambda^{\otimes *})$. Consider the section

$$\frac{g(p)}{p} \in \mathrm{H}^0(\mathcal{M}(2), \mathcal{O}_{\mathcal{M}(2)}) \cong \mathrm{H}^0(X, \mathcal{O}_X),$$

which can for E as above be written as $(e_{g(2)} - e_{g(1)})/(e_2 - e_1)$. For example, we have $g(p)/p = q/p$ for $g = \tau$, which equals t on X. For a scheme S with a map $f\colon S \to X$ or $f\colon S \to \mathcal{M}(2)$, we denote the torsor adjoining the square root of $f^*g(p)/p$ by $T_{f,g} \to S$.

For the next lemma, we recall that an object in $\mathrm{B}C_{2,X}(S)$ corresponds to a C_2-torsor $T \to S$ and a C_2-equivariant map $T \to X$, where X has the trivial C_2-action. Equivalently, an object can be described as a C_2-torsor $T \to S$ with a map $f\colon S \to X$. Let S_3 act on $\mathrm{B}C_{2,X}$ in the following way: $g \in S_3$ acts (from the right) on $(T, f) \in \mathrm{B}C_{2,X}(S)$ by setting $g(T)$ to be $(T \times_S T_{f,g})/C_2$ and the map $g(f)$ to be the composition $S \xrightarrow{f} X \xrightarrow{g} X$.

Lemma 4.7. *The natural map* $\mathcal{M}(2) \to \mathrm{B}C_{2,X}$ *induces an* S_3*-equivariant equivalence* $\mathrm{B}C_{2,X} \simeq \mathcal{M}(2)$.

Proof. As noted above, the map $\mathcal{M}(2) \to \mathrm{B}C_{2,X}$ classifying the torsor $\mathcal{M}(2)(\sqrt{p}) \to \mathcal{M}(2)$ is an equivalence by Proposition 4.5 (as this torsor is equivalent to $X \to \mathcal{M}(2)$). We have only to check the S_3-equivariance of this map.

Given an $f\colon S \to \mathcal{M}(2)$, the corresponding object in $\mathrm{B}C_{2,X}$ is the torsor $S(\sqrt{f^*p}) \to S$ together with $S \to \mathcal{M}(2) \to X$. The composition gf for $g \in S_3$ corresponds to the torsor $S(\sqrt{(gf)^*p}) \to S$ together with $S \to \mathcal{M}(2) \to X \xrightarrow{g} X$ as $\mathcal{M}(2) \to X$ is S_3-equivariant. As $(gf)^*p = f^*(g(p))$, we have $(gf)^*p = f^*p \cdot f^*(g(p)/p)$. Thus, we have a natural isomorphism $(S(\sqrt{p}) \times_S T_{f,g})/C_2 \xrightarrow{\cong} S(\sqrt{(gf)^*p})$. One can check that these isomorphisms are compatible (similarly to [Romagny 2005, Definition 2.1], although we do not have a strict S_3-action on $\mathrm{B}C_{2,X}$) so that one actually gets the structure of an S_3-equivariant map. □

Of particular import will be the action of S_3 on the units of X. Let ρ be the tautological permutation representation of S_3 on $\mathbb{Z}^{\oplus 3}$ and let $\tilde{\rho}$ be the kernel of the morphism

$$\rho \cong \mathrm{ind}_{C_2}^{S_3} \mathbb{Z} \to \mathbb{Z}$$

to the trivial representation, the adjoint to the identity.

Lemma 4.8. *For any connected normal noetherian scheme S over $\mathbb{Z}\left[\frac{1}{2}\right]$, there is an S_3-equivariant exact sequence*

$$0 \to \mathbb{G}_m(S) \to \mathbb{G}_m(X_S) \to \tilde{\rho} \to 0,$$

where S_3 acts on $\mathbb{G}_m(S)$ trivially and $\tilde{\rho}$ is additively generated by the images of t and $t-1$. This exact sequence is nonequivariantly split.

Proof. Denote by $\pi\colon X_S \to S$ the structure map. We have a map $f\colon \mathbb{Z}^2 \oplus \mathbb{G}_m \to \pi_*\mathbb{G}_{m,X_S}$ of sheaves on S, where f takes the two $\mathbb{Z}$-summands to t and $(t-1)$, respectively. We claim that this map is an isomorphism. It is enough to check this on affine connected opens $\operatorname{Spec} R$, where it follows from R being an integral domain (as it is normal). The nonequivariant statement follows.

Moreover, the action of S_3 on $\mathbb{G}_m(S)$ is trivial by definition. Set $\sigma = (1\,3\,2)$ and $\tau = (2\,3)$. If we choose the basis vectors

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

for $\tilde{\rho}$, we obtain exactly the same S_3-representation as on $\mathbb{G}_m(X_S)/\mathbb{G}_m(S) \cong \mathbb{Z}\{t, t-1\}$, where the latter denotes the free $\mathbb{Z}$-module on t and $t-1$ with elements thought of as $t^k(t-1)^l$. □

5. Beginning of the computation

Let S be a connected regular noetherian scheme over $\mathbb{Z}\left[\frac{1}{2}\right]$, let $\mathcal{M}_S$ be the moduli stack of elliptic curves over S, and let $\mathcal{M}(2)_S$ be the moduli stack of elliptic curves with full level 2 structure over S. The Leray–Serre spectral sequence for $\mathcal{M}(2)_S \to \mathcal{M}_S$ takes the form

$$\mathrm{E}_2^{p,q} : \mathrm{H}^p(S_3, \mathrm{H}^q(\mathcal{M}(2)_S, \mathbb{G}_m)) \Rightarrow \mathrm{H}^{p+q}(\mathcal{M}_S, \mathbb{G}_m), \tag{5.1}$$

with differentials d_r of bidegree $(r, 1-r)$. In this section, we will collect the basic tools to compute the E_2-term. We start with two brief remarks about the cohomology of S_3.

Lemma 5.2. *Let M be a trivial S_3-module. Then,*

$$\begin{aligned}
\mathrm{H}^1(S_3, M) &\cong M[2],\\
\mathrm{H}^2(S_3, M) &\cong M/2,\\
\mathrm{H}^3(S_3, M) &\cong M[6],\\
\mathrm{H}^4(S_3, M) &\cong M/6.
\end{aligned}$$

Proof. We use the Lyndon–Hochschild–Serre spectral sequence for

$$1 \to \mathbb{Z}/3 \to S_3 \to \mathbb{Z}/2 \to 1.$$

A reference is [Weibel 1994, Example 6.7.10]. On the E_2-page, $\mathrm{E}_2^{pq} = 0$ whenever $p > 0$ and $q > 0$ because the cohomology of $\mathbb{Z}/3$ is 3-torsion. Moreover, $\mathbb{Z}/2$ acts on $\mathrm{H}^q(\mathbb{Z}/3, M)$ by multiplication by -1 for $q \equiv 1, 2 \bmod 4$ and by 1 for $q \equiv 0, 3 \bmod 4$. □

The next lemma is about the cohomology of the reduced regular representation $\tilde{\rho}$ of S_3 introduced in Section 4.

Lemma 5.3. *Let M be an abelian group, and let $\tilde{\rho} \otimes M$ be an S_3-module through the action on $\tilde{\rho}$. Then,*

$$\begin{aligned}
\mathrm{H}^0(S_3, \tilde{\rho} \otimes M) &\cong M[3],\\
\mathrm{H}^1(S_3, \tilde{\rho} \otimes M) &\cong M/3,\\
\mathrm{H}^2(S_3, \tilde{\rho} \otimes M) &\cong 0\\
\mathrm{H}^3(S_3, \tilde{\rho} \otimes M) &\cong 0.
\end{aligned}$$

Proof. There is a short exact sequence of S_3-modules

$$0 \to \tilde{\rho} \otimes M \to \rho \otimes M \to M \to 0.$$

In the associated long exact sequence in cohomology, note that $\mathrm{H}^i(S_3, \rho \otimes M) \cong \mathrm{H}^i(C_2, M)$ by Shapiro's lemma, as $\rho \otimes M \cong \mathrm{ind}_{C_2}^{S_3} M$. The map

$$\mathrm{H}^i(S_3, \rho \otimes M) \cong \mathrm{H}^i(C_2, M) \to \mathrm{H}^i(S_3, M)$$

is the transfer. We obtain short exact sequences

$$0 \to \operatorname{coker} \mathrm{tr}_{C_2}^{S_3}(\mathrm{H}^{i-1}) \to \mathrm{H}^i(S_3, \tilde{\rho} \otimes M) \to \ker \mathrm{tr}_{C_2}^{S_3}(\mathrm{H}^i) \to 0.$$

Because $C_2 \to S_3$ has a retraction, the restriction map $\mathrm{H}^i(S_3, M) \to \mathrm{H}^i(C_2, M)$ is the projection to a direct summand. The transfer equals 3 times the inclusion of this summand as can easily be deduced from the equation $\mathrm{tr}_{C_2}^{S_3} \mathrm{res}_{C_2}^{S_3} = 3$. Thus, the transfer is multiplication by 3 on H^0, an isomorphism on H^1 and H^2 and the inclusion $M[2] \to M[6]$ on H^3. The lemma follows. □

These computations allow us to compute the E_2-term of the Leray–Serre spectral sequence

$$\mathrm{E}_2^{p,q} : \mathrm{H}^p(S_3, \mathrm{H}^q(\mathcal{M}(2)_S, \mathbb{G}_m)) \Rightarrow \mathrm{H}^{p+q}(\mathcal{M}_S, \mathbb{G}_m)$$

in a range. Using the results of the last two sections, we can analyze $\mathrm{H}^q(\mathcal{M}(2)_S, \mathbb{G}_m)$ in terms of $\mathrm{H}^q(X_S, \mathbb{G}_m)$. Especially Proposition 3.2 turns out to be useful as the short exact sequences in it are S_3-equivariant by naturality. Using additionally Lemma 4.8 for the first one, we obtain the S_3-equivariant extensions

$$0 \to \mathbb{G}_m(S) \to \mathbb{G}_m(\mathcal{M}(2)_S) \cong \mathbb{G}_m(X_S) \to \tilde{\rho} \to 0, \tag{5.4}$$

$$0 \to \mathrm{Pic}(X_S) \cong \mathrm{Pic}(S) \to \mathrm{Pic}(\mathcal{M}(2)_S) \to \mu_2(S) \to 0, \tag{5.5}$$

and

$$0 \to \mathrm{Br}'(X_S) \to \mathrm{Br}'(\mathcal{M}(2)_S) \to \mathrm{H}^1(X_S, \mu_2) \to 0. \tag{5.6}$$

The only point needing justification is that the pullback map $\mathrm{Pic}(S) \to \mathrm{Pic}(X_S)$ is an isomorphism. It is injective because X_S has an S-point. It is surjective as it factors through the isomorphism $\mathrm{Pic}(S) \to \mathrm{Pic}(\mathbb{A}^1_S)$ and since $j^*\colon \mathrm{Pic}(\mathbb{A}^1_S) \to \mathrm{Pic}(X_S)$ is surjective, where j denotes the inclusion $X_S \subseteq \mathbb{A}^1_S$. Indeed, given a line bundle $\mathcal{L}$ on X_S, we take a coherent subsheaf $\mathcal{F}$ of $j_*\mathcal{L}$ with $j^*\mathcal{F} \cong \mathcal{L}$. The double dual of $\mathcal{F}$ is a reflexive sheaf $\mathcal{L}'$ with $j^*\mathcal{L}'$ still isomorphic to $\mathcal{L}$. By [Hartshorne 1980, Proposition 1.9], $\mathcal{L}'$ is a line bundle.

The sequence (5.5) is S_3-equivariantly split and thus consists only of S_3-modules with the trivial action. Indeed, the morphism $S \to \operatorname{Spec} \mathbb{Z}\big[\frac{1}{2}\big]$ induces by pullback a morphism from the exact sequence

$$0 \to 0 \to \mathrm{Pic}(\mathcal{M}(2)) \to \mu_2\big(\mathbb{Z}\big[\tfrac{1}{2}\big]\big) \to 0,$$

where the splitting is clearly S_3-equivariant. As $\mu_2\big(\mathbb{Z}\big[\frac{1}{2}\big]\big) \to \mu_2(S)$ is an isomorphism for S connected, the result follows. These observations allow us to compute the $q = 0, 1$ lines of the Leray–Serre spectral sequence (5.1).

Lemma 5.7. *If S is a connected regular noetherian scheme over $\mathbb{Z}\big[\frac{1}{2}\big]$, then there are natural extensions*

$$0 \to \mathrm{H}^p(S_3, \mathbb{G}_m(S)) \to \mathrm{H}^p(S_3, \mathbb{G}_m(\mathcal{M}(2)_S)) \to \mathrm{H}^p(S_3, \tilde{\rho}) \to 0$$

for $0 \le p \le 3$, and natural isomorphisms

$$\mathrm{H}^p(S_3, \mathrm{H}^1(\mathcal{M}(2)_S, \mathbb{G}_m)) \cong \mathrm{H}^p(S_3, \mathrm{Pic}(S)) \oplus \mathrm{H}^p(S_3, \mu_2(S))$$

for all $p \ge 0$.

Proof. The first exact sequence follows from Lemmas 5.2 and 5.3 using that $\mathrm{H}^1(S_3, \tilde{\rho})$ is 3-torsion and $\mathrm{H}^2(S_3, \mathbb{G}_m(S))$ is 2-torsion. The direct sum decomposition follows from the fact that (5.5) is S_3-equivariantly split. □

The only necessary remaining group we need to understand for our computations is $\mathrm{H}^2(\mathcal{M}(2)_S, \mathbb{G}_m)^{S_3}$, which we analyze using the short exact sequence (5.6).

Lemma 5.8. *If S is a regular noetherian scheme over $\mathbb{Z}\big[\frac{1}{2}\big]$, then there is a canonical isomorphism $\mathrm{H}^1(S, \mu_2) \cong \mathrm{H}^1(X_S, \mu_2)^{S_3}$.*

Proof. Using the S_3-equivariant short exact sequence

$$0 \to \mathbb{G}_m(X_S)/2 \to \mathrm{H}^1(X_S, \mu_2) \to \mathrm{Pic}(X_S)[2] \to 0,$$

we get a long exact sequence

$$0 \to (\mathbb{G}_m(X_S)/2)^{S_3} \to \mathrm{H}^1(X_S, \mu_2)^{S_3} \to \mathrm{Pic}(X_S)[2]^{S_3} \to \mathrm{H}^1(S_3, \mathbb{G}_m(X_S)/2) \to \cdots .$$

As the canonical map $X_S \to S$ is S_3-equivariant, we obtain a map into this from the exact sequence

$$0 \to \mathbb{G}_m(S)/2 \to \mathrm{H}^1(S, \mu_2) \to \mathrm{Pic}(S)[2] \to 0.$$

As the maps $\mathbb{G}_m(S)/2 \to (\mathbb{G}_m(X_S)/2)^{S_3}$ and $\mathrm{Pic}(S)[2] \to \mathrm{Pic}(X_S)[2]$ are isomorphisms (using the exact sequence (5.4) and Lemma 5.3), the five lemma implies that $\mathrm{H}^1(S, \mu_2) \to \mathrm{H}^1(X_S, \mu_2)^{S_3}$ is an isomorphism as well. □

From (5.6), we obtain a long exact sequence

$$0 \to \mathrm{Br}'(X_S)^{S_3} \to \mathrm{Br}'(\mathcal{M}(2))^{S_3} \to \mathrm{H}^1(S, \mu_2) \to \mathrm{H}^1(S_3, \mathrm{Br}'(X_S)) \to \cdots . \tag{5.9}$$

Lemma 5.10. *Let S be a regular noetherian scheme over* $\mathrm{Spec}\,\mathbb{Z}[1/p]$ *for some prime p, and let $X_S = \mathbb{A}^1_S - \{0, 1\}$ as before. There is a noncanonically split exact sequence*

$$0 \to {}_p\,\mathrm{Br}'(S) \to {}_p\,\mathrm{Br}'(X_S) \to {}_p\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \to 0.$$

Proof. By Proposition 2.14 we have an exact sequence

$$0 \to {}_p\,\mathrm{Br}'(\mathbb{A}^1_S) \to {}_p\,\mathrm{Br}'(X_S) \to {}_p\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \to {}_p\mathrm{H}^3(\mathbb{A}^1_S, \mathbb{G}_m) \to {}_p\mathrm{H}^3(X_S, \mathbb{G}_m).$$

Because p is invertible on S, Proposition 2.5 implies that ${}_p\mathrm{H}^i(S, \mathbb{G}_m) \cong {}_p\mathrm{H}^i(\mathbb{A}^1_S, \mathbb{G}_m)$ for all $i \geq 0$. But, since X_S has an S-point, it follows that ${}_p\mathrm{H}^i(\mathbb{A}^1_S, \mathbb{G}_m) \cong {}_p\mathrm{H}^i(S, \mathbb{G}_m) \to {}_p\mathrm{H}^i(X_S, \mathbb{G}_m)$ is split injective for all i. □

Lemma 5.11. *For any prime p and any regular noetherian scheme over* $\mathrm{Spec}\,\mathbb{Z}\left[\frac{1}{p}\right]$, *there is a canonical isomorphism*

$$\mathrm{H}^q_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \cong \mathrm{H}^q_{\{0\}}(\mathbb{A}^1_S, \mathbb{G}_m) \oplus \mathrm{H}^q_{\{1\}}(\mathbb{A}^1_S, \mathbb{G}_m).$$

The action of S_3 on ${}_p\,\mathrm{Br}'(X_S)/{}_p\,\mathrm{Br}'(S)$ is isomorphic to

$$\tilde{\rho} \otimes {}_p\mathrm{H}^3_{\{0\}}(\mathbb{A}^1_S, \mathbb{G}_m) \cong \tilde{\rho} \otimes \mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p).$$

Proof. Given any étale sheaf $\mathcal{F}$, there is a canonical isomorphism

$$\mathrm{H}^0_{\{0,1\}}(\mathbb{A}^1_S, \mathcal{F}) \cong \mathrm{H}^0_{\{0\}}(\mathbb{A}^1_S, \mathcal{F}) \oplus \mathrm{H}^0_{\{1\}}(\mathbb{A}^1_S, \mathcal{F}),$$

as one sees by an easy diagram chase. By deriving this isomorphism, the first part of the lemma follows.

To prove the second statement, we compare the sequence of Lemma 5.10 with the long exact sequence for étale cohomology with supports coming from the open inclusion $X_S \subseteq \mathbb{P}^1_S$. Using the natural map of long exact sequences, we obtain a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & {}_p\operatorname{Br}'(\mathbb{P}^1_S) \arrow[r] \arrow[d, "\cong"] & {}_p\operatorname{Br}'(X_S) \arrow[r] \arrow[d, "="] & {}_p\mathrm{H}^3_{\{0,1,\infty\}}(\mathbb{P}^1_S, \mathbb{G}_m) \arrow[d] & \\ 0 \arrow[r] & {}_p\operatorname{Br}'(\mathbb{A}^1_S) \arrow[r] & {}_p\operatorname{Br}'(X_S) \arrow[r] & {}_p\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \arrow[r] & 0 \end{tikzcd}$$

with exact rows, where the left-hand vertical map is an isomorphism because it is injective (by Proposition 2.5(iv)), ${}_p\operatorname{Br}'(S) \to {}_p\operatorname{Br}'(\mathbb{A}^1_S)$ is an isomorphism, and there is an S-point of $\mathbb{A}^1_S \subseteq \mathbb{P}^1_S$.

Now, by Proposition 2.14,

$${}_p\mathrm{H}^3_{\{0,1,\infty\}}(\mathbb{P}^1_S, \mathbb{G}_m) \cong \bigoplus_{\{0,1,\infty\}} \mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p) \quad \text{and} \quad {}_p\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \cong \bigoplus_{\{0,1\}} \mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p).$$

With this description, the right-hand vertical map above is the natural projection away from the factor of $\mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p)$ corresponding to ∞. Let χ_0 and χ_1 be p-primary characters of S, i.e., elements of $\mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p)$. Then, as χ_0, χ_1 vary, the Azumaya algebras $(\chi_0, t) \otimes (\chi_1, t-1)$ give elements of $\operatorname{Br}(X_S)$ whose ramification classes (χ_0, χ_1) span ${}_p\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m)$. The ramification of such a class computed in $\mathrm{H}^3_{\{0,1,\infty\}}(\mathbb{P}^1_S, \mathbb{G}_m)$ is $(\chi_0, \chi_1, -\chi_0 - \chi_1)$. This follows from Proposition 2.16, the fact that $\mathrm{ram}_{(\pi)}(\chi, \pi^{-1}) = -\mathrm{ram}_{(\pi)}(\chi, \pi)$ in the notation of that proposition, and the fact that both t^{-1} and $(t-1)^{-1}$ are uniformizing parameters for the divisor at ∞ of $\mathbb{P}^1_S$. It follows that the image of ${}_p\operatorname{Br}'(X_S)$ inside ${}_p\mathrm{H}^3_{\{0,1,\infty\}}(\mathbb{P}^1_S, \mathbb{G}_m) \cong \bigoplus_{\{0,1,\infty\}} \mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p)$ can be identified with $\tilde{\rho} \otimes \mathrm{H}^1(S, \mathbb{Q}_p/\mathbb{Z}_p)$. □

We will analyze the implications for p-primary torsion for $p > 2$ in the next three sections. For the rest of this section, we will begin the study of the 2-primary torsion of $\operatorname{Br}'(\mathcal{M}_S)$ where S is a $\mathbb{Z}\left[\frac{1}{2}\right]$-scheme. By Lemmas 5.11 and 5.3, we know that ${}_2\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m)^{S_3} = 0$. Thus from Lemma 5.10, we see that ${}_2\operatorname{Br}'(S) \to {}_2\operatorname{Br}'(X_S)^{S_3}$ is an isomorphism. If we tensor the sequence (5.9) with $\mathbb{Z}_{(2)}$, we obtain (using Lemmas 5.2 and 5.3) the exact sequence

$$0 \to {}_2\operatorname{Br}'(S) \to {}_2\operatorname{Br}'(\mathcal{M}(2))^{S_3} \to \mathrm{H}^1(S, \mu_2) \xrightarrow{\partial} \mathrm{H}^1(S_3, {}_2\operatorname{Br}'(X_S)) \cong \operatorname{Br}'(S)[2] \to \cdots. \tag{5.12}$$

Here we recall that we denote for an abelian group A by $A[2]$ its 2-torsion, while ${}_2A$ denotes its 2-primary torsion. We want to analyze the boundary map ∂.

Lemma 5.13. *If $u \in \mathrm{H}^1(S, \mu_2)$, then $\partial(u)$ equals the Brauer class of the cyclic (quaternion) algebra $(-1, u)$.*

Proof. We assume that S is connected. Denote by $\pi\colon X_S \to \mathrm{B}C_{2,X_S}$ the projection and by $c\colon \mathrm{B}C_{2,X_S} \to X_S$ the canonical map to the coarse moduli space. Denote by

$$r\colon \operatorname{Br}'(\mathrm{B}C_{2,X_S}) \to \mathrm{H}^1(X_S, \mu_2)$$

the map obtained from the Leray spectral sequence. Finally, let

$$s\colon \mathrm{H}^1(X_S,\mu_2)\to \mathrm{Br}'(\mathrm{B}C_{2,X_S})$$

be given by $s(u)=[(\chi,u)]=[(\chi,c^*u)]$, where $\chi\in\mathrm{H}^1(\mathrm{B}C_{2,X_S},C_2)$ classifies π. We have $r(s(u))=u$ by Lemma 3.3.

Using [Serre 1997, Section 5.4], we can compute a crossed homomorphism representing $\partial(u)$ thus as

$$g\mapsto \pi^*(g(s(u))-s(u))\in \mathrm{Br}'(X_S).$$

Consider the subgroup $C_2=\langle(2\,3)\rangle\subset S_3$ and the C_2-equivariant morphism $z\colon S\to X_S$ classifying $y^2=x(x-1)(x+1)$ (i.e., $t=-1$). It follows from Lemmas 5.2 and 5.3 that the morphism $z^*\operatorname{res}^{S_3}_{C_2}$ induces an isomorphism $\mathrm{H}^1(S_3,{}_2\mathrm{Br}'(X_S))\to\mathrm{H}^1(C_2,{}_2\mathrm{Br}'(S))$. The isomorphism $\mathrm{H}^1(C_2,\mathrm{Br}'(S))\to\mathrm{Br}'(S)[2]$ is given by evaluating the crossed homomorphism at the nontrivial element $(2\,3)\in C_2$. Thus, the coboundary map $\partial:\mathrm{H}^1(S,\mu_2)\to\mathrm{Br}'(S)[2]$ sends u to

$$z^*\pi^*((2\,3)(s(u))-s(u)).$$

As the pullback of $X_S\to\mathrm{B}C_{2,X_S}$ along $\pi\circ z\colon S\to X_S\to\mathrm{B}C_{2,X_S}$ is the trivial C_2-torsor, $z^*\pi^*s(u)=(\pi z)^*(\chi,u)$ defines the trivial Brauer class. By Lemma 4.7, the action of $(2\,3)$ multiplies the torsor $X_S\to\mathrm{B}C_{2,X_S}$ with the torsor $\mathrm{B}C_{2,X_S}(\sqrt{t})\to\mathrm{B}C_{2,X_S}$. Thus, $z^*\pi^*(2\,3)(s(u))=(z^*t,u)=(-1,u)$. □

Summarizing, we obtain the following result.

Proposition 5.14. *Let S be a regular noetherian scheme over $\mathbb{Z}\left[\frac{1}{2}\right]$. We have an exact sequence*

$$0\to{}_2\mathrm{Br}'(S)\to{}_2\mathrm{Br}'(\mathcal{M}(2)_S)^{S_3}\to\mathrm{H}^1(S,\mu_2)\xrightarrow{\partial}\mathrm{Br}'(S)[2]$$

with $\partial(u)=[(-1,u)]$. The map ${}_2\mathrm{Br}'(S)\to{}_2\mathrm{Br}'(\mathcal{M}(2)_S)^{S_3}$ is noncanonically split.

Proof. The exact sequence is exactly (5.12). The identification of $\partial(u)$ follows from the previous lemma. For the splitting, choose an S-point $S\to\mathcal{M}(2)_S$. Then the composition $\mathrm{Br}'(\mathcal{M}(2)_S)^{S_3}\to\mathrm{Br}'(\mathcal{M}(2)_S)\to\mathrm{Br}'(S)$ provides the splitting. □

6. The p-primary torsion in $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ for primes $p\geq 5$

Before we proceed to study the 3-primary and 2-primary torsion, we will show in this section that for a large class of S there is no p-primary torsion for $p\geq 5$ in the Brauer group of $\mathcal{M}_S$. Lemma 5.3 implies the crucial fact that there are no S_3-invariant classes in ${}_p\mathrm{Br}'(X_S)$ ramified at $\{0,1\}$ when $p\neq 3$ is invertible on S. The main point of the following theorem is that this is true for $p\geq 5$ even for certain regular noetherian schemes where p is not a unit.

Theorem 6.1. *Let S be a regular noetherian scheme over $\mathbb{Z}$ and $p\geq 5$ prime. Assume that $S[1/(2p)]=S_{\mathbb{Z}[1/(2p)]}$ is dense in S and that $\mathcal{M}_S\to S$ has a section. Then the natural map ${}_p\mathrm{Br}'(S)\to{}_p\mathrm{Br}'(\mathcal{M}_S)$ is an isomorphism.*

Proof. Assume first that 2 is invertible on S. The only contribution to ${}_p\operatorname{Br}'(\mathcal{M}_S)$ in the Leray–Serre spectral sequence (5.1) occurs as

$$ {}_p(\mathrm{H}^2(\mathrm{B}C_{2,X_S}, \mathbb{G}_m)^{S_3}) = ({}_p\mathrm{H}^2(\mathrm{B}C_{2,X_S}, \mathbb{G}_m))^{S_3} \tag{6.2}$$

because H^i of S_3 for $i \geq 1$ can never have p-primary torsion for $p \geq 5$.

We will argue that the p-group (6.2) is isomorphic to ${}_p\operatorname{Br}'(S)$ for all primes $p \geq 5$ if additionally p is invertible on S. To do so, note first that

$$ {}_p\mathrm{H}^2(\mathrm{B}C_{2,X_S}, \mathbb{G}_m) \cong {}_p\operatorname{Br}'(X_S)$$

for $p \neq 2$ by Proposition 3.2. By Lemmas 5.10, 5.11 and 5.3, we see that ${}_p\operatorname{Br}'(S) \to {}_p\operatorname{Br}'(X_S)^{S_3}$ is an isomorphism.

This shows the theorem if $2p$ is invertible on S. Let now S be arbitrary regular noetherian such that $S[1/(2p)] \subset S$ is dense and $\mathcal{M}_S$ has an S-point. Consider the commutative diagram

$$\begin{tikzcd}
{}_p\operatorname{Br}'(\mathcal{M}_S) \arrow[r] \arrow[d] & {}_p\operatorname{Br}'(\mathcal{M}_{S[1/(2p)]}) \arrow[d, "\cong"] \\
{}_p\operatorname{Br}'(S) \arrow[r] & {}_p\operatorname{Br}'\big(S\big[\frac{1}{2p}\big]\big).
\end{tikzcd}$$

induced by the choice of an S-point of $\mathcal{M}_S$. As $\mathcal{M}_S$ has a cover by a scheme that is fppf over S and fppf morphisms are open [EGA IV$_2$ 1965, Theorem 2.4.6], $\mathcal{M}_S \to S$ is open as well. Thus, $\mathcal{M}_{S[1/(2p)]} \subset \mathcal{M}_S$ is dense and hence $\operatorname{Br}'(\mathcal{M}_S) \to \operatorname{Br}'(\mathcal{M}_{S[1/(2p)]})$ is injective by Proposition 2.5. This implies that $\operatorname{Br}'(\mathcal{M}_S) \to \operatorname{Br}'(S)$ is injective as well. As it is also split surjective, we see that it is an isomorphism. $\square$

Remark 6.3. In general, it is a subtle question to decide whether $\mathcal{M}_S$ has an S-point. For example for $S = \mathbb{Z}\big[\frac{1}{2}\big]$ or $S = \mathbb{Z}\big[\frac{1}{3}\big]$, there is such an S-point, but for $S = \mathbb{Z}\big[\frac{1}{5}\big]$ or $S = \mathbb{Z}\big[\frac{1}{29}\big]$ there is none (for this and other examples; see [Edixhoven et al. 1990, Corollary 1]). Nevertheless, sometimes one can still control the p-power torsion for $p \geq 5$ if there is no S-point, as the following corollary shows.

Corollary 6.4. *The Brauer groups* $\operatorname{Br}(\mathcal{M}) \subseteq \operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ *have only* 2 *and* 3*-primary torsion.*

Proof. Indeed, $\operatorname{Br}(\mathcal{M}) \subseteq \operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$, and there is no p-torsion in $\operatorname{Br}\big(\mathbb{Z}\big[\frac{1}{2}\big]\big) \cong \mathbb{Z}/2$ for $p \neq 2$. $\square$

7. The 3-primary torsion in $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/6]})$

The next theorem describes the 3-primary torsion in $\operatorname{Br}'(\mathcal{M}_S)$ in many cases.

Theorem 7.1. *Let S be a regular noetherian scheme. If* 6 *is a unit on* S, *then there is an exact sequence*

$$0 \to {}_3\operatorname{Br}'(S) \to {}_3\operatorname{Br}'(\mathcal{M}_S) \to \mathrm{H}^1(S, C_3) \to 0,$$

which is noncanonically split. The map ${}_3\operatorname{Br}'(\mathcal{M}_S) \to \mathrm{H}^1(S, C_3)$ *can be described as the composition of pullback to* X_S *and taking the ramification at the divisor* $\{0\}$ *in* $\mathbb{A}^1_S$ *defined by* t (*using Proposition 2.14*).

Proof. The 3-primary torsion in $\mathrm{Br}'(\mathcal{M}(2)_S) \cong \mathrm{Br}'(\mathrm{B}C_{2,X_S})$ is just the 3-primary torsion in $\mathrm{Br}'(X_S)$ by Proposition 3.2 as $\mathrm{H}^1_{pl}(X_S; \mu_2)_{(3)} = 0$. Similarly, since 2 is invertible in S, the Leray–Serre spectral sequence (5.1) together with the group cohomology computations of Lemmas 5.2 and 5.3 and the exact sequences (5.4)–(5.6) say that

$${}_3\,\mathrm{Br}'(\mathcal{M}_S) \cong ({}_3\mathrm{Br}'(X_S))^{S_3}.$$

Since 3 is invertible in S, we have a short exact sequence

$$0 \to {}_3\,\mathrm{Br}'(S) \to {}_3\,\mathrm{Br}'(X_S) \to {}_3\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \to 0$$

by Lemma 5.10.

The 0 and 1 sections are disjoint, so that there is an isomorphism of S_3-modules

$${}_3\mathrm{H}^3_{\{0,1\}}(\mathbb{A}^1_S, \mathbb{G}_m) \cong \bigoplus_{i=0,1} \mathrm{H}^1(S, \mathbb{Q}_3/\mathbb{Z}_3) \cong \tilde{\rho} \otimes \mathrm{H}^1(S, \mathbb{Q}_3/\mathbb{Z}_3)$$

by Proposition 2.14 and Lemma 5.11. Thus, Lemma 5.3 implies that the long exact sequence in S_3-cohomology takes the following form:

$$0 \to {}_3\,\mathrm{Br}'(S) \to {}_3\,\mathrm{Br}'(X_S)^{S_3} \to \mathrm{H}^1(\mathrm{Spec}\, S, C_3) \to \mathrm{H}^1(S_3, {}_3\,\mathrm{Br}'(S)).$$

However, the action of S_3 on ${}_3\,\mathrm{Br}'(S)$ is trivial, so the group on the right vanishes by Lemma 5.2. Since $\mathcal{M}_S$ has an S-point (because 2 is inverted), the splitting follows.

The map ${}_3\,\mathrm{Br}'(X_S) \to \bigoplus_{i=0,1} \mathrm{H}^1(S, \mathbb{Q}_3/\mathbb{Z}_3)$ takes the ramification at the divisors $\{0\}$ and $\{1\}$. At this point, we need to make the isomorphism $(\tilde{\rho} \otimes \mathrm{H}^1(S, \mathbb{Q}_3/\mathbb{Z}_3))^{S_3} \to \mathrm{H}^1(S, C_3)$ more explicit. By choosing the ordered basis $t, t-1$ of $\mathbb{G}_m(X_S)/\mathbb{G}_m(S) \cong \tilde{\rho}$, we see from the description of the action that $\mathbb{Z}/3 \cong (\tilde{\rho}/3)^{S_3} \subseteq \tilde{\rho}/3$ is generated by $t(t-1)$. Indeed, $\sigma(t(t-1)) = -t^{-2}(t-1)$ and $\tau(t(t-1)) = -t^{-2}(t-1)$ for $\sigma = (1\,3\,2)$ and $\tau = (2\,3)$ and thus

$$\sigma(t(t-1)) \equiv t(t-1) \equiv \tau(t(t-1)) \in \mathbb{G}_m(X_S)/3.$$

This implies that $(\tilde{\rho} \otimes \mathrm{H}^1(S, \mathbb{Q}_3/\mathbb{Z}_3))^{S_3} \to \mathrm{H}^1(S, C_3)$ can be identified with projection onto the first coordinate. Thus, ${}_3\,\mathrm{Br}'(S_R)^{S_3} \to \mathrm{H}^1(S, C_3)$ takes the ramification at the divisor $\{0\}$. □

We want to be more specific about the Azumaya algebras arising from $\mathrm{H}^1(S, C_3)$. For that purpose consider the section $\Delta \in \mathrm{H}^0(\mathcal{M}, \lambda^{\otimes 12})$, which is defined as follows. Given an elliptic curve E over S, we can write it Zariski locally in Weierstrass form. Consider its discriminant $\Delta_E \in \mathcal{O}(S)$ and its invariant differential $\omega \in \Omega^1_{E/R}(E) \cong \lambda(R)$. It is easy to see by [Silverman 2009, Table 3.1] that $\Delta = \Delta_E \omega^{\otimes 12}$ is a section of $\lambda^{\otimes 12}$, which is invariant under coordinate changes. Thus, Δ defines a section of $\lambda^{\otimes 12}$ on $\mathcal{M}$.

Lemma 7.2. *By Construction 2.8, we can associate with the line bundle $\mathcal{L} = \lambda^{\otimes 4}$ and the trivialization $\Delta\colon \mathcal{O}_{\mathcal{M}} \to \mathcal{L}^{\otimes 3}$ the μ_3-torsor $\mathcal{M}(\sqrt[3]{\Delta}) \to \mathcal{M}$ whose class in $\mathrm{H}^1(\mathcal{M}, \mu_3)$ we denote by $[\Delta]_3$. If S is a regular noetherian scheme and 6 is a unit on S, then the composite*

$$\mathrm{H}^1(S, C_3) \to \mathrm{H}^1(\mathcal{M}_S, C_3) \xrightarrow{\cup(-[\Delta]_3)} \mathrm{H}^2(\mathcal{M}_S, \mu_3) \to {}_3\,\mathrm{Br}'(\mathcal{M}_S)$$

is a section of the map ${}_3\,\mathrm{Br}'(\mathcal{M}_S) \to \mathrm{H}^1(S, C_3)$ of Theorem 7.1.

Remark 7.3. Informally, this section associates with $\chi \in \mathrm{H}^1(S, C_3)$ the symbol algebra $[(\chi, \Delta^{-1})_3]$.

Proof. The pullback of Δ to X is the discriminant of the universal Legendre curve, which is $16t^2(t-1)^2$ (using the standard trivialization of λ on X given by $dx/(2y)$). For $\chi \in \mathrm{H}^1(S, C_3)$ the pullback of $[(\chi, \Delta^{-1})_3]$ to $\mathrm{Br}'(X_S)$ is thus $[(\chi, 4t(t-1))_3]$. As $4t(t-1)$ is a uniformizer for the local ring of $\mathbb{A}^1_S$ at $t=0$, Proposition 2.16 implies the result. □

Remark 7.4. While this map $\mathrm{H}^1(S, C_3) \to \mathrm{Br}'(\mathcal{M}_S)$ is defined whether or not 6 is a unit on S, without this assumption we do not know that $\mathrm{H}^1(S, C_3)$ is the cokernel of ${}_3\mathrm{Br}'(S) \to {}_3\mathrm{Br}'(\mathcal{M}_S)$.

Corollary 7.5. *When $S = \mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{6}\big]$, there is an isomorphism ${}_3\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/6]}) \cong \mathbb{Q}_3/\mathbb{Z}_3 \oplus \mathbb{Z}/3$. The 3-torsion subgroup is generated by classes σ and θ, which can be described as follows. Let $\chi \in \mathrm{H}^1\big(\mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{6}\big], C_3\big)$ be the character of the Galois extension $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$ of $\mathbb{Q}$. Then $\sigma = [(\chi, 6)_3]$ and $\theta = [(\chi, 16\Delta^{-1})_3]$, which pulls back to $[(\chi, t(t-1))_3]$ on $X_{\mathbb{Z}[1/6]}$.*

Proof. We claim first that $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$ is the only cyclic cubic extension L of $\mathbb{Q}$ that ramifies at most at 2 and 3. This can either be deduced from [Hasse 1948, I. Section 1.2] or shown as follows. By the Kronecker–Weber theorem, any cyclic cubic extension L of $\mathbb{Q}$ has to embed into a cyclotomic extension $\mathbb{Q}(\zeta_n)$ and is more precisely its fixed field under a normal subgroup $H \subset (\mathbb{Z}/n)^\times$ of index 3. As H contains all elements of 2-power order, we can assume that n is odd and thus L does not ramify at 2 but only at 3. Proposition 3.1 of [Lemmermeyer 2005] shows that L is unique and must be $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$. This implies that $\mathrm{H}^1\big(\mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{6}\big], C_3\big) \cong \mathbb{Z}/3$.

Using that ${}_3\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{6}\big]\big) \cong \mathbb{Q}_3/\mathbb{Z}_3$, the structure of the Brauer group $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/6]})[3]$ follows from Theorem 7.1. The description of θ follows directly from the last lemma (where we have modified the section by an element of ${}_3\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{6}\big]\big)$ for convenience).

Last we need to show that $[(\chi, 6)_3]$ is nonzero in $\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{6}\big]\big)[3] \cong \mathbb{Z}/3$. It suffices to check that $(\chi, 6)$ is ramified at the prime (2). Note that the minimal polynomial of $\zeta_9 + \bar{\zeta}_9$ is $w^3 + w + 1$. By Proposition 2.16, the ramification at (2) in $\mathrm{H}^1(\mathbb{F}_2, C_3) \cong \mathbb{Z}/3$ is the class of the extension $w^3 + w + 1$ over $\mathbb{F}_2$. Since this polynomial is irreducible (it has no solutions in $\mathbb{F}_2$ and it has degree 3), it follows that the ramification is nonzero. □

Corollary 7.6. *Let $R = \mathbb{Z}\big[\frac{1}{2}\big]$ or $R = \mathbb{Z}$. Then the order of ${}_3\mathrm{Br}(\mathcal{M}_R)$ is either 1 or 3.*

Proof. Using the injectivity of $\mathrm{Br}(\mathcal{M}) \to \mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$, it suffices to prove this when $R = \mathbb{Z}\big[\frac{1}{2}\big]$. Now, we claim that no nonzero class $\alpha \in {}_3\mathrm{Br}\big(\mathbb{Z}\big[\frac{1}{6}\big]\big) \subseteq {}_3\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/6]})$ extends to $\mathcal{M}_{\mathbb{Z}[1/2]}$. Indeed, we can take the Legendre curve $y^2 = x(x-1)(x-2)$, which defines a point $\mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{2}\big] \to \mathcal{M}$. Using the commutative diagram

$$\begin{tikzcd} \mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{6}\big] \arrow[r] \arrow[d] & \mathcal{M}_{\mathbb{Z}[1/6]} \arrow[r] \arrow[d] & \mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{6}\big] \\ \mathrm{Spec}\,\mathbb{Z}\big[\frac{1}{2}\big] \arrow[r] & \mathcal{M}_{\mathbb{Z}[1/2]}, & \end{tikzcd}$$

we see that if $\alpha \in {}_3\operatorname{Br}(\mathbb{Z}[\frac{1}{6}])$ did extend to $\mathcal{M}_{\mathbb{Z}[1/2]}$, then it would be zero in the Brauer group of $\operatorname{Spec}\mathbb{Z}[\frac{1}{6}]$, as it would extend to ${}_3\operatorname{Br}(\mathbb{Z}[\frac{1}{2}]) = 0$. However, these classes are all nonzero in $\operatorname{Br}(\mathbb{Z}[\frac{1}{6}])$ since the composition at the top of the commutative diagram is the identity for any $\mathbb{Z}[\frac{1}{6}]$-point of the moduli stack.

Now, if $\alpha = \beta + m\theta$ extends, where $\beta \in {}_3\operatorname{Br}(\mathbb{Z}[\frac{1}{6}])$, then $3\alpha = 3\beta$ also extends. Hence, it must be that α has order at most 3. In particular, this means that every class of ${}_3\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ is actually 3-torsion and hence this group is a subgroup of $(\mathbb{Z}/3)^2$. But, we have already seen that σ does not extend. So, it is a proper subgroup, and hence it has order at most 3. □

Proposition 7.7. *Suppose there are Legendre curves $E_i : y^2 = x(x-1)(x-t_i)$ over $\operatorname{Spec}\mathbb{Z}_3[\zeta_3]$ for $i = 1, 2$ such that*

$$[(\chi, t_1(t_1-1))] \neq 0 \quad \textit{and} \quad [(\chi, t_2(t_2-1))] = 0$$

in $\operatorname{Br}(\mathbb{Q}_3(\zeta_3))[3] = \mathbb{Z}/3$, where χ denotes the pullback of the Galois extension $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$ of $\mathbb{Q}$ to $\mathbb{Q}_3(\zeta_3)$. Then ${}_3\operatorname{Br}(\mathcal{M}[\frac{1}{2}]) = 0$.

Proof. Suppose that $\alpha = a\sigma + b\theta$ is a linear combination of the classes found in Corollary 7.5 where we can assume that $b \in \{1, 2\}$ since σ does not extend. Suppose that α extends to $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$. We can pull back α along the two $\mathbb{Q}_3(\zeta_3)$-points of $\mathcal{M}$ defined by E_i and compute the ramifications in $\operatorname{Br}(\mathbb{Q}_3(\zeta_3))$. Let $k = [(\chi, t_1(t_1-1))]$. For $i = 1$, we get $a + bk$ and for $i = 2$ we get a. Since k is nonzero, these cannot be simultaneously zero modulo 3. But the two maps $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]}) \to \operatorname{Br}(\mathbb{Q}_3(\zeta_3))$ factor over $\operatorname{Br}(\mathbb{Z}_3[\zeta_3]) = 0$, which is a contradiction. Thus, α cannot extend to $\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$. □

8. The ramification of the 3-torsion

Our aim in this section is to show that ${}_3\operatorname{Br}(\mathcal{M}[\frac{1}{2}]) = 0$ using Proposition 7.7.

Lemma 8.1. *The natural inclusion $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9, \zeta_3) \to \mathbb{Q}(\zeta_9)$ is an isomorphism.*

Proof. The left-hand side is a subfield of $\mathbb{Q}(\zeta_9)$ that is strictly larger than $\mathbb{Q}(\zeta_9 + \bar{\zeta}_9)$ and thus is equal to $\mathbb{Q}(\zeta_9)$. □

We will need the following lemma to aid our Hilbert symbol calculations below.

Lemma 8.2. *Consider the cyclotomic field $\mathbb{Q}(\zeta)$ with $\zeta = \zeta_3$ and $\pi = 1 - \zeta$ and denote by* Tr *the trace for $\mathbb{Q}(\zeta)$ over $\mathbb{Q}$. Then*

$$\operatorname{Tr}(\pi^{6k+l}) = \begin{cases} (-1)^{3k} \cdot 3^{3k} \cdot 2 & \textit{if } l = 0, \\ (-1)^{3k} \cdot 3^{3k+1} & \textit{if } l = 1, \\ (-1)^{3k} \cdot 3^{3k+1} & \textit{if } l = 2, \\ 0 & \textit{if } l = 3, \\ (-1)^{3k+1} \cdot 3^{3k+2} & \textit{if } l = 4, \\ (-1)^{3k+1} \cdot 3^{3k+3} & \textit{if } l = 5, \end{cases}$$

Proof. We have $\pi^2 = (1-\zeta)^2 = -3\zeta$ and thus $\pi^{6k} = (-3)^{3k}$. Therefore, we have just to compute $\operatorname{Tr}(\pi^l)$ for $l = 0, \dots, 5$, which is easily done. □

We come to a key arithmetic point in our proof, where we compute the Hilbert symbol at the prime 3 of certain degree 3 cyclic algebras. By Proposition 2.17, this will allow us to check whether certain cyclic algebras are zero in the Brauer group.

Lemma 8.3. *Consider the cyclotomic field $\mathbb{Q}_3(\zeta)$ with $\zeta = \zeta_3$ and $\pi = 1-\zeta$ the uniformizer. Then we have*

$$\left(\frac{\zeta,\, t(t-1)}{\pi}\right) = \zeta^{1-b^2}$$

in $\mu_3(\mathbb{Q}_3(\zeta))$, where $t = 2 + b\pi$ with $b \in \mathbb{Z}_3$.

Proof. We use the formula of Artin and Hasse (see [Neukirch 1999, Theorem V.3.8]) to compute this Hilbert symbol. By this formula, we have

$$\left(\frac{\zeta,\, a}{\pi}\right) = \zeta^{\mathrm{Tr}(\log a)/3},$$

where $a \in 1+\mathfrak{p}$ (for $\mathfrak{p} \subset \mathbb{Z}_3[\zeta_3]$ the maximal ideal) and Tr the trace for $\mathbb{Q}_3(\zeta)$ over $\mathbb{Q}_3$.

This formula directly applies to $t-1 = 1+b\pi$. We have

$$\log(t-1) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{(b\pi)^i}{i}.$$

Again, it follows easily from Lemma 8.2 that $\mathrm{Tr}(\pi^i)/i$ is divisible by 9 for $i \geq 3$. Thus,

$$\frac{\mathrm{Tr}(\log(t-1))}{3} \equiv \frac{\mathrm{Tr}(b\pi)}{3} - \frac{\mathrm{Tr}(b^2\pi^2)}{6} \equiv b - \frac{b^2}{2} \mod 3$$

and $\left(\frac{\zeta,t-1}{\pi}\right) = \zeta^{b-b^2/2}$.

To compute $\left(\frac{\zeta,t}{\pi}\right)$ note that $\left(\frac{\zeta,t}{\pi}\right) = \left(\frac{\zeta,-t}{\pi}\right)\left(\frac{\zeta,-1}{\pi}\right) = \left(\frac{\zeta,-t}{\pi}\right)$. Indeed, $\left(\frac{\zeta,-1}{\pi}\right)^2 = 1$ and hence also $\left(\frac{\zeta,-1}{\pi}\right) = 1$ (in $\mu_3(\mathbb{Q}_3(\zeta_3))$). We have $-t = 1 + (-3-b\pi)$. Thus,

$$\log(-t) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{(-3-b\pi)^i}{i}.$$

It follows easily from Lemma 8.2 that $\mathrm{Tr}((-3-b\pi)^i))/i$ is divisible by 9 for $i \geq 3$. Thus,

$$\frac{\mathrm{Tr}(\log(-t))}{3} \equiv \frac{\mathrm{Tr}(-3-b\pi)}{3} - \frac{\mathrm{Tr}((-3-b\pi)^2)}{6} \equiv -2-b-\frac{b^2}{2} \mod 3.$$

Thus, $\left(\frac{\zeta,t}{\pi}\right) = \left(\frac{\zeta,-t}{\pi}\right) = \zeta^{-2-b-b^2/2}$. It follows that

$$\left(\frac{\zeta,\, t(t-1)}{\pi}\right) = \left(\frac{\zeta,\, t}{\pi}\right)\left(\frac{\zeta,\, t-1}{\pi}\right) = \zeta^{-2-b^2} = \zeta^{1-b^2},$$

as desired. □

Theorem 8.4. *We have*

$${}_3\operatorname{Br}(\mathcal{M}) = {}_3\operatorname{Br}(\mathcal{M}_{\mathbb{Z}[1/2]}) = 0.$$

$$\begin{tikzcd}
{}_2\operatorname{Br}'(S)\oplus\operatorname{Pic}(S)[2]\oplus G \arrow[drr] & & & \\
\operatorname{Pic}(S)_{(2)}\oplus\mu_2(S) \arrow[drr] & \operatorname{Pic}(S)[2]\oplus\mu_2(S) \arrow[drr] & \operatorname{Pic}(S)/2\oplus\mu_2(S) & \\
\mathbb{G}_m(S)_{(2)} & \mu_2(S) & \mathbb{G}_m(S)/2 & \mu_2(S)
\end{tikzcd}$$

Figure 3. The E_2-page of the Leray–Serre spectral sequence computing $\mathrm{H}^i(\mathcal{M}_S, \mathbb{G}_m)_{(2)}$ for $i \leq 2$.

Proof. By Proposition 7.7 and Proposition 2.17, it suffices to find two Legendre curves E_1 and E_2 over $\mathbb{Z}_3[\zeta_3]$ with corresponding classes $[(\chi, t_1(t_1-1))] \neq 0$ and $[(\chi, t_2(t_2-1))] = 0$. The associated condition on the Hilbert symbols is $\left(\frac{\zeta, t_1(t_1-1)}{\pi}\right) \neq 1$ and $\left(\frac{\zeta, t_2(t_2-1)}{\pi}\right) = 1$. (Recall here that χ is the character associated with adjoining $\zeta_9+\bar{\zeta}_9$ which over $\mathbb{Q}_3(\zeta_3)$ is isomorphic to $\mathbb{Q}_3(\zeta_9)$ by Lemma 8.1.) Take $t_i = 2 + b_i\pi$, where $b_1 = 0$ and $b_2 = 1$. Consider the two elliptic curves

$$E_1 : y^2 = x(x-1)(x-2) \quad \text{and} \quad E_2 : y^2 = x(x-1)(x-(2+\pi)).$$

The previous lemma says that we have

$$\left(\frac{\zeta,\, t_1(t_1-1)}{\pi}\right) = \left(\frac{\zeta,\, 2(2-1)}{\pi}\right) = \zeta \neq 1 \quad \text{and} \quad \left(\frac{\zeta,\, t_2(t_2-1)}{\pi}\right) = \left(\frac{\zeta,\, (2+\pi)(1+\pi)}{\pi}\right) = \zeta^0 = 1.$$

This completes the proof. $\square$

9. The 2-primary torsion in $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$

Throughout this section, let S denote a connected regular noetherian scheme over $\operatorname{Spec}\mathbb{Z}\left[\frac{1}{2}\right]$. Given a stack X over S, let $\overline{\operatorname{Br}}'(X) = \operatorname{coker}(\operatorname{Br}'(S) \to \operatorname{Br}'(X))$.

Theorem 9.1. *Let S be a regular noetherian scheme over $\mathbb{Z}\left[\frac{1}{2}\right]$ with $\operatorname{Pic}(S) = 0$. There is a natural exact sequence*

$$0 \to \mathbb{G}_m(S)/2 \to {}_2\overline{\operatorname{Br}}'(\mathcal{M}_S) \to G \to 0,$$

where $G \subset \mathbb{G}_m(S)/2$ is the subgroup of all those u with $[(-1, u)] = 0 \in \operatorname{Br}(S)$.

We will prove the theorem after several preliminaries. Figure 3 shows a small part of the 2-local Leray–Serre spectral sequence (5.1) for the S_3-Galois cover $\mathcal{M}(2)_S \to \mathcal{M}_S$. The description follows from Lemmas 5.2, 5.3, 5.7 and Proposition 5.14.

From now on, we will localize everything in this section implicitly at 2.

Proposition 9.2. *Let S be a connected regular noetherian scheme over $\mathbb{Z}\left[\frac{1}{2}\right]$. The differential $d_2^{0,1}$ in the Leray–Serre spectral sequence of Figure 3 always vanishes and $d_2^{0,2}$ and $d_3^{0,2}$ vanish if $\operatorname{Pic}(S) = 0$.*

Proof. The map

$$\operatorname{Pic}(\mathcal{M}_S) \to \mathrm{E}_2^{0,1} \cong \operatorname{Pic}(S)_{(2)} \oplus \mu_2(S)$$

is surjective as $-1 \in \mu_2(S)$ can be realized as $\lambda^{\otimes 6}$. This implies that there can be no differential originating from $\mathrm{E}_2^{0,1}$. Moreover, ${}_2\operatorname{Br}'(S)$ splits off from $\operatorname{Br}'(\mathcal{M}_S)$, so the differentials $d_2^{0,2}$ and $d_3^{0,3}$ vanish on ${}_2\operatorname{Br}'(S)$.

Now assume $\operatorname{Pic}(S) = 0$. Then also $\operatorname{Pic}(S)[2] = 0$ and $\operatorname{Pic}(S)/2 = 0$. So, we are concerned with the vanishing of $d_2^{0,2} : G \to \mu_2(S)$ and $d_3^{0,2} : G \to \mu_2(S)$. However, by pulling back the spectral sequence to a geometric point $\bar{x}$ of S, we find $G = \mathbb{G}_m(\bar{x})/2 = 0$, while $\mu_2(\bar{x}) \cong \mathbb{Z}/2$. This implies that $d_2^{0,2}$ and $d_3^{0,2}$ vanish. □

To resolve the differential $d_2^{1,1}$ and solve possible extension issues we will first consider schemes S over $\mathbb{Z}\left[\frac{1}{2}, i\right]$. In this case we can compare the Leray–Serre spectral sequence considered above with the Leray–Serre spectral sequence for the C_2-Galois cover $\mathrm{B}C_{2,S} \to \mathrm{B}C_{4,S}$.

Proposition 9.3. *If S is a regular noetherian $\mathbb{Z}\left[\frac{1}{2}, i\right]$-scheme, then* $\operatorname{Br}'(\mathcal{M}_S) \cong \operatorname{Br}'(\mathrm{B}C_{4,S})$.

Proof. Consider the elliptic curve $E : y^2 = x(x-1)(x+1)$ over $\mathbb{Z}\left[\frac{1}{2}, i\right]$ with discriminant 64. It has an automorphism η of order 4 given by $y \mapsto iy$ and $x \mapsto -x$, which defines a map $\mathrm{B}C_{4,\mathbb{Z}[1/2,i]} \to \mathcal{M}_{\mathbb{Z}[1/2]}$. The 2-torsion points of E are $(0,0)$, $(1,0)$ and $(-1,0)$; taking them in this order defines a full level 2 structure. We can base change this elliptic curve together with its level structure to an arbitrary $\mathbb{Z}\left[\frac{1}{2}, i\right]$-scheme S. This results in pullback squares

$$\begin{tikzcd} \coprod_{S_3} S \arrow[r] \arrow[d] & \coprod_{S_3/C_2} \mathrm{B}C_{2,S} \arrow[r] \arrow[d] & \mathcal{M}(2)_S \arrow[d] \\ S \arrow[r] & \mathrm{B}C_{4,S} \arrow[r] & \mathcal{M}_S. \end{tikzcd} \tag{9.4}$$

Here we use that η acts on the scheme $\coprod_{S_3} S$ of level structures on E_S by multiplication with the cycle $(2\,3) \in S_3$ and in particular η^2 acts trivially. Thus, the stack quotient $\left(\coprod_{S_3} S\right)/C_4$ is equivalent to $\coprod_{S_3/C_2} \mathrm{B}C_{2,S}$. More precisely, the S_3-Galois cover $\coprod_{S_3/C_2} \mathrm{B}C_{2,S} \to \mathrm{B}C_{4,S}$ is induced along an inclusion $C_2 \to S_3$ from the C_2-Galois cover $\mathrm{B}C_{2,S} \to \mathrm{B}C_{4,S}$. The right square is indeed cartesian as can be checked after base change along the étale cover $S \to \mathrm{B}C_{4,S}$.

In the Leray–Serre spectral sequence

$$\mathrm{E}_2^{p,q} = \mathrm{H}^p\left(S_3, \mathrm{H}^q\left(\coprod_{S_3/C_2} \mathrm{B}C_{2,S}, \mathbb{G}_m\right)\right) \Rightarrow \mathrm{H}^{p+q}(\mathrm{B}C_{4,S}, \mathbb{G}_m),$$

the S_3-modules $\mathrm{H}^q\left(\coprod_{S_3/C_2} \mathrm{B}C_{2,S}\right)$ are all induced up from C_2. Thus, the spectral sequence is isomorphic to the Leray–Serre spectral sequence for the C_2-Galois cover $\mathrm{B}C_{2,S} \to \mathrm{B}C_{4,S}$.

The Leray–Serre spectral sequence computing $\mathrm{H}^{p+q}(\mathrm{B}C_{4,S}, \mathbb{G}_m)$ from the $\mathbb{G}_m$-cohomology of $\mathrm{B}C_{2,S}$ is displayed in Figure 4. The computation follows from Proposition 3.2 together with the fact that C_2 acts trivially on the cohomology of $\mathrm{B}C_{2,S}$ (as indeed the morphism $t\colon \mathrm{B}C_{2,S} \to \mathrm{B}C_{2,S}$ for $t \in C_2$ the generator is the identity; only the natural transformation $\mathrm{id} \to t^2$ is not the identity).

By the considerations above, the pullback square (9.4) induces a map

$$\mathrm{H}^p(S_3, \mathrm{H}^q(\mathcal{M}(2)_S, \mathbb{G}_m)) \to \mathrm{H}^p(C_2, \mathrm{H}^q(\mathrm{B}C_{2,S}, \mathbb{G}_m)).$$

$$\begin{tikzcd}[column sep=small]
\mathrm{Br}'(S)\oplus\mathrm{Pic}(S)[2]\oplus\mathbb{G}_m(S)/2 \arrow[drr] & & & \\
\mathrm{Pic}(S)\oplus\mu_2(S) & \mathrm{Pic}(S)[2]\oplus\mu_2(S) \arrow[drr] & \mathrm{Pic}(S)/2\oplus\mu_2(S) & \\
\mathbb{G}_m(S) & \mu_2(S) & \mathbb{G}_m(S)/2 & \mu_2(S)
\end{tikzcd}$$

Figure 4. Part of the Leray–Serre spectral sequence for $\mathrm{B}C_{2,S}\to\mathrm{B}C_{4,S}$.

Note first that $G=\mathbb{G}_m(S)/2$ in our case as -1 is a square. If we identify $\mathcal{M}(2)_S$ with $\mathrm{B}C_{2,X_S}$ this map on cohomology groups is induced by the maps $S\to X_S$ (classifying the Legendre curve E_S) and $C_2\to S_3$. This induces an isomorphism of spectral sequences for $p+q\le 3$ and $q\le 1$ for $p+q=3$ by Figure 3. □

Corollary 9.5. *Let S be a regular noetherian scheme over $\mathbb{Z}\big[\frac{1}{2}\big]$. The restriction of the differential $d_2^{1,1}$ to $\mu_2(S)$ in the Leray–Serre spectral sequence for $\mathcal{M}(2)_S\to\mathcal{M}_S$ defines an isomorphism $\mu_2(S)\xrightarrow{\cong}\mu_2(S)$, while $d_3^{0,2}=0$.*

Proof. Consider $S'=\operatorname{Spec}\mathbb{Z}\big[\frac{1}{2},i\big]$. By Proposition 3.2, we see that

$$\mathrm{Br}'(\mathrm{B}C_{4,S'})\cong\mathrm{Br}'(S')\oplus\mathrm{Pic}(S')[4]\oplus\mathbb{G}_m(S')/4\cong\mathbb{G}_m(S')/4,$$

since the Brauer and Picard groups of $\mathbb{Z}\big[\frac{1}{2},i\big]$ are zero. Hence, $\mathrm{Br}'(\mathrm{B}C_{4,S'})\cong\mathbb{Z}/4\oplus\mathbb{Z}/4$, with generators given as i and $1+i$. In Figure 4, we see that the only way to have a group of order 16 in the abutment $\mathrm{Br}'(\mathrm{B}C_{4,S'})$ is that $d_2^{1,1}:\mu_2(S')\to\mu_2(S')$ is an isomorphism, both in the Leray–Serre spectral sequence for $\mathrm{H}^*(\mathrm{B}C_{4,S'},\mathbb{G}_m)$ and for $\mathrm{H}^*(\mathcal{M}_{S'},\mathbb{G}_m)$. It follows that this differential is already an isomorphism in the Leray–Serre spectral sequence for $\mathrm{H}^*(\mathcal{M}_{\mathbb{Z}[\,unfrac12]},\mathbb{G}_m)$ by naturality. This in turn implies by naturality that $d_2^{1,1}|_{\mu_2(S)}:\mu_2(S)\to\mu_2(S)$ in the Leray–Serre spectral sequence for $\mathrm{H}^*(\mathcal{M}_S,\mathbb{G}_m)$ is an isomorphism for any regular noetherian $\mathbb{Z}\big[\frac{1}{2}\big]$-scheme S. As the target of $d_3^{0,2}$ is already zero on E_3, the differential $d_3^{0,2}$ must vanish. □

Finally, we prove the theorem from the beginning of the section.

Proof of Theorem 9.1. The claim follows from the determination of the differentials in the range pictured in Figure 3. □

We want to be more specific about the Brauer group classes coming from $\mathbb{G}_m(S)/2$. Recall the section $\Delta\in\mathrm{H}^0(\mathcal{M},\lambda^{\otimes 12})$ from Section 7. As in Construction 2.8, we can define the C_2-torsor $\mathcal{M}(\sqrt{\Delta})=\mathbf{Spec}_{\mathcal{M}_{\mathbb{Z}[1/2]}}\big(\bigoplus_{i\in\mathbb{Z}}\lambda^{\otimes 6i}/(\Delta-1)\big)\to\mathcal{M}_{\mathbb{Z}[1/2]}$ that adjoins a square root of Δ to $\mathcal{M}_{\mathbb{Z}[1/2]}$. For a unit $u\in\mathbb{G}_m\big(\mathbb{Z}\big[\frac{1}{2}\big]\big)$, we denote by (Δ,u) the symbol (quaternion) algebra associated with this torsor.

Proposition 9.6. *Let S denote a connected regular noetherian scheme over* $\operatorname{Spec}\mathbb{Z}\big[\frac{1}{2}\big]$. *Then the map*

$$\mathbb{G}_m(S)/2\to\mathrm{Br}'(\mathcal{M}_S)$$

from the Leray–Serre spectral sequence sends u to $[(u,\Delta)_2]$.

Proof. We consider the Leray–Serre spectral sequence

$$\mathrm{H}^p(C_3, \mathrm{H}^q(\mathcal{M}(2)_S, \mathbb{G}_m))_{(2)} \Rightarrow \mathrm{H}^{p+q}(\mathcal{M}(2)_S/C_3, \mathbb{G}_m)_{(2)},$$

where $\mathcal{M}(2)/C_3$ denotes the stack quotient by the subgroup $C_3 \subset S_3$. Its E_2-term is clearly concentrated in the column $p=0$. From Lemma 5.7 and Proposition 5.14, it is easy to see that $\mathrm{H}^0(C_3, \mathrm{H}^q(\mathcal{M}(2), \mathbb{G}_m)_{(2)} \cong \mathrm{H}^0(S_3, \mathrm{H}^q(\mathcal{M}(2), \mathbb{G}_m)_{(2)})$. We can now consider the further Leray–Serre spectral sequence

$$\mathrm{H}^p(C_2, \mathrm{H}^q(\mathcal{M}(2)_S/C_3, \mathbb{G}_m))_{(2)} \Rightarrow \mathrm{H}^{p+q}(\mathcal{M}_S, \mathbb{G}_m)_{(2)},$$

and we see that it has the same E_2-term as the Leray–Serre spectral sequence for the S_3-cover $\mathcal{M}(2)_S \to \mathcal{M}_S$ in the range depicted in Figure 3.

We claim that the C_2-torsor $\mathcal{M}(2)_S/C_3 \to \mathcal{M}_S$ agrees with $\mathcal{M}_S(\sqrt{\Delta}) \to \mathcal{M}_S$. For this it suffices to show that Δ becomes a square on $\mathcal{M}(2)_S/C_3$. With $p, q \in \mathrm{H}^0(\mathcal{M}(2), \lambda^{\otimes 2})$ as in the discussion after Corollary 4.6, we have $\Delta = 16p^2q^2(p-q)^2$. The C_3-action permutes p, $(-q)$ and $(q-p)$ cyclically so that $4pq(p-q)$ is a C_3-invariant section of $\lambda^{\otimes 6}$ whose square is indeed Δ.

Now the statement follows from Corollary 3.6. □

The following is one of our main results. We recall the convention that everything is implicitly 2-local so that $\mathrm{Br}(\mathcal{M}_R)$ for a ring R denotes really $\mathrm{Br}(\mathcal{M}_R)_{(2)}$.

Proposition 9.7. *Let P be a set of prime numbers including 2 and denote by $\mathbb{Z}_P \subset \mathbb{Q}$ the subset of all fractions where the denominator is only divisible by primes in P. Then*

$$\mathrm{Br}(\mathcal{M}_{\mathbb{Z}_P}) \cong \mathrm{Br}(\mathbb{Z}_P) \oplus \bigoplus_{\substack{p \in P \cup \{-1\},\\ p \equiv 3 \bmod 4}} \mathbb{Z}/2 \oplus \bigoplus_{\substack{p \in P,\\ p \not\equiv 3 \bmod 4}} \mathbb{Z}/4.$$

Proof. First we have to compute the subgroup

$$G \subset \mathbb{G}_m(\mathbb{Z}_P)/2 \cong \bigoplus_{P \cup \{-1\}} \mathbb{F}_2.$$

By Proposition 2.17, a quaternion algebra (a, b) ramifies at p if and only if the Hilbert symbol $\left(\frac{a,b}{p}\right)$ equals -1. By [Neukirch 1999, Theorem V.3.6], $\left(\frac{-1,-1}{p}\right) = -1$ if and only if $p = 2, \infty$ and $\left(\frac{-1,q}{p}\right) = -1$ if and only if $q \equiv 3 \bmod 4$ and $p = 2, q$ (for q a prime number). We see that G has an $\mathbb{F}_2$-basis given by the primes not congruent to 3 mod 4.

We obtain a diagram:

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{G}_m(\mathbb{Z}_P)/2 \arrow[r] \arrow[d] & \overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P}) \arrow[r] \arrow[d] & G \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathbb{G}_m(\mathbb{Z}_P[i])/2 \arrow[r] & \overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P[i]}) \arrow[r] & \mathbb{G}_m(\mathbb{Z}_P[i])/2 \arrow[r] & 0
\end{tikzcd}$$

By Proposition 9.3, $\overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}[1/2,i]}) \cong \mathbb{G}_m(\mathbb{Z}_P[i])/4$. As the map $G \to \mathbb{G}_m(\mathbb{Z}_P[i])/2$ is injective, we see that none of the nonzero lifts of elements of G to $\overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P})$ are 2-torsion. The proposition follows. □

This shows the 2-local part of the computation of $\mathrm{Br}(\mathcal{M}_{\mathbb{Q}})$ and $\mathrm{Br}(\mathcal{M}_{\mathbb{Z}[1/2]})$ in Theorem 1.1, while the 3-local part was already contained in Theorems 7.1 and 8.4 and the p-local part for $p > 3$ in Theorem 6.1.

Remark 9.8. We can describe all the Brauer classes in $\overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P})$ explicitly when P is again a set of prime numbers including 2. We already saw in the last two propositions that $\overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P})[2]$ has an $\mathbb{F}_2$-basis given by $[(p, \Delta)_2]$, where $p \in P \cup \{-1\}$. When $p \in S$ and either $p = 2$ or $p \equiv 1 \bmod 4$, we will give explicit elements of order 4 in the Brauer group of $\mathcal{M}_{\mathbb{Z}_P}$ generating the $\mathbb{Z}/4$-subgroups of the proposition.

To describe the 4-torsion we start with a small observation. Given a cyclic algebra (χ, υ), where χ is a C_4-torsor and υ a μ_4-torsor, $2[(\chi, \upsilon)]$ is represented by (χ', υ'), where χ' and υ' are obtained from χ and υ via the morphisms $C_4 \to C_2$ and $\mu_4 \to \mu_2$. Concretely, this means that χ' are the C_2-fixed points of χ and that if υ is given by adjoining the 4-th root of a section u of $\mathcal{L}^{\otimes 4}$, then υ' is given by adjoining a square root of u, a section of $(\mathcal{L}^{\otimes 2})^{\otimes 2}$.

For primes $p \equiv 1 \bmod 4$, we construct a C_4-Galois extension L of $\mathbb{Q}$ whose C_2-fixed points are $\mathbb{Q}(\sqrt{p})$. As $\sqrt{p}$ is the Gauss sum $\sum_{a=1}^{p-1} \left(\frac{a}{p}\right) \zeta_p^a$, we see that $\mathbb{Q}(\sqrt{p}) \subset \mathbb{Q}(\zeta_p)$. The Galois group of the $\mathbb{Q}$-extension $\mathbb{Q}(\zeta_p)$ is cyclic of order $p-1$, which is divisible by 4. Thus, it has a unique cyclic subextension L of degree 4 whose C_2-fixed points are $\mathbb{Q}(\sqrt{p})$. Note that L is only ramified at p. Explicitly, L is generated by the Gauss sum $\sum_{a=1}^{p-1} \varphi(a)\zeta_p^a$, where $\varphi\colon (\mathbb{Z}/p)^\times \to \mu_4(\mathbb{C})$ is a surjective character.

For $p = 2$, we take $L = \mathbb{Q}(\zeta_{16} + \bar{\zeta}_{16})$ instead, which is the unique C_4-Galois subextension of $\mathbb{Q}(\zeta_{16})$ over $\mathbb{Q}$. If we denote for $p = 2$ or $p \equiv 1 \bmod 4$ the character of $L/\mathbb{Q}$ by χ, these define C_4-Galois covers of $\mathcal{M}_S$ by pullback and we abuse notation and write χ also for these covers. The cup product $[(\chi, \Delta)_4]$ in $\mathrm{Br}'(\mathcal{M}_{\mathbb{Z}_P})$ is a class such that $2[(\chi, \Delta)_4] = [(p, \Delta)_2]$ and thus has exact order 4. It follows that the classes $[(\chi, \Delta)_4]$ give a basis of the 4-torsion of $\overline{\mathrm{Br}}(\mathcal{M}_{\mathbb{Z}_P})$.

10. The Brauer group of $\mathcal{M}$

In this section, we will complete the computation of $\mathrm{Br}(\mathcal{M})$. In the last section, we saw that

$$\mathrm{Br}\bigl(\mathcal{M}\bigl[\tfrac{1}{2}\bigr]\bigr) \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/4$$

with generators $\alpha = [(-1, -1)_2]$, $\beta = [(-1, \Delta)_2]$ and $\frac{1}{2}\gamma$ for $\gamma = [(2, \Delta)_2]$. We will study the ramification of these classes for certain elliptic curves and the reader can find more information about these curves in *The L-functions and modular forms database* [LMFDB 2013].

As above, we will write $\binom{a,b}{2} \in \mu_2(\mathbb{Q}_2)$ for the Hilbert symbol in $\mathbb{Q}_2$ at the prime 2. Hence, $\binom{a,b}{2} = \pm 1$. Recall from Proposition 2.17 that if $\chi \in \mathrm{H}^1(\mathbb{Q}_2, C_2) \cong \mathrm{H}^1(\mathbb{Q}_2, \mu_2)$ corresponds to a unit $v \in \mathbb{G}_m(\mathbb{Q}_2)/2$, then the degree 2 cyclic algebra (χ, u) has class $\binom{u,v}{2}$ in $\mathrm{Br}(\mathbb{Q}_2)[2] \cong \mathbb{Z}/2 \cong \mu_2(\mathbb{Q}_2)$. Since $\mathrm{Br}(\mathbb{Z}_2) = 0$, the Hilbert symbol measures the ramification along (2) in $\operatorname{Spec} \mathbb{Z}_2$.

Proposition 10.1. *Every nonzero linear combination of α, β, $\frac{1}{2}\gamma$ is ramified along* (2), *so these linear combinations are not in the image of* $\mathrm{Br}(\mathcal{M}) \to \mathrm{Br}\bigl(\mathcal{M}\bigl[\frac{1}{2}\bigr]\bigr)$.

Proof. It suffices to prove this for all seven nonzero linear combinations of α, β, γ. Indeed, if all these linear combinations are ramified, then any linear combination $r\alpha + s\beta + \frac{1}{2}\gamma$ is ramified as well. As explained in the introduction around the diagram (1.4), it suffices to construct for each nonzero linear combination ρ an elliptic curve $\operatorname{Spec}\mathbb{Z}_2 \to \mathcal{M}$ such that the pullback of ρ to $\operatorname{Spec}\mathbb{Q}_2$ is nonzero in $\operatorname{Br}(\mathbb{Q}_2)$. Indeed, $\operatorname{Br}(\mathbb{Z}_2) = 0$.

Let

$$E_1 : y^2 + y = x^3 - x^2,$$

the elliptic curve with Cremona label `11a3`. This curve has discriminant -11, which is a unit, so we get an elliptic curve over $\mathbb{Z}_2$, and two associated Brauer classes, $(2, -11)$ and $(-1, -11)$ over $\mathbb{Q}_2$. We can ask what the ramification is. The Hilbert symbol in this case is computed as follows [Serre 1973, Chapter III]. Given $a = 2^\alpha u$ and $b = 2^\beta v \in \mathbb{G}_m(\mathbb{Q}_2)$, where $u, v \in \mathbb{G}_m(\mathbb{Z}_2)$, we have

$$\left(\frac{a, b}{2}\right) = (-1)^{\epsilon(u)\epsilon(v)+\alpha\omega(v)+\beta\omega(u)},$$

where $\epsilon(u) \equiv (u-1)/2$ and $\omega(u) \equiv (u^2-1)/8$.

Hence,

$$\left(\frac{2, -11}{2}\right) = (-1)^{\omega(-11)} = (-1)^{15} = -1.$$

Hence, $(\Delta, 2)$ is ramified at 2. Similarly,

$$\left(\frac{-1, -11}{2}\right) = (-1)^{\epsilon(-1)\epsilon(-11)} = (-1)^{6} = 1.$$

The curve

$$E_2 : y^2 + xy = x^3 - 2x^2 + x$$

has Cremona label `15a8` and discriminant -15. This time, the Hilbert symbols are

$$\left(\frac{2, -15}{2}\right) = 1 \quad \text{and} \quad \left(\frac{-1, -15}{2}\right) = 1.$$

The curve

$$E_3 : y^2 + xy + y = x^3 - x^2$$

has Cremona label `53a1` and discriminant -53. In this case, the Hilbert symbols are

$$\left(\frac{2, -53}{2}\right) = -1 \quad \text{and} \quad \left(\frac{-1, -53}{2}\right) = -1.$$

Now, let $\rho = \alpha + k\beta + m\gamma$ with k and m integers. The elliptic curve E_2 gives a point $x : \operatorname{Spec}\mathbb{Q}_2 \to \mathcal{M}$ where $v_2(x^*\rho) = -1$. It follows that all classes of the form ρ are ramified along (2). Now, E_3 proves that β and γ are ramified along (2). Finally, E_1 proves that $\beta + \gamma$ is ramified along (2). □

Theorem 10.2. $\operatorname{Br}(\mathcal{M}) = 0$.

Proof. This follows from Corollary 6.4, Theorem 8.4, and Propositions 9.7 and 10.1. □

11. The Brauer group of $\mathcal{M}$ over $\mathbb{F}_q$ with q odd

As another application of our methods, we compute the Brauer group of $\mathcal{M}_{\mathbb{F}_q}$ when $q = p^n$ and p is an odd prime. There is remarkable regularity in this case, which is possibly surprising based on what happens for number fields.

Theorem 11.1. *Let $q = p^n$ where p is an odd prime. Then,* $\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q}) \cong \mathbb{Z}/12$.

Proof. Recall that $\mathrm{Br}(\mathbb{F}_q) = 0$. Thus, by Theorem 9.1, there is an extension $0 \to \mathbb{F}_q^\times/2 \to {}_2\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q}) \to \mathbb{F}_q^\times/2 \to 0$. Since q is odd, $\mathbb{F}_q^\times/2 \cong \mathbb{Z}/2$. The remainder of Section 9, especially Remark 9.8, implies that the extension is nonsplit so that in fact ${}_2\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q}) \cong \mathbb{Z}/4$.

Now, let $\ell \geq 3$ be a prime, which we *do not* assume is different from p. The possible terms contributing to ${}_\ell\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})$ in the Leray–Serre spectral sequence for $\mathcal{M}(2)_{\mathbb{F}_q} \to \mathcal{M}_{\mathbb{F}_q}$ in $\mathbb{G}_m$ cohomology, besides ${}_\ell\mathrm{Br}(X_{\mathbb{F}_q})^{S_3}$, are $\mathrm{H}^1(S_3, {}_\ell\mathrm{Pic}(\mathcal{M}(2)_{\mathbb{F}_q})$ and $\mathrm{H}^2(S_3, {}_\ell\mathbb{G}_m(\mathcal{M}(2)_{\mathbb{F}_q})) \cong \mathrm{H}^2(S_3, {}_\ell\mathbb{G}_m(X_{\mathbb{F}_q}))$. The first of these is zero since ${}_\ell\mathrm{Pic}(\mathcal{M}(2)_{\mathbb{F}_q}) \cong {}_\ell\mathrm{Pic}(X_{\mathbb{F}_q}) = 0$ for ℓ odd. The second has no odd primary torsion. This follows from the exact sequence $0 \to \mathbb{G}_m(\mathbb{F}_q) \to \mathbb{G}_m(X_{\mathbb{F}_q}) \to \tilde{\rho} \otimes \mathbb{Z} \to 0$ together with Lemmas 5.2 and 5.3.

Thus, we see that ${}_\ell(\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})) \cong {}_\ell\mathrm{Br}(X_{\mathbb{F}_q})^{S_3}$ for ℓ odd. So, it suffices to compute the Brauer group of $X_{\mathbb{F}_q}$ as an S_3-module. In general, our argument in the rest of the paper relies fundamentally on Lemma 5.11, which requires ℓ to be invertible to analyze the ramification map as in Proposition 2.14. However, in this case, $X_{\mathbb{F}_q}$ is a curve over a finite field, so the ramification theory simplifies drastically. Consider the commutative diagram of exact ramification sequences due to [Grothendieck 1968c, Proposition 2.1]:

$$\begin{tikzcd}
0 \arrow[r] & \mathrm{Br}(\mathbb{P}^1_{\mathbb{F}_q}) \arrow[r] \arrow[d] & \mathrm{Br}(\eta) \arrow[r] \arrow[d, "="] & \bigoplus_{x \in (\mathbb{P}^1_{\mathbb{F}_q})^{(1)}} \mathbb{Q}/\mathbb{Z} \arrow[r] \arrow[d] & \mathbb{Q}/\mathbb{Z} \arrow[r] & 0 \\
0 \arrow[r] & \mathrm{Br}(X_{\mathbb{F}_q}) \arrow[r] & \mathrm{Br}(\eta) \arrow[r] & \bigoplus_{x \in (X_{\mathbb{F}_q})^{(1)}} \mathbb{Q}/\mathbb{Z} & &
\end{tikzcd}$$

where η is the generic point of $X_{\mathbb{F}_q}$. Note that the exactness at the right in the top sequence is due to the fact (see [Gille and Szamuely 2006, Corollary 6.5.4]) that $\bigoplus_{x \in (\mathbb{P}^1_{\mathbb{F}_q})^{(1)}} \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$ is given by summation in $\mathbb{Q}/\mathbb{Z}$. Since $\mathrm{Br}(\mathbb{P}^1_{\mathbb{F}_q}) = 0$ by [Grothendieck 1968c, Remarques 2.5.b], we see that $\mathrm{Br}(X_{\mathbb{F}_q}) \subseteq \mathrm{Br}(\eta)$ is the subgroup consisting of classes ramified only at $0, 1, \infty$. Using the top row of the diagram, it follows that $\mathrm{Br}(X_{\mathbb{F}_q})$ fits into an exact sequence

$$0 \to \mathrm{Br}(X_{\mathbb{F}_q}) \to \bigoplus_{0,1,\infty} \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z} \to 0,$$

from which it follows that

$$\mathrm{Br}(X_{\mathbb{F}_q}) \cong \tilde{\rho} \otimes \mathbb{Q}/\mathbb{Z}.$$

By Lemma 5.3, we find that ${}_\ell\mathrm{Br}(X_{\mathbb{F}_q})^{S_3} = 0$ if $\ell \geq 5$ and ${}_3\mathrm{Br}(X_{\mathbb{F}_q})^{S_3} \cong \mathbb{Z}/3$. This proves the theorem. $\square$

Remark 11.2. We can again be more specific about the Azumaya algebras representing the classes in $\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})$ with q odd. Let T be a $\mathbb{F}_q$-scheme (or stack) with an elliptic curve E of discriminant Δ. Let χ_m be the pullback of a character of the Galois extension $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$ to T. We claim that there is a generator $a \in \mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})$ such that the pullback of a to $\mathrm{Br}(T)$ agrees with $[(\chi_{12}, \Delta)]$. Informally, $a = [(\chi_{12}, \Delta)]$ in the universal case $T = \mathcal{M}_{\mathbb{F}_q}$, where more precisely we should replace here Δ by the μ_{12}-torsor corresponding to $\Delta \in \Gamma(\lambda^{\otimes 12})$ via Construction 2.8.

First we consider the 3-torsion. The proof of Lemma 7.2 applies here to show that $[(\chi_3, \Delta)]$ is indeed the pullback of a generator of $\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})[3]$. Moreover, Proposition 9.6 implies that the unique 2-torsion element $6a \in \mathrm{Br}(\mathcal{M}_{\mathbb{F}_q})$ pulls back to $[(\chi_2, \Delta)]$ as $\mathbb{F}_{q^2}$ agrees with $\mathrm{F}_q[\sqrt{x}]$ for an arbitrary nonsquare x in $\mathbb{F}_q$. As in Remark 9.8 we see that $6[(\chi_{12}, \Delta)] = [(\chi_2, \Delta)] \neq 0$ and $4[(\chi_{12}, \Delta)] = [(\chi_3, \Delta)] \neq 0$. Thus, $[(\chi_{12}, \Delta)]$ is indeed a generator of $\mathrm{Br}(\mathcal{M}_{\mathbb{F}_q}) \cong \mathbb{Z}/12$.

Finally, we also treat the easier case of an algebraically closed base.

Proposition 11.3. *Let k be an algebraically closed field of characteristic not* 2. *Then* $\mathrm{Br}(\mathcal{M}_k) = 0$.

Proof. By Theorem 9.1, ${}_2\,\mathrm{Br}(\mathcal{M}_k) = 0$. As in the last proof we see that ${}_\ell(\mathrm{Br}(\mathcal{M}_k)) \cong {}_\ell\,\mathrm{Br}(X_k)^{S_3}$ for ℓ an odd prime. By Tsen's theorem, $\mathrm{Br}(\eta)$ vanishes for η the generic point of X_k. By [Grothendieck 1968c, Proposition 2.1] we obtain that $\mathrm{Br}(X_k) = 0$. □

Remark 11.4. In an earlier version of this paper we suggested using the $\mathrm{GL}_2(\mathbb{Z}/3)$-cover $\mathcal{M}(3) \to \mathcal{M}_{\mathbb{Z}[1/3]}$ to determine $\mathrm{Br}(\mathcal{M}_k)$ also for algebraically closed fields k of characteristic 2. Combining this approach with a new idea, Minseon Shin [2019] has proved in the meanwhile that $\mathrm{Br}(\mathcal{M}_k) \cong \mathbb{Z}/2$ for such fields k.

Acknowledgments

We thank the Hausdorff Research Institute for Mathematics, UIC, and Universität Bonn for hosting one or the other author in 2015 and 2016 while this paper was being written. Furthermore, we thank Asher Auel, Akhil Mathew, and Vesna Stojanoska for their ideas at the beginning of this project and Peter Scholze and Yichao Tian for the suggestion to use fppf-cohomology and the idea for the splitting in Proposition 2.9. Finally, we thank the anonymous referees for their helpful suggestions.

References

[Auel et al. 2014] A. Auel, M. Bernardara, and M. Bolognesi, "Fibrations in complete intersections of quadrics, Clifford algebras, derived categories, and rationality problems", *J. Math. Pures Appl.* (9) **102**:1 (2014), 249–291. MR Zbl

[Česnavičius 2019] K. Česnavičius, "Purity for the Brauer group", *Duke Math. J.* **168**:8 (2019), 1461–1486. MR Zbl

[Chan and Ingalls 2005] D. Chan and C. Ingalls, "The minimal model program for orders over surfaces", *Invent. Math.* **161**:2 (2005), 427–452. MR Zbl

[Cornell et al. 1997] G. Cornell, J. H. Silverman, and G. Stevens (editors), *Modular forms and Fermat's last theorem* (Boston, 1995), Springer, 1997. MR Zbl

[Deligne and Rapoport 1973] P. Deligne and M. Rapoport, "Les schémas de modules de courbes elliptiques", pp. 143–316 in *Modular functions of one variable, II* (Antwerp, 1972), edited by P. Deligne and W. Kuyk, Lecture Notes in Math. **349**, Springer, 1973. MR Zbl

[Deninger 1988] C. Deninger, "A proper base change theorem for nontorsion sheaves in étale cohomology", *J. Pure Appl. Algebra* **50**:3 (1988), 231–235. MR Zbl

[Edixhoven et al. 1990] B. Edixhoven, A. de Groot, and J. Top, "Elliptic curves over the rationals with bad reduction at only one prime", *Math. Comp.* **54**:189 (1990), 413–419. MR Zbl

[EGA IV$_2$ 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR Zbl

[Fuchs 1970] L. Fuchs, *Infinite abelian groups, I*, Pure Appl. Math. **36**, Academic Press, New York, 1970. MR Zbl

[Fujiwara 2002] K. Fujiwara, "A proof of the absolute purity conjecture (after Gabber)", pp. 153–183 in *Algebraic geometry* (Azumino, Japan, 2000), edited by S. Usui et al., Adv. Stud. Pure Math. **36**, Math. Soc. Japan, Tokyo, 2002. MR Zbl

[Fulton and Olsson 2010] W. Fulton and M. Olsson, "The Picard group of $\mathcal{M}_{1,1}$", *Algebra Number Theory* **4**:1 (2010), 87–104. MR Zbl

[Gabber 1981] O. Gabber, "Some theorems on Azumaya algebras", pp. 129–209 in *Groupe de Brauer* (Plans-sur-Bex, Switzerland, 1980), edited by M. Kervaire and M. Ojanguren, Lecture Notes in Math. **844**, Springer, 1981. MR Zbl

[Gabber and Kelly 2015] O. Gabber and S. Kelly, "Points in algebraic geometry", *J. Pure Appl. Algebra* **219**:10 (2015), 4667–4680. MR Zbl

[Gille and Quéguiner-Mathieu 2011] P. Gille and A. Quéguiner-Mathieu, "Formules pour l'invariant de Rost", *Algebra Number Theory* **5**:1 (2011), 1–35. MR Zbl

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Stud. Adv. Math. **101**, Cambridge Univ. Press, 2006. MR Zbl

[Grothendieck 1968a] A. Grothendieck, "Le groupe de Brauer, I: Algèbres d'Azumaya et interprétations diverses", pp. 46–66 in *Dix exposés sur la cohomologie des schémas*, Adv. Stud. Pure Math. **3**, North-Holland, Amsterdam, 1968. MR Zbl

[Grothendieck 1968b] A. Grothendieck, "Le groupe de Brauer, II: Théorie cohomologique", pp. 67–87 in *Dix exposés sur la cohomologie des schémas*, Adv. Stud. Pure Math. **3**, North-Holland, Amsterdam, 1968. MR Zbl

[Grothendieck 1968c] A. Grothendieck, "Le groupe de Brauer, III: Exemples et compléments", pp. 88–188 in *Dix exposés sur la cohomologie des schémas*, Adv. Stud. Pure Math. **3**, North-Holland, Amsterdam, 1968. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Grad. Texts in Math. **52**, Springer, 1977. MR Zbl

[Hartshorne 1980] R. Hartshorne, "Stable reflexive sheaves", *Math. Ann.* **254**:2 (1980), 121–176. MR Zbl

[Hasse 1948] H. Hasse, "Arithmetische Bestimmung von Grundeinheit und Klassenzahl in zyklischen kubischen und biquadratischen Zahlkörpern", *Abh. Deutsch. Akad. Wiss. Berlin. Math.-Nat. Kl.* **1948**:2 (1948), 95 pp. MR Zbl

[de Jong 2005] A. J. de Jong, "A result of Gabber", preprint, 2005, Available at http://www.math.columbia.edu/~dejong/.

[Katz and Mazur 1985] N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Ann. of Math. Stud. **108**, Princeton Univ. Press, 1985. MR Zbl

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik (3) **39**, Springer, 2000. MR Zbl

[Lemmermeyer 2005] F. Lemmermeyer, "Kronecker–Weber via Stickelberger", *J. Théor. Nombres Bordeaux* **17**:2 (2005), 555–558. MR Zbl

[Lieblich 2008] M. Lieblich, "Twisted sheaves and the period-index problem", *Compos. Math.* **144**:1 (2008), 1–31. MR Zbl

[Lieblich 2011] M. Lieblich, "Period and index in the Brauer group of an arithmetic surface", *J. Reine Angew. Math.* **659** (2011), 1–41. MR Zbl

[LMFDB 2013] LMFDB Collaboration, "The L-functions and modular forms database", electronic reference, 2013, Available at http://www.lmfdb.org.

[Mathew and Stojanoska 2016] A. Mathew and V. Stojanoska, "The Picard group of topological modular forms via descent theory", *Geom. Topol.* **20**:6 (2016), 3133–3217. MR Zbl

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Math. Series **33**, Princeton Univ. Press, 1980. MR Zbl

[Mumford 1965] D. Mumford, "Picard groups of moduli problems", pp. 33–81 in *Arithmetical algebraic geometry* (West Lafayette, IN, 1963), edited by O. F. G. Schilling, Harper & Row, New York, 1965. MR Zbl

[Neukirch 1999] J. Neukirch, *Algebraic number theory*, Grundlehren der Math. Wissenschaften **322**, Springer, 1999. MR Zbl

[Neukirch et al. 2000] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, Grundlehren der Math. Wissenschaften **323**, Springer, 2000. MR Zbl

[Parimala and Srinivas 1992] R. Parimala and V. Srinivas, "Analogues of the Brauer group for algebras with involution", *Duke Math. J.* **66**:2 (1992), 207–237. MR Zbl

[Romagny 2005] M. Romagny, "Group actions on stacks and applications", *Michigan Math. J.* **53**:1 (2005), 209–236. MR Zbl

[Saltman 1999] D. J. Saltman, *Lectures on division algebras*, CBMS Region. Conf. Ser. Math. **94**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Serre 1973] J.-P. Serre, *A course in arithmetic*, Grad. Texts in Math. **7**, Springer, 1973. MR Zbl

[Serre 1997] J.-P. Serre, *Galois cohomology*, Springer, 1997. MR Zbl

[SGA 4_3 1973] M. Artin, A. Grothendieck, and J. L. Verdier, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (Séminaire de Géométrie Algébrique du Bois Marie 1963–1964), Lecture Notes in Math. **305**, Springer, 1973. MR Zbl

[Shin 2019] M. Shin, "The Brauer group of the moduli stack of elliptic curves over algebraically closed fields of characteristic 2", *J. Pure Appl. Algebra* **223**:5 (2019), 1966–1999. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Grad. Texts in Math. **106**, Springer, 2009. MR Zbl

[Stacks 2018] T. Stacks Project Authors, "*The Stacks project*", 2018, Available at http://stacks.math.columbia.edu.

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. **38**, Cambridge Univ. Press, 1994. MR Zbl

Communicated by Martin Olsson
Received 2017-11-10 Revised 2020-03-26 Accepted 2020-05-04

antieau@northwestern.edu *Department of Mathematics, Northwestern University, Chicago, IL, United States*

f.l.m.meier@uu.nl *Mathematical Institut, Universiteit Utrecht, Utrecht, Netherlands*

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2335

Modular forms from Noether–Lefschetz theory

François Greer

We enumerate smooth rational curves on very general Weierstrass fibrations over hypersurfaces in projective space. The generating functions for these numbers lie in the ring of classical modular forms. The method of proof uses topological intersection products on a period stack and the cohomological theta correspondence of Kudla and Millson for special cycles on a locally symmetric space of orthogonal type. The results here apply only in base degree 1, but heuristics for higher base degree match predictions from the topological string partition function.

1. Introduction

Locally symmetric spaces of noncompact type are special Riemannian manifolds which serve as classifying spaces for (torsion-free) arithmetic groups. As such, their geometry has been studied intensely from several different perspectives. By a well known theorem of Baily and Borel, if such a manifold admits a parallel complex structure, then it is a complex quasiprojective variety. In this paper, we study more general indefinite orthogonal groups, which act on Hodge structures of even weight, and draw some conclusions about holomorphic curve counts.

Let Λ be an integral lattice inside $\mathbb{R}^{2,l}$, and Γ a congruence subgroup of $O(\Lambda)$. The Baily–Borel theorem [1966] implies that the double quotient $\Gamma\backslash O(2,l)/O(2)\times O(l)$ is a quasiprojective variety. These Hermitian symmetric examples have played a central role in classical moduli theory. For instance, moduli spaces of polarized K3 surfaces, cubic fourfolds, and holomorphic symplectic varieties are all contained within these Baily–Borel varieties as Zariski open subsets. Automorphic forms provide natural compactifications for these moduli spaces and bounds on their cohomology.

One can interpret $O(2,l)/O(2)\times O(l)$ as the set of 2-planes in $\mathbb{R}^{2,l}$ on which the pairing is positive definite. The presence of the integral lattice Λ allows us to define a sequence of $\mathbb{R}$-codimension 2 submanifolds, indexed by $n\in\mathbb{Q}_{>0}$ and $\alpha\in\Lambda^\vee/\Lambda$, where $\Lambda^\vee$ is the dual lattice of covectors taking integral values on Λ.

$$C_{n,\alpha} := \Gamma\backslash\Bigg(\bigcup_{\substack{v\in\Lambda^\vee,\ v+\Lambda=\alpha\\ (v,v)=-n}} v^\perp\Bigg)\subset \Gamma\backslash O(2,l)/O(2)\times O(l).$$

Borel [1969] showed that there are finitely many Γ-orbits of lattice vectors with fixed norm, so the above union is finite in the quotient space. Each $C_{n,\alpha}$ is isomorphic to a locally symmetric space for $O(2,l-1)$,

MSC2010: 14NXX.
Keywords: modular forms, elliptic fibrations, Calabi–Yau threefolds, rational curves.

so it is actually an algebraic subvariety of $\mathbb{C}$-codimension 1 called a *Heegner divisor*. The classes of these divisors in the Picard group satisfy nontrivial relations from the Howe theta correspondence between orthogonal and symplectic groups:

Theorem 1 [Borcherds 1999]. *The formal q-series with coefficients in*

$$\mathrm{Pic}_{\mathbb{Q}}(\Gamma\backslash O(2,l)/O(2)\times O(l))\otimes\mathbb{Q}[\Lambda^\vee/\Lambda]$$

given by

$$e(V^\vee)e_0+\sum_{n,\alpha}[C_{2n,\alpha}]e_\alpha q^n$$

transforms like a $\mathbb{Q}[\Lambda^\vee/\Lambda]$-valued modular form with respect to the Weil representation of the metaplectic group $\mathrm{Mp}_2(\mathbb{Z})$. Here $\{e_\alpha\}$ denotes the standard basis for $\mathbb{Q}[\Lambda^\vee/\Lambda]$, and $e(V^\vee)$ is the Euler class of the (dual) tautological bundle of positive definite 2-planes.

In other words, the q-series above lies in the finite dimensional subspace

$$\mathrm{Mod}\left(1+\frac{l}{2},\mathrm{Mp}_2(\mathbb{Z}),\mathbb{Q}[\Lambda^\vee/\Lambda]\right)\otimes\mathrm{Pic}_{\mathbb{Q}}(\Gamma\backslash O(2,l)/O(2)\times O(l)).$$

Theorem 1 has been used to describe the Picard group of moduli spaces [Greer et al. 2015], and also has applications to enumerative geometry, initiated by [Maulik and Pandharipande 2013]. In this paper, we move beyond the Hermitian symmetric space to more general symmetric spaces of orthogonal type. These no longer have a complex structure, and their arithmetic quotients are no longer algebraic, but they still have a theta correspondence, and thus an analogous modularity statement for special cycles in singular cohomology:

Theorem 2 [Kudla and Millson 1990]. *Assume for convenience of this exposition that $\Lambda\subset\mathbb{R}^{p,l}$ is even and unimodular. Then the formal q-series*

$$e(V^\vee)+\sum_{n\geq 1}[C_{2n}]q^n\in\mathbb{Q}[[q]]\otimes_{\mathbb{Q}}H^p(\Gamma\backslash O(p,l)/O(p)\times O(l),\mathbb{Q})$$

lies in the finite-dimensional subspace of modular forms:

$$\mathrm{Mod}\left(\frac{p+l}{2},\mathrm{SL}_2(\mathbb{Z})\right)\otimes H^p(\Gamma\backslash O(p,l)/O(p)\times O(l),\mathbb{Q}).$$

Here $e(V^\vee)$ is the Euler class of the dual tautological bundle of p-planes. See Remark 35 for a justification of the level group $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 2 will be used to enumerate smooth rational curves on certain elliptically fibered varieties $X\to Y$. We give a general formula which applies to Weierstrass fibrations over hypersurfaces in projective space. The answers are honest counts, not virtual integrals, and are expressed in terms of q-expansions of modular forms.

All period domains D for smooth projective surfaces with positive geometric genus admit smooth proper fibrations

$$D \to O(p,l)/O(p) \times O(l).$$

The Noether–Lefschetz loci in D are the preimages of special subsymmetric spaces of $\mathbb{R}$-codimension p. This provides a valuable link between moduli theory and the cohomology of locally symmetric spaces. We expect the ideas developed in this paper to compute algebraic curve counts on a broad class of varieties. The results are consistent with general conjectures in [Oberdieck and Pixton 2019] for elliptic fibrations.

Let $Y \subset \mathbb{P}^{m+1}$ be a smooth hypersurface of degree d and dimension $m \geq 2$. For an ample line bundle $\mathcal{L} = \mathcal{O}_Y(k)$, a Weierstrass fibration over Y is a hypersurface

$$X \subset \mathbb{P}(\mathcal{L}^{\otimes -2} \oplus \mathcal{L}^{\otimes -3} \oplus \mathcal{O}_Y)$$

cut out by a global Weierstrass equation (see Section 2 for details). For general choice of coefficients, X is smooth of dimension $m+1$, and the morphism $\pi : X \to Y$ is flat with generic fiber of genus one. Since π admits a section $i : Y \to X$, the generic fiber is actually an *elliptic* curve.

The second homology group of X is given by

$$H_2(X, \mathbb{Z}) \simeq H_2(Y) \oplus \mathbb{Z}f = \mathbb{Z}\ell + \mathbb{Z}f,$$

where ℓ is the line class on Y pushed forward via i, and f is the class of a fiber. We begin by posing the naive:

Question 3. How many smooth rational curves are there on X in the homology class $\ell + nf$?

The deformation theory of curves $C \subset X$ allows us to estimate when this question has a finite answer. The expected dimension of the moduli space of curves in X is given by the Hirzebruch–Riemann–Roch formula:

$$h^0(C, N_{C/X}) - h^1(C, N_{C/X}) = \int_C c_1(T_X) + (1-g)(\dim X - 3).$$

The adjunction formula gives $-c_1(T_X) = K_X = \pi^*(K_Y + c_1(\mathcal{L}))$.

Remark 4. Since K_X is pulled back from Y, we have $K_X \cdot f = 0$, so our dimension estimate is independent of n. This feature holds more generally for any morphism π with K-trivial fibers.

We expect a finite answer to Question 3 whenever

$$0 = -K_X \cdot (\ell + nf) + (m-2) \quad \Longleftrightarrow \quad k = 2m - d.$$

Recall that $\mathcal{L} = \mathcal{O}_Y(k)$ was the ample line bundle used to construct the Weierstrass fibration $X \to Y$, so for the rest of the paper, we require that $k = 2m - d > 0$. Note that X is Calabi–Yau if and only if $\dim(X) = 3$. Our main result is:

Theorem 5. *A very general Weierstrass model $X \to Y$ constructed using $\mathcal{L} = \mathcal{O}_Y(k)$ contains finitely many smooth rational curves in the class $\ell + nf$, whose count we denote $r_X(n)$. For $k \leq 4$, the generating series is given by*

$$\sum_{n \geq 1} r_X(n) q^n = \varphi(q) - \Theta(q),$$

where $\varphi(q) \in \mathrm{Mod}(6k-2, \mathrm{SL}_2(\mathbb{Z}))$, and $\Theta(q) \in \mathbb{Q}[\theta_{A_1}, \theta_{A_2}, \theta_{A_3}]_{<k}$, a polynomial of weighted degree $< k$.

Recall that for a lattice A, the associated theta series is given by

$$\theta_A(q) = \sum_{v \in A} q^{(v,v)/2},$$

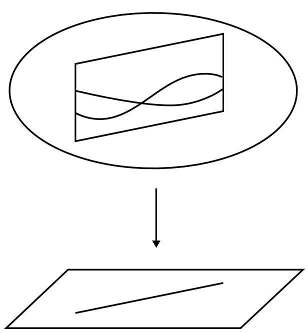
and we assign the weight ρ to the series θ_{A_ρ} for the root lattice A_ρ.

In short, the curve counts $r_X(n)$ are controlled by a finite amount of data, since $\mathrm{Mod}(6k-2, \mathrm{SL}_2(\mathbb{Z}))$ and $\mathbb{Q}[\theta_{A_1}, \theta_{A_2}, \dots]_{<k}$ are finite-dimensional $\mathbb{Q}$-vector spaces. The series can be explicitly computed when $k \leq 3$. We record some numerical examples in Section 7 to illustrate the scope of Theorem 5.

Remark 6. We expect that Theorem 5 can be extended to $k \leq 8$, but the statement is less tidy and involves the root systems D_4, E_6, and E_7. For $k > 8$, the elliptic surfaces involved will have singularities worse than ADE, so more sophisticated techniques are needed.

The argument proceeds roughly as follows. The curves C that we wish to count have the property that $\pi(C) \subset Y$ is a line. In other words, they can be viewed as sections of the elliptic fibration

$$\pi^{-1}(\pi(C)) \to \pi(C) \simeq \mathbb{P}^1.$$



As the line $L \subset Y$ varies, this construction produces a family of elliptic surfaces over the Fano variety of lines in Y:

$$\nu : \mathscr{S} \to F(Y).$$

We wish to count points $[L] \in F(Y)$ such that $\mathscr{S}_{[L]} = \pi^{-1}(L)$ contains a section curve other than the identity section, i.e., a nontrivial Mordell–Weil group. The Shioda–Tate sequence expresses the Mordell–Weil group of an elliptic surface in terms of its Néron–Severi lattice and the sublattice $V(S)$ spanned by vertical classes and the identity section:

$$0 \to V(S) \to \mathrm{NS}(S) \to \mathrm{MW}(S/\mathbb{P}^1) \to 0. \tag{1}$$

The period domain for a given class of elliptic surfaces is related to a locally symmetric space, whence the modular form $\varphi(q)$, which counts surfaces with jumping Picard rank. To obtain the counts $r_X(n)$ we subtract contributions $\Theta(q)$ from surfaces with jumping $V(S)$, which are precisely those with A_ρ singularities. The terms in the difference formula of Theorem 5 are matched to the groups in the short exact sequence (1).

The paper is organized as follows. In Section 2, we review the basic theory of elliptic fibrations and set up the tools for proving transversality of intersections in moduli. In Section 3, we review the theory of period domains and lattices in the cohomology of elliptic surfaces. Section 4 is devoted to the deformation and resolution of A_ρ singularities in families of surfaces, and we introduce the monodromy stack of such a family. Section 5 explains how Noether–Lefschetz intersection numbers on the period stack satisfy a modularity statement from Theorem 2. In Section 6, we use the fact that $k \leq 4$ to classify the singularities which occur in the family ν at various codimensions, and compute their degrees in terms of Schubert intersections. Finally, Section 7 explains how to compute the modular form $\varphi(q)$ when $k \leq 3$, and the general form of the correction term $\Theta(q)$.

2. Elliptic fibrations

We begin by reviewing the Weierstrass equation for elliptic curves in $\mathbb{P}^2$:

$$y^2 z = x^3 + Axz^2 + Bz^3. \tag{2}$$

This cubic curve has a flex point at $[0:1:0]$, which serves as the identity of a group law in the smooth case. The curve is singular if and only if the right-hand side has a multiple root, which occurs when $\Delta = 4A^3 + 27B^2 = 0$. These singular curves are all isomorphic to the nodal cubic, except for when $A = B = 0$, which corresponds to the cuspidal cubic.

To replicate this construction in the relative case, let Y be a smooth projective variety, and $\mathcal{L} \in \operatorname{Pic}(Y)$ an ample line bundle. We form the $\mathbb{P}^2$ bundle

$$\mathbb{P}(\mathcal{L}^{\otimes -2} \oplus \mathcal{L}^{\otimes -3} \oplus \mathcal{O}_Y) \to Y.$$

The same Weierstrass equation (2) makes sense for x, y, z fiber coordinates, and

$$A \in H^0(Y, \mathcal{L}^{\otimes 4}), \quad B \in H^0(Y, \mathcal{L}^{\otimes 6}).$$

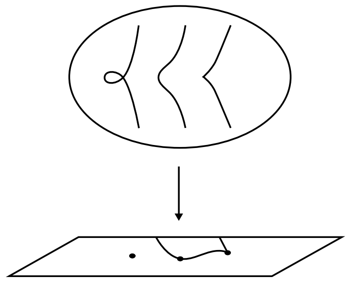
Let $X \subset \mathbb{P}(\mathcal{L}^{\otimes -2} \oplus \mathcal{L}^{\otimes -3} \oplus \mathcal{O})$ be the solution of the global Weierstrass equation and $\pi : X \to Y$ the morphism to the base. The fibers of π are elliptic curves in Weierstrass form, and there is a global section $i : Y \to X$ given in coordinates by $[0:1:0]$, which induces a group law on each smooth fiber. Now,

$$\Delta = 4A^3 + 27B^2 \in H^0(Y, \mathcal{L}^{\otimes 12})$$

cuts out a hypersurface in Y whose generic fiber is a nodal cubic. The singularities of Δ occur along the smooth complete intersection $(A) \cap (B)$, and are analytically locally isomorphic to

$$(\text{cusp}) \times \mathbb{C}^{m-2}.$$

The case of $m = 2$ and $d = 1$ is pictured below, with three fibers drawn over points in $\mathbb{P}^2$ in different singularity strata of Δ.



By the adjunction formula applied to $X \subset \mathbb{P}(\mathcal{L}^{\otimes -2} \oplus \mathcal{L}^{\otimes -3} \oplus \mathcal{O}_Y)$,

$$\begin{aligned} K_X &= (K_{\mathbb{P}(\mathcal{L}^{\otimes -2} \oplus \mathcal{L}^{\otimes -3} \oplus \mathcal{O})} + [X])|_X \\ &= (\pi^* K_Y - 5\pi^* c_1(\mathcal{L}) - 3\zeta) + (3\zeta + 6\pi^* c_1(\mathcal{L})) \\ &= \pi^*(K_Y + c_1(\mathcal{L})), \end{aligned}$$

so the relative dualizing sheaf $\omega_{X/Y}$ is isomorphic to $\pi^*\mathcal{L}$. By the adjunction formula applied to $Y \subset X$ through the section i,

$$K_Y = (K_X + [Y])|_Y = K_Y + c_1(\mathcal{L}) + c_1(N_{Y/X}),$$

so the normal bundle $N_{Y/X}$ is isomorphic to $\mathcal{L}^\vee$.

Definition 7. The parameter space for Weierstrass fibrations over Y is given by the weighted projective space

$$W(Y, \mathcal{L}) := \big(H^0(Y, \mathcal{L}^{\otimes 4}) \oplus H^0(Y, \mathcal{L}^{\otimes 6}) - \{0\}\big)/\mathbb{C}^\times,$$

where $\mathbb{C}^\times$ acts on the direct summands with weight 2 and 3, respectively.

In this paper, X is always a general member of $W(Y, \mathcal{O}(k))$, where $Y \subset \mathbb{P}^{m+1}$ is a smooth hypersurface of degree d, and $k = 2m - d$. Since we are interested in counting curves in X which lie over lines in Y, we will also consider elliptic surfaces $S \in W(\mathbb{P}^1, \mathcal{O}(k))$. For convenience, we gather some properties of these surfaces.

Proposition 8. *For $S \in W(\mathbb{P}^1, \mathcal{O}(k))$ a smooth surface, its Hodge numbers are*

$$h^1(S, \mathcal{O}_S) = 0, \quad h^2(S, \mathcal{O}_S) = k - 1, \quad h^1(S, \Omega_S) = 10k.$$

As a result, $e(S) = 12k$, which is the number of singular fibers, and the canonical line bundle is given by

$$\omega_S \simeq \pi^* \mathcal{O}_{\mathbb{P}^1}(k-2).$$

Proof. These are standard computations using Noether's formula and the Leray spectral sequence for π; see Lecture III of [Miranda 1989]. □

Theorem 9. *Any elliptic surface $S \to \mathbb{P}^1$ is birational to a Weierstrass surface. Furthermore, there is a bijection between (isomorphism classes of) smooth relatively minimal surfaces and Weierstrass fibrations with rational double points.*

Proof. This uses Kodaira's classification of singular fibers; see Lecture II of [Miranda 1989]. □

It will be convenient for us to take a further quotient of $W(\mathbb{P}^1, \mathcal{O}(k))$ to account for changes of coordinates on the base.

Definition 10. The moduli space for Weierstrass surfaces over $\mathbb{P}^1$ is given by the (stack) quotient

$$\mathcal{W}_k := W(\mathbb{P}^1, \mathcal{O}(k))/\,\mathrm{PGL}(2).$$

Remark 11. Miranda [1981] showed that $[S] \in W(\mathbb{P}^1, \mathcal{O}(k))$ is GIT stable with respect to PGL(2) if and only if it has rational double points, so $\mathcal{W}_k$ has a quasiprojective coarse space with good modular properties.

For any variety $Y \subset \mathbb{P}^{m+1}$, the locus of lines contained in Y is called the Fano scheme[1] of Y, and is denoted

$$F(Y) \subset \mathbb{G}(1, m+1).$$

Theorem 12. *For a general hypersurface $Y \subset \mathbb{P}^{m+1}$ of degree d, the Fano scheme is smooth of dimension $2m - d - 1 = k - 1$, for $k > 0$.*

Proof. To study the general behavior, we construct an incidence correspondence

$$\Omega = \{(L, Y) : L \subset Y\} \subset \mathbb{G}(1, m+1) \times \mathbb{P}^N.$$

The first projection $\Omega \to \mathbb{G}(1, m+1)$ is surjective and has linear fibers of dimension $N - d - 1$, so Ω is smooth and irreducible of dimension $N + 2m - d - 1$. The second projection has fiber $F(Y)$ over $[Y] \in \mathbb{P}^N$. To get the desired dimension, it suffices to show that a general hypersurface Y contains a line, so that the second projection $\Omega \to \mathbb{P}^N$ is surjective. This can be done by constructing a smooth hypersurface containing a line whose normal bundle is balanced as in [Eisenbud and Harris 2016]. □

Moreover, if we vary the hypersurface Y, then $F(Y)$ varies freely inside $\mathbb{G}(1, m+1)$. To be precise:

Definition 13. Let $Z \to B$ be a submersion of complex manifolds, and let $f : Z \to P$ be a family of immersions $\{f_b : Z_b \to P\}$. This deformation is called *freely movable* if for any $x_0 \in Z_0$ and any $v \in T_{f(x_0)}P$, there exists a 1-parameter subfamily $T \subset B$ and a section $x(t) \in Z_t$ such that $x(0) = x_0$ and $\frac{d}{dt}\big|_{t=0}(f \circ x) = v$.

The second projection $\Omega \to \mathbb{P}^N$ from Theorem 12 is generically smooth, and the family of embeddings $\Omega \to \mathbb{G}(1, m+1)$ is freely movable because for any line $L \in F(Y_0)$ and tangent vector $v \in T_{[L]}\mathbb{G}$, there is a curve of lines $\{L_t\}$ in the direction v. Since every line lies on a hypersurface, there is a deformation Y_t such that $[L_t] \in F(Y_t)$. This property is useful for proving transversality statements, using:

[1] There is a natural scheme structure on $F(Y)$ coming from its defining equations in the Grassmannian, but for our purposes Y is general, so $F(Y)$ is a variety.

Lemma 14. *Let $(Z \to B, f : Z \to P,)$ be freely movable, and fix some subvariety $\Pi \subset P$. Then for general $b \in B$, Z_b intersects Π transversely.*

Proof. We argue using local holomorphic coordinates. Suppose that $\Pi \subset P$ is simply $\mathbb{C}^\ell \subset \mathbb{C}^n$, and that $f_b : D^r \to P$ is an embedding. Assume for the sake of contradiction that the locus $\Sigma \subset D^r \times B$ where f_b is not transverse to Π surjects onto B. We may choose $0 \in B$ such that $\Sigma \to B$ does not have a multiple fiber over 0. If f_0 is nontransverse to Π at $p \in D^r$, then we have

$$\mathbb{C}^\ell + df_0(T_p D^r) \subsetneq \mathbb{C}^n.$$

Let $\vec{v}$ be a vector outside the subspace above, and use the hypothesis of free movability to find a subfamily $f_t : D^r \to P$ and section $x(t) \in D^r$ such that $x(0) = p$ and $\frac{d}{dt}\big|_{t=0}(f \circ x) = v$. Since Σ does not have a multiple fiber, there exists another section $y(t) \in D^r$ such that $y(0) = p$ and $f_t(y(t))$ meets Π nontransversely. Now

$$f_t(y(t)) - f_0(p) = (f_t(y(t)) - f_t(x(t))) + (f_t(x(t)) - f_0(p)).$$

At first order in t, the left-hand side lies in $\mathbb{C}^\ell$, the first term on the right-hand side lies in the image of df_0, and the second term on the right-hand side lies in the span of $\vec{v}$. This contradicts our choice of v. Since transversality is Zariski open, we obtain the statement for general $b \in B$. □

Any smooth curve $C \subset X$ with class $\ell + nf$ maps isomorphically to a line $L \subset Y$. The preimage $\pi^{-1}(L)$ will contain C, so to set up the enumerative problem, we consider the family of all elliptic surfaces over lines in Y.

Definition 15. Let $U \to F(Y) \times Y$ be the universal line, and form the fibered product

$$\mathscr{S} := X \times_Y U.$$

The natural morphism $\nu : \mathscr{S} \to F(Y)$ is flat by base change and composition:

$$\begin{tikzcd} & \mathscr{S} \arrow[r] \arrow[dl, "\nu"'] \arrow[d, "\pi'"] & X \arrow[d, "\pi"] \\ F(Y) & U \arrow[l] \arrow[r] & Y \end{tikzcd}$$

The family ν will be our primary object of study. Its fiber over a line $[L] \in F(Y)$ is simply $\pi^{-1}(L)$. Proposition 38 shows that for $k \leq 4$, the fibers of ν have no worse than isolated A_ρ singularities. We have an associated moduli map to the Weierstrass moduli space

$$\mu_X : F(Y) \to \mathcal{W}_k$$

by restricting the global Weierstrass equation from Y to L.

Lemma 16. *The map μ_X is an immersion for general Y and $X \in W(Y, \mathcal{O}(k))$.*

Proof. This is a statement about unordered point configurations on $\mathbb{P}^1$. The argument is rather technical and is relegated to Appendix B. □

If we fix Y and vary $X \in W(Y, \mathcal{O}(k))$, we obtain a family of immersions μ_X.

Proposition 17. *The family of immersions given by $F(Y) \times W(Y, \mathcal{O}(k)) \to \mathcal{W}_k$ is freely movable in the sense of Definition 13.*

Proof. This follows from surjectivity of the restriction map

$$H^0(Y, \mathcal{O}(4k)) \oplus H^0(Y, \mathcal{O}(6k)) \to H^0(L, \mathcal{O}(4k)) \oplus H^0(L, \mathcal{O}(6k)).$$

There is no need to vary the line L, only $[A : B] \in W(Y, \mathcal{O}(k))$. □

By Lemma 14, we may assume after deformation that μ_X is transverse to any fixed subvariety of $\mathcal{W}_k$. This will be applied to the Noether–Lefschetz loci inside $\mathcal{W}_k$. The intersection of μ_X with the discriminant divisor in $\mathcal{W}_k$ is responsible for the correction term $\Theta(q)$ in Theorem 5.

3. Noether–Lefschetz theory

Let $S \in \mathcal{W}_k$ be an elliptic surface. Its Picard rank is automatically ≥ 2 because its Néron–Severi group $\mathrm{NS}(S)$ contains the fiber class f and the class of the identity section z. Any section class has self-intersection $-k$ by adjunction, so $\mathrm{NS}(S)$ contains the rank 2 lattice:

$$\langle f, z \rangle = \begin{pmatrix} 0 & 1 \\ 1 & -k \end{pmatrix}.$$

We refer to this sublattice as the polarization, and it comes naturally from the elliptic fibration structure.

Remark 18. Except for the K3 case ($k = 2$), the elliptic fibration structure on S is canonical because K_S is a nonzero multiple of the fiber class.

Definition 19. An elliptic surface $S \in \mathcal{W}_k$ is called Noether–Lefschetz special if its relatively minimal resolution has Picard rank > 2.

Theorem 20 [Cox 1990]. *All components of the Noether–Lefschetz locus in $\mathcal{W}_k$ are reduced of codimension* $k - 1$, *except for the discriminant divisor*, *which is codimension* 1.

Noether–Lefschetz theory is the study of Picard rank jumping in families of surfaces. The short exact sequence of Shioda–Tate for an elliptic surface clarifies the two potential sources of jumping:

$$0 \to V(S) \to \mathrm{NS}(S) \to \mathrm{MW}(S/\mathbb{P}^1) \to 0. \tag{3}$$

Here $V(S)$ is the sublattice spanned by the zero section class and all vertical classes, and $\mathrm{MW}(S/\mathbb{P}^1)$ is the Mordell–Weil group of the generic fiber, which is an elliptic curve over $\mathbb{C}(\mathbb{P}^1)$. If S is a smooth Weierstrass fibration, then all fibers are integral, so $V(S)$ is simply the polarization sublattice. The group $\mathrm{MW}(S/\mathbb{P}^1)$ will be torsion-free in the cases that concern us (see Appendix A), so the intersection form on $\mathrm{NS}(S)$ splits the short exact sequence (3). In particular, the orthogonal projection $\Pi : \mathrm{NS}(S) \to V(S)^\perp$ induces an isomorphism of groups

$$\mathrm{MW}(S/\mathbb{P}^1) \to V(S)^\perp.$$

Lemma 21. *If $\sigma \in \mathrm{NS}(S)$ is the class of a section curve, then its orthogonal projection to $V(S)^{\perp}$ has self-intersection*

$$-2(z\cdot\sigma+k).$$

Proof. Since $\mathrm{MW}(S/\mathbb{P}^1)$ is torsion-free, σ is orthogonal to all the exceptional curves. The projection can be computed by applying Gram–Schmidt to the polarization sublattice $\langle f, z\rangle$. ☐

Lemma 22. *If σ is a section curve, and σ^{*m} its m-th power with respect to Mordell–Weil group law, then the class of σ^{*m} in $\mathrm{NS}(S)$ is given by*

$$[\sigma^{*m}] = m\sigma-(m-1)z+(z\cdot\sigma+k)m(m-1)f.$$

In particular $(\sigma^{*m})\cdot z$ grows quadratically with m.

Proof. The class on the generic fiber is computed using the Abel–Jacobi map for elliptic curves. To determine the coefficient of f, use the fact that any section curve has self-intersection $-k$ in $\mathrm{NS}(S)$. ☐

Lemma 23. *If σ is the class of a section curve, and $\iota : S \to X$ is the inclusion morphism, then*

$$\iota_*(\sigma)=\ell+(z\cdot\sigma+k)f \in H_2(X,\mathbb{Z}).$$

Proof. The class can be computed by intersecting with complementary divisors. The global section $i(Y)\subset X$ has normal bundle $\mathcal{O}_Y(-k)$, whence the shift. ☐

Setting $\mathrm{NS}_0(S) := \langle f, z\rangle^{\perp}\subset \mathrm{NS}(S)$ and $V_0(S) := \mathrm{NS}_0(S)\cap V(S)$, the sequence

$$0\to V_0(S)\to \mathrm{NS}_0(S)\to \mathrm{MW}(S/\mathbb{P}^1)\to 0$$

is also split exact. The lattice $V_0(S)$ will be a root lattice spanned by the classes of exceptional curves.

To set up the Noether–Lefschetz jumping phenomenon, we consider the full polarized cohomology lattice

$$\Lambda(S) := \langle f, z\rangle^{\perp}\subset H^2(S,\mathbb{Z}).$$

Theorem 24. *As abstract lattices, $\Lambda(S)\simeq H^{\oplus 2k-2}\oplus E_8(-1)^{\oplus k}$, where H denotes the rank 2 hyperbolic lattice, and $E_8(-1)$ denotes the E_8 lattice with signs reversed.*

Proof. By Poincaré duality, the pairing on $H^2(S,\mathbb{Z})$ is unimodular, and the Hodge index theorem gives its signature to be $(2k-1, 10k-1)$. The polarization sublattice $\langle f, z\rangle$ is unimodular, so its orthogonal complement is as well. The Wu formula for Stiefel–Whitney classes reads

$$\alpha\cdot\alpha\equiv\alpha\cdot K_S \pmod 2,$$

so $\alpha\cdot\alpha\in 2\mathbb{Z}$ for $\alpha\in\langle z, f\rangle^{\perp}$. By the classification of indefinite unimodular lattices, there is a unique even lattice of signature $(2k-2, 10k-2)$, namely the one above. ☐

By the Lefschetz (1, 1) theorem, we have

$$\mathrm{NS}_0(S) \simeq (H^{2,0}(S) \oplus H^{0,2}(S))^{\perp}_{\mathbb{R}} \cap \Lambda(S),$$

so the Picard rank jumping can be detected from the Hodge structure of S. There is a period domain which parametrizes polarized[2] Hodge structures on the abstract lattice Λ. A weight 2 Hodge structure can be interpreted as a representation of the Deligne torus $\mathbb{S}^1 = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{C}^\times$, valued in the orthogonal group $O(\Lambda_{\mathbb{R}})$:

$$\psi : \mathbb{S}^1 \to O(\Lambda_{\mathbb{R}}),$$

such that $\psi(t) = t^2$ for $t \in \mathbb{R}^\times$. The Hodge decomposition comes from extending linearly to $\Lambda_{\mathbb{C}}$, and setting

$$H^{p,q}(\psi) = \{v \in \Lambda_{\mathbb{C}} : \psi(z) \cdot v = z^p \bar{z}^q v\}.$$

For fixed Hodge numbers, the group $O(\Lambda_{\mathbb{R}})$ acts transitively via conjugation on the set of such representations. This realizes the relevant period domain as a homogeneous space:

$$D \simeq O(2k-2, 10k-2)/\pm U(k-1) \times O(10k-2).$$

This can alternatively be viewed as an open orbit inside a complex flag variety, so in particular it has a complex structure. It contains a sequence of Noether–Lefschetz loci, which are given by

$$\widetilde{\mathrm{NL}}_{2n} := \bigcup_{\beta\in\Lambda,\, (\beta,\beta)=-2n} \{\psi \in D : H^{2,0}(\psi) \subset \beta^{\perp}\},$$

each of which is simultaneously a homogeneous space

$$\widetilde{\mathrm{NL}}_{2n} \simeq O(2k-2, 10k-3)/\pm U(k-1) \times O(10k-3)$$

and a complex submanifold of $\mathbb{C}$-codimension $k-1$. These loci parametrize Hodge structures on Λ which potentially come from a surface S with $\mathrm{NS}_0(S) \neq 0$, since $\beta \in \mathrm{NS}_0(S)$.

The family $\mathscr{S} \to F(Y)$ is generically smooth, so we have a holomorphic period map

$$j : F(Y) \dashrightarrow \Gamma\backslash D,$$

defined away from the singular locus, where Γ is the image of monodromy for the smooth family, which lies in the arithmetic group $O(\Lambda) \subset O(\Lambda_{\mathbb{R}})$.

Proposition 25. *The period map j is an immersion away from the singular locus.*

Proof. Combine Lemma 16 with the infinitesimal Torelli theorem of M. Saito [1983] for deformations of smooth elliptic surfaces. □

By Proposition 38, singularities in the fibers of ν are ADE type when $k \le 4$, so the local monodromy of the smooth family is finite order. This allows us to extend j over all of $F(Y)$ on general grounds [Schmid 1973]. The extension can be understood explicitly in terms of a simultaneous resolution (see Theorem 26). The period image of a singular surface is the Hodge structure of its minimal resolution.

[2]For the rest of the paper, all Hodge structures are assumed to be polarized.

The Noether–Lefschetz numbers of the family $\mathscr{S} \to F(Y)$ are morally the intersections of $j_*[F(Y)]$ with

$$\mathrm{NL}_{2n} := \Gamma\backslash\widetilde{\mathrm{NL}}_{2n} \subset \Gamma\backslash D.$$

However, since Γ contains torsion elements, the period space $\Gamma\backslash D$ has singularities. To compute the topological intersection product we consider instead the smooth analytic stack quotient $[\Gamma\backslash D]$. The period map j does not lift to this stack, so in Section 4 we construct a stack $\mathfrak{F}(Y)$ with coarse space $F(Y)$, admitting a map

$$\mathrm{j} : \mathfrak{F}(Y) \to [O(\Lambda)\backslash D].$$

lifting the classical period map, which extends to all of $\mathfrak{F}(Y)$.

Lastly, we note that the Noether–Lefschetz loci $\mathrm{NL}_{2n} \subset [O(\Lambda)\backslash D]$ are irreducible after fixing the divisibility of the lattice vector. This follows from a uniqueness theorem for embeddings of rank 1 lattices into a unimodular lattice; see Theorem 1.1.2 of [Nikulin 1979]. The locus $\mathrm{NL}_{2n} \subset [O(\Lambda)\backslash D]$ decomposes into components indexed by $m \in \mathbb{N}$ such that $m^2|n$. Let $v_m \in \Lambda$ be a lattice vector of self-intersection $2n/m^2$ so that mv_m has self-intersection $2n$. Then we can write

$$\mathrm{NL}_{2n} = \bigcup_{m^2|n} O(\Lambda)\backslash\{\psi \in D : H^{2,0}(\psi) \subset v_m^{\perp}\}.$$

4. Simultaneous resolution

In this section, we study flat families of surfaces with rational double points, focusing on the A_ρ case.

Theorem 26 [Brieskorn 1968]. *Let $\pi : X \to B$ be a flat family of surfaces over a smooth variety B, such that each fiber X_b has at worst ADE singularities. Then after a finite base change $B' \to B$ in the category of analytic spaces, the new family $\pi' : X' \to B'$ admits a* ***simultaneous resolution****, a proper birational morphism $\widetilde{X}' \to X'$ which restricts to a minimal resolution on each fiber of π.*

Étale locally it suffices to consider a versal family. In the case of the A_ρ singularity $(x^2 + y^2 = z^{\rho+1})$, this family is given by

$$x^2 + y^2 = z^{\rho+1} + s_1 z^{\rho-1} + s_2 z^{\rho-2} + \cdots + s_\rho \tag{4}$$

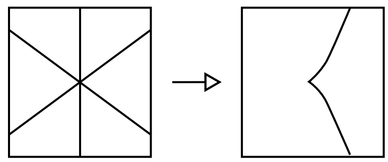
in the deformation coordinate $\vec{s} \in S = \mathbb{C}^\rho$. The base change required is given by the elementary symmetric polynomials

$$s_i = \sigma_{i+1}(\vec{t}),$$

where $\vec{t} \in T = \operatorname{Spec} \mathbb{C}[t_0, t_1, \ldots, t_\rho]/\sum t_i \simeq \mathbb{C}^\rho$. The cover is Galois with deck group $\mathfrak{S}_{\rho+1}$, branched over the discriminant hypersurface. After base change, the equation can be factored

$$x^2 + y^2 = \prod_{i=0}^{\rho} (z + t_i).$$

Singularities in the fibers occur over the big diagonal in T, which is the hyperplane arrangement dual to the root system $A_\rho \simeq \mathbb{Z}^\rho \subset \mathbb{C}^n$. The $\mathfrak{S}_3$ covering $T \to S$ in the case of A_2 is pictured below.



A simultaneous resolution can be obtained by taking a small resolution of the total space. This is far from unique; one can write the total space as an affine toric variety, and then choose any simplicial subdivision of the single cone:

$$\mathbb{Z}_+\langle \vec{e}_1, \vec{e}_2, \dots, \vec{e}_{\rho+2}, \vec{e}_1 - \vec{e}_2 + \vec{e}_3, \vec{e}_1 - \vec{e}_2 + \vec{e}_4, \dots, \vec{e}_1 - \vec{e}_2 + e_{\rho+2}\rangle \subset \mathbb{Z}^{\rho+2}.$$

The smooth fiber in both families is diffeomorphic to a resolved A_ρ singularity, and therefore has cohomology lattice

$$H^2(X_s, \mathbb{Z}) \simeq A_\rho(-1).$$

The Gauss–Manin local system on $S - \Delta$ corresponds to the representation

$$\pi_1(S - \Delta) \simeq \mathrm{Br}_{\rho+1} \to \mathfrak{S}_{\rho+1} \to O(A_\rho),$$

where $\mathrm{Br}_{\rho+1}$ denotes the braid group, and $\mathfrak{S}_{\rho+1}$ is the Weyl group of A_ρ acting by reflections. The base change morphism $T \to S$ can be interpreted as the quotient $\mathbb{C}^\rho \to \mathbb{C}^\rho/\mathfrak{S}_{\rho+1} \simeq \mathbb{C}^\rho$ by extending the Weyl group action $\mathbb{C}$-linearly to $A_\rho \otimes \mathbb{C}$. The quotient map is ramified along the dual hyperplane arrangement and branched over the discriminant Δ.

Artin framed Theorem 26 in the language of representable functors.

Theorem 27 [Artin 1974]. *For $X \to B$ as above, let $\mathrm{Res}_{X/B}$ be the functor from* $\mathrm{Sch}/B \to \mathrm{Set}$ *which sends*

$$[B' \to B] \mapsto \{\textit{simultaneous resolutions } \widetilde{X}' \to X' = X \times_B B'\}.$$

Then $\mathrm{Res}_{X/B}$ is represented by a locally quasiseparated algebraic space.

The space is often not separated, even in the case of an ordinary double point A_1:

Example 28. Let $X \to \mathbb{C}$ be the versal deformation of A_1: $x^2 + y^2 + z^2 = t$. To build a simultaneous resolution, we base change by $t \mapsto t^2$, and then take a small resolution of the threefold singularity $x^2 + y^2 + z^2 = t^2$. There are two choices of small resolution, differing by the Atiyah flop. Hence, the algebraic space $\mathrm{Res}_{X/\mathbb{C}}$ is isomorphic to $\mathbb{A}^1$, with the étale equivalence relation

$$\mathscr{R} = \Delta \cup \{(x, -x) : x \neq 0\} \subset \mathbb{A}^1 \times \mathbb{A}^1.$$

To do intersection theory, we want a nicer base for the simultaneous resolution, namely a smooth Deligne–Mumford stack. From the perspective of periods we only need a resolution at the level of cohomology lattices. Let $\Delta \subset B$ be the discriminant locus of the family, and $j : U \hookrightarrow B$ its complement. Assuming that U is nonempty, we have a Gauss–Manin local system $R^2\pi_{U*}(\mathbb{Z})$ on $U(\mathbb{C})$ whose stalk at $b \in U$ is $H^2(X_b, \mathbb{Z})$ equipped with the cup product pairing.

Remark 29. In our situation, we will consider instead the primitive cohomology $R^2\pi_{U*}(\mathbb{Z})_{\text{prim}} \subset R^2\pi_{U*}(\mathbb{Z})$ whose stalk at $b \in U$ is isomorphic to the orthogonal complement of the polarization sublattice $\langle f, z\rangle$ as defined in Section 3. This is well-defined over U because $\langle f, z\rangle$ is monodromy invariant.

The pushforward sheaf

$$\mathcal{H} := j_* R^2\pi_{U*}(\mathbb{Z})_{\text{prim}}$$

is a constructible sheaf on $B(\mathbb{C})$ whose stalk at $b \in \Delta(\mathbb{C})$ consists of classes invariant under the local monodromy action.

Definition 30. Let Λ be the stalk of $\mathcal{H}$ over a smooth point $b \in U(\mathbb{C})$. A cohomological simultaneous resolution is an embedding $\mathcal{H} \hookrightarrow \mathbb{L}$ into a local system $\mathbb{L}$ on $B(\mathbb{C})$ with stalk Λ.

If $X \to B$ admits a simultaneous resolution, then the local monodromy action is trivial so $\mathcal{H}$ is a local system already. If it does not, then the cohomological resolutions are representable by a stack:

Definition 31. The monodromy stack $\mathfrak{B}$ over B is the following category fibered in groupoids. An object of $\mathfrak{B}$ is given by a pair

$$(f : B' \to B, i' : f^*\mathcal{H} \hookrightarrow \mathbb{L}'),$$

where $\mathbb{L}'$ is a local system on $B'(\mathbb{C})$ with stalk Λ. A morphism from $(B', f, \mathbb{L}', i')$ to $(B'', g, \mathbb{L}'', i'')$ is a map $h : B' \to B''$ such that $f = g \circ h$, and an isomorphism $\phi : h^*\mathbb{L} \to \mathbb{L}$ such that $i' = \phi \circ h^* i''$.

To check that $\mathfrak{B}$ is a stack for the étale topology on B, we must verify that isomorphisms form a sheaf, and that objects satisfy descent. Both of these follow from the corresponding facts for local systems and the fact that étale morphisms of $\mathbb{C}$-schemes induce local isomorphisms on their underlying analytic spaces. Automorphisms of an object $(B', f, \mathbb{L}', i')$ are the automorphisms of $\mathbb{L}'$ which fix $i'(f^*\mathcal{H})$. When B' is a point p, then $\mathcal{H}_p$ is the space of local invariant cycles, so the automorphism group is generated by reflections in the vanishing cycle classes.

Theorem 32. *The stack $\mathfrak{B}$ is Deligne–Mumford.*

Proof. We have natural morphisms

$$\mathrm{Res}_{X/B} \to \mathfrak{B} \to B,$$

since a simultaneous resolution induces a cohomological one. Hence, $\mathfrak{B}$ is an algebraic stack by Theorem 27. The Deligne–Mumford property follows from the fact that simple surface singularities have finite monodromy. $\square$

Proposition 33. *If $X \to S$ is the versal family of the A_ρ singularity as described in* (4), *then its monodromy stack is isomorphic to*

$$[T/\mathfrak{S}_{\rho+1}],$$

which has coarse space S.

Proof. We define an equivalence of categories fibered over Sch/S. An object of $[T/\mathfrak{S}_{\rho+1}](Y)$ consists of a principal $\mathfrak{S}_{\rho+1}$-bundle $E \to Y$ with an equivariant map $\tilde{f} : E \to T$. Composing with the coarse space map $[T/\mathfrak{S}_{\rho+1}] \to S$ produces a map $f : Y \to S$. Let $\mathbb{L}$ be the sheaf of sections of the associated bundle

$$E \times_{\mathfrak{S}_{\rho+1}} A_\rho \to Y,$$

which has fiber A_ρ. We can describe $\mathcal{H}$ as the sheaf of sections of

$$T \times_{\mathfrak{S}_{\rho+1}} A_\rho \to S.$$

To be single-valued in a neighborhood of $s \in S$, a section of $\mathcal{H}$ must send s to the class of (t, a), where a is fixed by all $g \in \mathrm{Stab}_{\mathfrak{S}_{\rho+1}}(t)$. Since the action of $\mathfrak{S}_{\rho+1}$ is generated by reflections in the roots $R(A_\rho) \subset A_\rho$, this condition is equivalent to

$$a \in (t^\perp \cap R(A_\rho))^\perp.$$

With this in mind, we define

$$\widetilde{F} := \left\{(e, a) : a \in \left(\tilde{f}(e)^\perp \cap R(A_\rho)\right)^\perp\right\} \subset E \times A_\rho.$$

This is $\mathfrak{S}_{\rho+1}$-invariant, so it descends to

$$F \subset E \times_{\mathfrak{S}_{\rho+1}} A_\rho$$

over Y. The equivariant map $\tilde{f} : E \to T$ induces a map

$$E \times_{\mathfrak{S}_{\rho+1}} A_\rho \to (T \times_{\mathfrak{S}_{\rho+1}} A_\rho) \times_S Y. \tag{5}$$

If $\mathcal{F}$ is the sheaf of sections of $F \to Y$, then (5) gives an isomorphism

$$\mathcal{F} \to f^*\mathcal{H}.$$

Morphisms in $[T/\mathfrak{S}_{\rho+1}]$ are Cartesian diagrams of principal bundles with commuting equivariant maps, which induce morphisms of the above data. This defines a functor from $[T/\mathfrak{S}_{\rho+1}]$ to the monodromy stack. Recall that the action of $\mathfrak{S}_{\rho+1}$ on A_ρ gives an equivalence from the category of principal $\mathfrak{S}_{\rho+1}$-bundles to the category of A_ρ-local systems. The data of commuting maps to T corresponds to the coincidence of cohomology subsheaves, so our functor is fully faithful.

To show essential surjectivity, let $(f : Y \to S, i : f^*\mathcal{H} \hookrightarrow \mathbb{L})$ be an object of the monodromy stack. The bundle E of Weyl chambers in the stalks of $\mathbb{L}$ is a principal $\mathfrak{S}_{\rho+1}$-bundle over Y. Let $\Delta' \subset \Delta$ be the smallest singular stratum containing the image of f, which corresponds to a partition of $\rho+1$, or equivalently a conjugacy class of subgroup $G \subset \mathfrak{S}_{\rho+1}$. The general stalk of $f^*\mathcal{H}$ is isomorphic to the invariant sublattice $(A_\rho)^G$. We can form the associated bundle

$$\bar{E} := E \times_{\mathfrak{S}_{\rho+1}} \mathfrak{S}_{\rho+1}/G.$$

The fact that $f^*\mathcal{H}$ extends to a local system over Y implies that it is trivial, so $\bar{E}$ is a trivial bundle. Any choice of lift $Y \to T$ gives rise to an equivariant map from $\bar{E} \to T$. Lifting this to an equivariant map $E \to T$ is automatic because each G-coset maps to the same point of T. □

More generally, if $X \to B$ has only A_ρ singularities in the fibers, then étale locally on B, we have a morphism

$$B \to \prod_j S_j$$

to the versal bases of the isolated singularities in the central fiber. There is an embedding of the associated root lattice $A \subset \Lambda$, which induces an embedding of the Weyl group $W(A) \subset O(\Lambda)$. The monodromy of the family lies in $W(A)$, which implies that the diagram

$$\begin{tikzcd} \mathfrak{B} \arrow[r] \arrow[d] & \prod_j [T_j/\mathfrak{S}_{\rho_j+1}] \arrow[d] \\ B \arrow[r] & \prod_j S_j \end{tikzcd}$$

of stacks is 2-Cartesian. In particular, B is the coarse space of $\mathfrak{B}$. If the family has maximal variation at each singularity then $\mathfrak{B}$ is smooth.

We apply Theorem 32 to the family $\nu : \mathscr{S} \to F(Y)$ to obtain a Deligne–Mumford stack $\mathfrak{F}(Y)$ such that the (primitive) Gauss–Manin system on the smooth locus extends to a local system $\mathbb{L}$ on all of $\mathfrak{F}(Y)$ with stalk Λ. Let $\mathfrak{E}$ be the principal $O(\Lambda)$-bundle on $\mathfrak{F}(Y)$ of isomorphisms from $\mathbb{L}$ to the constant sheaf $\underline{\Lambda}$. There is an equivariant map from $\mathfrak{E} \to D$ sending a point in $\mathfrak{E}$ to the Hodge structure on Λ obtained by identifying the stalk of $\mathbb{L}$ with Λ via the isomorphism. This data gives the desired stacky period map

$$\mathrm{j} : \mathfrak{F}(Y) \to [O(\Lambda)\backslash D]$$

extending $j : F(Y) \dashrightarrow O(\Lambda)\backslash D$.

5. Modularity statement

The work of Kudla and Millson produces a modularity statement for intersection numbers in a general class of locally symmetric spaces M of orthogonal type. In this section, we summarize the material in [Kudla and Millson 1990],[3] and adapt it to our situation.

Let M be the double quotient $\Gamma\backslash O(p,l)/K$ of an orthogonal group on the left by a torsion-free arithmetic subgroup preserving an even unimodular lattice $\Lambda \subset \mathbb{R}^{p+l}$, and on the right by a maximal compact subgroup. This is automatically a manifold, since any torsion-free discrete subgroup acts freely on the compact cosets. We can interpret $O(p,l)/K$ as an open subset of the real Grassmannian $\mathrm{Gr}(p, p+l)$

[3] We match the notation of [Kudla and Millson 1990] for the most part, but all instances of positive and negative definiteness are switched.

consisting of those p-planes $Z \subset \mathbb{R}^{p+l}$ on which the form is positive definite. For any negative definite line $\langle v\rangle \subset \mathbb{R}^{p+l}$, set

$$\widetilde{C}_{\langle v\rangle} := \{Z \in O(p,l)/K : Z \subset \langle v\rangle^{\perp}\}$$

which is $\mathbb{R}$-codimension p. Indeed, the normal bundle to $\widetilde{C}_{\langle v\rangle}$ has fiber at Z equal to $\mathrm{Hom}(Z, \langle v\rangle)$. While the image of $\widetilde{C}_{\langle v\rangle}$ in M may be singular, it can always be resolved if we instead quotient by a finite index normal subgroup of Γ. Furthermore, [Kudla and Millson 1990] gives a coherent way of orienting the $\widetilde{C}_{\langle v\rangle}$, so that it makes sense to take their classes in the Borel–Moore homology group $H^{BM}_{pl-p}(M) \simeq H^p(M)$.

For any positive integer n, the action of Γ on the lattice vectors in Λ of norm $-2n$ has finitely many orbits [Borel 1969]. Choose orbit representatives $\{v_1, \dots, v_k\}$, and set

$$\widetilde{C}_{2n} := \bigcup_{i=1}^{k} \widetilde{C}_{\langle v_i\rangle}.$$

The image of $\widetilde{C}_{2n}$ in the arithmetic quotient M is denoted by C_{2n}. Locally, $\widetilde{C}_n$ is a union of smooth (real) codimension p cycles meeting pairwise transversely, one for each lattice vector of norm $-2n$ orthogonal to Z. We quote the following result directly from [Kudla and Millson 1990], in the case of $\mathrm{Sp}(1) \simeq \mathrm{SL}(2)$:

Theorem 34. *For any homology class $\alpha \in H_p(M)$, the series*

$$\alpha \cap e(V^{\vee}) + \sum_{n=1}^{\infty} \alpha \cap [C_{2n}]\, e^{2\pi i n\tau}$$

is a classical modular form for $\tau \in \mathbb{H}$ of weight $(p+l)/2$. The constant term is the integral of the Euler class of the (*dual*) *tautological bundle of p-planes.*

Remark 35. The statement in [Kudla and Millson 1990] does not specify the level, but when the lattice Λ is even and unimodular, we will see that the level group is the full modular group $\mathrm{SL}_2(\mathbb{Z})$. The proof of Theorem 34 proceeds by considering the functional $\Theta : \mathscr{S}(\Lambda \otimes \mathbb{R}) \to \mathbb{C}$ on Schwartz functions given by a sum of delta functions supported at lattice points:

$$\Theta = \sum_{v \in \Lambda} \delta_v.$$

Now $\mathscr{S}(\Lambda \otimes \mathbb{R})$ admits an action of $O(\Lambda \otimes \mathbb{R}) \times \mathrm{Mp}_2(\mathbb{R})$, where the first factor acts by precomposition, and the second factor acts via the Weil representation. Explicitly, the generators S and T of $\mathrm{Mp}_2(\mathbb{Z})$ act via

$$(T \cdot f)(x) = e^{\pi i (x,x)} f(x) \quad \text{and} \quad (S \cdot f)(x) = \frac{e^{\mathrm{sgn}(\Lambda)\pi i/4}}{\sqrt{\det(\Lambda^{\vee})}} \hat{f}(x),$$

where $\hat{f}$ denotes the Fourier transform. Since $\Gamma \subset O(\Lambda)$, we see that Θ is Γ-invariant. Observe that Θ is also $\mathrm{Mp}_2(\mathbb{Z})$-invariant, using the Poisson summation formula combined with the fact that Λ is even (for T), unimodular (for S), and $\mathrm{sgn}(\Lambda)$ is divisible by 8. Kudla and Millson then defined the composition

$$\theta : H^*_{cts}(O(\Lambda_{\mathbb{R}}), \mathscr{S}(\Lambda_{\mathbb{R}})) \xrightarrow{\text{res}} H^*(\Gamma, \mathscr{S}(\Lambda_{\mathbb{R}})) \xrightarrow{\Theta} H^*(\Gamma, \mathbb{C}) \simeq H^*(M, \mathbb{C}).$$

The cohomological theta correspondence is a morphism:

$$H_c^i(M,\mathbb{C})\otimes H_{cts}^{a-i}(O(\Lambda_{\mathbb{R}}),\mathscr{S}(\Lambda_{\mathbb{R}}))\to\mathscr{C}^\infty(\mathrm{Mp}_2(\mathbb{R})),\qquad (\mu,\nu)\mapsto\varphi(g)=\int_M\mu\wedge\theta(g\cdot\nu),$$

for $g\in\mathrm{Mp}_2(\mathbb{R})$. The action of $\mathrm{Mp}_2(\mathbb{R})$ on $H^*_{cts}(O(\Lambda_{\mathbb{R}}),\mathscr{S}(\Lambda_{\mathbb{R}}))$ induces an action of its complexified Lie algebra, which splits into $\mathfrak{h}\oplus\mathfrak{p}_+\oplus\mathfrak{p}_-$ (the Cartan, holomorphic, and antiholomorphic parts). If ν is annihilated by $\mathfrak{p}_-$, and $\mathfrak{h}$ acts on ν with weight w, then φ descends to a holomorphic modular form of weight $w/2$ for $\mathrm{Mp}_2(\mathbb{Z})$. If m is even, then this modular form descends to $\mathrm{SL}_2(\mathbb{R})$. Kudla and Millson constructed special classes ν such that the resulting φ has only positive Fourier coefficients controlled by intersection numbers with the special cycles C_{2n}.

To apply this statement to our situation, we note that the further quotient map

$$g:D\to O(2k-2,10k-2)/O(2k-2)\times O(10k-2)$$

is a smooth proper fiber bundle with fiber $SO(2k-2)/U(k-1)$. Given a positive definite real $(2k-2)$-plane $Z\subset\Lambda_{\mathbb{R}}$, a polarized Hodge structure is given by a choice of splitting

$$Z_{\mathbb{C}}\simeq H^{0,2}\oplus H^{2,0}\subset\Lambda_{\mathbb{C}}$$

into a pair of conjugate complex subspaces, isotropic with respect to the form. A fiber of g over $[Z]$ corresponds to this choice, which does not affect the orthogonal complement of Z. Thus, the Noether–Lefschetz loci in D are pulled back from the symmetric space:

$$\widetilde{\mathrm{NL}}_{2n}=g^{-1}(\widetilde{C}_{2n}).$$

The constant term of the series can be interpreted in terms of the Hodge bundle on $\Gamma\backslash D$. Indeed, if V is the tautological bundle of p-planes,

$$g^*V\otimes\mathbb{C}=V^{0,2}\oplus V^{2,0},$$

where each summand is isomorphic to g^*V as a real vector bundle. There is a natural complex structure on $V^{2,0}$ coming from the Hodge filtration, so we can take its Chern class:

$$c_{\mathrm{top}}(V^{2,0})=g^*e(V).$$

Theorem 34 is only proved for Γ torsion-free, so that M is a manifold. For our application, we need to allow torsion elements, since ADE singularities have finite order monodromy. For convenience, we will take $\Gamma=O(\Lambda)$.

Lemma 36. *$O(\Lambda)$ contains a finite index normal subgroup which is torsion-free.*

Proof. This is a well-known result of Selberg; see for instance Corollary 17.7 in [Borel 1969]. □

We denote the torsion-free subgroup by $\Gamma_{\mathrm{tf}} \subset O(\Lambda)$. The analytic stack $[O(\Lambda)\backslash D]$ can be realized as the quotient of a complex manifold by a finite group $G \simeq O(\Lambda)/\Gamma_{\mathrm{tf}}$. The spaces described above are related by

$$\begin{tikzcd} & \bar{D} := \Gamma_{\mathrm{tf}}\backslash D \arrow[dl, "h"'] \arrow[dr, "g"] & \\ {[O(\Lambda)\backslash D]} & & M \end{tikzcd}$$

where h is a G-cover, and g is a proper fiber bundle. The modularity statement carries over to intersections on the stack as follows. If $\alpha \in H_p([O(\Lambda)\backslash D], \mathbb{Q})$ is a rational homology class, and $\overline{\mathrm{NL}}_{2n} \subset \bar{D}$ is the Noether–Lefschetz locus in the manifold $\bar{D}$,

$$\alpha \cap h_*[\overline{\mathrm{NL}}_{2n}] = h^*\alpha \cap [\overline{\mathrm{NL}}_{2n}] = h^*\alpha \cap g^*[C_{2n}] = g_*h^*\alpha \cap [C_{2n}],$$

by repeated applications of the push-pull formula, which is valid for smooth stacks of Deligne–Mumford type in the sense of [Behrend 2004]. Since the locus $\overline{\mathrm{NL}}_{2n}$ inside $\bar{D}$ is $O(\Lambda)$-invariant, its pushforward under h acquires a multiplicity of $|G|$. This overall factor can be divided out from the generating series. With these adjustments, we modify Theorem 34 to fit our Hodge theoretic situation:

Theorem 37. *For any homology class $\alpha \in H_p([O(\Lambda)\backslash D], \mathbb{Q})$, the series*

$$\varphi(q) = \alpha \cap c_{\mathrm{top}}(\lambda^\vee) + \sum_{r=1}^{\infty} \alpha \cap [\mathrm{NL}_{2n}]\, q^n$$

is a modular form of weight $6k-2$ and level $\mathrm{SL}_2(\mathbb{Z})$. The constant term is the integral of the top Chern class of the dual Hodge bundle.

We apply this statement to the $\alpha = \mathrm{j}_*[\mathfrak{F}(Y)]$, as defined in Section 4. To compute the intersection product, we spread out the period map to a section of a smooth fiber bundle over $\mathfrak{F}(Y)$. Using the principal $O(\Lambda)$-bundle $\mathfrak{E} \to \mathfrak{F}(Y)$ in the definition of j, we set

$$\bar{D}(\mathbb{L}) := \bar{D} \times_{O(\Lambda)} \mathfrak{E} = [(\bar{D} \times \mathfrak{E})/O(\Lambda)],$$

which admits a section $s : \mathfrak{F}(Y) \to \bar{D}(\mathbb{L})$ coming from the graph of the period map. The Noether–Lefschetz loci can be spread out similarly

$$\overline{\mathrm{NL}}_{2n}(\mathbb{L}) := \overline{\mathrm{NL}}_{2n} \times_{O(\Lambda)} \mathfrak{E}.$$

By Lemma 2.1 of [Kudla and Millson 1990], after further shrinking Γ_{tf}, we may assume that $\overline{\mathrm{NL}}_{2n} \subset \bar{D}$ has only normal crossing singularities. As a result, the map from the normalization $\overline{\mathrm{NL}}_{2n}(\mathbb{L})^{\mathrm{norm}} \to \bar{D}(\mathbb{L})$ is a local regular embedding. The section $s : \mathfrak{F}(Y) \to \bar{D}(\mathbb{L})$ is also local regular embedding of stacks, and it fits into the 2-Cartesian square:

$$\begin{tikzcd} W \arrow[r, "s'"] \arrow[d, "g'"'] & \overline{\mathrm{NL}}_{2n}(\mathbb{L})^{\mathrm{norm}} \arrow[d, "g"] \\ \mathfrak{F}(Y) \arrow[r, "s"] & \bar{D}(\mathbb{L}) \end{tikzcd}$$

The desired intersection number is given by

$$\mathrm{j}_*[\mathfrak{F}(Y)] \cap [\mathrm{NL}_{2n}] = s_*[\mathfrak{F}(Y)] \cap [\overline{\mathrm{NL}}_{2n}(\mathbb{L})^{\mathrm{norm}}] = \deg s^![\overline{\mathrm{NL}}_{2n}(\mathbb{L})^{\mathrm{norm}}] = \deg g^![\mathfrak{F}(Y)].$$

We use Vistoli's formalism [1989] to write the Gysin map $g^!$ in terms of a normal cone. Let $N = (s')^* N_g$, a vector bundle containing the normal cone $C_{W/\mathfrak{F}(Y)}$. Since $W \subset \mathfrak{F}(Y)$ is a closed substack of a Deligne–Mumford stack, and N is the dual Hodge bundle on W, we are in the algebraic setting of Vistoli, so the intersection number is given by $0^!_N[C_{W/\mathfrak{F}(Y)}]$. This can be computed étale locally on $\mathfrak{F}(Y)$. If $f : Z \to \mathfrak{F}(Y)$ is an étale morphism from a scheme, we can form the fibered product

$$\overline{D}(\mathbb{L})_Z := Z \underset{\mathfrak{F}(Y)}{\times} \overline{D}(\mathbb{L}) \to Z$$

which also admits a section $s_Z : Z \to \overline{D}(\mathbb{L})_Z$. Lastly, we form $W_Z = Z \times_{\mathfrak{F}(Y)} W$ and $\overline{\mathrm{NL}}_{2n}(\mathbb{L})_Z = \overline{D}(\mathbb{L})_Z \times_{\overline{D}(\mathbb{L})} \overline{\mathrm{NL}}_{2n}(\mathbb{L})$. This spaces fit into the cubic diagram below, where each face is 2-Cartesian.

$$\begin{tikzcd}
W_Z \arrow[rr] \arrow[dr, "f'"'] \arrow[dd] & & \overline{\mathrm{NL}}_{2n}(\mathbb{L})_Z \arrow[dr] \arrow[dd, "g_Z" near start] & \\
& W \arrow[rr] \arrow[dd] & & \overline{\mathrm{NL}}_{2n}(\mathbb{L}) \arrow[dd, "g"] \\
Z \arrow[rr, "s_Z" near end] \arrow[dr, "f"'] & & \overline{D}(\mathbb{L})_Z \arrow[dr] & \\
& \mathfrak{F}(Y) \arrow[rr, "s"] & & \overline{D}(\mathbb{L})
\end{tikzcd}$$

The front and lateral sides are 2-Cartesian by construction, the bottom uses the fact that s is a monomorphism, and the top and back are proven by repeated application of the universal property. All the diagonal morphisms are étale, by stability of the étale property under base change. The normal cones are related by

$$C_{W_Z/Z} = (f')^* C_{W/\mathfrak{F}(Y)}.$$

Assuming that Z covers the support of W in $\mathfrak{F}(Y)$, we have

$$\deg g_Z^![Z] = \deg(f) \cdot \deg g^![\mathfrak{F}(Y)].$$

We use this formula in Section 7 to compute the contributions of 0-dimensional intersections to the $\Theta(q)$ correction term.

6. Discriminant hypersurfaces

Consider the cuspidal hypersurface cut out by $\Delta = 4A^3 + 27B^2$ of degree $12k$ inside $\mathbb{P}^{m+1}$. The intersection multiplicity of a line $L \subset \mathbb{P}^{m+1}$ with Δ dictates the singularities in the surface $\pi^{-1}(L)$. Mildly singular elliptic surfaces still have pure Hodge structures (given by a minimal resolution), and the presence of exceptional curves makes these surfaces Noether–Lefschetz special. The contributions of such singular surfaces to the modular form $\varphi(q)$ are enough to determine it uniquely. In this section, we compute those contributions using classical enumerative geometry.

Proposition 38. *The fibers of $\nu : \mathscr{S} \to F(Y)$ have isolated rational double points of type A_ρ.*

Proof. The surface $\pi^{-1}(L)$ can only be singular at the singular points of the cubic fibers, since elsewhere it is locally a smooth fiber bundle.

- If L intersects Δ_{sm} at a point p with multiplicity μ, then the local equation of $\pi^{-1}(L)$ near the node in the fiber $\pi^{-1}(p)$ is
$$x^2 + y^2 = t^\mu.$$
This is an $A_{\mu-1}$ singularity.[4] We refer to such lines as Type I.
- If L intersects $\Delta_{\text{sing}} = (A) \cap (B)$ at a point p, then the fiber $\pi^{-1}(p)$ has a cusp. The local equation of $\pi^{-1}(L)$ there depends on the intersection multiplicity α (resp. β) of L the hypersurface (A) (resp. (B)):
$$x^3 + y^2 = t^\alpha x + t^\beta.$$
This singularity can be wild for large values of α and β, but the codimension of this phenomenon is $\alpha + \beta - 1$. For $k \le 4$, only A_1 and A_2 singularities can occur for $L \in F(Y)$ since the latter has dimension $k - 1$. We refer to such lines as Type II.

α	β	Type
$n \geq 1$	1	A_0
1	$n > 1$	A_1
2	2	A_2

Lastly, we must rule out the possibility of lines $L \subset Y$ lying completely inside Δ, because otherwise $\pi^{-1}(L)$ would not be normal. For this, consider the incidence correspondence
$$\Omega = \{(L, [A : B]) : (4A^3 + 27B^2)|_L = 0\} \subset F(Y) \times W(\mathbb{P}^{m+1}, \mathcal{O}(k)).$$
The first projection $\Omega \to F(Y)$ is surjective and has irreducible fibers. To see this, note that if $(4A^3 + 27B^2)|_L = 0$, then
$$A|_L = -3f^2, \quad B|_L = 2f^3$$
for some $f \in H^0(\mathbb{P}^1, \mathcal{O}(2k))$, by unique factorization. The codimension of this locus in $W(\mathbb{P}^1, \mathcal{O}(k))$ is greater than
$$8k > k - 1 = \dim F(Y).$$
Hence, the second projection is not dominant, so for general A and B, there are no lines on Y contained inside the discriminant hypersurface. $\square$

[4]By convention, A_0 means a smooth point.

We introduce tangency schemes to record how these A_ρ singularities appear,

Definition 39. Given a partition μ of $12k$, let

$$T_\mu(\Delta) := \overline{\left\{L : \Delta_{sm} \cap L = \sum \mu_j p_j\right\}} \subset \mathbb{G}(1, m+1).$$

Proposition 40. *For general A and B, $T_\mu(\Delta)$ has codimension*

$$\sum_{j=1}^{l} (\mu_j - 1)$$

when the latter is $\leq k \leq 4$.

Proof. For the partition $(\mu_1, 1, \ldots, 1)$, consider the incidence correspondence

$$\Omega_{\mu_1} = \{(L, p, [A : B]) : p \in L,\ \mathrm{mult}_p(A^3 + B^2)|_L \geq \mu_1\} \subset U \times W(\mathbb{P}^{m+1}, \mathcal{O}(k)),$$

The first projection $\Omega_{\mu_1} \to U$ has fiber cut out by μ_1 equations on W, which we wish to be independent. Setting $A(t) = a_0 + a_1 t + \cdots$ and $B(t) = b_0 + b_1 t + \cdots$ for t the uniformizer at p, the differential of the multiplicity condition is given by

$$\begin{pmatrix} 3a_0^2 & 0 & 0 & 0 & \cdots & 2b_0 & 0 & 0 & \cdots & 0 \\ 6a_0a_1 & 3a_0^2 & 0 & 0 & \cdots & 2b_1 & 2b_0 & 0 & \cdots & 0 \\ 6a_0a_2 + 3a_1^2 & 6a_0a_1 & 3a_0^2 & 0 & \cdots & 2b_2 & 2b_1 & 2b_0 & \cdots & 0 \\ \cdots & & & & & \cdots & & & & \end{pmatrix}$$

The rows are independent unless $a_0 = b_0 = 0$. In this case, comparing the t-valuations of A, B, and $A^3 + B^2$, we see that $A(t)$ (resp. $B(t)$) is actually divisible by $t^{\lceil \mu_1/3 \rceil}$ (resp. $t^{\lceil \mu_1/2 \rceil}$) when $\mu_1 \leq 6$. There is an entire irreducible component

$$\Omega_{\mu_1,\mathrm{II}} = \left\{(L, p, A, B) : p \in L,\ \mathrm{mult}_p(A) \geq \lceil \mu_1/3 \rceil,\ \mathrm{mult}_p(B) \geq \lceil \mu_1/2 \rceil\right\} \subset \Omega$$

whose fiber a pair (L, p) is a linear subspace of W. For $\mu_1 < 6$, Ω_{μ_1} is equidimensional with two irreducible components of dimension $\dim \mathbb{G} + 1 + \dim W - \mu$:

$$\Omega = \Omega_{\mu_1,\mathrm{I}} \cup \Omega_{\mu_1,\mathrm{II}}.$$

Note that $\Omega_{\mu_1,\mathrm{II}}$ consists of lines with points which lie in the *singular locus* of Δ, where $A = B = 0$. Since we are ultimately interested in $T_\mu(\Delta)$, which is defined in terms of tangency at *smooth* points of Δ, we focus on $\Omega_{\mu_1,\mathrm{I}}$. Consider the second projection $\Omega_{\mu_1,\mathrm{I}} \to W$. It is dominant by a dimension count, so the general fiber must have codimension $\mu_1 - 1$.

The case of a general partition μ is similar; gather only the multiplicities μ_j $(1 \leq j \leq l)$ which are greater than 1, and consider

$$\begin{aligned} \Omega_\mu = \left\{(L, p_1, \ldots, p_l, A, B) : p_j \in L,\ (A^2{+}B^3)|_L = \sum \mu_j p_j\right\} &\subset (U \times_{\mathbb{G}} \cdots \times_{\mathbb{G}} U) \times W(\mathbb{P}^{m+1}, \mathcal{O}(k)) \\ &=: \mathscr{U} \times W. \end{aligned}$$

Consider the fiber at $(L, p_1, \dots, p_l) \in \mathscr{U}$ for distinct points $p_j \in L$. If A and B do not simultaneously vanish at any of the points, then the matrix of differentials can be row reduced to a matrix with blocks of the form

$$\frac{1}{n!}\partial_x^n(1, x, x^2, x^3, \dots)|_{x=p_j} \qquad (0 \le n \le \mu_j).$$

This is a generalized Vandermonde matrix which has independent rows since the points are distinct. If A and B vanish simultaneously at some p_j, then in fact they vanish maximally, which is a linear condition on W of the expected codimension. Hence, the general fiber of $\Omega_\mu \to \mathscr{U}$ has 2^l irreducible components, which collapse over the big diagonal. Since the monodromy is trivial, we conclude that Ω_μ itself has 2^l components. We are only interested in one of them, $\Omega_{\mu,I}$, which gives the desired codimension for $T_\mu(\Delta)$. □

Next, we will show that Type II lines $L \in F(Y)$ are always limits of Type I lines (when $k \le 4$), so we can effectively ignore them. Trailing 1's in the partitions are suppressed for convenience.

Lemma 41. *Let $J_{\alpha,\beta}(A, B) \subset \mathbb{G}(1, m+1)$ be the locus lines L meeting (A) (resp. (B)) with multiplicity α (resp. β) at a common point $p \in (A) \cap (B)$. Then we have*

$$J_{1,2}(A, B),\ J_{1,3}(A, B) \subset T_2(\Delta); \quad J_{2,2}(A, B) \subset T_3(\Delta).$$

These are the only Type II lines which remain after intersecting with $F(Y)$.

Proof. Since $J_{1,2}(A, B)$ is codimension 2, we intersect Δ with a general $\mathbb{P}^2 \subset \mathbb{P}^{m+1}$ containing p. The Type II lines are those which pass through the cusps of $\mathbb{P}^2 \cap \Delta$ in the preferred (cuspidal) direction. They lie in the dual plane curve to $\mathbb{P}^2 \cap \Delta$. Since $J_{2,2}(A, B)$ and $J_{1,3}(A, B)$ are codimension 3, we intersect Δ with a general $\mathbb{P}^3 \subset \mathbb{P}^{m+1}$ containing p. The lines in $J_{2,2}(A, B)$ are those tangent to the curve $\mathbb{P}^3 \cap (A) \cap (B)$. They are always limits of flex lines at smooth points of $\mathbb{P}^3 \cap \Delta$, by a local calculation. The lines in $J_{1,3}(A, B)$ are those meeting $\mathbb{P}^3 \cap (A) \cap (B)$, while being flex to $\mathbb{P}^3 \cap (B)$. They are always limits of tangent lines at smooth points of $\mathbb{P}^3 \cap \Delta$, by a local calculation. □

With these dimension results in hand, we turn to degree computations. The main tools are the Plücker formulas relating degree, # of nodes, and # of cusps for a plane curve (d, δ, c) with those three numbers for the dual curve (d^*, δ^*, c^*). Note that δ^* (resp. c^*) is the number of bitangent (resp. flex) lines to the original curve.

Proposition 42. *The class of $T_2(\Delta)$ in $H_{4m-2}(\mathbb{G}(1, m+1))$ is Poincaré dual to*

$$12k(6k-1) \cdot \sigma_1 \in H^2(\mathbb{G}(1, m+1)).$$

Proof. We intersect $[T_2(\Delta)]$ with the complementary Schubert class σ^1, which is represented by a pencil of lines in a general $\mathbb{P}^2 \subset \mathbb{P}^{n+1}$. Since $\mathbb{P}^2 \cap \Delta$ is a curve of degree $d = 12k$ with $c = 4k \cdot 6k$ cusps, the intersection number is given by the Plücker formula for the dual degree:

$$d^* = d(d-1) - 3c. \qquad \square$$

Proposition 43. *The class of $T_3(\Delta)$ in $H_{4m-4}(\mathbb{G}(1, m+1))$ is Poincaré dual to*

$$24k(10k-3)\cdot\sigma_{11} + 24k(6k-1)(4k-1)\cdot\sigma_2 \in H^4(\mathbb{G}(1, m+1)).$$

Proof. First, we intersect $[T_3(\Delta)]$ with the class σ^{11}, which is represented by all the lines in a general $\mathbb{P}^2 \subset \mathbb{P}^{n+1}$. The intersection number is again given by the Plücker formula for flex lines applied to $\mathbb{P}^2 \cap \Delta$:

$$c^* = 3d(d-2) - 8c.$$

Next, we intersect with the class σ^2, which is represented by all lines through a point in a general $\mathbb{P}^3 \subset \mathbb{P}^{n+1}$. To compute this number we find the top Chern class of a bundle of principal parts. Consider the universal line $U \to \mathbb{G}(1, 3) \times \mathbb{P}^3$, and let $\mathcal{P}$ denote the bundle whose fiber at a point (L, p) is the space of 2nd order germs at p of sections of $\mathcal{O}_L(12k)$. The surface $\mathbb{P}^3 \cap \Delta$ induces a section of $\mathcal{P}$, by restriction, vanishing at each flex. By standard arguments in [Eisenbud and Harris 2016], $\mathcal{P}$ admits a filtration with successive quotients given by

$$\pi_2^*\mathcal{O}_{\mathbb{P}^3}(12k),\ \pi_2^*\mathcal{O}_{\mathbb{P}^3}(12k)\otimes\Omega^1_{U/\mathbb{G}},\ \pi_2^*\mathcal{O}_{\mathbb{P}^3}(12k)\otimes\mathrm{Sym}^2\Omega^1_{U/\mathbb{G}}.$$

By the Whitney sum formula, its top Chern class is given by

$$\begin{aligned} c_3(\mathcal{P}) &= (12k\zeta)((12k-2)\zeta+\sigma_1)((12k-4)\zeta+2\sigma_1) \\ &= 96k(6k-1)(3k-1)\zeta^3 + 48k(9k-2)\zeta^2\sigma_1 + 24k\zeta\sigma_1^2, \end{aligned}$$

where ζ denotes the relative hyperplane class on $U \to \mathbb{G}(1, 3)$ as a projective bundle. We intersect this class with $\sigma_2 \in H^4(\mathbb{G}(1, 3))$ to get the number of flex lines for Δ passing through a general point in $\mathbb{P}^3$.

$$c_3(\mathcal{P})\cdot\sigma_2 = 96k(6k-1)(3k-1) + 48k(9k-2) + 24k.$$

This number is too high, since lines meeting $(A)\cap(B)$ tangent to (B) have vanishing principal part, and each contribute multiplicity 8 by a local calculation. If $\theta : (B) \to \mathbb{P}^{3*}$ is the Gauss map associated the hypersurface (B), the number of such lines is

$$\deg(\theta|_{(A)\cap(B)}) = 4k\cdot 6k(6k-1).$$

Subtracting this cuspidal correction from the Chern class gives the desired number. □

Proposition 44. *The class of $T_{2,2}(\Delta)$ in $H_{4m-4}(\mathbb{G}(1, m+1))$ is Poincaré dual to*

$$108k(3k-1)(8k^2-1)\cdot\sigma_{11} + 36k(6k-1)(4k-1)(3k-1)\cdot\sigma_2 \in H^4(\mathbb{G}(1, m+1)).$$

Proof. The intersection of $[T_{2,2}(\Delta)]$ with the class σ^{11} is given by the Plücker formula for bitangent lines applied to $\mathbb{P}^2 \cap \Delta$:

$$\delta^* = \frac{d^*(d^*-1) - d - 3c^*}{2}.$$

Next, we intersect with the class σ^2, which counts bitangent lines to $S = \mathbb{P}^3 \cap \Delta$ passing through a general point $p \in \mathbb{P}^3 \subset \mathbb{P}^{n+1}$. Consider the projection from p

$$\Pi_p : S \to \mathbb{P}^2.$$

The normalization of S is a smooth surface $\widetilde{S}$, and we write

$$\widetilde{\Pi}_p : \widetilde{S} \to \mathbb{P}^2$$

for the composition. The ramification curve $R = R' \cup R''$ has two components: R' is the preimage of the cusp curve, and R'' is the closure of the ramification locus for $\Pi_p : S_{sm} \to \mathbb{P}^2$. These lie over the branch locus $B = B' \cup B'' \subset \mathbb{P}^2$. The degree of B'' was already computed in Proposition 42,

$$\deg(B'') = 12k(6k-1),$$

so we know its arithmetic genus. Nodes of B'' correspond to a bitangent lines to S through p, and cusps of B'' correspond to a flex lines to S through p, which we already counted in Section 6. There are no worse singularities in B'' for a general choice of p, so it suffices to compute

$$\delta(B'') = p_a(B'') - g(R'').$$

The Riemann–Hurwitz formula says that

$$K_{\widetilde{S}} = \widetilde{\Pi}_p^* K_{\mathbb{P}^2} + R = -3H + R.$$

Realizing $\widetilde{S}$ inside the blow up of $\mathbb{P}^3$ along $\Sigma = \mathbb{P}^3 \cap (A) \cap (B)$, the adjunction formula reads

$$K_{\widetilde{S}} = (12k-4)H - E = (12k-4)H - 2R'.$$

Combining these equations, we deduce that

$$R'' = (12k-1)H - 3R'.$$

This is enough to determine the genus of R'', using

$$H \cdot H = 12k; \quad H \cdot R' = 24k^2; \quad R' \cdot R' = 48k^3.$$

The latter can be computed inside the projective bundle $\mathbb{P}N_{\Sigma/\mathbb{P}^3}$, where R' is the class of $\mathbb{P}N_{\Sigma/(B)}$. As an additional check, observe that

$$R' \cdot R'' = 24k^2(6k-1) = \deg(\theta|_{(A)\cap(B)}),$$

which agrees with the cuspidal correction from Section 6. Finally, we use the genus formula on R'':

$$2g-2 = (K_{\widetilde{S}} + R'') \cdot R''; \quad \delta(B'') = 12k(9k-1)(6k-1)(4k-1).$$

Subtracting the flex line count from Section 6 leaves the desired number. $\square$

In the sequel, we will use the notation

$$t_\mu := [T_\mu(\Delta)] \cdot F(Y)$$

when this intersection is 0-dimensional, that is,

$$\sum_{j=1}^{l} (\mu_j - 1) = k - 1.$$

By the results of Section 4, the stack $\mathfrak{F}(Y)$ will have isotropy group

$$\prod_{j=1}^{l} \mathfrak{S}_{\mu_j}$$

at these isolated points.

7. Counting curves

We begin by discussing the case $k = 1$, which is trivial because the period domain is a point. Since $F(Y)$ is 0-dimensional, let $N_m = \#F(Y)$, and we may assume that for each $[L] \in F(Y)$, the rational elliptic surface $S = \pi^{-1}(L)$ is smooth. Every class in the polarized Néron–Severi lattice $\mathrm{NS}_0(S) \simeq E_8$ is the orthogonal projection of a section curve. The degree shifts from Lemmas 21 and 23 match, so we have

$$\sum_{n \geq 1} r_X(n) q^n = N_m \theta_{E_8}(q) - N_m$$

and $\theta_{E_8}(q)$ is the modular form of weight 4, as desired.

For $2 \leq k \leq 4$, recall the family of elliptic surfaces $\nu : \mathscr{S} \to F(Y)$ defined by the diagram

$$\begin{tikzcd} & \mathscr{S} \arrow[r] \arrow[dl, "\nu"'] \arrow[d, "\pi'"] & X \arrow[d, "\pi"] \\ F(Y) & U \arrow[l] \arrow[r] & Y \end{tikzcd}$$

The period map $j : F(Y) \to O(\Lambda)\backslash D$ lifts to an immersion of stacks $j : \mathfrak{F}(Y) \to [O(\Lambda)\backslash D]$ of Deligne–Mumford type. Theorem 37 applied to $\alpha = j_*[\mathfrak{F}(Y)]$ yields a classical modular form $\varphi(q)$ of weight $6k-2$ and full level. Our first task is to determine this modular form. We start by computing the constant term as the top Chern class of the dual Hodge bundle:

Definition 45. The Hodge bundle of the family $\nu : \mathscr{S} \to F(Y)$ is defined as the pushforward of the relative dualizing sheaf:

$$\lambda := \nu_*(\omega_{\mathscr{S}/F(Y)}).$$

Proposition 46. *Let S be the tautological rank 2 bundle on $F(Y)$. Then*

$$\lambda \simeq \mathrm{Sym}^{k-2}(S^\vee) \otimes \mathcal{O}_{F(Y)}(\sigma_1).$$

Proof. Writing ν as the composition $u \circ \pi'$, we compute the pushforward in stages.

$$\begin{aligned}\pi'_*(\omega_{\mathscr{S}/F(Y)}) &= \pi'_*(\omega_{\mathscr{S}/U} \otimes \pi'^*\omega_{U/F(Y)}) \\ &= \pi'_*(\omega_{\mathscr{S}/U}) \otimes \omega_{U/F(Y)} \\ &= \mathcal{O}_U(k\zeta) \otimes \omega_{U/F(Y)} \\ &= \mathcal{O}_U(k\zeta) \otimes \mathcal{O}_U(-2\zeta) \otimes u^*\mathcal{O}_{F(Y)}(\sigma_1)\end{aligned}$$

using the fact that π' is a Weierstrass fibration, and u is the restriction to $F(Y)$ of the projective bundle $\mathbb{P}(S) \to \mathbb{G}(1, m+1)$. Next, we compute

$$u_*(\mathcal{O}_U((k-2)\zeta) \otimes u^*\mathcal{O}_{F(Y)}(\sigma_1)) = \mathrm{Sym}^{k-2}(S^\vee) \otimes \mathcal{O}_{F(Y)}(\sigma_1) \qquad \square$$

The top Chern class of λ is enough to determine the constant term of $\varphi(q)$ when $k = 2$. Indeed:

Proposition 47. *For* $k = 2, 3, 4$, *the space of modular forms* $\mathrm{Mod}(6k-2, \mathrm{SL}_2(\mathbb{Z}))$ *has dimension* 1, 2, 2, *respectively.*

Proof. This follows from the presentation of the ring $\mathrm{Mod}(\bullet, \mathrm{SL}_2(\mathbb{Z}))$ as a free polynomial ring on the Eisenstein series E_4 and E_6. $\square$

The positive degree terms of $\varphi(q)$ are Noether–Lefschetz intersections. By the split sequence of Shioda–Tate, there are two sources of jumping Picard rank in the family of Hodge structures.

- Resolved singular surfaces have nontrivial $V_0(S)$. If e is the class of an exceptional curve, then $\langle f, z, e\rangle$ has intersection matrix
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -k & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$
- Surfaces with extra sections have nontrivial Mordell–Weil group $\mathrm{MW}(S/\mathbb{P}^1)$. If σ is the class of a section, then $\langle f, z, \sigma\rangle$ has intersection matrix
$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & -k & z\cdot\sigma \\ 1 & z\cdot\sigma & -k \end{pmatrix}.$$

By the shift in Lemma 21, the Mordell–Weil jumping starts to contribute at order q^k, which matches the shift in Lemma 23 for the rational curve class on X. Thus, terms of order $< k$ are determined by enumerating the singular surfaces.

Remark 48. Every vector in the sublattice $\mathrm{MW}(S/\mathbb{P}^1)$ corresponds to the class of some section curve. On the other hand, classes in $V_0(S)$ are less geometric: they are arbitrary $\mathbb{Z}$-linear combinations of exceptional curve classes, which are accounted for by the theta series $\Theta(q)$.

For $k=3$, we can compute the full generating series. The q^1 term is given by

$$[\varphi]_1 = \mathrm{j}_*[\mathfrak{F}(Y)]\cdot[\mathrm{NL}_2],$$

which is an excess intersection along $T_2(\Delta)\cap F(Y)$. If $v\in\Lambda$ lies in $Z^\perp$ with $v^2=-2$, then any integer multiple $mv\in\Lambda$ lies in $Z^\perp$ with $(mv)^2=-2m^2$, and the corresponding component of NL_{2m^2} meets $\mathfrak{F}(Y)$ with isomorphic normal cone. As a result this excess intersection contributes $\theta_1(q)$ to the generating series $\varphi(q)$. Next, the q^2 term is given by

$$[\varphi]_2 = \mathrm{j}_*[\mathfrak{F}(Y)]\cdot[\mathrm{NL}_4],$$

which is a 0-dimensional intersection along $T_{2,2}(\Delta)\cap F(Y)$. Since $\mathfrak{F}(Y)$ has isotropy group $\mathfrak{S}_2\times\mathfrak{S}_2$ there, we can compute

$$[\varphi]_2 = \tfrac{1}{4}t_{2,2}.$$

This completely determines the modular form $\varphi(q)$, so we can solve for $[\varphi]_1$. The only remaining singular surfaces which contribute are the 0-dimensional intersection along $T_3(\Delta)\cap F(Y)$. Since $\mathfrak{F}(Y)$ has isotropy group $\mathfrak{S}_3$ there, we have

$$\varphi(q) = \tfrac{1}{2}[\varphi]_1\theta_1(q) + \tfrac{1}{4}t_{2,2}\big(\theta_1(q)^2-2\theta_1(q)\big) + \tfrac{1}{6}t_3\big(\theta_2(q)-3\theta_1(q)\big) + \sum_{n\ge 3} r_X(n)q^n$$

For each successive singular stratum, we subtract the root lattice vectors which are limits of previous strata. There are 3 copies of A_1 in A_2, and there are 2 copies of A_1 in $A_1\times A_1$, so we subtract the double counted exceptional classes.

For $k=4$, there are too many undetermined excess intersections (denoted a_i) to determine $\varphi(q)$, but we still have the general form

$$\varphi(q) = \Theta(q) + \sum_{n\ge 4} r_X(n)q^n,$$

where the theta correction term is given by

$$\begin{aligned}\Theta(q) = a_1\theta_1(q) + a_2\big(\theta_1(q)^2-2\theta_1(q)\big) + a_3\big(\theta_2(q)-3\theta_1(q)\big) + t_4\big(\theta_3(q)-4\theta_2(q)-3\theta_1(q)^2+18\theta_1(q)\big)\\ + t_{2,2,2}\big(\theta_1(q)^3-3\theta_1(q)\big) + t_{2,3}\big(\theta_1(q)\theta_2(q)-4\theta_1(q)\big).\end{aligned}$$

We conclude with two examples of the full computation:

Example 49. $m=d=2$.

The quadric surface $Y\subset\mathbb{P}^3$ is isomorphic to $\mathbb{P}^1\times\mathbb{P}^1$, and the Weierstrass model X is a Calabi–Yau threefold sometimes called the STU model. Its curve counts were computed previously, and can be found in [Klemm et al. 2010].

$$\varphi(q) = -2E_4E_6;\quad \Theta(q) = -266+264\,\theta_1.$$

Example 50. $m = 2$, $d = 1$.

The hyperplane $Y \subset \mathbb{P}^3$ is isomorphic to $\mathbb{P}^2$, and the Weierstrass model X is a Calabi–Yau threefold, which can also be realized as the resolution of the weighted hypersurface

$$X_{18} \subset \mathbb{P}(1, 1, 1, 6, 9).$$

The topological string partition function of X is computed in [Huang et al. 2015], but our formula is the first in the mathematical literature.

$$\varphi(q) = \tfrac{31}{48} E_4^4 + \tfrac{113}{48} E_4 E_6^2; \quad \Theta(q) = 47253 - 93582\,\theta_1 + 46008\,\theta_1^2 + 324\,\theta_2.$$

Example 51. $m = 3$, $d = 3$.

The hypersurface $Y \subset \mathbb{P}^4$ is a cubic threefold, and the Weierstrass model Y is a (non-CY) fourfold.

$$\varphi(q) = \tfrac{433}{16} E_4^4 + \tfrac{1439}{16} E_4 E_6^2; \quad \Theta(q) = 2089215 - 4107510\,\theta_1 + 1969272\,\theta_1^2 + 49140\,\theta_2.$$

8. Future directions

Theorem 5 gives a generating series for true counts of smooth rational curves on X lying over lines in Y. One can ask how this relates to the generating series for genus 0 Gromov–Witten invariants, which virtually count nodal rational curves. When X is a threefold, the work of Oberdieck and Shen [2020] on stable pairs invariants of elliptic fibrations implies via the GW/pairs correspondence [Pandharipande and Pixton 2014] that

$$\sum_{n=0}^{\infty} \mathrm{GW}_0^X(\ell + nf) q^n = \varphi(q) \cdot \eta(q)^{-12k},$$

where $\eta(q)$ is the Dedekind eta function. We can recover this formula in the case $k = 2$, and for general k the $\Theta(q)$ correction term from Theorem 5 yields formulas for the local contributions of A_ρ singular surfaces $S \subset X$ to the stable pairs moduli space; compare with [Toda 2015].

Counting curves in base degree $e > 1$ presents challenges involving degenerations of Hodge structures. The analog of the family ν in this situation is

$$\begin{tikzcd} & \mathscr{S} \arrow[r] \arrow[dl, "\nu"'] \arrow[d, "\pi'"] & X \arrow[d, "\pi"] \\ \overline{M}_0(Y, e) & \mathcal{C} \arrow[l] \arrow[r] & Y \end{tikzcd}$$

where the Fano variety is replaced by the Kontsevich space of stable maps. At nodal curves

$$C_1 \cup C_2 \to Y,$$

the fiber of ν is a surface with normal crossings,

$$S = \pi^{-1}(C_1) \cup \pi^{-1}(C_2),$$

which does not have a pure Hodge structure. Thus, the associated period map

$$\overline{M}_0(Y, e) \dashrightarrow \Gamma\backslash D$$

does not extend over all of $\overline{M}_0(Y, e)$. Since the Noether–Lefschetz loci in $\Gamma\backslash D$ are noncompact, we must compactify the period map to a larger target in order to have a topological intersection product. Such a target $(\Gamma\backslash D)^*$ is provided in [Green et al. 2017], which satisfies the Borel Extension property for all period maps coming from algebraic families. The boundary of the partial completion $(\Gamma\backslash D)^*$ consists of products of period spaces of lower dimension, whose Noether–Lefschetz classes satisfy a modularity statement. Motivated by this observation and computations in the Hermitian symmetric case, we make a conjecture for higher base degrees.

Let X be a Weierstrass fibration in $W(Y, \mathcal{O}(k))$, where $Y \subset \mathbb{P}^{m+1}$ is a hypersurface of degree d, and

$$k = \left(\frac{e+1}{e}\right)m - d + \left(2 - \frac{2}{e}\right).$$

Note: $m \equiv 2 \pmod{e}$ is equivalent to integrality of this expression.

Conjecture 52. *Let $r_X(e, n)$ be the number of smooth rational curves on X in the homology class $e\ell + nf$. Then for $ke \leq 4$ or $m = 2$,*

$$\sum_{n\geq 1} r_X(e, n)q^n = \varphi(q) - \Theta(q),$$

where $\varphi(q) \in \mathrm{QMod}(\bullet, \mathrm{SL}_2(\mathbb{Z}))$, and $\Theta(q) \in \mathbb{Q}[\theta_1, \theta_3, \theta_3]_{<ke}$.

Recall that *quasimodular forms* are an enlargement of the algebra of modular forms to include the Eisenstein series E_2:

$$\mathrm{QMod}(\bullet, \mathrm{SL}_2(\mathbb{Z})) = \mathbb{Q}[E_2, E_4, E_6].$$

Appendix A: Torsion in the Mordell–Weil group

Let $S \to \mathbb{P}^1$ be a regular minimal elliptic surface. The vertical sublattice

$$V_0(S) \subset \mathrm{NS}_0(S)$$

is a direct sum of ADE root lattices, one for each singular fiber. The discriminant group d_i of each root lattice is the component group of the Néron model for the degeneration. A torsion element of $\mathrm{MW}(S/\mathbb{P}^1)$ restricts to a nonidentity component on some fiber, so we have an embedding of finite abelian groups

$$\mathrm{TMW}(S/\mathbb{P}^1) \hookrightarrow \bigoplus_i d_i$$

Furthermore, $\mathrm{TMW}(S/\mathbb{P}^1)$ is totally isotropic with respect to the quadratic form on the discriminant group. For the surfaces appearing in this paper (with $k \leq 4$), the discriminant group is sufficiently small that there are no nontrivial isotropic subgroups, so $\mathrm{MW}(S/\mathbb{P}^1)$ is torsion-free.

Appendix B: Configurations of points on a line

We study configurations of $6k$ unordered points on $\mathbb{P}^1$. The moduli space of point configurations is given by

$$M_{6k} := \mathbb{P}^{6k} / \mathrm{PGL}(2).$$

We will often ignore phenomena in codimension $\geq k$, since $\dim F(Y) = k-1$ and $F(Y) \subset \mathbb{G}(1, m+1)$ is freely movable. Away from codimension k, there are at least

$$6k - 2(k-1) = 4k+2$$

singleton points. In particular, all such configurations are GIT-stable. For a fixed general hypersurface $B \subset \mathbb{P}^{m+1}$ of degree $6k$, we have a morphism

$$\phi_B : \mathbb{G}(1, m+1) \to M_{6k},$$

given by intersecting with B. First we show that when m is large, ϕ_B has large rank.

Lemma 53. *If $2m \geq 4k$, then $d\phi$ has rank $\geq 2k$ at lines L meeting B in $\geq 4k+2$ reduced points.*

Proof. Consider the incidence correspondence

$$\Omega := \{(B, L) : \mathrm{rank}(d\phi_B) < 2k\} \subset \mathbb{P}^N \times \mathbb{G}(1, m+1).$$

The second projection $\Omega \to \mathbb{G}(1, m+1)$ is dominant, and we study the fiber of this morphism. Assume that B intersects L transversely at $q, o, p_1, p_2, \ldots, p_{4k}$. Pick coordinates on $\mathbb{P}^{m+1}$ such that $\mathbb{T}_q B \simeq \mathbb{P}^m$ is the hyperplane at infinity, o is the origin, and $T_o B$ is orthogonal to L. The remaining points are nonzero scalars, so we assume that $p_1 = 1$. Pick coordinates on $\mathbb{G}(1, m+1)$ near L by taking pencils based at $q \in L$ and $o \in L$ in a set of m general $\mathbb{P}^2$'s containing L. At each marked point p_i, the transverse tangent space $T_{p_i} B$ is given by some slope vector

$$\vec{\lambda}_i \in \mathbb{C}^m.$$

The $(4k-1)\times(2m)$ matrix for $d\phi$ restricted to these points is given by

$$d\phi = \begin{pmatrix} p_2\lambda_1^1 - \lambda_2^1 & \lambda_1^1 - \lambda_2^1 & p_2\lambda_1^2 - \lambda_2^2 & \lambda_1^2 - \lambda_2^2 & \cdots & p_2\lambda_1^m - \lambda_2^m & \lambda_1^m - \lambda_2^m \\ p_3\lambda_1^1 - \lambda_3^1 & \lambda_1^1 - \lambda_3^1 & p_3\lambda_1^2 - \lambda_3^2 & \lambda_1^2 - \lambda_3^2 & \cdots & p_3\lambda_1^m - \lambda_3^m & \lambda_1^m - \lambda_3^m \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{4k}\lambda_1^1 - \lambda_{4k}^1 & \lambda_1^1 - \lambda_{4k}^1 & p_{4k}\lambda_1^2 - \lambda_{4k}^2 & \lambda_1^2 - \lambda_{4k}^2 & \cdots & p_{4k}\lambda_1^m - \lambda_{4k}^m & \lambda_1^m - \lambda_{4k}^m \end{pmatrix}.$$

The coordinates of $\vec{\lambda}_i$ are free, and the small rank variety is cut out by the $(2k)\times(2k)$ minors of this matrix. This is a linear section of a determinantal variety, defined by a 1-generic matrix of linear forms in the sense of [Eisenbud 1988]. By the principal result of [Eisenbud 1988], it has the expected codimension

$$c = (2k)(2m - 2k + 1) > 2m.$$

Hence, the projection $\Omega \to \mathbb{P}^N$ is not dominant for dimension reasons. $\square$

To understand the tangent space to $F(Y)$ inside $\mathbb{G}(1, m+1)$, we use the short exact sequence of normal bundles

$$0 \to N_{L/Y} \to N_{L/\mathbb{P}^{m+1}} \to N_{Y/\mathbb{P}^{m+1}}|_L \to 0,$$
$$0 \to H^0(N_{L/Y}) \to H^0(N_{L/\mathbb{P}^{m+1}}) \to H^0(N_{Y/\mathbb{P}^{m+1}}).$$

The tangent space $T_{[L]}F(Y) \simeq H^0(N_{L/Y})$ can be identified with the kernel of

$$H^0(L, \mathcal{O}(1)^{\oplus m}) \to H^0(L, \mathcal{O}(d)).$$

This linear map can be understood as follows: for a set of m general $\mathbb{P}^2$'s containing L, consider $Y \cap \mathbb{P}^2_i = L \cup C_i$. The residual $C_i \cap L$ is a section of $g_i \in H^0(L, \mathcal{O}(d-1))$, and together they give the map. In coordinates, the matrix looks like

$$\begin{pmatrix} g_{11} & 0 & g_{21} & 0 & \cdots & 0 \\ g_{12} & g_{11} & g_{22} & g_{21} & \cdots & g_{m1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{1d} & g_{1(d-1)} & g_{2d} & g_{2(d-1)} & \cdots & g_{m(d-1)} \\ 0 & g_{1d} & 0 & g_{2d} & \cdots & g_{md} \end{pmatrix}.$$

Only the residuals matter when determining $H^0(N_{L/Y})$, and any set of residuals forms $g_i \in H^0(L, \mathcal{O}(d-1))$ comes from a hypersurface Y containing L.

Proposition 54. *When $2m \geq 4k$, the morphism*

$$\mu_X : F(Y) \to M_{6k}$$

is an immersion for general Y and $X \in W(Y, \mathcal{O}(k))$.

Proof. Consider the incidence correspondence

$$\Omega := \{(Y, L) : L \subset Y,\ d\mu_X \text{ is not an immersion}\} \subset \mathbb{P}^N \times \mathbb{G}(1, m).$$

The second projection $\Omega \to \mathbb{G}(1, m+1)$ is surjective, and we study the fiber of this morphism. The condition $Y \subset L$ is codimension $d+1$, so we need k more independent conditions for the codimension to exceed $2m$. For this, we use another incidence correspondence: let $K = \ker(d\phi_{[L]}) \subset \mathbb{C}^{2m}$ which has dimension $\leq 2m - 2k$ by Lemma 53, and consider

$$\Omega' := \{(M, w) : w \subset \ker(M)\} \subset \mathbb{P}^{md-1} \times \mathbb{P}K,$$

where $\mathbb{C}^{md} \subset \mathrm{Hom}(\mathbb{C}^{2m}, \mathbb{C}^{d+1})$ is the subspace of matrices which come from residual forms. It suffices to prove that $\pi_1(\Omega') \subset \mathbb{P}^{md-1}$ has codimension $\geq k$. The fibers of the second projection $\Omega' \to \mathbb{P}K$ are cut out by $d+1$ linear conditions in md variables. In terms of the coordinates of w, the conditions are

$$\begin{pmatrix} w_1 & 0 & 0 & 0 & \cdots & w_3 & 0 & 0 & 0 & \cdots \\ w_2 & w_1 & 0 & 0 & \cdots & w_4 & w_3 & 0 & 0 & \cdots \\ 0 & w_2 & w_1 & 0 & \cdots & 0 & w_4 & w_3 & 0 & \cdots \\ & & \ddots & \ddots & & & & \ddots & \ddots & \end{pmatrix}.$$

The degeneracy loci of this matrix are high codimension in $\mathbb{P}K$ by explicit calculation with minors, so we have the dimension count:

$$\begin{aligned}\dim \Omega' &= \dim \mathbb{P}K + (md-1) - (d+1)\\ &\le (2m-2k-1) + (md-1) - (2m-k+1),\\ \operatorname{codim} \pi_1(\Omega') &\ge (2m-k+1) - (2m-2k-1) = k+2.\end{aligned}$$

□

Proposition 55. *When $2m < 4k$, the morphism*

$$\mu_X : F(Y) \to \mathcal{W}_k$$

is an immersion for general Y and $X \in W(Y, \mathcal{O}(k))$.

Proof. Since $2m - 6k - 1 < 2m - 4k - 1 < 0$, general forms $A \in H^0(\mathbb{P}^{m+1}, 4k)$ and $B \in H^0(\mathbb{P}^{m+1})$ do not vanish on any line. Hence, we have a morphism

$$\phi_{A,B} : \mathbb{G}(1, m+1) \to M_{4k} \times M_{6k}$$

which we claim is an immersion on $F(Y)$. Since $F(Y) \subset \mathbb{G}(1, m)$ is freely movable, we may assume that each line intersects A and B at $\ge 8k+2$ reduced points. Consider the incidence correspondence

$$\Omega = \{([A : B], L) : d\mu_{[L]} \text{ is not injective}\} \subset W(\mathbb{P}^{m+1}, \mathcal{O}(k)) \times \mathbb{G}(1, m+1).$$

The fiber of the second projection $\Omega \to \mathbb{G}(1, m+1)$ is a linear section of a determinantal variety, as in Lemma 53. By [Eisenbud 1988], it has the expected codimension:

$$c = (8k+2) - 2m + 1 > 2m.$$

Hence, the projection $\Omega \to W(\mathbb{P}^{m+1}, \mathcal{O}(k))$ is not dominant for dimension reasons. □

Acknowledgements

The author is deeply indebted to his advisor Jun Li for guidance throughout the project, and to Jim Bryan for providing the initial question. He has also benefitted from conversations with Philip Engel, Tony Feng, Zhiyuan Li, Davesh Maulik, Georg Oberdieck, Arnav Tripathy, Ravi Vakil, and Abigail Ward.

References

[Artin 1974] M. Artin, "Algebraic construction of Brieskorn's resolutions", *J. Algebra* **29**:2 (1974), 330–348. MR Zbl

[Baily and Borel 1966] W. L. Baily, Jr. and A. Borel, "Compactification of arithmetic quotients of bounded symmetric domains", *Ann. of Math.* (2) **84**:3 (1966), 442–528. MR Zbl

[Behrend 2004] K. Behrend, "Cohomology of stacks", pp. 249–294 in *Intersection theory and moduli* (Trieste, Italy, 2002), edited by E. Arbarello et al., ICTP Lect. Notes **19**, Abdus Salam Int. Cent. Theoret. Phys., Trieste, Italy, 2004. MR Zbl

[Borcherds 1999] R. E. Borcherds, "The Gross–Kohnen–Zagier theorem in higher dimensions", *Duke Math. J.* **97**:2 (1999), 219–233. MR Zbl

[Borel 1969] A. Borel, *Introduction aux groupes arithmétiques*, Publ. Inst. Math. Univ. Strasbourg **15**, Hermann, Paris, 1969. MR Zbl

[Brieskorn 1968] E. Brieskorn, "Die Auflösung der rationalen Singularitäten holomorpher Abbildungen", *Math. Ann.* **178** (1968), 255–270. MR Zbl

[Cox 1990] D. A. Cox, "The Noether–Lefschetz locus of regular elliptic surfaces with section and $p_g \geq 2$", *Amer. J. Math.* **112**:2 (1990), 289–329. MR Zbl

[Eisenbud 1988] D. Eisenbud, "Linear sections of determinantal varieties", *Amer. J. Math.* **110**:3 (1988), 541–575. MR Zbl

[Eisenbud and Harris 2016] D. Eisenbud and J. Harris, *3264 and all that: a second course in algebraic geometry*, Cambridge Univ. Press, 2016. MR Zbl

[Green et al. 2017] M. Green, P. Griffiths, R. Laza, and C. Robles, "Completion of period mappings and ampleness of the Hodge bundle", preprint, 2017. arXiv

[Greer et al. 2015] F. Greer, Z. Li, and Z. Tian, "Picard groups on moduli of K3 surfaces with Mukai models", *Int. Math. Res. Not.* **2015**:16 (2015), 7238–7257. MR Zbl

[Huang et al. 2015] M.-x. Huang, S. Katz, and A. Klemm, "Topological string on elliptic CY 3-folds and the ring of Jacobi forms", *J. High Energy Phys.* **10** (2015), art. id. 125. MR Zbl

[Klemm et al. 2010] A. Klemm, D. Maulik, R. Pandharipande, and E. Scheidegger, "Noether–Lefschetz theory and the Yau–Zaslow conjecture", *J. Amer. Math. Soc.* **23**:4 (2010), 1013–1040. MR Zbl

[Kudla and Millson 1990] S. S. Kudla and J. J. Millson, "Intersection numbers of cycles on locally symmetric spaces and Fourier coefficients of holomorphic modular forms in several complex variables", *Inst. Hautes Études Sci. Publ. Math.* **71** (1990), 121–172. MR Zbl

[Maulik and Pandharipande 2013] D. Maulik and R. Pandharipande, "Gromov–Witten theory and Noether–Lefschetz theory", pp. 469–507 in *A celebration of algebraic geometry* (Cambridge, 2011), edited by B. Hassett et al., Clay Math. Proc. **18**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Miranda 1981] R. Miranda, "The moduli of Weierstrass fibrations over $\mathbb{P}^1$", *Math. Ann.* **255**:3 (1981), 379–394. MR Zbl

[Miranda 1989] R. Miranda, *The basic theory of elliptic surfaces*, Ph.D. thesis, Universita di Pisa, 1989, Available at https://tinyurl.com/rickmirphd.

[Nikulin 1979] V. V. Nikulin, "Integral symmetric bilinear forms and some of their applications", *Izv. Akad. Nauk SSSR Ser. Mat.* **43**:1 (1979), 111–177. In Russian; translated in *Math. USSR-Izv.* **14**:1 (1980), 103–167. MR Zbl

[Oberdieck and Pixton 2019] G. Oberdieck and A. Pixton, "Gromov–Witten theory of elliptic fibrations: Jacobi forms and holomorphic anomaly equations", *Geom. Topol.* **23**:3 (2019), 1415–1489. MR Zbl

[Oberdieck and Shen 2020] G. Oberdieck and J. Shen, "Curve counting on elliptic Calabi–Yau threefolds via derived categories", *J. Eur. Math. Soc.* **22**:3 (2020), 967–1002. MR Zbl

[Pandharipande and Pixton 2014] R. Pandharipande and A. Pixton, "Gromov–Witten/pairs descendent correspondence for toric 3-folds", *Geom. Topol.* **18**:5 (2014), 2747–2821. MR Zbl

[Saito 1983] M.-H. Saito, "On the infinitesimal Torelli problem of elliptic surfaces", *J. Math. Kyoto Univ.* **23**:3 (1983), 441–460. MR Zbl

[Schmid 1973] W. Schmid, "Variation of Hodge structure: the singularities of the period mapping", *Invent. Math.* **22** (1973), 211–319. MR Zbl

[Toda 2015] Y. Toda, "S-duality for surfaces with A_n-type singularities", *Math. Ann.* **363**:1-2 (2015), 679–699. MR Zbl

[Vistoli 1989] A. Vistoli, "Intersection theory on algebraic stacks and on their moduli spaces", *Invent. Math.* **97**:3 (1989), 613–670. MR Zbl

Communicated by Richard Borcherds
Received 2019-01-31 Revised 2020-03-27 Accepted 2020-04-29

fgreer@scgp.stonybrook.edu *Simons Center for Geometry and Physics, Stony Brook University, Stony Brook, NY, United States*

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2369

Quadratic Chabauty for (bi)elliptic curves and Kim's conjecture

Francesca Bianchi

We explore a number of problems related to the quadratic Chabauty method for determining integral points on hyperbolic curves. We remove the assumption of semistability in the description of the quadratic Chabauty sets $\mathcal{X}(\mathbb{Z}_p)_2$ containing the integral points $\mathcal{X}(\mathbb{Z})$ of an elliptic curve of rank at most 1. Motivated by a conjecture of Kim, we then investigate theoretically and computationally the set-theoretic difference $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$. We also consider some algorithmic questions arising from Balakrishnan and Dogra's explicit quadratic Chabauty for the rational points of a genus-two bielliptic curve. As an example, we provide a new solution to a problem of Diophantus which was first solved by Wetherell.

Computationally, the main difference from the previous approach to quadratic Chabauty is the use of the p-adic sigma function in place of a double Coleman integral.

1. Introduction

Let (E, O) be an elliptic curve over $\mathbb{Q}$ and fix an odd prime p of good reduction. Denote by $\mathcal{E}$ the minimal regular model of E and by $\mathcal{X}$ the complement of the origin in $\mathcal{E}$.

When E has Mordell–Weil rank 1 and the Tamagawa number of $E/\mathbb{Q}$ is trivial at all primes, Kim [2010a] described an explicit locally analytic function on $\mathcal{X}(\mathbb{Z}_p)$ which vanishes on the set $\mathcal{X}(\mathbb{Z})$ of global integral points.

Subsequently, Balakrishnan, Dan-Cohen, Kim and Wewers [Balakrishnan et al. 2018] generalised the result to arbitrary semistable elliptic curves of rank 1 and gave a similar p-adic characterisation of $\mathcal{X}(\mathbb{Z})$ when E is semistable and has rank 0.

The discussion fits into Kim's nonabelian Chabauty programme as introduced in [Kim 2005; 2009]. In particular, Kim constructed a sequence of subsets of p-adic points

$$\mathcal{X}(\mathbb{Z}_p) \supset \mathcal{X}(\mathbb{Z}_p)_1 \supset \mathcal{X}(\mathbb{Z}_p)_2 \supset \cdots \supset \mathcal{X}(\mathbb{Z}).$$

The p-adic locally analytic functions from [Balakrishnan et al. 2018] are essentially those that define $\mathcal{X}(\mathbb{Z}_p)_2$, the set of *cohomologically global points of level* 2, in the larger $\mathcal{X}(\mathbb{Z}_p)$.

The subscript n in $\mathcal{X}(\mathbb{Z}_p)_n$ indicates a particular quotient U_n of the unipotent p-adic étale fundamental group U of $\mathcal{X}_{\overline{\mathbb{Q}}}$ (at a tangential base point). The set $\mathcal{X}(\mathbb{Z}_p)_n$ is then defined in terms of certain "unipotent Kummer maps" from $\mathcal{X}(\mathbb{Z})$ and $\mathcal{X}(\mathbb{Z}_q)$, at every prime q, to global and local cohomology sets with

MSC2010: primary 11D45; secondary 11G50, 11Y50, 14H52.
Keywords: quadratic Chabauty, p-adic heights, integral points on hyperbolic curves.

U_n-coefficients, respectively, in a way that generalises to objects with nonabelian étale fundamental group the role played by $\mathbb{Q}_p$-Selmer groups in our understanding of rational points on abelian varieties.

Despite its abstract cohomological definition, the set $\mathcal{X}(\mathbb{Z}_p)_n$ is believed to be computable in practice [Balakrishnan et al. 2018] as a union of intersections of zero loci of locally analytic functions defined in terms of iterated p-adic integrals. Unfortunately, such a characterisation is yet to be provided for $n \geq 3$.

Nevertheless, the explicit description of $\mathcal{X}(\mathbb{Z}_p)_2$ in the rank 0 semistable case given in [Balakrishnan et al. 2018] was already sufficient to collect some computational evidence for the following special case of a conjecture of Kim (see [Balakrishnan et al. 2018, §3.1]).

Conjecture 1.1 (Kim, 2012). *For sufficiently large n, we have*

$$\mathcal{X}(\mathbb{Z}_p)_n = \mathcal{X}(\mathbb{Z}).$$

Indeed, the authors of [loc. cit.] verified the equality

$$\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z})$$

for the prime $p = 5$ and for all the 256 semistable elliptic curves of rank 0 for which they computed $\mathcal{X}(\mathbb{Z}_p)_2$. An additional test that was performed by the same authors was that of fixing $\mathcal{X}$ and varying the prime p: once again, no point in $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ was found. No other study of the difference $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ appears in the literature, hence motivating the following two questions:

Question 1.2. Does there exist any elliptic curve of rank 0 for which $\mathcal{X}(\mathbb{Z}_p)_2$ contains at least one point which is not in $\mathcal{X}(\mathbb{Z})$?

Question 1.3. What geometric or algebraic properties should a point in $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ satisfy?

One goal of the present paper is to give answers to these questions, with the idea that elliptic curves should serve as a test case for a conjecture that is in fact formulated by Kim in much greater generality than how we stated it here, and that as such would have striking applications if it were to hold. Indeed, $\mathcal{X}$ could be replaced by a suitable $\mathbb{Z}$-model $\mathcal{C}$ of any hyperbolic curve over $\mathbb{Q}$ with good reduction at p. In particular, the conjecture would give an effective approach towards finding the set of rational points on a curve of genus $g \geq 2$.

In the elliptic curve case, the conjecture might not have direct Diophantine interest, in the sense that there already exist algorithms for the computation of integral points on elliptic curves [Smart 1994; Pethő et al. 1999; Stroeker and Tzanakis 1994], and the rank 0 and 1 instances which we will explore are particularly well understood. However, the known explicit versions of nonabelian Chabauty for curves of higher genus (cf. [Balakrishnan et al. 2016; 2019a; Balakrishnan and Dogra 2018]) all generalise the explicit description of $\mathcal{X}(\mathbb{Z}_p)_2$ for elliptic curves of rank at most 1. Therefore, a conceptual understanding of the zero sets of the p-adic equations defining $\mathcal{X}(\mathbb{Z}_p)_2$ is essential to hope to achieve something similar in more complicated settings.

In general, even finiteness of $\mathcal{C}(\mathbb{Z}_p)_n$ for n large enough is only conjectural (but see [Kim 2010b; Coates and Kim 2010; Ellenberg and Hast 2017; Balakrishnan and Dogra 2018] for results in this direction, and

[Kim 2009] for a proof assuming the Bloch–Kato conjecture). If $g \geq 2$ and, for a given n, $\mathcal{C}(\mathbb{Z}_p)_n$ is finite and explicitly computable to arbitrary p-adic precision, then the Mordell–Weil sieve could be used to try to provably extrapolate $\mathcal{C}(\mathbb{Z})$ from $\mathcal{C}(\mathbb{Z}_p)_n$. However, the Mordell–Weil sieve is not guaranteed to terminate. Thus, finiteness of $\mathcal{C}(\mathbb{Z}_p)_n$ would not be sufficient to imply an effective version of Faltings's theorem.

Suppose now that $\mathcal{C}(\mathbb{Z}_p)_n$ is finite for n sufficiently large. One reason for expecting that the inclusion $\mathcal{C}(\mathbb{Z}) \subset \mathcal{C}(\mathbb{Z}_p)_n$ should eventually become sharp is explained in [Balakrishnan et al. 2018, §1.8]: assuming some well known motivic conjectures, the number of algebraically independent locally analytic functions vanishing on $\mathcal{C}(\mathbb{Z}_p)_n$ is strictly increasing in n (for $n \gg 0$). See also [Balakrishnan et al. 2018, §1, §3.4] for the philosophy behind Conjecture 1.1 (in its general form) and for its relationship with the conjectural finiteness of the Tate–Shafarevich group and with Grothendieck's section conjecture.

For an elliptic curve of rank 1, finiteness of $\mathcal{X}(\mathbb{Z}_p)_n$ can only hold at level $n \geq 2$. On the other hand, for a rank 0 elliptic curve, $\mathcal{X}(\mathbb{Z}_p)_1$ is finite and there are two independent equations defining $\mathcal{X}(\mathbb{Z}_p)_2$ (see Theorem 1.6 below), hence justifying why Question 1.2 had proved itself arduous. We show, however, that the answer to the question is negative. More precisely, we prove the following two theorems (see also Theorem 4.10).

Theorem 1.4. *There exist infinitely many rank* 0 *elliptic curves for which*

$$\mathcal{X}(\mathbb{Z}) \subsetneq \mathcal{X}(\mathbb{Z}_p)_2$$

for infinitely many good primes p.

Theorem 1.5. *There exists exactly one rank* 0 *elliptic curve of conductor at most* 30000 *for which*

$$\mathcal{X}(\mathbb{Z}) \subsetneq \mathcal{X}(\mathbb{Z}_p)_2$$

for all primes p of good (*ordinary and supersingular*) *reduction. This is the curve 8712.u5* [LMFDB 2019].

When we analyse these results in conjunction with Question 1.3, it will become apparent that they should not be considered as negative evidence for Conjecture 1.1. We will return below to discussing Theorems 1.4, 1.5 and answers to Question 1.3 in the context of the methods we develop in order to prove them. For the reader's convenience, let us first digress to write down the equations defining $\mathcal{X}(\mathbb{Z}_p)_2$. In fact, the very first goal of this article is to extend the explicit description of $\mathcal{X}(\mathbb{Z}_p)_2$ to an arbitrary elliptic curve of rank 0 and at the same time correct a slight imprecision in the analogous statement in the semistable case [Balakrishnan et al. 2018, Theorem 1.12] (see Remark 2.6).

Before stating the theorem, we introduce some additional notation, which is convenient to maintain similar to [Balakrishnan et al. 2018]. Let $\mathcal{E}$ be described by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{1}$$

and let S be the set of primes at which E has bad reduction.

For each $q \in S$, define the set $W_q \subset \mathbb{Q}_p$ as follows. If the Tamagawa number at q is 1, let

$$W_q = \{0\};$$

in all other cases,

$$W_q = \begin{cases} W_q^{\mathrm{bad}} & \text{if } q = 2 \text{ and } E \text{ is split multiplicative at } q, \\ W_q^{\mathrm{bad}} \cup \{0\} & \text{otherwise,} \end{cases}$$

where W_q^{bad} is the finite subset of $\mathbb{Q}_p$ described in Table 1 (with $F = \mathbb{Q}$ and $v = (q_v) = (q)$); in particular, W_q^{bad} only depends on the reduction type of E at q. Let

$$W = \prod_{q \in S} W_q$$

and, if $w \in W$, write $\|w\| = \sum_{q \in S} w_q$. Let b be the integral tangent vector at the origin which is dual to $\omega(O)$, where

$$\omega = \frac{dx}{2y + a_1 x + a_3}$$

and let $\eta = x\omega$. Furthermore, for $z \in \mathcal{X}(\mathbb{Z}_p)$ write

$$\mathrm{Log}(z) = \int_b^z \omega \quad \text{and} \quad D_2(z) = \int_b^z \omega\eta,$$

where the integrals are Coleman integrals.

Theorem 1.6. *Suppose that E has rank 0 and the p-primary part of the Tate–Shafarevich group is finite.*

(1) *If, for at least one of $q \in \{2, 3\}$, the reduction of E at q is good and $\bar{E}(\mathbb{F}_q) = \{O\}$, or if E has split multiplicative reduction of Kodaira type I_1 at 2, then*

$$\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z}) = \varnothing.$$

(2) *Otherwise,*

$$\mathcal{X}(\mathbb{Z}_p)_2 = \bigcup_{w \in W} \phi(w),$$

where

$$\phi(w) = \{z \in \mathcal{X}(\mathbb{Z}_p) : \mathrm{Log}(z) = 0, 2D_2(z) + \|w\| = 0\}.$$

We also remove the assumption of semistable reduction in the rank 1 case [Balakrishnan et al. 2018, Proposition 5.12]. Assume E has good ordinary[1] reduction at p. Let $\boldsymbol{E}_2$ be the Katz p-adic weight 2 Eisenstein series [Katz 1976] and let

$$C = \frac{a_1^2 + 4a_2 - \boldsymbol{E}_2(E, \omega)}{12}. \tag{2}$$

[1] Although not explicitly stated in [Balakrishnan et al. 2018], their statement also holds only when p is ordinary. However, a similar result holds in the supersingular case; see Remark 2.8.

Let $h_p : E(\mathbb{Q}) \to \mathbb{Q}_p$ be $(-2p)$ times the p-adic height of [Mazur et al. 2006] and define

$$c = \frac{h_p(z_0)}{\mathrm{Log}(z_0)^2}$$

for a nontorsion point $z_0 \in E(\mathbb{Q})$.

Theorem 1.7. *Suppose that E has rank 1 and that p is a prime of good ordinary reduction.*

(1) *If, for at least one of $q \in \{2, 3\}$, the reduction of E at q is good and $\bar{E}(\mathbb{F}_q) = \{O\}$, or if E has split multiplicative reduction of Kodaira type I_1 at 2, then*

$$\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z}) = \varnothing.$$

(2) *Otherwise,*

$$\mathcal{X}(\mathbb{Z}) \subset \mathcal{X}(\mathbb{Z}_p)_2' := \bigcup_{w \in W} \psi(w),$$

where

$$\psi(w) = \{z \in \mathcal{X}(\mathbb{Z}_p) : 2D_2(z) + C(\mathrm{Log}(z))^2 + \|w\| = c(\mathrm{Log}(z))^2\}.$$

According to [Balakrishnan et al. 2018], the set $\bigcup_{w \in W} \psi(w)$ should equal $\mathcal{X}(\mathbb{Z}_p)_2$: hence the notation $\mathcal{X}(\mathbb{Z}_p)_2'$. Section 2 is devoted to the proofs of Theorems 1.6 and 1.7.

The equations defining the sets of p-adic points of the two theorems can be given an elementary interpretation as linear relations amongst $\mathbb{Q}_p$-valued quadratic functions on $E(\mathbb{Q})$, dictated by the assumptions on the rank. More precisely, any global p-adic height $E(\mathbb{Q}) \to \mathbb{Q}_p$ (of Bernardi, Coleman–Gross, Mazur–Tate) vanishes identically if the rank is 0, and is a scalar multiple of $\mathrm{Log}^2|_{E(\mathbb{Q})}$ if the rank is 1. To go from here to a p-adic approximation of the global integral points, one invokes the decomposition of the p-adic height on $E(\mathbb{Q}) \setminus \{O\}$ as a sum, over the nonarchimedean primes q, of local p-adic heights $\lambda_q : E(\mathbb{Q}_q) \setminus \{O\} \to \mathbb{Q}_p$. Indeed, the restriction of λ_q to $\mathcal{X}(\mathbb{Z}_q) \supset \mathcal{X}(\mathbb{Z})$ has finite image for all $q \neq p$, zero image for almost all $q \neq p$, and is locally analytic for $q = p$.

This point of view is crucial in our investigation, in Section 3, of what points could arise in $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ in rank 0. Recall that no example of an elliptic curve of rank 0 and a prime p for which $\mathcal{X}(\mathbb{Z}_p)_2 \supsetneq \mathcal{X}(\mathbb{Z})$ was previously known. A careful study of the Mazur–Tate and Bernardi local p-adic heights allows us to deduce possible obstructions to the sharpness of $\mathcal{X}(\mathbb{Z}_p)_2$, and to give necessary and sufficient conditions for a point in $\mathcal{X}(\mathbb{Z}_p) \setminus \mathcal{X}(\mathbb{Z})$ to belong to $\mathcal{X}(\mathbb{Z}_p)_2$.

First note that a point in $\mathcal{X}(\mathbb{Z}_p)_2$ is algebraic, since it is in the zero set of the abelian logarithm Log. Our sufficient conditions then come from studying how automorphisms of $E/\overline{\mathbb{Q}}$ affect the values of the local p-adic heights at certain algebraic points, and from an analysis of noncyclotomic p-adic heights over nontotally real number fields. A combination of these two phenomena explains the appearance of extra points at level 2 in the family of quadratic twists of the modular curve $X_0(49)$ (see Proposition 3.14). As an application, in Section 3C we prove Theorem 1.4.

As regards necessary conditions for $\mathcal{X}(\mathbb{Z}_p)_2$ to contain parasite points, we prove a sort of "p-adic height saturation" condition (see the discussion in Section 3D):

Theorem 1.8 (Theorem 3.18). *Let $E/\mathbb{Q}$ and p be as in Theorem 1.6. Suppose that $z \in \mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$. Then z is the localisation of a torsion point P over a number field K and, for each rational prime q, the value $\lambda_{\mathfrak{q}}(P)$ of the local height at $\mathfrak{q}$ of the cyclotomic p-adic height of E/K is independent of the prime $\mathfrak{q} \mid q$ of K.*

We then present in Section 4 the computations of the sets $\mathcal{X}(\mathbb{Z}_p)_2$ for all the elliptic curves over $\mathbb{Q}$ of rank 0 and conductor less than or equal to 30000 and for some choices of p. We propose a slightly different but equivalent way of computing the set $\mathcal{X}(\mathbb{Z}_p)_2$, compared to the one used in [Balakrishnan et al. 2018]. In particular, our method does not rely on general algorithms to compute double Coleman integrals, but rather uses Bernardi's and Mazur and Tate's description of the p-adic height on an elliptic curve to express the double Coleman integrals in terms of p-adic sigma functions.

Our computations (run on `SageMath` [2017–2019]) suggest that the failure of sharpness of $\mathcal{X}(\mathbb{Z}_p)_2$ is still to be considered a rare phenomenon, which we were always able to explain using the sufficient conditions of Section 3. Extra points become even more exceptional if we allow the prime p to vary. In particular, we prove Theorem 1.5 (see Theorem 4.10).

In future work, it would be interesting to verify whether Conjecture 1.1 holds at level 3 for the curves and primes for which we found $\#\mathcal{X}(\mathbb{Z}_p)_2 > \#\mathcal{X}(\mathbb{Z})$.

When E has rank 1, the set described in Theorem 1.7(2) is generally larger than $\mathcal{X}(\mathbb{Z})$. Naively, this is because $\mathcal{X}(\mathbb{Z}_p)_2'$ is cut out by the vanishing of one function only. In Section 5A, we ask what algebraic points can belong to $\mathcal{X}(\mathbb{Z}_p)_2' \setminus \mathcal{X}(\mathbb{Z})$. In Section 5B we compute $\mathcal{X}(\mathbb{Z}_p)_2'$ for all the 14783 rank 1 elliptic curves of conductor at most 5000; for each curve, we let p be the smallest prime greater than or equal to 5 at which the curve has good ordinary reduction.

Finally, in Section 6 we apply some of our techniques for elliptic curves to the computation of rational points on certain genus 2 curves C over $\mathbb{Q}$. Indeed, when C admits degree 2 maps to two elliptic curves over $\mathbb{Q}$, each of rank 1, Balakrishnan and Dogra [2018] described a $\mathbb{Q}_p$-valued locally analytic function on $C(\mathbb{Q}_p)$ vanishing on $C(\mathbb{Q})$. This is defined using local p-adic heights on each elliptic curve. We explain how one can replace direct computations of double Coleman integrals with computations involving the p-adic sigma function and division polynomials also in this situation. We make the computation explicit for a curve arising from a problem from the *Arithmetica* of Diophantus and use it to give an alternative proof to the one given by Wetherell in his thesis [1997] of the fact that the curve has exactly 8 rational points.

Balakrishnan and Dogra implemented their method numerically in some examples. However, algorithmically, their approach was still based on a case-by-case study. This consideration applies especially to a preliminary step which consists in the computation of two finite subsets of $\mathbb{Q}_p$ (which play the role of the set $\|W\|$ above). By combining results of [Balakrishnan and Dogra 2018] with properties of local p-adic heights on elliptic curves, in Section 6A we offer a more general and applicable numerical approach to the method. For example, we give an algorithm that takes as input a bielliptic curve C whose associated elliptic curves have Mordell–Weil rank 1 together with a good prime p and outputs a finite set of p-adic points containing $C(\mathbb{Q})$ (i.e., we remove the preliminary computation step). In doing so, we also provide a more elementary proof and approach to the explicit result of [Balakrishnan and Dogra 2018].

The code used for the computations in this article is available at [Bianchi 2019].

2. Description of $\mathcal{X}(\mathbb{Z}_p)_2$

2A. ***The p-adic height and its local components.*** Let p be an odd prime and extend the usual p-adic logarithm $\log : 1 + p\mathbb{Z}_p \to \mathbb{Q}_p$ to $\mathbb{Q}_p^\times$ via $\log(p) = 0$.

Let E be an elliptic curve over $\mathbb{Q}$ as in Section 1 and assume E has good reduction at p. We will sometimes need to consider the base-change of E to a number field F (whose ring of integers is denoted $\mathcal{O}_F$); thus, we do not restrict the following definitions to $E(\mathbb{Q})$. Let

$$h_p : E(F) \to \mathbb{Q}_p$$

be a cyclotomic[2] p-adic height of Coleman–Gross (see [Coleman and Gross 1989] and [Balakrishnan and Besser 2015]). The use of the indefinite article here is due to the dependence of h_p on a choice that will be made explicit in Section 2A2. The function h_p is quadratic, i.e., satisfies the relation

$$h_p(mP) = m^2 h_p(P) \quad \text{for all } m \in \mathbb{Z} \text{ and } P \in E(F),$$

and is defined as a sum of local heights, one for each nonarchimedean prime of F. In particular, we have[3] $h_p(O) = 0$ and for $P \in E(F) \setminus \{O\}$

$$h_p(P) = \frac{1}{[F : \mathbb{Q}]} \sum_v n_v \lambda_v(P),$$

where the sum is over all finite primes v of F, $n_v = [F_v : \mathbb{Q}_v]$ and

$$\lambda_v : E(F_v) \setminus \{O\} \to \mathbb{Q}_p$$

is a *p-adic local Néron function at* v.

Let q_v be the norm of v and $|\cdot|_v$ be the normalised absolute value corresponding to v. That is, if $x \in F_v^\times$, we have

$$|x|_v = q_v^{-\mathrm{ord}_v(x)/n_v},$$

where the valuation ord_v is such that $\mathrm{ord}_v(F_v^\times) = \mathbb{Z}$.

[2] If the space of continuous idele class characters $\mathbb{A}_F^\times / F^\times \to \mathbb{Q}_p$ has dimension larger than 1, we will see in Section 3B that one can define other types of p-adic heights.

[3] Here we choose to normalise the p-adic height in such a way that it becomes independent of the choice of the field F containing the coordinates of P; note that this is not the case in many other articles, such as [Mazur et al. 2006].

2A1. The p-adic local Néron function at a nonarchimedean prime $v \nmid p$ is equal to the real local Néron function $\widehat{\lambda}_v$ at v with the p-adic logarithm in place of the real one. Thus, any reference that we provide for $\widehat{\lambda}_v$ can be applied also to our setting. For instance, analogously to the real case, for $v \nmid p$, the following properties determine a unique function $\lambda_v : E(F_v) \setminus \{O\} \to \mathbb{Q}_p$:

(i) λ_v is continuous on $E(F_v) \setminus \{O\}$ and bounded on the complement of any neighbourhood of O with respect to the v-adic topology.

(ii) $\lim_{P \to O}(\lambda_v(P) - \log |x(P)|_v)$ exists.

(iii) λ_v satisfies the *quasiparallelogram law*: for all $P, Q \in E(F_v)$ such that $P, Q, P \pm Q \neq O$, we have

$$\lambda_v(P+Q) + \lambda_v(P-Q) = 2\lambda_v(P) + 2\lambda_v(Q) - 2\log|x(P) - x(Q)|_v. \tag{3}$$

Uniqueness follows from topological reasons. For existence, it suffices to show that the p-adic analogue of $\widehat{\lambda}_v$ obtained as described above satisfies (i)–(iii) (see [Silverman 1994, VI, Exercise 6.3]).

We also have

(iv) For all $P \in E(F_v)$ and all $m \geq 1$ with $mP \neq O$,

$$\lambda_v(mP) = m^2\lambda_v(P) - 2\log|f_m(P)|_v,$$

where f_m is the m-th division polynomial of E (see for instance [Silverman 2009, III, Exercise 3.7] for the definition of f_m). We say that λ_v is *quasiquadratic*.

Moreover, uniqueness implies the following key fact:

(v) If ψ is an automorphism of E defined over F_v, then $\lambda_v(\psi(P)) = \lambda_v(P)$ for all $P \in E(F_v)$.

See also [Bernardi 1981] for a more general transformation property under isogeny.

We wish to determine which values λ_v can attain on $\mathcal{X}(\mathcal{O}_v)$, where $\mathcal{O}_v$ is the ring of integers of F_v and $v \nmid p$. For this, it will be convenient to assume that $\mathcal{E}$ is minimal at v. If that is not the case, we can always switch to a minimal equation at v and use the following (see [Cremona et al. 2006, Lemma 4]):

$$\lambda_v = \lambda_v^{\min} + \tfrac{1}{6}\log|\Delta/\Delta^{\min}|_v, \tag{4}$$

where Δ denotes the discriminant and the superscript min has the obvious meaning. See also Remark 3.3.

So assume for the rest of this subsection that λ_v is computed with respect to a minimal model at the prime v. Denote by $\bar{E}_{ns}(\mathbb{F}_v)$ the group of nonsingular points of the reduction of E modulo v and let

$$E_0(F_v) = \{P \in E(F_v) : \bar{P} \in \bar{E}_{ns}(\mathbb{F}_v)\}.$$

If E has good reduction at v, we may also write $\bar{E}(\mathbb{F}_v)$ for $\bar{E}_{ns}(\mathbb{F}_v)$.

Lemma 2.1. *Suppose $v \nmid p$. Then*

(i) *if $P \in E_0(F_v) \setminus \{O\}$, $\lambda_v(P) = \log(\max\{1, |x(P)|_v\})$;*

(ii) *if $P \notin E_0(F_v)$, $\lambda_v(P)$ depends exclusively on the image of P in $E(F_v)/E_0(F_v)$.*

Proof. See [Silverman 1988] or [Cremona et al. 2006, Proposition 5]. ☐

Proposition 2.2. *If $v \nmid p$ and $[E(F_v) : E_0(F_v)] = 1$, then*

$$\lambda_v(\mathcal{X}(\mathcal{O}_v)) = \begin{cases} \{0\} & \textit{if } \#\bar{E}_{ns}(\mathbb{F}_v) > 1 \\ \varnothing & \textit{otherwise.} \end{cases}$$

Proof. By Lemma 2.1(i), $\lambda_v(\mathcal{X}(\mathcal{O}_v)) \subseteq \{0\}$, with equality if and only if $\mathcal{X}(\mathcal{O}_v) \neq \varnothing$. Let

$$E_1(F_v) = \{P \in E(F_v) : \bar{P} = \bar{O}\};$$

in particular, $\mathcal{X}(\mathcal{O}_v) \cap E_1(F_v)$ is empty and every point in $E_0(F_v) \setminus E_1(F_v)$ comes from a point in $\mathcal{X}(\mathcal{O}_v)$. According to [Silverman 2009, VII, Proposition 2.1], the sequence

$$0 \to E_1(F_v) \to E_0(F_v) \to \bar{E}_{ns}(\mathbb{F}_v) \to 0$$

is exact, which proves the proposition. □

We now give an elementary necessary condition for $\#\bar{E}_{ns}(\mathbb{F}_v) = 1$. We show it is also a sufficient condition in all cases except when v is of good reduction.

Lemma 2.3. *The group $\bar{E}_{ns}(\mathbb{F}_v)$ has cardinality at least 2 in all of the following cases*:

(1) *E has additive or nonsplit multiplicative reduction at v.*

(2) *E has good reduction at v and $q_v > 4$.*

(3) *E has split multiplicative reduction at v and $q_v > 2$.*

Conversely, if E has split multiplicative reduction at v and $q_v = 2$, then

$$\#\bar{E}_{ns}(\mathbb{F}_v) = 1.$$

Proof. If E has additive reduction at v, then $\bar{E}_{ns}(\mathbb{F}_v) \cong \mathbb{F}_v^+$ always contains at least two elements. If the reduction is nonsplit multiplicative, then

$$\bar{E}_{ns}(\mathbb{F}_v) \cong \{a \in k^\times : N_{k/\mathbb{F}_v}(a) = 1\},$$

where $[k : \mathbb{F}_v] = 2$ and $N_{k/\mathbb{F}_v}$ is the field norm of $k/\mathbb{F}_v$. Thus, if $q_v > 2$, then the statement is clear; if $q_v = 2$, then k is the splitting field of $x^4 - x$ over $\mathbb{F}_2$ and each element in $k^\times$ has norm 1 over $\mathbb{F}_2$.

When E has good reduction at v and $q_v > 4$, the Hasse bound yields $\#\bar{E}(\mathbb{F}_v) > 1$. Finally, if the reduction is split multiplicative, then $\bar{E}_{ns}(\mathbb{F}_v) \cong \mathbb{F}_v^\times$. □

Proposition 2.4. *If $v \nmid p$ and $[E(F_v) : E_0(F_v)] \neq 1$, then*

$$n_v\lambda_v(\mathcal{X}(\mathcal{O}_v)) = \begin{cases} W_v^{\mathrm{bad}} & \textit{if } q_v = 2 \textit{ and } E \textit{ is split multiplicative at } v, \\ W_v^{\mathrm{bad}} \cup \{0\} & \textit{otherwise,} \end{cases}$$

where W_v^{bad} is defined in Table 1.

Proof. By a similar argument to the proof of Proposition 2.2, each non trivial coset of $E(F_v)/E_0(F_v)$ is represented by an element in $\mathcal{X}(\mathcal{O}_v)$ and there exists at least one point in $\mathcal{X}(\mathcal{O}_v)$ which reduces to a nonsingular point in $\bar{E}_{ns}(\mathbb{F}_v)$ if and only if $\#\bar{E}_{ns}(\mathbb{F}_v) > 1$.

Kodaira symbol	$[E(F_v):E_0(F_v)]$	W_v^{bad}
$\mathrm{I}_n (n \geq 2)$	2 (nonsplit)	$\left\{-\frac{n}{4}\log q_v\right\}$
	n (split)	$\left\{-\frac{i(n-i)}{n}\log q_v : 1 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor\right\}$
III	2	$\left\{-\frac{1}{2}\log q_v\right\}$
IV	3	$\left\{-\frac{2}{3}\log q_v\right\}$
I_0^*	2 or 4	$\{-\log q_v\}$
$\mathrm{I}_n^* (n \geq 1)$	2	$\{-\log q_v\}$
	4	$\left\{-\log q_v, -\frac{n+4}{4}\log q_v\right\}$
IV^*	3	$\left\{-\frac{4}{3}\log q_v\right\}$
III^*	2	$\left\{-\frac{3}{2}\log q_v\right\}$

Table 1. The sets W_v^{bad}.

By Lemma 2.1(i), if $P \in \mathcal{X}(\mathcal{O}_v)$ reduces to a nonsingular point, then $\lambda_v(P) = 0$; by Lemma 2.3, such P exists unless $q_v = 2$ and E is split multiplicative at 2.

Therefore, by Lemma 2.1(ii), it suffices to show that W_v^{bad} coincides exactly with the values of $n_v\lambda_v$ on $E(F_v)/E_0(F_v) \setminus \{0\}$. For this, we use the work of Cremona, Prickett and Siksek [Cremona et al. 2006] for the local heights of the real canonical height. The proof of [loc. cit., Proposition 6] can be used verbatim here with the p-adic logarithm in place of the real one and Table 1 is nothing but the translation of [Cremona et al. 2006, Table 2] to the p-adic setting. □

2A2. The p-adic local Néron function at a prime $v \mid p$ is not unique: it depends on a choice of subspace $N_v \subset H^1_{\text{dR}}(E/F_v)$ complementary to the space of holomorphic differentials (see [Coleman and Gross 1989]). Let ξ_v be the one-form of the second kind with a double pole at O and no others, representative of the class in N_v dual to ω with respect to the cup product (i.e., such that $[\omega] \cup [\xi_v] = 1$). Let $\mathrm{tr}_{F_v/\mathbb{Q}_p}$ denote the field trace. Then by [Balakrishnan and Besser 2015, Theorem 4.1], for all $P \in E(F_v) \setminus \{O\}$ one has

$$\lambda_v(P) = \frac{1}{n_v}\,\mathrm{tr}_{F_v/\mathbb{Q}_p}\left(2\int_b^P \omega\xi_v\right).$$

In particular,

$$\xi_v = \eta + \gamma\omega \quad \text{for some } \gamma \in F_v$$

and hence

$$\lambda_v(P) = \frac{1}{n_v}\,\mathrm{tr}_{F_v/\mathbb{Q}_p}(2D_2(P) + \gamma\,\mathrm{Log}(P)^2).$$

In [Balakrishnan and Besser 2015, Corollary 3.2], it is shown that if E has good ordinary reduction at v and N_v is the unit root eigenspace of Frobenius, then λ_v is related to the logarithm of the v-adic sigma function of Mazur and Tate [1991].

In fact, it is easy to see that their proof shows the following stronger result.

Proposition 2.5. *Let $x(t)$ be the Laurent series expansion of x in terms of the parameter for the formal group $t = -x/y$. Let $\sigma_v^{(\gamma)}(t) = t + \cdots \in F_v[[t]]$ be the unique odd*[4] *function satisfying*

$$x(t) + \gamma = -\frac{d}{\omega}\left(\frac{1}{\sigma_v^{(\gamma)}}\frac{d\sigma_v^{(\gamma)}}{\omega}\right)$$

and let V be a neighbourhood of O on which $\sigma_v^{(\gamma)}$ converges. Then, for all $P \in V \setminus \{O\}$, we have

$$\lambda_v(P) = -\frac{2}{n_v}\operatorname{tr}_{F_v/\mathbb{Q}_p}(\log_v(\sigma_v^{(\gamma)}(P))),$$

where $\log_v : F_v^\times \to F_v$ *extends* log.

For our applications, we may assume that there is an isomorphism $F_v \simeq \mathbb{Q}_p$, which is now fixed. Since λ_v is not unique, we will use the following convention. If the reduction is good *ordinary* at each prime v above p, we choose N_v to be the unit root eigenspace of Frobenius, i.e.,

$$\gamma = C,$$

where C is defined in (2). If P belongs to the formal group at v, then Proposition 2.5 says that

$$\lambda_v(P) = -2\log(\sigma_p(P)),$$

where σ_p is the Mazur–Tate p-adic sigma function. Furthermore, in this case the global p-adic height coincides with the p-adic height of Mazur–Tate.

If E is good *supersingular* at some prime $v \mid p$, we let, for each $v \mid p$,

$$\gamma = \frac{a_1^2 + 4a_2}{12},$$

so that $\sigma_p^{(\gamma)}$ is the p-adic sigma function of Bernardi [1981]. This choice of γ gives a power series $\sigma_v^{(\gamma)}(t)$ with coefficients in F (and in fact $\mathbb{Q}$ in our case), which is related to the Taylor expansion $\sigma(z)$ of the complex Weierstrass sigma function by the change of variables $z = L_v(t)$, where L_v is the formal group logarithm. Unlike the p-adic sigma function of Mazur and Tate, the one of Bernardi does not converge on the whole formal group over $\overline{F_v}$, as it may not have p-adically integral coefficients, as a power series in t. However, since we are assuming that $F_v \simeq \mathbb{Q}_p$, the function $\sigma_p^{(\gamma)}$ converges on all the points P of the formal group whose coordinates are defined over F_v, since these satisfy $\operatorname{ord}_v(t(P)) > 1/(p-1)$.

In both the ordinary and supersingular cases, λ_v satisfies (see [Coleman and Gross 1989; Mazur and Tate 1991; Bernardi 1981]):

[4] Odd as a function on a subset of the formal group and not as a function of t.

(i) λ_v is locally analytic on $\mathcal{X}(\mathcal{O}_v)$.

(ii) For all $P \in E(F_v)$ and all $m \geq 1$ with $mP \neq O$,

$$\lambda_v(mP) = m^2\lambda_v(P) - \frac{2}{n_v}\operatorname{tr}_{F_v/\mathbb{Q}_p}(\log_v(f_m(P))).$$

(iii) For all P, $Q \in E(F_v)$ such that P, Q, $P \pm Q \neq O$,

$$\lambda_v(P+Q) + \lambda_v(P-Q) = 2\lambda_v(P) + 2\lambda_v(Q) - \frac{2}{n_v}\operatorname{tr}_{F_v/\mathbb{Q}_p}(\log_v(x(P) - x(Q))).$$

(iv) If ψ is an automorphism of E defined over F_v, then $\lambda_v(\psi(P)) = \lambda_v(P)$ for all $P \in E(F_v)$. In view of the assumption that $F_v \simeq \mathbb{Q}_p$ and by Deuring's criterion, at supersingular primes this is simply saying that λ_v is an even function. At ordinary primes, let ζ be the root of unity such that $\psi^*(\omega) = \zeta\omega$. Then $f_m(\psi(P)) = \zeta^{1-m^2} f_m(P)$ by [Mazur and Tate 1991, Appendix I, Proposition 2], and the Mazur–Tate p-adic sigma function satisfies $\sigma_p(\psi(P)) = \zeta\sigma_p(P)$ if P is in the formal group [Mazur and Tate 1991, §3]. The claim then follows since $\log(\zeta) = 0$. Note that invariance under any automorphism would also hold if we used the Bernardi sigma function to define the local heights at ordinary primes, since for curves of j-invariant 0 or 1728 the weight 2 Eisenstein series vanishes, so the Bernardi and Mazur–Tate p-adic sigma function are equal.

We also remark that if L/F is a finite field extension, w is a prime of L above v, where v is any prime of F, and $P \in E(F_v)$, then

$$\lambda_v(P) = \lambda_w(P).$$

2B. ***Proof of Theorems 1.6 and 1.7.*** It would be pointless to reproduce here the whole proofs, as they are straightforward from Section 2A and the proofs in [Balakrishnan et al. 2018]. Thus we content ourselves with giving a sketch and correcting a few imprecisions in [loc. cit., Theorem 1.12].

We start with some notation and we refer the reader to [Kim 2005; 2009; Balakrishnan et al. 2018] for more details. Let $T = S \cup \{p\}$ and denote by G_T the Galois group of the maximal extension of $\mathbb{Q}$ unramified outside T. For a prime q, write G_q for the absolute Galois group of $\mathbb{Q}_q$. For $q \in T$, G_q may be identified with a subgroup of G_T. For $q \notin T$, this is not possible; however, we may still define maps $G_q \to G_T$ which are trivial on the inertia subgroup $I_q \leq G_q$.

Let U be the unipotent p-adic étale fundamental group of $\mathcal{X}_{\overline{\mathbb{Q}}}$ at b and U_n the quotient of U by the n-th level of its central series.

For each prime q and $n \geq 1$, we have commutative diagrams

$$\begin{tikzcd} \mathcal{X}(\mathbb{Z}) \arrow[r] \arrow[d] & \mathcal{X}(\mathbb{Z}_q) \arrow[d, "j_q^n"] \\ H^1_f(G_T, U_n) \arrow[r, "\mathrm{loc}_q^n"] & H^1(G_q, U_n) \end{tikzcd}$$

Here the H^1 are cohomology sets and $H^1_f(G_T, U_n) = (\mathrm{loc}^n_p)^{-1}(H^1_f(G_p, U_n))$, where $H^1_f(G_p, U_n)$ is the subset of $H^1(G_p, U_n)$ of crystalline U_n-torsors. We are interested in determining

$$\mathcal{X}(\mathbb{Z}_p)_n = (j^n_p)^{-1}(\mathrm{loc}^n_p(\mathrm{Sel}^n(\mathcal{X}))),$$

where the Selmer scheme $\mathrm{Sel}^n(\mathcal{X})$ is defined as

$$\mathrm{Sel}^n(\mathcal{X}) = \bigcap_{q \neq p} (\mathrm{loc}^n_q)^{-1}(\mathrm{Im}\, j^n_q).$$

From now on, we will focus on $n = 2$ and will drop the superscript n from the maps j_q and loc_q.

Proof of Theorem 1.6. If $\mathcal{X}(\mathbb{Z}_q)$ is empty for some q, then $\mathcal{X}(\mathbb{Z}_p)_2$ is trivially empty. Lemma 2.3 shows that this occurs precisely when E has good reduction at q, where $q = 2$ or 3, and $\bar{E}(\mathbb{F}_q) = \{O\}$, or when E has split multiplicative reduction of type I_1 at $q = 2$. This shows (1).

We may now suppose that $\mathcal{X}(\mathbb{Z}_q) \neq \varnothing$ for all q (including $q = p$). Since $E(\mathbb{Q})$ has rank 0 and the p-primary part of the Tate–Shafarevich group is finite, by Lemma 5.2 in [Balakrishnan et al. 2018],

$$\mathrm{Sel}^2(\mathcal{X}) \subset H^1_f(G_T, \mathbb{Q}_p(1)),$$

where $H^1_f(G_T, \mathbb{Q}_p(1)) \subset H^1_f(G_T, U_2)$ via $0 \to \mathbb{Q}_p(1) \to U_2 \to V_p(E) \to 0$.

For $q \neq p$, we have

$$j_q : \mathcal{X}(\mathbb{Z}_q) \to H^1(G_q, U_2) \simeq H^1(G_q, \mathbb{Q}_p(1)) \simeq \mathbb{Q}_p,$$

the last map being $c \mapsto \log\chi \cup c \in H^2(G_q, \mathbb{Q}_p(1)) \simeq \mathbb{Q}_p$, where χ is the p-adic cyclotomic character and $\cup$ is the cup product (note $\log\chi \in H^1(G_q, \mathbb{Q}_p)$). The middle bijection is proved in [Balakrishnan et al. 2018, §4.1.5]. Thus, finding the image of j_q is equivalent to finding

$$\{\phi_q(z) := \log\chi \cup j_q(z) : z \in \mathcal{X}(\mathbb{Z}_q)\}.$$

Theorem 4.1.6 in [Balakrishnan et al. 2018] shows that

$$2\phi_q : \mathcal{X}(\mathbb{Z}_q) \to \mathbb{Q}_q, \qquad z \mapsto 2(\log\chi \cup j_q(z))$$

is the restriction to $\mathcal{X}(\mathbb{Z}_q)$ of a p-adic local Néron function in the sense of Section 2A1 and must thus be equal to the function λ_q.

In particular, for each $q \neq p$, the set $2\phi_q(\mathcal{X}(\mathbb{Z}_q))$ is the finite set described by Propositions 2.2 and 2.4.

The cup product $\log\chi \cup c$, for $c \in H^1_f(G_p, \mathbb{Q}_p(1))$, and local reciprocity also yield an isomorphism

$$\psi_p : H^1_f(G_p, \mathbb{Q}_p(1)) \xrightarrow{\sim} H^2(G_p, \mathbb{Q}_p(1)) \simeq \mathbb{Q}_p$$

and we get a commutative diagram

$$\begin{tikzcd}
H^1_f(G_T, \mathbb{Q}_p(1)) \arrow[r, hook, "\bigoplus_{q \in T} \mathrm{loc}_q"] \arrow[d, "g = \log\chi \cup \cdot"] & H^1_f(G_p, \mathbb{Q}_p(1)) \oplus \bigoplus_{q \in S} H^1(G_q, \mathbb{Q}_p(1)) \arrow[d, "\wr"] \\
H^2(G_T, \mathbb{Q}_p(1)) \arrow[r, hook, "\bigoplus_{q \in T} \mathrm{loc}_q"] & \bigoplus_{q \in T} H^2(G_q, \mathbb{Q}_p(1)) \simeq \bigoplus_{q \in T} \mathbb{Q}_p
\end{tikzcd}$$

On the other hand, by global class field theory and Hilbert's Theorem 90, the image of $H^2(G_T, \mathbb{Q}_p(1))$ in the bottom row is the kernel of the map

$$\bigoplus_{q \in T} \mathbb{Q}_p \to \mathbb{Q}_p, \qquad (a_q) \to \sum_q a_q$$

and by dimension considerations, one concludes that the map g is in fact also an isomorphism.

From above we know that the image of $\bigoplus_{q\in S} j_q(\mathcal{X}(\mathbb{Z}_q))$ in $\bigoplus_{q \in S} \mathbb{Q}_p$ is precisely $\left(\frac{1}{2}\right) \bigoplus_{q\in S} W_q$, where

- if $[E(\mathbb{Q}_q) : E_0(\mathbb{Q}_q)] = 1$, then $W_q = \{0\}$;
- if $[E(\mathbb{Q}_q) : E_0(\mathbb{Q}_q)] \neq 1$, then

$$W_q = \begin{cases} W_q^{\text{bad}} & \text{if } q = 2 \text{ and } E \text{ is split multiplicative at } q, \\ W_q^{\text{bad}} \cup \{0\} & \text{otherwise} \end{cases}$$

  and W_q^{bad} is defined in Proposition 2.4.

Let $W = \prod_{q \in S} W_q$. It follows from the above that for every $w = (w_q)_{q\in S} \in W$ there exists a unique $c \in H^1_f(G_T, \mathbb{Q}_p(1))$ with $2(\log\chi \cup \mathrm{loc}_q(c)) = w_q$ for every $q \in S$. Further, this satisfies $2\psi_p(\mathrm{loc}_p(c)) = -\|w\|$. On the other hand, if $q \notin T$ and $c \in H^1_f(G_T, \mathbb{Q}_p(1))$ then $\mathrm{loc}_q(c) = 0$ and $\mathrm{Im}(j_q) = \{0\}$ by Proposition 2.2. Therefore,

$$\mathrm{Sel}^2(\mathcal{X}) = \bigcap_{q\in S} \mathrm{loc}_q^{-1}(\mathrm{Im}\, j_q)$$

and

$$\mathrm{loc}_p(\mathrm{Sel}^2(\mathcal{X})) = \bigcup_{w\in W} \{c \in H^1_f(G_p, \mathbb{Q}_p(1)) : 2\psi_p(c) + \|w\| = 0\}.$$

It remains to compute the preimage of this set under j_p. As is shown in [Balakrishnan et al. 2018, §5.7], we find

$$\mathcal{X}(\mathbb{Z}_p)_2 = \{z \in \mathcal{X}(\mathbb{Z}_p) : \mathrm{Log}(z) = 0, 2D_2(z) + \|w\| = 0\};$$

indeed, the condition $\mathrm{Log}(z) = 0$ is equivalent to requiring that

$$j_p(z) \in H^1_f(G_p, \mathbb{Q}_p(1)) \subset H^1_f(G_p, U_2)$$

and the other condition comes from the explicit formula

$$\psi_p(j_p(z)) = D_2(z) \quad \text{for } j_p(z) \in H^1_f(G_p, \mathbb{Q}_p(1)). \qquad \square$$

Remark 2.6. The corrections to the proof in [Balakrishnan et al. 2018] made here are the following. First of all, if $\mathcal{X}(\mathbb{Z}_q)$ is empty for some q, the proof does not hold. Of course this is a trivial case (treated in (1)), but it is not clear that the union given in [loc. cit., Theorem 1.12] should be empty. In fact, in Example 4.2 we find a curve satisfying the hypotheses of Theorem 1.6(1), but for which $\bigcup_{w\in W} \phi(w) \neq \varnothing$.

Secondly, if the reduction type at q is nonsplit multiplicative of type I_m, with $m > 2$, not all the values in their sets W_q will be attained by a point in $E(\mathbb{Q}_q) \setminus E_0(\mathbb{Q}_q)$. Therefore, if a prime in S is nonsplit multiplicative, their statement should just be an inclusion of $\mathcal{X}(\mathbb{Z}_p)_2$ into the union of the $\Psi(w)$. One should note, however, that it seems like this was taken care of in the computations when the Tamagawa number at q is 1, but not when it is 2 (and hence m is even).

For the same reasons, if $q = 2$ is a prime of split multiplicative reduction of type I_m, with $m > 0$, the element 0 should not be included in W_q.

We remark that in all the examples they provided the set they computed turned out to be equal to $\mathcal{X}(\mathbb{Z})$ and hence to $\mathcal{X}(\mathbb{Z}_p)_2$.

Remark 2.7. The proof of [Balakrishnan et al. 2018, Theorem 1.12] is rather technical. However, for an elliptic curve of any rank, denoting by $\mathcal{X}(\mathbb{Z})_{\mathrm{tors}}$ the set of points of $\mathcal{X}(\mathbb{Z})$ of finite order, the easier statement

$$\mathcal{X}(\mathbb{Z})_{\mathrm{tors}} \subseteq \bigcup_{w \in W} \{z \in \mathcal{X}(\mathbb{Z}_p) : \mathrm{Log}(z) = 0, 2D_2(z) + \|w\| = 0\} \tag{5}$$

is elementary to prove. Indeed, the condition $\mathrm{Log}(z) = 0$ cuts out the torsion points in $\mathcal{X}(\mathbb{Z}_p)$. On the other hand, let

$$\gamma = \begin{cases} C & \text{if } E \text{ is ordinary at } p, \\ \frac{1}{12}(a_1^2 + 4a_2) & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{cases} \mathrm{Log}(z) = 0, \\ 2D_2(z) + \|w\| = 0 \end{cases} \iff \begin{cases} \mathrm{Log}(z) = 0 \\ 2D_2(z) + \gamma\,\mathrm{Log}(z)^2 + \|w\| = 0 \end{cases} \iff \begin{cases} \mathrm{Log}(z) = 0 \\ \lambda_p(z) + \|w\| = 0 \end{cases}$$

and, for $z \in \mathcal{X}(\mathbb{Z})$, $h_p(z) = \lambda_p(z) + \|w\|$ for some $w \in W$, where h_p and λ_p are the global and local p-adic heights of Section 2A. In particular, if $z \in \mathcal{X}(\mathbb{Z})_{\mathrm{tors}}$, we have $h_p(z) = 0$. In fact, we could have also obtained a height function by setting the local height at p to be the dilogarithm $2D_2(z)$.

Proof of Theorem 1.7. The proof of part (1) is identical to the proof of Theorem 1.6(1). The proof of part (2) is straightforward from Section 2B and the proof of [Balakrishnan et al. 2018, Proposition 5.12]: the idea is that any two quadratic functions on the rank-one $E(\mathbb{Q})$ must be linearly dependent. Note that in the semistable case our statement is slightly different, as our set W is smaller if there are primes of nonsplit multiplicative reduction of type I_m, with $m > 2$ and also if $q = 2$ is a prime of split multiplicative reduction (cf. Remark 2.6). □

Remark 2.8. Theorem 1.7 is a consequence of the quadraticity of the p-adic height and of the square of the elliptic curve logarithm. Of course, that the latter function is quadratic follows from the linearity of the logarithm. We remark that Log^2 is in fact the p-adic height attached to the basis element ω of the Dieudonné module of E, in the language of generalised p-adic heights (see for instance [Stein and Wuthrich 2013, §4]), whereas the p-adic height h_p of Mazur–Tate is the one attached to an eigenvector with unit eigenvalue under the action of Frobenius. We could remove the assumption that p is ordinary in the statement of Theorem 1.7 if we replaced C with $\frac{1}{12}(a_1^2 + 4a_2)$ and let h_p be the global p-adic height that we defined in Section 2A when p is supersingular.

3. Obstructions to $\mathcal{X}(\mathbb{Z}) = \mathcal{X}(\mathbb{Z}_p)_2$ in rank 0.

We now derive some criteria for $\mathcal{X}(\mathbb{Z}_p)_2 \supsetneq \mathcal{X}(\mathbb{Z})$. In Section 4, we will compute $\mathcal{X}(\mathbb{Z}_p)_2$ for several curves and provide explicit examples for the results of this section. Since a necessary condition for $z \in \mathcal{X}(\mathbb{Z}_p)_2$ is that $\mathrm{Log}(z) = 0$, which can only occur if $z \in \mathcal{X}(\mathbb{Z}_p)_{\mathrm{tors}}$, after having fixed all appropriate embeddings, we must have

$$\mathcal{X}(\mathbb{Z}_p)_2 \subset \mathcal{E}(\overline{\mathbb{Z}})_{\mathrm{tors}} = E(\overline{\mathbb{Q}})_{\mathrm{tors}}.$$

In Section 3D we derive a stronger necessary condition, which roughly says that if a point lies in $\mathcal{X}(\mathbb{Z}_p)_2$, then its local heights cannot distinguish it from a point defined over $\mathbb{Q}$. To motivate the intuition behind this, it is more natural to first investigate sufficient conditions. In particular, we consider two reasons why extra points could arise in $\mathcal{X}(\mathbb{Z}_p)_2$: invariance of local heights under automorphism (Section 3A) and existence of noncyclotomic local heights over certain number fields (Section 3B). Sometimes, a combination of the two is needed, as is the case in Proposition 3.14, which provides us with infinitely many curves over $\mathbb{Q}$ with points over a quartic field appearing in $\mathcal{X}(\mathbb{Z}_p)_2$ for suitable choices of p. In Section 3C, we use this to deduce Theorem 1.4.

We start by proving an elementary fact: any obstruction to $\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z})$ must come from points defined over number fields larger than $\mathbb{Q}$.

Proposition 3.1. *Suppose that E satisfies the assumptions of Theorem 1.6(2) and that p is an odd prime of good reduction. Then*

$$\mathcal{X}(\mathbb{Z}_p)_2 \cap \mathcal{E}(\mathbb{Z}) = \mathcal{X}(\mathbb{Z}).$$

Proof. Suppose $P \in \mathcal{X}(\mathbb{Z}_p)_2 \cap \mathcal{E}(\mathbb{Z})$. In particular, $\mathrm{Log}(P) = 0$, so P is torsion and hence $h_p(P) = 0$. On the other hand, since $P \in \mathcal{X}(\mathbb{Z}_p)_2$,

$$\lambda_p(P) + \|w\| = 0$$

for some $w \in W$. Therefore,

$$\sum_{q \neq p} \lambda_q(P) = \|w\|.$$

By definition, we have

$$\sum_{q \neq p} \lambda_q(z) = \sum_{q \neq p} \alpha_q \log q, \qquad \|w\| = \sum_{q \neq p} \beta_q \log q$$

for some $\alpha_q, \beta_q \in \mathbb{Q}$, $\alpha_q = 0$ for all but finitely many q, $\beta_q = 0$ for all $q \notin S$ and $\beta_q \leq 0$ for all q. Thus

$$\log\Bigg(\prod_{q \neq p} q^{d(\alpha_q - \beta_q)}\Bigg) = 0,$$

for some nonzero integer d such that $d(\alpha_q - \beta_q) \in \mathbb{Z}$ for all q. This implies that $\alpha_q = \beta_q$ for all q, since the kernel of the p-adic logarithm is the subgroup of $\mathbb{Q}_p^\times$ generated by p and by the roots of unity. Suppose that z is not integral at q. Then by Lemma 2.1(i), $\alpha_q > 0$, but $\beta_q \leq 0$, a contradiction. $\square$

Remark 3.2. According to [Silverman 2009, VII Application 3.5], if $P \in \mathcal{E}(\mathbb{Z})_{\mathrm{tors}}$ then P is integral at all primes except possibly at 2 if P is 2-torsion. Thus the only q for which the proof of Proposition 3.1 is nonempty is $q = 2$. However, note that, with minor changes, the same proof shows the perhaps less trivial fact that $\mathcal{X}(\mathbb{Z}_p)_2' \cap \mathcal{E}(\mathbb{Z}) = \mathcal{X}(\mathbb{Z})$ in rank 1.

Remark 3.3. Unlike in Section 2A, given a prime v of a number field, henceforth the notation λ_v will be used for the local height at v computed with respect to the model $\mathcal{E}$, which may not be minimal at v. The translation with the values computed with respect to a minimal model (Lemma 2.1 and Proposition 2.4) is given by (4).

3A. ***Automorphisms.*** Recall that local heights are even functions. Therefore, if K is a quadratic field with $\mathrm{Gal}(K/\mathbb{Q}) = \langle \tau \rangle$ and $z \in \mathcal{X}(\mathcal{O}_K)$ satisfies $\tau(z) = -z$, then

$$h_p(z) = \sum_q \lambda_{\mathfrak{q}}(z) = \lambda_{\mathfrak{p}}(z) + \sum_{q \in S} \lambda_{\mathfrak{q}}(z),$$

where $\mathfrak{q}$ (resp. $\mathfrak{p}$) is any prime of K above q (resp. p). Intuitively, in terms of local p-adic heights, the point z behaves as if it were defined over $\mathbb{Q}$; if furthermore z is a torsion point and p is split in K, then z will give rise to a point in $\mathcal{X}(\mathbb{Z}_p)_2$, provided that $\lambda_{\mathfrak{q}}(z) \in W_q$ at every $q \in S$. If the j-invariant of E is different from 0 and 1728, the automorphism group of $E/\overline{\mathbb{Q}}$ is generated by $z \mapsto -z$. On the other hand, if $j(E) \in \{0, 1728\}$, we can use the invariance of our local heights under any automorphism (cf. property (v) in Section 2A1 and property (iv) in Section 2A2) to generalise the above example as follows.

Proposition 3.4. *Suppose that E satisfies the assumptions of Theorem 1.6 and that p is an odd prime of good reduction. Let K be a Galois extension of $\mathbb{Q}$, such that there is an embedding $\rho : K \hookrightarrow \mathbb{Q}_p$. Extend ρ to a map $\mathcal{E}(\mathcal{O}_K) \hookrightarrow \mathcal{E}(\mathbb{Z}_p)$. Let $z \in \mathcal{X}(\mathcal{O}_K)_{\mathrm{tors}}$ and suppose that for every $\tau \in \mathrm{Gal}(K/\mathbb{Q})$ there exists $\psi_\tau \in \mathrm{Aut}(E/\overline{\mathbb{Q}})$ such that $\tau(z) = \psi_\tau(z)$.*

(1) *For each rational prime q, let $\mathfrak{q}$ be one (any) prime of K above q and let $\lambda_{\mathfrak{q}}$ be the local height at $\mathfrak{q}$ with respect to the model $\mathcal{E}$. If*

$$\sum_{q \in S} \lambda_{\mathfrak{q}}(z) = \|w\|$$

for some $w \in W$, then $\rho(z) \in \mathcal{X}(\mathbb{Z}_p)_2$.

(2) *In particular, if $z = \psi'(P)$ for some $\psi' \in \mathrm{Aut}(E/\overline{\mathbb{Q}})$ and some $P \in \mathcal{X}(\mathbb{Z})$, then $\rho(z) \in \mathcal{X}(\mathbb{Z}_p)_2$.*

Proof. The assumption that z is a torsion point implies that $h_p(z) = 0$ and $\mathrm{Log}(z) = 0$. Since for $\tau \in \mathrm{Gal}(K/\mathbb{Q})$ there is an automorphism ψ_τ of E which acts on z in the same way as τ and local heights are invariant under automorphisms, for each prime $\mathfrak{q}$ of K we have

$$\lambda_{\mathfrak{q}}(z) = \lambda_{\mathfrak{q}}(\psi_\tau(z)) = \lambda_{\mathfrak{q}}(\tau(z)) = \lambda_{\tau^{-1}(\mathfrak{q})}(z).$$

Therefore,

$$0 = [K : \mathbb{Q}] h_p(z) = [K : \mathbb{Q}] \biggl(\lambda_p(\rho(z)) + \sum_{q \in S} \lambda_{\mathfrak{q}}(z) \biggr)$$

and (1) follows. For (2), since $z = \psi'(P)$, we have, similarly to above,

$$\lambda_{\mathfrak{q}}(z) = \lambda_{\mathfrak{q}}(\psi'(P)) = \lambda_q(P).$$

In particular, the hypothesis of (1) is satisfied. $\square$

We now list a few consequences of Proposition 3.4. See Section 4 for explicit examples.

Corollary 3.5. *Suppose that E satisfies the assumptions of Theorem 1.6 and that p is an odd prime of good ordinary reduction. Suppose that*

$$\mathcal{E} : y^2 + a_3 y = x^3 + a_6 \quad \textit{for some } a_6 \in \mathbb{Z} \textit{ and } a_3 \in \{0, 1\},$$

and that there exists $y_0 \in \mathbb{Z}$ such that $a_6 - y_0^2 - a_3 y_0$ is a cube in $\mathbb{Z}$ and the points over $\overline{\mathbb{Q}}$ with y-coordinate equal to y_0 have finite order. Then $s(x) = x^3 + a_6 - y_0^2 - a_3 y_0$ splits completely in $\mathbb{Q}_p$ and for each root $\alpha \in \mathbb{Z}_p$ of $s(x)$, $\pm(\alpha, y_0) \in \mathcal{X}(\mathbb{Z}_p)_2$.

Remark 3.6. If in the corollary we have $a_3 = 1$, then $E(\mathbb{Q})_{\mathrm{tors}}$ is isomorphic to a subgroup of $\mathbb{Z}/3\mathbb{Z}$, since $E(\mathbb{Q})[2] = \{O\}$ and E has good reduction at 2 with $\#\bar{E}(\mathbb{F}_2) = 3$. By looking at the third division polynomial for E, it is then straightforward to check that Corollary 3.5 applies nontrivially only if $4a_6 = -(27n^6 + 1)$ for some $n \in \mathbb{Z}$, $n \equiv 1 \bmod 2$. All such curves are isomorphic over $\mathbb{Q}$ to the elliptic curve 27.a3 [LMFDB 2019]. When $a_3 = 0$, there are infinitely many curves nonisomorphic over $\mathbb{Q}$ for which the corollary applies with $y_0 = 0$: see for example Section 3C. There is also at least one curve for which the Corollary applies to points of order 6, namely 36.a4 (see Table 2).

Proof. Since E has vanishing j-invariant, its automorphism group $\mathrm{Aut}(E/\overline{\mathbb{Q}})$ is a cyclic group of order 6 generated by $\psi : E \to E$, $\psi(x, y) = (\zeta x, -y - a_3)$, for a primitive third root of unity ζ.

Let $x_0 \in \mathbb{Z}$ such that $x_0^3 = y_0^2 + a_3 y_0 - a_6$. We may assume that $y_0^2 + a_3 y_0 - a_6$ is nonzero, as otherwise the statement of the corollary is trivial. Thus $s(x)$ has three distinct roots x_0, ζx_0 and $\zeta^2 x_0$ in $\overline{\mathbb{Z}}$.

Note also that, by Deuring's criterion [Lang 1973, Chapter 13, Theorem 12], the primes of good ordinary reduction for E split completely in $\mathbb{Q}(\zeta)$, so $s(x)$ splits completely over $\mathbb{Q}_p$. Successively applying ψ to $(x_0, y_0) \in \mathcal{X}(\mathbb{Z})$ and localising at p we obtain all points of the form $\pm(\alpha, y_0)$. The corollary then follows from Proposition 3.4(2). $\square$

The following corollary to Proposition 3.4 is a special case of the motivating example of the beginning of this subsection.

Corollary 3.7. *Suppose that E satisfies the assumptions of Theorem 1.6 and that p is an odd prime of good reduction. Let K be a quadratic field, in which p splits. Fix an embedding $\rho : K \hookrightarrow \mathbb{Q}_p$ and let τ be the nontrivial element in* $\mathrm{Gal}(K/\mathbb{Q})$. *Assume that no prime in S ramifies in K and that, if $q \in S$ is inert, then either E has Kodaira symbol I_0^* at q with Tamagawa number at least 2 or E has maximal Tamagawa number for its Kodaira symbol. Then*

$$\mathcal{X}(\mathbb{Z}_p)_2 \supset \mathcal{X}(\mathbb{Z}) \cup \{\rho(z) \in \mathcal{X}(\mathbb{Z}_p) : z \in \mathcal{X}(\mathcal{O}_K)_{\mathrm{tors}},\ \tau(z) = -z\}.$$

Proof. Since no prime in S ramifies in $K/\mathbb{Q}$, Tate's algorithm [Silverman 1994, IV, §9] shows that the equation for $\mathcal{E}$ defines a global minimal model for the base change E/K and that the Kodaira symbol at $\mathfrak{q} \mid q$ is the same as the Kodaira symbol at q. The Tamagawa number does not change if q is split in K; if q is inert, by assumption the Tamagawa number is unvaried, except possibly if the Kodaira symbol is I_0^*.

If q splits in K, fix a prime $\mathfrak{q}$ above it and an isomorphism $\rho_q : K_{\mathfrak{q}} \simeq \mathbb{Q}_q$. Let $z \in \mathcal{X}(\mathcal{O}_K)_{\mathrm{tors}}$ such that $\tau(z) = -z$. With the notation as in Proposition 3.4 and by Proposition 2.4, we have

$$\sum_{q\in S} \lambda_{\mathfrak{q}}(z) = \sum_{\substack{q\in S\\ q \text{ split}}} \lambda_q(\rho_q(z)) + \sum_{\substack{q\in S\\ q \text{ inert}}} \lambda_{\mathfrak{q}}(z) = \|w\|$$

for some $w \in W$. For the last step note that Proposition 2.4 gives the values of $2\lambda_{\mathfrak{q}}(z)$ for $\mathfrak{q}$ inert. However, the norm of $\mathfrak{q}$ is q^2. The corollary then follows from Proposition 3.4(1) with $\psi = -\mathrm{id} \in \mathrm{Aut}(E/\overline{\mathbb{Q}})$. □

Remark 3.8. Another source of quadratic points in $\mathcal{X}(\mathbb{Z}_p)_2$ comes from elliptic curves with j-invariant equal to 1728. Suppose that E satisfies the assumptions of Theorem 1.6, that p is an odd prime of good reduction and that

$$\mathcal{E} : y^2 = x^3 + a_4 x \quad \text{for some } a_4 \in \mathbb{Z},\ -a_4 \notin \mathbb{Z}^2.$$

Let $z \in \{(\pm\sqrt{-a_4}, 0)\}$ and $K = \mathbb{Q}(\sqrt{-a_4})$ be its field of definition. Let $\psi \in \mathrm{Aut}(E/\overline{\mathbb{Q}})$ be defined by $\psi(x, y) = (-x, iy)$. Then $\psi(z) = \tau(z)$, where $\mathrm{Gal}(K/\mathbb{Q}) = \langle\tau\rangle$. Therefore, under suitable conditions on how the reduction types change in $K/\mathbb{Q}$ and on the splitting of p in K, the localisations of the points z appear in $\mathcal{X}(\mathbb{Z}_p)_2$.

The following corollary explains how points over biquadratic extensions can show up in $\mathcal{X}(\mathbb{Z}_p)_2$ when the j-invariant is zero. For ease of notation, we assume that the a_3-coefficient in the equation defining $\mathcal{E}$ is zero, but this assumption could be removed.

Corollary 3.9. *Suppose that E satisfies the assumptions of Theorem 1.6 and that p is an odd prime of good ordinary reduction. Suppose that*

$$\mathcal{E} : y^2 = x^3 + a_6 \quad \textit{for some } a_6 \in \mathbb{Z}$$

and that there exists $x_0 \in \mathbb{Z}$ such that the points over $\overline{\mathbb{Q}}$ with x-coordinate equal to x_0 have finite order. Assume that p splits in $\mathbb{Q}(\sqrt{x_0^3 + a_6})$. Let $K = \mathbb{Q}(\sqrt{-3}, \sqrt{x_0^3 + a_6})$. For each rational prime q, let $\mathfrak{q}$ be one (any) prime of K above q and $\lambda_{\mathfrak{q}}$ the local height at $\mathfrak{q}$ with respect to the model $\mathcal{E}$. Let $\beta \in \mathbb{Z}_p$ be a root of $t(y) = y^2 - x_0^3 - a_6$. If

$$\sum_{q\in S} \lambda_{\mathfrak{q}}(x_0, \beta) = \|w\|$$

for some $w \in W$, then for each root $\alpha \in \mathbb{Z}_p$ of $s(x) = x^3 - x_0^3$ and for each root $\beta \in \mathbb{Z}_p$ of $t(y) = y^2 - x_0^3 - a_6$, we have $(\alpha, \beta) \in \mathcal{X}(\mathbb{Z}_p)_2$.

Proof. If $x_0^3+a_6$ is a square in $\mathbb{Z}$, the statement is precisely Corollary 3.5. Thus, we may assume that either

(i) K has degree 4 over $\mathbb{Q}$, or

(ii) $K=\mathbb{Q}(\sqrt{-3})=\mathbb{Q}(\sqrt{x_0^3+a_6})$.

Let $\zeta\in K$ be a primitive third root of unity. The automorphism group $\operatorname{Aut}(E/K)$ is generated by $\psi:E\to E$, $\psi(x,y)=(\zeta x,-y)$. In case (i), the Galois group of K over $\mathbb{Q}$ is generated by two elements: σ, whose fixed field is $\mathbb{Q}(\sqrt{x_0^3+a_6})$ and τ, whose fixed field is $\mathbb{Q}(\sqrt{-3})$. In case (ii), the Galois group is generated by $\sigma:\sqrt{-3}\mapsto-\sqrt{-3}$. Let $P=(a,b)$ where $a\in K$ is a root of $s(x)$ and $b\in K$ is a root of $t(y)$. Then in (i)

$$\sigma(P)\in\{P,-\psi(P),\psi^2(P)\},\qquad \tau(P)=-P.$$

Similarly, in case (ii), we have

$$\sigma(P)\in\{-P,\psi(P),-\psi^2(P)\}.$$

Therefore, we may apply Proposition 3.4(1). □

3B. *Noncyclotomic p-adic heights.* The set $\mathcal{X}(\mathbb{Z}_p)_2$ is a finite set of p-adic points containing $\mathcal{X}(\mathbb{Z})$. After having fixed a choice of a subspace of $H^1_{\mathrm{dR}}(E/\mathbb{Q}_p)$ complementary to the space of holomorphic forms, there is only one Coleman–Gross global height pairing on $E(\mathbb{Q})$, up to multiplication by a constant. The definition of $\mathcal{X}(\mathbb{Z}_p)_2$ depends on this height function. Nevertheless, when analysing what points could arise in the set $\mathcal{X}(\mathbb{Z}_p)_2\setminus\mathcal{X}(\mathbb{Z})$, we should bear in mind that other global height functions may exist on $E(F)$, where F is a number field, and that these also vanish on $E(F)_{\mathrm{tors}}$. In particular, suppose that there exists at least one embedding $\rho:F\hookrightarrow\mathbb{Q}_p$. It may happen that, for some $w\in W$ and some $Q\in\mathcal{X}(\mathcal{O}_F)$,

$$2D_2(\rho(Q))+\gamma\operatorname{Log}(\rho(Q))^2+\|w\|=h_p^{\star}(Q)$$

for some noncyclotomic global height $h_p^{\star}$. Then, if Q is in addition a torsion point, we have $\rho(Q)\in\mathcal{X}(\mathbb{Z}_p)_2$.

In order to introduce these more general types of heights, we need to recall the definition and properties of an idele class character.

Definition 3.10. Let $\mathbb{A}_F^{\times}$ be the group of ideles of F. An idele class character is a continuous homomorphism

$$\chi=\sum_{\mathfrak{q}}\chi_{\mathfrak{q}}:\mathbb{A}_F^{\times}/F^{\times}\to\mathbb{Q}_p;$$

here the sum is over all places of F.

We list some properties of an idele class character χ (see [Balakrishnan et al. 2019b] for more details).

(PI) The local character $\chi_{\mathfrak{q}}$ is trivial at an archimedean place $\mathfrak{q}$. Thus, henceforth $\mathfrak{q}$ will always denote a finite prime.

(PII) At a prime $\mathfrak{q}$ not above p, the local character $\chi_{\mathfrak{q}}$ vanishes on the units $\mathcal{O}_{\mathfrak{q}}^{\times}$. Thus, the value of $\chi_{\mathfrak{q}}$ at a uniformiser determines $\chi_{\mathfrak{q}}$ completely.

(PIII) At a prime $\mathfrak{p}$ above p, the restriction of the character $\chi_{\mathfrak{p}}$ to $\mathcal{O}_{\mathfrak{p}}^{\times}$ equals the composition

$$\mathcal{O}_{\mathfrak{p}}^{\times} \xrightarrow{\log_{\mathfrak{p}}} F_{\mathfrak{p}} \xrightarrow{t_{\mathfrak{p}}} \mathbb{Q}_p$$

for some $\mathbb{Q}_p$-linear map $t_{\mathfrak{p}}$. Here $\log_{\mathfrak{p}}$ is the restriction to $\mathcal{O}_{\mathfrak{p}}^{\times}$ of the extension of log to $F_{\mathfrak{p}}^{\times}$.

(PIV) The character χ is completely determined by the trace maps $(t_{\mathfrak{p}})_{\mathfrak{p}|p}$ and, conversely, a tuple of $\mathbb{Q}_p$-linear maps $(t_{\mathfrak{p}} : F_{\mathfrak{p}} \to \mathbb{Q}_p)_{\mathfrak{p}|p}$ gives an idele class character χ if and only if

$$\sum_{\mathfrak{p}|p} t_{\mathfrak{p}}(\log_{\mathfrak{p}}(\rho_{\mathfrak{p}}(\epsilon))) = 0 \quad \text{for all } \epsilon \in \mathcal{O}_F^{\times}, \tag{6}$$

where $\rho_{\mathfrak{p}} : F \hookrightarrow F_{\mathfrak{p}}$ is the completion (see [Balakrishnan et al. 2019b] for a proof).

In particular, it suffices to check that (6) is satisfied for a set of fundamental units and (PIV) gives a concrete method for classifying all idele class characters for a given number field F. The maximal number of independent characters is at least $r_2 + 1$, where r_2 is the number of conjugate pairs of nonreal embeddings of F into $\mathbb{C}$ (with equality if Leopoldt's conjecture holds for F).

For instance, for any number field F, the *cyclotomic* idele class character is the idele class character corresponding to the tuple of trace maps $(\mathrm{tr}_{F_{\mathfrak{p}}/\mathbb{Q}_p})_{\mathfrak{p}|p}$. When $F = \mathbb{Q}$ (or F is a totally real abelian number field), this is the only nontrivial idele class character, up to multiplication by a scalar. The p-adic height we have considered so far is implicitly associated to this character.

More generally though, we can define a p-adic height as a composition of two maps: firstly, we associate to a point $P \in E(F)$ an idele $i(P)$ and, secondly, we apply to $i(P)$ an idele class character χ. We denote the corresponding local and global heights by $\lambda_{\mathfrak{q}}^{\chi}$ and h_p^{χ}, respectively. The theory of local heights that we outlined in the cyclotomic case in Section 2A goes through unvaried at the primes $\mathfrak{q} \nmid p$, after replacing the p-adic logarithm with $-\chi_{\mathfrak{q}}/n_{\mathfrak{q}}$. At the primes $\mathfrak{p} \mid p$, we may assume here that we always work with points not in the residue disk of the point at infinity.[5] So let $z \in \mathcal{X}(\mathcal{O}_{\mathfrak{p}})$ and $m \in \mathbb{N}$ such that mz is in the domain of convergence of $\sigma_{\mathfrak{p}}^{(\gamma)}$. Then

$$\lambda_{\mathfrak{p}}^{\chi}(z) = -\frac{2}{m^2}\frac{\chi_{\mathfrak{p}}}{n_{\mathfrak{p}}}\left(\frac{\sigma_{\mathfrak{p}}^{(\gamma)}(mz)}{f_m(z)}\right) = -\frac{2}{m^2 n_{\mathfrak{p}}} t_{\mathfrak{p}}\left(\log_{\mathfrak{p}}\left(\frac{\sigma_{\mathfrak{p}}^{(\gamma)}(mz)}{f_m(z)}\right)\right), \tag{7}$$

since

$$\mathrm{ord}_{\mathfrak{p}}(\sigma_{\mathfrak{p}}^{(\gamma)}(mz)) = \mathrm{ord}_{\mathfrak{p}}(x(mz)y(mz)^{-1}) = \mathrm{ord}_{\mathfrak{p}}(f_m(z))$$

(see Section 4A and the proof of Theorem 3.18 for how to interpret (7) when $mz = O$). We will omit χ from our notation when using the cyclotomic character.

Example 3.11. Let F be an imaginary quadratic field in which p splits. Then by (PIV), any pair of $\mathbb{Q}_p$-linear maps $\mathbb{Q}_p \to \mathbb{Q}_p$ gives rise to an idele class character. In particular, choosing (id, $-$id) gives the so-called *anticyclotomic* character.

[5] There is a subtlety in the disk at infinity which has to do with the choice of branch of the $\mathfrak{p}$-adic logarithm. See also [Balakrishnan et al. 2019b, Remark 2.1].

We now give an instance of how the existence of noncyclotomic heights for imaginary quadratic fields can give rise to points in $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$.

Proposition 3.12. *Suppose that E satisfies the assumptions of Theorem 1.6 and that p is an odd prime of good reduction. Let K be an imaginary quadratic field in which p splits. Fix an embedding $\rho : K \hookrightarrow \mathbb{Q}_p$. Suppose that $z \in \mathcal{X}(\mathcal{O}_K)_{\mathrm{tors}}$ has good reduction at all primes that split in K. Then*

$$2D_2(\rho(z)) + \sum_{q \in S} \lambda_{\mathfrak{q}}(z) = 0,$$

where $\mathfrak{q}$ is a prime of K above q. In particular, if $\sum_{q\in S} \lambda_{\mathfrak{q}}(z) = \|w\|$ for some $w \in W$, then $\rho(z) \in \mathcal{X}(\mathbb{Z}_p)_2$.

Proof. It suffices to show that $2D_2(\rho(z)) + \sum_{q\in S} \lambda_{\mathfrak{q}}(z)$ is the value at z of a height function on $E(K)$, since then the assumption that z is a torsion point will imply the vanishing. The height function that we are after is the one corresponding to an idele class character $\mathbb{A}_K^\times / K^\times \to \mathbb{Q}_p$ which vanishes on $\mathcal{O}_{\bar{\mathfrak{p}}}^\times$, if $\mathfrak{p}$ is the prime corresponding to the embedding ρ. Indeed, with the notation of (PIV), consider the idele class character corresponding to $(\mathrm{id} : K_{\mathfrak{p}} \simeq \mathbb{Q}_p \to \mathbb{Q}_p, 0 : K_{\bar{\mathfrak{p}}} \simeq \mathbb{Q}_p \to \mathbb{Q}_p)$. Then

$$\lambda_{\mathfrak{p}}(z) = \lambda_{\mathfrak{p}}^{\chi}(z) \quad \text{and} \quad \lambda_{\bar{\mathfrak{p}}}^{\chi}(z) = 0.$$

Further, since χ factors through $\mathbb{A}_K^\times / K^\times$ and in view of (PII), if there is a unique prime $\mathfrak{q}$ above q, we have

$$\chi_{\mathfrak{q}}(q) = -\chi_{\mathfrak{p}}(q) - \chi_{\bar{\mathfrak{p}}}(q) = -\log(q),$$

so that $2\lambda_{\mathfrak{q}}^{\chi} = \lambda_{\mathfrak{q}}$ for all primes which are either inert or ramified. Thus $2h_p^{\chi}(z) = 2D_2(\rho(z)) + \sum_{q\in S} \lambda_{\mathfrak{q}}(z)$. □

In some cases, extra points in $\mathcal{X}(\mathbb{Z}_p)_2$ are explained by a combination of automorphisms and noncyclotomic idele class characters, as in Proposition 3.14. Before we state it and prove it, we first need an auxiliary lemma.

Lemma 3.13. *Let F be a number field and let L be a finite extension of F. Suppose that $\chi : \mathbb{A}_F^\times / F^\times \to \mathbb{Q}_p$ is an idele class character determined by the tuple of $\mathbb{Q}_p$-linear maps $(t_{\mathfrak{p}} : F_{\mathfrak{p}} \to \mathbb{Q}_p)_{\mathfrak{p}|p}$. Then the tuple $(t_{\mathfrak{q}}^L : L_{\mathfrak{q}} \to \mathbb{Q}_p)_{\mathfrak{q}|p}$, defined by $t_{\mathfrak{q}}^L = t_{\mathfrak{p}} \circ \mathrm{tr}_{L_{\mathfrak{q}}/F_{\mathfrak{p}}}$ for $\mathfrak{q} \mid \mathfrak{p}$, determines an idele class character $\chi^L : \mathbb{A}_L^\times / L^\times \to \mathbb{Q}_p$ such that $\chi^L|_{\mathbb{A}_F^\times/F^\times} = [L : F]\chi$.*

Proof. Each $t_{\mathfrak{q}}^L$ is $\mathbb{Q}_p$-linear as a composition of $\mathbb{Q}_p$-linear maps. We need to check that (PIV) is satisfied. If $\epsilon \in \mathcal{O}_L^\times$, then

$$\begin{aligned}\sum_{\mathfrak{q}|p} t_{\mathfrak{q}}^L(\log_{\mathfrak{q}}(\rho_{\mathfrak{q}}(\epsilon))) &= \sum_{\mathfrak{p}|p} t_{\mathfrak{p}} \circ \Big(\sum_{\mathfrak{q}|\mathfrak{p}} \mathrm{tr}_{L_{\mathfrak{q}}/F_{\mathfrak{p}}} \circ \log_{\mathfrak{q}}(\rho_{\mathfrak{q}}(\epsilon)) \Big) = \sum_{\mathfrak{p}|p} t_{\mathfrak{p}} \circ \log_{\mathfrak{p}} \Big(\prod_{\mathfrak{q}|\mathfrak{p}} N_{L_{\mathfrak{q}}/F_{\mathfrak{p}}}(\rho_{\mathfrak{q}}(\epsilon)) \Big) \\ &= \sum_{\mathfrak{p}|p} t_{\mathfrak{p}} \circ \log_{\mathfrak{p}} \big(\rho_{\mathfrak{p}}(N_{L/F}(\epsilon)) \big) = 0,\end{aligned}$$

since $N_{L/F}(\epsilon) \in \mathcal{O}_F^\times$. By construction, the resulting idele class character χ^L restricts to $[L : F]\chi$ on $\mathbb{A}_F^\times / F^\times$. □

Proposition 3.14. *Let d be a nonzero square-free integer and let E^d be the quadratic twist of $X_0(49)$ by d; assume that E^d satisfies the assumptions of Theorem 1.6 and let $\mathcal{X}^d$ be the complement of the origin in the minimal regular model of E^d. Let $p \nmid 7d$ be an odd prime with at least 3 primes lying above it in $L = \mathbb{Q}[x]/(x^4+7d^2)$. Then*

$$\mathcal{X}^d(\mathbb{Z}_p)_2 \supseteq \mathcal{X}^d(\mathbb{Z}) \cup \{\pm\rho(Q)\}$$

for some $Q \in \mathcal{X}^d(\mathcal{O}_L)$ of order 4 and for every embedding $\rho : L \hookrightarrow \mathbb{Q}_p$.

Remark 3.15. The proposition also holds in rank 1 if we replace $\mathcal{X}^d(\mathbb{Z}_p)_2$ with $\mathcal{X}^d(\mathbb{Z}_p)_2'$.

Proof. The elliptic curve $E = X_0(49)$ has reduced minimal model

$$\mathcal{E} : y^2 + xy = x^3 - x^2 - 2x - 1; \tag{8}$$

however, since we are considering quadratic twists of E, it is more convenient to work (at least until we introduce heights) with the model

$$\mathcal{E}_{\text{short}} : y^2 = x^3 - 2835x - 71442,$$

as then the twist E^d of E by the nonzero square-free integer d admits the Weierstrass equation

$$\mathcal{E}^d_{\text{short}} : y^2 = x^3 - 2835d^2x - 71442d^3.$$

Recall that E^d has complex multiplication by $K = \mathbb{Q}(a)$, where a is a root of x^2+7. Over $K[x, y]$, the fourth division polynomial f_4^d of $\mathcal{E}^d_{\text{short}}$ has the factorisation

$$f_4^d(x, y) = 4y(x-9ad)(x+9ad)(x+(-18a+63)d)(x+(18a+63)d)(x^2-126xd-5103d^2).$$

In particular, since

$$x^3 - 2835d^2x - 71442d^3 = (x-63d)(x+(-9/2a+63/2)d)(x+(9/2a+63/2)d),$$

all the points of order 2 are defined over K. As for the points of order 4, we see that, as a polynomial in x, $f_4^d(x, y)/y$ has two roots in $\mathbb{Q}(\sqrt{7})$ and four roots in K. Substituting the latter roots into the equation for $\mathcal{E}^d_{\text{short}}$, we find that the y-coordinates of the points with $x = 9ad$ and $x = -(18a+63)d$ are defined over $\mathbb{Q}(\sqrt{-ad})$, whereas those with $x = -9ad$ and $x = -(-18a+63)d$ are over $\mathbb{Q}(\sqrt{ad})$.

Therefore, over the quartic field $L = K[x]/(x^2-ad) = K(b) \cong \mathbb{Q}[x]/(x^4+7d^2)$, $E^d(L)[4] \cong \mathbb{Z}/2\times\mathbb{Z}/4$. Let $\mathrm{Gal}(L/K) = \langle\bar{\tau}\rangle$ and let

$$Q_{\text{short}} = (18b^2 - 63d, \pm(54b^3 - 378bd)) \in \mathcal{E}^d_{\text{short}}(L)[4], \qquad P_{\text{short}} = (63d, 0).$$

Then Q_{short} satisfies

$$\bar{\tau}(Q_{\text{short}}) = -Q_{\text{short}}. \tag{9}$$

Let Q be the image of Q_{short} in a minimal model $\mathcal{E}^d$ for E^d over $\mathbb{Z}$ and let P be the image of P_{short}. Note that

- if $d \equiv 1 \mod 4$, we may apply to $\mathcal{E}^d_{\text{short}}$ the change of variables

$$x \mapsto 36x - 9d, \qquad y \mapsto 216y + 108x$$

to obtain the integral model

$$\mathcal{E}^d_1 : y^2 + xy = x^3 - \frac{3d+1}{4}x^2 - 2d^2x - d^3.$$

The discriminant of $\mathcal{E}^d_1$ is $\Delta = -7^3d^6$, so by [Silverman 2009, VII, Remark 1.1] $\mathcal{E}^d_1$ is a minimal model for $E^d/\mathbb{Q}$ and we may set $\mathcal{E}^d = \mathcal{E}^d_1$. Then $x(P) = 2d \in \mathbb{Z}$ and $x(Q) \in \mathcal{O}_K$.

- if $d \equiv 2, 3 \mod 4$, then we may take

$$\mathcal{E}^d : y^2 = x^3 - 3dx^2 - 32d^2x - 64d^3,$$

which has discriminant $\Delta = -2^{12} \cdot 7^3 \cdot d^6$. Minimality of $\mathcal{E}^d$ at the primes different from 2 follows as in the case $d \equiv 1 \mod 4$. At the prime 2, it can be deduced following Tate's algorithm. We have $x(P) = 8d \in \mathbb{Z}$ and $x(Q) \in \mathcal{O}_K$.

Now, let p be an odd prime of good reduction for E^d which splits in K. By Deuring's criterion, this is equivalent to requiring that p is a prime of good ordinary reduction. Let $p\mathcal{O}_K = \mathfrak{p}\overline{\mathfrak{p}}$. The prime p is unramified also in L, since we are assuming that it is of good reduction. We suppose furthermore that $\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1\overline{\mathfrak{q}_1}$, $\overline{\mathfrak{p}}\mathcal{O}_L = \mathfrak{q}_2$ or $\overline{\mathfrak{p}}\mathcal{O}_L = \mathfrak{q}_2\overline{\mathfrak{q}_2}$, for some primes $\mathfrak{q}_1$ and $\mathfrak{q}_2$ of L. These conditions, together, are equivalent to those of the statement of the proposition.

By Lemma 3.13, the idele class character on $\mathbb{A}_K^\times/K^\times$ which is trivial on $\mathcal{O}_{\overline{\mathfrak{p}}}^\times$ (and which we used also in the proof of Proposition 3.12) extends to an idele class character on $\mathbb{A}_L^\times/L^\times$. In particular, the tuple of linear maps

$$\Big(\mathrm{id}_{L_{\mathfrak{q}_1} \simeq \mathbb{Q}_p}, \mathrm{id}_{L_{\overline{\mathfrak{q}_1}} \simeq \mathbb{Q}_p}, 0 : \prod_{\mathfrak{q} \mid \overline{\mathfrak{p}}} L_{\mathfrak{q}} \to \mathbb{Q}_p\Big)$$

determines an idele class character χ. Consider the associated global height h_p^χ on $E^d(L)$ with local heights λ_v^χ with respect to the model $\mathcal{E}^d$. Since Q is a torsion point, we must have

$$h_p^\chi(Q) = 0.$$

It follows from (9) and the definition of χ that

$$\lambda^\chi_{\mathfrak{q}_1}(Q) = \lambda^\chi_{\overline{\mathfrak{q}_1}}(Q) = \lambda_{\mathfrak{q}_1}(Q) \quad \text{and} \quad \lambda^\chi_{\mathfrak{q}}(Q) = 0 \quad \text{for } \mathfrak{q} \mid \overline{\mathfrak{p}}.$$

Furthermore, using (PII) and the fact that χ is trivial on $L^\times$, we find that, for a fixed rational prime ℓ,

$$\sum_{v \mid \ell} \chi_v(\ell) = -2 \log \ell. \tag{10}$$

Since $P \in \mathcal{X}^d(\mathbb{Z})$, where $\mathcal{X}^d$ is the complement of the origin in $\mathcal{E}^d$, in order to prove the proposition it then suffices to show that

$$\frac{1}{2}\sum_{v\nmid p} n_v \lambda_v^{\chi}(Q) = \sum_{\ell\nmid p} \lambda_\ell(P),$$

where the left sum runs over primes of L and the right sum over rational primes.

In view of Lemma 3.13 and (4), we are allowed to perform isomorphisms over extensions of L to calculate local heights. In particular, the change of variables $(x, y) \mapsto (36dx-9d, 216d\sqrt{d}y+108d\sqrt{d}x)$, defined over $\mathbb{Q}(\sqrt{d})$, maps $\mathcal{E}^d_{\text{short}}$ to (8), which has discriminant -7^3. Under this isomorphism,

$$x(P_{\text{short}}) \mapsto 2 \in \mathbb{Z} \quad \text{and} \quad x(Q_{\text{short}}) \mapsto \frac{a-3}{2} \in \mathcal{O}_K.$$

Therefore, the local heights of P and Q away from $7p$ are trivial when computed with respect to (8). Using (10) and letting $d' = 7^{-\operatorname{ord}_7(d)}d$, we then have

$$\frac{1}{2}\sum_{v\nmid 7p} n_v \lambda_v^{\chi}(Q) = -\frac{1}{12[F:L]}\sum_{w\nmid 7p} \chi_w^F(\Delta^{-1}) = \begin{cases} -\log d' & \text{if } d \equiv 1 \bmod 4, \\ -\log 4d' & \text{otherwise,} \end{cases}$$

where $F = L(\sqrt{d})$, the second sum runs over the primes w of F and the character χ^F is the idele class character of F obtained from χ as in Lemma 3.13. Similarly, to calculate heights of P away from $7p$ we may base change to $\mathbb{Q}(\sqrt{d})$ and get

$$\sum_{\ell\nmid 7p} \lambda_\ell(P) = \frac{1}{2}\sum_{u\nmid 7p} n_u \lambda_u(P) = \frac{1}{2}\sum_{u\nmid 7p} \frac{n_u}{6}\log(|\Delta|_u) = \begin{cases} -\log d' & \text{if } d \equiv 1 \bmod 4, \\ -\log 4d' & \text{otherwise,} \end{cases}$$

where u runs over primes of $\mathbb{Q}(\sqrt{d})$.

It remains to calculate the local contributions at primes above 7. For this, it is convenient to work with $\mathcal{E}^d_{\text{short}}$ which is minimal over $\mathbb{Z}_7$. Let v be the unique prime above 7 in L. Then $\operatorname{ord}_v(\Delta) = 12 + 24\operatorname{ord}_7(d)$. On the other hand, by [Silverman 2009, VII, Exercise 7.2], a minimal equation at v has discriminant of valuation at most 11. Therefore, E^d has good reduction at v and, since we are not in characteristic 2 or 3, a minimal equation at v is obtained from $\mathcal{E}^d_{\text{short}}$ via $(x, y) \mapsto (\pi_7^{\operatorname{ord}_v(\Delta)/6}x, \pi_7^{\operatorname{ord}_v(\Delta)/4}y)$, where π_7 is a uniformiser at v. Under such a change of variables, P_{short} and Q_{short} are mapped to v-adically integral points. Comparing discriminants as above, we then conclude that

$$\frac{1}{2}\sum_{v\nmid p} n_v \lambda_v^{\chi}(Q) = \sum_{\ell\nmid p} \lambda_\ell(P) = \begin{cases} -\frac{1}{2}\log 7 - \log d & \text{if } d \equiv 1 \bmod 4, \\ -\frac{1}{2}\log 7 - \log 4d & \text{otherwise.} \end{cases} \qquad \square$$

3C. ***Proof of Theorem 1.4.*** In this subsection we explain how Theorem 1.4 can be deduced either from Corollary 3.5 or from Proposition 3.14. For an elliptic curve E over $\mathbb{Q}$, denote by $L(E, s)$ its complex L-function.

Theorem 3.16 [Waldspurger 1981; Vignéras 1981; Murty and Murty 1997, Chapter 6, Theorem 1.1]. *Let E be an elliptic curve over $\mathbb{Q}$. There exist infinitely many nonzero square-free integers d such that the quadratic twist E^d of E by d satisfies $L(E^d, 1) \neq 0$.*

Theorem 3.17 [Kolyvagin 1988]. *Let E be an elliptic curve over $\mathbb{Q}$ such that $L(E, 1) \neq 0$. Then the rank of $E(\mathbb{Q})$ is zero and the Tate–Shafarevich group of $E/\mathbb{Q}$ is finite.*

It follows from Theorems 3.16 and 3.17 that there are infinitely many twists of $X_0(49)$ satisfying the hypotheses of Proposition 3.14. For each such curve, by Chebotarev's density theorem, there are infinitely many primes for which the proposition holds.

We now see how Corollary 3.5 also provides us with an alternative proof of Theorem 1.4. Consider the elliptic curve E with label 36.a3 [LMFDB 2019] which has reduced minimal equation

$$\mathcal{E} : y^2 = x^3 - 27.$$

We have $\mathcal{E}(\mathbb{Z}) = \mathcal{E}(\mathbb{Z})[2] = \{O, (3, 0)\}$. It follows that every quadratic twist E^d of E by a nonzero square-free integer d satisfies $E^d(\mathbb{Q})[2] \cong \mathbb{Z}/2\mathbb{Z}$. The equation

$$\mathcal{E}^d : y^2 = x^3 - 27d^3$$

has discriminant equal to $-2^4 \cdot 3^9 \cdot d^6$ and is hence globally minimal, except if $3 \mid d$, in which case we apply $(x, y) \mapsto (9x, 27y)$ to obtain a minimal model. Thus, the point of exact order 2 of E^d defined over $\mathbb{Q}$ is integral. Therefore, by Theorems 3.16 and 3.17, there exist infinitely many d for which Corollary 3.5 holds with $y_0 = 0$.

3D. ***A necessary condition: quadratic saturation.*** In Sections 3A and 3B we proved sufficient conditions for a point in $\mathcal{X}(\mathbb{Z}_p)$ to belong to $\mathcal{X}(\mathbb{Z}_p)_2$. We now prove the necessary condition given by Theorem 1.8, which we restate here for the reader's convenience.

Theorem 3.18. *Let $E/\mathbb{Q}$ and p be as in Theorem 1.6. Suppose that $z \in \mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$. Then z is the localisation of a torsion point P over a number field K and, for each rational prime q, the value $\lambda_{\mathfrak{q}}(P)$ of the local height $\lambda_{\mathfrak{q}}$ is independent of the prime $\mathfrak{q} \mid q$ of K.*

Proof. As we observed at the beginning of this section, a point $z \in \mathcal{X}(\mathbb{Z}_p)_2$ is necessarily the p-adic localisation of a point $P \in \mathcal{E}(\overline{\mathbb{Z}})_{\text{tors}}$. Let K be the minimal number field over which the coordinates of P are defined. Let $m \in \mathbb{Z}$, $m \neq 0$, such that $mP = O$. Since $z \in \mathcal{X}(\mathbb{Z}_p)$, there exists an embedding ψ of K into $\mathbb{Q}_p$ under which P is mapped to z, i.e., a prime $\mathfrak{p}_0$ of K such that $\mathfrak{p}_0 \mid p$ and

$$\lambda_{\mathfrak{p}_0}(P) = \lambda_p(z) = -\|w\| =: -\sum_{q \in S} \alpha_q \log q \quad \text{for some } w \in W, \alpha_q \in \mathbb{Q}.$$

For any prime $\mathfrak{p} \mid p$ of K, the value $\lambda_{\mathfrak{p}}(P)$ can be computed as follows. Let $x(t), y(t) \in K[[t]]$ be coordinates around P, i.e., $P = (x(0), y(0))$, the power series $x(t), y(t)$ converge in the intersection over $\mathfrak{p} \mid p$ of small enough $\mathfrak{p}$-adic neighbourhoods of P and t vanishes to order 1 at P. Let $Q(t) = m(x(t), y(t)) \in E(K((t)))$. Since $K[[t]]$ is a complete DVR with residue field K, by [Wuthrich 2004, Proposition 1] the t-adic valuation of $-x(Q(t))/y(Q(t))$ equals the one of $f_m(x(t), y(t))$. More precisely, since f_m vanishes to order 1 at every point of order dividing m, we have

$$\begin{aligned} -x(Q(t))/y(Q(t)) &= at + O(t^2) \quad \text{for some } a \in K^{\times}, \\ f_m(x(t), y(t)) &= ct + O(t^2) \quad \text{for some } c \in K^{\times}. \end{aligned}$$

Since $\sigma_p^{(\gamma)}(T) = T + O(T^2)$, by Section 2A2(ii) we then have

$$\begin{aligned} \lambda_{\mathfrak{p}}(P) &= \lim_{t\to 0} -\frac{2}{n_{\mathfrak{p}} m^2} \operatorname{tr}_{K_{\mathfrak{p}}/\mathbb{Q}_p}\left(\log_{\mathfrak{p}}\left(\frac{\sigma_p^{(\gamma)}(-x(Q(t))/y(Q(t)))}{f_m(x(t), y(t))}\right)\right) \\ &= -\frac{2}{n_{\mathfrak{p}} m^2} \operatorname{tr}_{K_{\mathfrak{p}}/\mathbb{Q}_p} \log_{\mathfrak{p}}\left(\frac{a}{c}\right), \end{aligned}$$

where $\log_{\mathfrak{p}}$ is an extension of log to $K_{\mathfrak{p}}^{\times}$.

In particular, if d is the least common multiple of the denominators of the α_q, then

$$\log\left(\psi\left(\frac{a}{c}\right)^{2d}\right) = \log\left(\prod_{q\in S} q^{d\alpha_q m^2}\right).$$

Since $\mathfrak{p}$ is a prime of good reduction, we also have $\operatorname{ord}_{\mathfrak{p}}(a/c) = 0$ (strictly speaking we could also avoid using this fact, since the branch of the logarithm corresponding to the cyclotomic character vanishes at p), so

$$\left(\frac{a}{c}\right)^{2d} = \zeta \prod_{q\in S} q^{d\alpha_q m^2}$$

for some root of unity $\zeta \in K$. Thus

$$\lambda_{\mathfrak{p}}(P) = -\frac{1}{d n_{\mathfrak{p}} m^2} \operatorname{tr}_{K_{\mathfrak{p}}/\mathbb{Q}_p} \log_{\mathfrak{p}}\left(\zeta \prod_{q\in S} q^{d\alpha_q m^2}\right) = -\sum_{q\in S} \alpha_q \log q$$

is independent of $\mathfrak{p}$.

Let now $\mathfrak{q}$ be a prime not above p. By Section 2A1(ii), (iv), we have

$$\lim_{R\to P} \frac{1}{m^2}(\lambda_{\mathfrak{q}}(mR) + 2\log|f_m(R)|_{\mathfrak{q}}) = \lim_{R\to P} \lambda_{\mathfrak{q}}(R) = \lambda_{\mathfrak{q}}(P).$$

Since mR is in the formal group at $\mathfrak{q}$, then

$$\begin{aligned} \lambda_{\mathfrak{q}}(P) &= \lim_{R\to P} \frac{1}{m^2}\left(\log|x(mR)|_{\mathfrak{q}} |f_m(R)|_{\mathfrak{q}}^2\right) \\ &= \lim_{R\to P} \frac{1}{m^2} \log\left(\left|\frac{x(mR)}{y(mR)}\right|_{\mathfrak{q}}^{-2} |f_m(R)|_{\mathfrak{q}}^2\right) \\ &= \frac{1}{m^2}\log\left(\left|\frac{c}{a}\right|_{\mathfrak{q}}^2\right) = \frac{1}{d}\log\left(\left|\prod_{q\in S} q^{-d\alpha_q}\right|_{\mathfrak{q}}\right), \end{aligned}$$

which completes the proof. ∎

Corollary 3.19. *If $z \in \mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ is the localisation of a point P defined over K, we have $P \in \mathcal{X}(\mathcal{O}_K)$.*

Proof. The proof is similar to that of Proposition 3.1. Note that a torsion point defined over an arbitrary number field can fail to be integral at $\mathfrak{q}$ only if its order is q^n for some n (where q is the norm of $\mathfrak{q}$) and $\operatorname{ord}_{\mathfrak{q}}(q) \geq q^n - q^{n-1}$ (cf. [Silverman 2009, VIII, Theorem 7.1]). □

Theorem 3.18 is in some sense a natural analogue of a conjecture of Stoll for the classical abelian Chabauty method [Stoll 2006, Conjecture 9.5], which appears in an unpublished draft of [Stoll 2007]. Let us restrict to the case when C is a hyperelliptic curve over $\mathbb{Q}$ of genus g, whose Jacobian J has rank $g-2$ over $\mathbb{Q}$ (for the conjecture in its full generality see [Stoll 2006]). Suppose that $\iota : C \hookrightarrow J$ is an embedding such that $\iota(C)$ generates J and that J is simple. Stoll's conjecture predicts the existence of a finite subscheme $Z \subset J$ and a set R of primes which has density 1 in the set of all primes such that, for each $\ell \in R$, we have

$$\overline{J(\mathbb{Q})} \cap \iota(C(\mathbb{Q}_\ell)) \subset Z(\mathbb{Q}_\ell),$$

where $\overline{J(\mathbb{Q})}$ is the ℓ-adic closure of $J(\mathbb{Q})$ in $J(\mathbb{Q}_\ell)$.

In [Balakrishnan et al. 2019c] some evidence for Stoll's conjecture was collected when $g = 3$, with the scheme Z being the intersection of $\iota(C)$ with the saturation of $J(\mathbb{Q})$, that is

$$Z = \{\iota(P) \in J : n\iota(P) \in J(\mathbb{Q}) \text{ for some } n \in \mathbb{Z}_{\geq 1}\}.$$

In our setting of an elliptic curve of rank 0, it is clear that $\mathcal{X}(\mathbb{Z}_p)_2$ should be contained in the saturation of the Mordell–Weil group $E(\mathbb{Q})$, since $E(\mathbb{Q}) = E(\mathbb{Q})_{\text{tors}}$ and the equation $\operatorname{Log} z = 0$ cuts out the torsion points in $E(\mathbb{Q}_p)$. However, this is not a very strong requirement in the elliptic curve case, because the curve and the Jacobian are identified.

Theorem 3.18 asserts that the extra constraint $\lambda_p(z) = -\|w\|$, for some $w \in W$, leads to another type of "saturation", in the sense that we have to consider those points for which the local heights behave as if the point were defined over $\mathbb{Q}$.

Note that Stoll's conjecture assumes that $g - r + l \geq 3$, where $l = 1$ is the level, in the sense that the Chabauty–Coleman method computes the cohomologically global points of level 1. In the situation discussed here we have $g = 1$, $r = 0$, $l = 2$, i.e., $g - r + l = 3$, so one could naively hope that for rank 1 similar conjectures could be formulated at level 3.

4. Algorithm and computations in rank 0

4A. ***The algorithm.*** We wish to explicitly compute the sets $\phi(w)$ and $\psi(w)$ from Theorems 1.6 and 1.7. Each of $D_2(z)$ and $\operatorname{Log}(z)$ are locally analytic functions on $\mathcal{X}(\mathbb{Z}_p)$. In other words, given a point $\bar{P} \in \bar{E}(\mathbb{F}_p) \setminus \{O\}$ and a fixed point $P \in \mathcal{X}(\mathbb{Z}_p)$ reducing to $\bar{P}$ modulo p, one can pick a uniformiser $t \in \mathbb{Q}_p(E)$ at P, which reduces to a uniformiser at $\bar{P}$. Then, for each $Q \in \mathcal{X}(\mathbb{Z}_p)$ in the residue disk of P, we have

$$\operatorname{Log}(Q) = f_P(t(Q)) \quad \text{and} \quad D_2(Q) = g_P(t(Q))$$

for some $f_P(x), g_P(x) \in \mathbb{Q}_p[[x]]$ convergent at all $x \in \mathbb{Z}_p$ with $|x|_p < 1$.

On the other hand, let $\gamma = C$ if p is of good ordinary reduction and $\gamma = \frac{1}{12}(a_1^2 + 4a_2)$ if p is of good supersingular reduction. By Proposition 2.5 and Section 2A2(ii), provided that $mQ \neq O$, where $m = \#\bar{E}(\mathbb{F}_p)$, then

$$2D_2(Q) + \gamma \operatorname{Log}(Q)^2 = -\frac{2}{m^2} \log\left(\frac{\sigma_p^{(\gamma)}(mQ)}{f_m(Q)}\right). \tag{11}$$

Since there are finitely many[6] points in each residue disk satisfying $mQ = O$, the local expansion of the right-hand side of (11) in terms of the local parameter t holds in the whole residue disk. In fact, the local expansion of $\sigma_p^{(\gamma)}(mQ)$ and $f_m(Q)$ have precisely the same zeros with the same multiplicity 1 and two p-adic power series which agree at infinitely many points in $\mathbb{Z}_p$ of absolute value less than 1 are equal by the p-adic Weierstrass preparation theorem [Koblitz 1984, Chapter IV, §4, Theorem 14]. Note that we already used this in the proof of Theorem 3.18.

By the same observation as in Remark 2.7, we obtain a way of computing the intersections of $\phi(w)$ and $\psi(w)$ with each residue disk using local expansions of the p-adic sigma function (of Mazur–Tate or Bernardi) and the m-th division polynomial,[7] in place of the double Coleman integral $D_2(z)$.

The function $\operatorname{Log} : \mathcal{X}(\mathbb{Z}_p) \to \mathbb{Q}_p$ is odd; the function $\lambda_p : \mathcal{X}(\mathbb{Z}_p) \to \mathbb{Q}_p$ is even (cf. property (iv) in Section 2A2). Therefore,

$$z \in \phi(w) \iff -z \in \phi(w)$$

and it will thus suffice to consider residue disks up to $\bar{P} \mapsto -\bar{P}$. The same holds for $\psi(w)$.

We also notice that different models can be used for computing the p-adic heights and the single Coleman integrals. In fact, we defined local p-adic heights using an integral minimal model and, for instance, there is an implementation for the Mazur–Tate p-adic sigma function in `SageMath` due to Harvey [2008] (see also [Mazur et al. 2006]). On the other hand, for Coleman computations on `SageMath` (see [Balakrishnan et al. 2010]), one requires the elliptic curve to be described by a Weierstrass model whose a_1 and a_3 coefficients are zero and there is no requirement on minimality; the only requirement on integrality is $\mathbb{Z}_p$-integrality. To avoid explicit Coleman integration computations, we could also work directly with the formal logarithm.

4B. ***Examples for Section 3.*** In the examples that follow, as well as in the ones of the next sections, we avoid making distinctions between the curve $E/\mathbb{Q}$ and the model $\mathcal{E}/\mathbb{Z}$. The Weierstrass equations that we work with are always minimal and reduced, unless stated otherwise.

Example 4.1 (Corollary 3.7, Proposition 3.12). Consider the rank 0 elliptic curve 17.a1 [LMFDB 2019]

$$E : y^2 + xy + y = x^3 - x^2 - 91x - 310 \tag{12}$$

and the prime $p = 5$ of good ordinary reduction.

[6] In fact, at most one.

[7] Computationally, it is more convenient to take m to be the order of $\bar{P}$ in $\bar{E}(\mathbb{F}_p)$, i.e., to choose potentially different values of m for different residue disks.

Since none of the conditions of Theorem 1.6(1) are satisfied, we need to explicitly compute $\mathcal{X}(\mathbb{Z}_p)_2$ as a union of $\phi(w)$. The curve has split multiplicative reduction at 17 with Kodaira symbol I_1 and good reduction everywhere else: thus, $W = \{0\}$. We find

$$\mathcal{X}(\mathbb{Z}_p)_2 = \{(-5, 2 \pm \rho(i))\},$$

where $\rho : \mathbb{Z}[i] \hookrightarrow \mathbb{Z}_p$ is a fixed embedding. If a priori our computations only return approximations of p-adic points, by Corollary 3.7 the p-adic points found are the localisations of the points over $\mathbb{Z}[i]$ listed above. Let us nevertheless explain it in detail for this example.

The Weierstrass equation (12) defines a global minimal model also for the base-change $E/\mathbb{Q}(i)$ and the prime p splits in $K = \mathbb{Q}(i)$. We extend ρ to a map $E(K) \hookrightarrow E(\mathbb{Q}_p)$. The point

$$Q = (-5, 2 + i) \in E(K)$$

is integral with respect to the global minimal model above and satisfies

$$4Q = O.$$

Thus, since the reduction types at the bad primes of $E/\mathbb{Q}(i)$ are the same as over $\mathbb{Q}$, we have

$$0 = h_p(Q) = \tfrac{1}{2}(\lambda_{\mathfrak{p}_1}(Q) + \lambda_{\mathfrak{p}_2}(Q)),$$

where $p\mathbb{Z}[i] = \mathfrak{p}_1\mathfrak{p}_2$ and explicitly (without loss of generality)

$$\lambda_{\mathfrak{p}_1}(Q) = \lambda_p(\rho(Q)), \qquad \lambda_{\mathfrak{p}_2}(Q) = \lambda_p(\rho(\tau(Q))),$$

where $\mathrm{Gal}(K/\mathbb{Q}) = \langle \tau \rangle$, so $\tau(Q) = (-5, 2 - i)$. On the other hand, $\tau(Q) = -Q$ and λ_p is an even function. Therefore

$$0 = \lambda_p(\rho(Q)) = \lambda_p(\rho(\tau(Q))).$$

Note that Proposition 3.12 also explains why the p-adic localisation of the point Q belongs to $\mathcal{X}(\mathbb{Z}_p)_2$.

Example 4.2 (Proposition 3.12, 3.4(1)). Consider the elliptic curve 121.d3 [LMFDB 2019]

$$E : y^2 + y = x^3 - x^2 - 40x - 221 \tag{13}$$

and the prime $p = 5$, at which E has good ordinary reduction. We have

$$\#\bar{E}(\mathbb{F}_2) = 1$$

and thus

$$\mathcal{X}(\mathbb{Z}_p)_2 = \varnothing$$

by Theorem 1.6(1). On the other hand, we can still compute $\bigcup_{w \in W} \phi(w)$ of Theorem 1.7(2). The curve has additive reduction of type I_1^* at 11 with Tamagawa number 2 and has good reduction everywhere else. Therefore,

$$W = \{0, -\log 11\}.$$

We find that $\phi(0) = \varnothing$, but

$$\phi(-\log 11) = \left\{\left(-7, \rho\left(\tfrac{1}{2}(-1 \pm 11\sqrt{-11})\right)\right), \left(4, \rho\left(\tfrac{1}{2}(-1 \pm 11\sqrt{-11})\right)\right)\right\},$$

where $\rho : \mathbb{Z}[(1+\sqrt{-11})/2] \hookrightarrow \mathbb{Z}_p$ is a fixed embedding.

Let $K = \mathbb{Q}(\sqrt{-11})$. The prime 11 ramifies in $K/\mathbb{Q}$ and E/K has split multiplicative reduction of type I_2 at $\mathfrak{q}$, where $\mathfrak{q}^2 = 11$. Let

$$Q \in \left\{\left(-7, \tfrac{1}{2}(-1 \pm 11\sqrt{-11})\right), \left(4, \tfrac{1}{2}(-1 \pm 11\sqrt{-11})\right)\right\} \subset E(K).$$

The point Q has order 5. Unlike in Example 4.1, the Weierstrass equation (13) is minimal at all primes except at $\mathfrak{q}$ and hence we cannot use straightforwardly the explicit formulae for the local height at $\mathfrak{q}$ given in Section 2A1. The curve E/K admits the global minimal model

$$E^{\min} : y^2 + y = x^3 - x^2$$

and the image of Q in $E^{\min}(K)$ has good reduction at $\mathfrak{q}$, so that $\lambda_{\mathfrak{q}}^{\min}(Q) = 0$. Equation (4) then yields $\lambda_{\mathfrak{q}}(Q) = -\log 11$. Therefore, by Proposition 3.12, we have

$$0 = \lambda_p(\rho(Q)) + \lambda_{\mathfrak{q}}(Q) = \lambda_p(\rho(Q)) - \log 11.$$

Similarly to Example 4.1, the appearance of $\rho(Q)$ in $\phi(-\log 11)$ is also justified by Proposition 3.4(1).

Example 4.3 (Remark 3.8). Consider the elliptic curve 14112.q1 [LMFDB 2019]

$$E : y^2 = x^3 - 9261x \tag{14}$$

and the prime $p = 5$, which is of good ordinary reduction. Note that p splits in $K = \mathbb{Q}(\sqrt{21})$ and by Remark 3.8, the localisations of the point $Q^{\pm} = (\pm 21\sqrt{21}, 0)$ belong to $\mathcal{X}(\mathbb{Z}_p)_2$ provided that (1/2) times the sum of its local heights at bad primes is in $\|W\|$. Both at 3 and 7, the curve has bad reduction of additive type III^* with Tamagawa number 2; at 2 the curve has reduction of type III with Tamagawa number 2. Thus, $W = W_2 \times W_3 \times W_7$, with $W_q = \left\{0, -\tfrac{3}{2}\log q\right\}$ for each $q \in \{3, 7\}$ and $W_2 = \left\{0, -\tfrac{1}{2}\log 2\right\}$.

A global minimal model for the base-change of E to K is given by $y^2 = x^3 - 21x$. Furthermore, 2 is inert in K and its reduction type does not change. The primes 3 and 7 become of type I_0^* with Tamagawa number 4. By Propositions 2.4 and 3.4(1) (see also Remark 3.8), we then find that $Q^{\pm}$ is indeed in $\mathcal{X}(\mathbb{Z}_p)_2$.

Our computation of $\mathcal{X}(\mathbb{Z}_p)_2$ recovers precisely the integral points and the ones coming from $Q^{\pm}$.

Example 4.4 (Proposition 3.4(1)). Consider the elliptic curve 11025.y2 [LMFDB 2019] whose reduced minimal model is

$$E : y^2 + y = x^3 + 15006$$

and let $p = 13$, which is the smallest prime of good ordinary reduction for E. Note that E has vanishing j-invariant. We find that

$$\mathcal{X}(\mathbb{Z}_p)_2 = \{\pm(0, 122)\} \cup \{\pm(\zeta_3^i \sqrt[3]{120050}, 367) : 0 \le i \le 2\}, \tag{15}$$

where ζ_3 is a primitive third root of unity and we assume that we have fixed an embedding of $\mathbb{Q}(\zeta_3, \sqrt[3]{120050})$ into $\mathbb{Q}_p$. As usual, the equality (15) is deduced from computations combined with theoretical results. In this particular example, the theory needed is that the Galois group of $\mathbb{Q}(\zeta_3, \sqrt[3]{120050})/\mathbb{Q}$ acts on the order-6 points $\pm(\zeta_3^i\sqrt[3]{120050}, 367)$ by automorphisms. We leave the reader to check that these points also have the right local heights at bad primes.

Example 4.5 (Corollary 3.5, 3.9). Consider the elliptic curve 900.g3 [LMFDB 2019] with reduced minimal model

$$E : y^2 = x^3 - 3375$$

and the prime 19, which is of good ordinary reduction for E. We have

$$\mathcal{X}(\mathbb{Z}_{19})_2 = \{(\zeta_3^i 15, 0), (-\zeta_3^i 30, \pm\sqrt{-30375}) : 0 \le i \le 2\}$$

where ζ_3 is a primitive third root of unity and we assume that we have fixed an embedding of $\mathbb{Q}(\zeta_3, \sqrt{-30375})$ into $\mathbb{Q}_{19}$.

4C. *Large-scale data.* Using the database [LMFDB 2019], we could run the code on all the 86213 elliptic curves over $\mathbb{Q}$ of rank 0 and conductor less than or equal to 30000; for each curve we let p be the smallest prime ≥ 5 of good ordinary reduction.[8]

Out of these, we found exactly 470 pairs (E, p) for which $\mathcal{X}(\mathbb{Z}_p)_2 \supsetneq \mathcal{X}(\mathbb{Z})$. The 10 such pairs with E of conductor ≤ 100 are listed in Table 2.

We summarise the results of the computations in Propositions 4.6, 4.7, 4.9.

Proposition 4.6. *Let E be an elliptic curve of rank* 0 *and conductor less than or equal to* 30000 *and let $p \ge 5$ be the smallest prime of good ordinary reduction. Assume that $j(E) \notin \{0, 1728, -3375\}$. Then $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ is either empty or consists of localisations of points defined over the ring of integers of a quadratic field K on which the Galois group of $K/\mathbb{Q}$ acts as multiplication by ± 1.*

Proposition 4.7. *Let E be an elliptic curve of rank* 0 *and conductor less than or equal to* 30000 *and let $p \ge 5$ be the smallest prime of good ordinary reduction. If $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$ contains localisations of points defined over a quadratic field K, these points satisfy the hypotheses of Proposition 3.4, i.e., the nontrivial element of* $\mathrm{Gal}(K/\mathbb{Q})$ *acts on them in the same way as an automorphism of E.*

Remark 4.8. In view of the large data collected, it might have been tempting to expect that Propositions 4.6 and 4.7 would be true for arbitrary prime and conductor. It turns out that this is not the case: varying the prime for the curve 8712.u5 [LMFDB 2019] (which does not have CM), we found some quadratic points on which Galois does not act by automorphisms (cf. Example 4.11). We note nevertheless that in the latter case the extra points were explained by Proposition 3.12.

[8] We could have allowed p to equal 3 and used the method of [Balakrishnan 2016] to compute the quantities involved in the 3-adic heights. Some computations with supersingular primes were carried out for Section 4D.

LMFDB	p	SS/CM	$\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z})$	Order	Explanation
17.a1	5	SS	$(-5, 2\pm i)$	4	Cor. 3.7/Prop. 3.12
27.a3	7	$j=0$	$\pm\big(\frac{1}{2}(-3\pm 3\sqrt{-3}), 4\big)$	3	Cor. 3.5/Prop. 3.12
32.a4	5	$j=1728$	$(-2, \pm 4i)$	4	Prop. 3.4(1)/Prop. 3.12
36.a3	7	$j=0$	$\big(\frac{1}{2}(-3\pm 3\sqrt{-3}), 0\big)$	2	Cor. 3.5/Prop. 3.12
36.a4	7	$j=0$	$\big(\frac{1}{2}(1\pm\sqrt{-3}), 0\big)$	2	Cor. 3.5/Prop. 3.12
			$\pm(-1\pm\sqrt{-3}, 3)$	6	Cor. 3.5/Prop. 3.12
49.a2	11	$j=-3375$	$\pm\big(\frac{1}{2}(-7b^2+25), \frac{1}{4}(-49b^3+7b^2-49b-25)\big)$	4	Prop. 3.14
49.a4	11	$j=-3375$	$\pm\big(\frac{1}{2}(b^2-3), \frac{1}{4}(b^3-b^2-7b+3)\big)$	4	Prop. 3.14
75.b4	11	–	$(27, -14\pm 5\sqrt{5})$	4	Prop. 3.4(1)
75.b6	11	–	$\big(12, \frac{1}{2}(-13\pm 25\sqrt{5})\big)$	4	Prop. 3.4(1)
			$\big(2, -\frac{3}{2}(1\pm 5\sqrt{5})\big)$	4	Prop. 3.4(1)
75.b7	11	–	$\big(2, \frac{1}{2}(-3\pm 5\sqrt{5})\big)$	4	Prop. 3.4(1)

Table 2. All curves of rank 0 and conductor ≤ 100 for which $\mathcal{X}(\mathbb{Z}_p)_2 \supsetneq \mathcal{X}(\mathbb{Z})$ ($p \geq 5$ smallest good ordinary prime); b satisfies $x^4+7=0$. The curve is given in the first column as an LMFDB label [LMFDB 2019]. In the third column, SS means "semistable" and "–" neither semistable nor CM.

We now turn to the extra points in our data defined over number fields of degree[9] at least equal to 3. By Proposition 4.6, these can only show up if E has complex multiplication and, in fact, its j-invariant is one of 0, 1728, -3375. There was only one curve beside the one of Example 4.4 where cubic points were recovered, namely the curve 19881.g2 [LMFDB 2019]. As in Example 4.4, the curve has j-invariant equal to zero and the appearance of these points in $\mathcal{X}(\mathbb{Z}_p)_2$ is explained by Proposition 3.4.

Finally, we recovered points defined over number fields of degree 4 on the curve 14112.q2 ($j = 1728$) [LMFDB 2019], which is explained by Proposition 3.4, and on all the twists of the modular curve $X_0(49)$, as predicted by Proposition 3.14. In fact the inclusion in the statement of Proposition 3.14 is an equality in all the following cases.

Proposition 4.9. *Let E be an elliptic curve of rank* 0, *conductor less than or equal to* 30000 *and* $j(E) = -3375$. *Then* $\mathcal{X}(\mathbb{Z}_{11})_2 = \mathcal{X}(\mathbb{Z}) \cup \{\pm Q\}$, *where Q has order* 4 *and comes from a point over the ring of integers of the smallest number field L over which* $E(L)[4] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$.

[9]Note that by degree we mean the degree of the smallest number field over which a point in $\mathcal{X}(\mathbb{Z}_p)_2$ is defined and not the degree of the number field containing all the coordinates of the points in $\mathcal{X}(\mathbb{Z}_p)_2$, which could be larger.

4D. ***Variation of the prime.*** In what follows, we assume that the rank of the elliptic curve is equal to zero. However, the discussion could easily go through word for word with $\mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{tors}}$ in place of $\mathcal{X}(\mathbb{Z}_p)_2$.

We have established that there exist curves of rank 0 and primes p for which $\mathcal{X}(\mathbb{Z}_p)_2 \supsetneq \mathcal{X}(\mathbb{Z})$. The next question we ask is whether there always exists a (not necessarily ordinary) prime for which the cohomologically global points of level 2 are precisely the global integral points. It turns out that the answer is negative, as we will see in Example 4.11 below.

Let us first gather some intuition on what is happening. Recall that, after having fixed all appropriate embeddings,

$$\phi(w) \subset \mathcal{E}(\overline{\mathbb{Z}})_{\mathrm{tors}} = E(\overline{\mathbb{Q}})_{\mathrm{tors}}.$$

Therefore, if $P \in E(F)_{\mathrm{tors}}$ for some minimal number field F, by picking p such that $[F_v : \mathbb{Q}_p] > 1$ for all $v \mid p$, we can guarantee that $P \notin \mathcal{X}(\mathbb{Z}_p)_2$. For instance in Example 4.1, if we pick $p' = 7$, which is of good ordinary reduction and which is inert in $\mathbb{Q}(i)$, we find $\mathcal{X}(\mathbb{Z}_{p'})_2 = \mathcal{X}(\mathbb{Z}) = \varnothing$.

Note that, in view of Corollary 3.5 and Section 3C, there exist (infinitely many) curves for which $\mathcal{X}(\mathbb{Z}_p)_2$ is strictly larger than $\mathcal{X}(\mathbb{Z})$ for all odd primes p of good *ordinary* reduction. More generally, if E has complex multiplication by the quadratic field K and there exist points defined over K and satisfying the assumptions of Proposition 3.4(1), then these points will show up in $\mathcal{X}(\mathbb{Z}_p)_2$ for any good ordinary odd prime p by Deuring's criterion. On the other hand, Deuring's criterion also implies that the good supersingular primes cannot split in K.

We ran the code on all the 470 curves of Section 4C for which we had found some extra points: this time, we varied the good ordinary prime until we found a prime for which no extra points showed up or we proved that such prime does not exist. If a good ordinary prime p for which $\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z}_p)$ does not exist, we repeated the calculations with supersingular primes. We summarise the results in the following theorem (which includes also the statement of Theorem 1.5).

Theorem 4.10. *Let E be an elliptic curve over $\mathbb{Q}$ of rank 0 and conductor less than or equal to 30000. Then there exists a good ordinary odd prime p for which $\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z})$, unless*:

(1) *E is 32.a4* ($j = 1728$) [LMFDB 2019];

(2) *E is one of the 20 elliptic curves of rank 0 with $j = 0$ and $\mathbb{Z}/2\mathbb{Z} \subset E(\mathbb{Q})$ or E is 27.a3* ($j = 0$) [LMFDB 2019];

(3) *E is 8712.u5* [LMFDB 2019].

Moreover, in cases (1) and (2), $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z}) \neq \varnothing$ for all good ordinary odd primes p, but there exists a supersingular prime p for which $\mathcal{X}(\mathbb{Z}_p)_2 = \mathcal{X}(\mathbb{Z}_p)$; in case (3) $\mathcal{X}(\mathbb{Z}_p)_2 \setminus \mathcal{X}(\mathbb{Z}) \neq \varnothing$ for all good (*ordinary and supersingular*) *primes p.*

Proof. That there exists a good ordinary prime for which Conjecture 1.1 holds at level 2 for all curves not in (1), (2) and (3) is shown computationally. The assertion of cases (1) and (2) follows from the discussion before the statement of the theorem together with explicit computations of $\mathcal{X}(\mathbb{Z}_p)_2$ at some supersingular primes p. Finally, we treat the curve 8712.u5 in detail in Example 4.11. □

Example 4.11. Consider the elliptic curve 8712.u5, given by

$$E : y^2 = x^3 + 726x + 9317. \tag{16}$$

We have $S = \{2, 3, 11\}$: in particular, the reduction is of type III with Tamagawa number 2 at 2, of type I_1^* with Tamagawa number 4 at 3 and of type I_0^* with Tamagawa number 2 at 11. Thus $W = W_2 \times W_3 \times W_{11}$, where

$$W_2 = \left\{0, -\tfrac{1}{2}\log 2\right\}, \quad W_3 = \left\{0, -\log 3, -\tfrac{5}{4}\log 3\right\}, \quad W_{11} = \{0, -\log 11\}.$$

Consider

$$A := \left\{(-44, \pm 99\sqrt{-11}), (22, \pm 33\sqrt{33}), \left(\tfrac{11}{2}(1 \pm 3\sqrt{-3}), 0\right)\right\} \subset \mathcal{X}(\overline{\mathbb{Z}})_{\mathrm{tors}}.$$

If $p \notin S$, then p splits in at least one of $\mathbb{Q}(\sqrt{-11})$, $\mathbb{Q}(\sqrt{33})$ and $\mathbb{Q}(\sqrt{-3})$ and therefore $A \cap \mathcal{X}(\mathbb{Z}_p) \neq \varnothing$ (after having fixed embeddings). The fact that

$$A \cap \mathcal{X}(\mathbb{Z}_p)_2 \neq \varnothing,$$

then follows from Proposition 3.4(1) (for the points over $\mathbb{Q}(\sqrt{-11})$ and $\mathbb{Q}(\sqrt{33})$), Proposition 3.12 (for the points over $\mathbb{Q}(\sqrt{-3})$) and the following table, which shows how the reduction changes at the primes in S. The symbol "–" means that the reduction type has not changed. In the last column, there are the possible values of $\lambda_{\mathfrak{q}}(\mathcal{X}(\mathcal{O}_{\mathfrak{q}}))$ where $\mathcal{O}_{\mathfrak{q}}$ is the ring of integers of the completion $K_{\mathfrak{q}}$ at a prime $\mathfrak{q}$ above q in the field K. We briefly explain how the table is computed. When the prime q splits in K, there is nothing to show: $\lambda_{\mathfrak{q}}(\mathcal{X}(\mathcal{O}_{\mathfrak{q}})) = W_q$ if $\mathfrak{q} \mid q$. If q is inert, by Tate's algorithm, (16) is minimal at $\mathfrak{q} \mid q$ and the Kodaira symbol is unchanged. Once we know the Tamagawa number at $\mathfrak{q}$, we can find $\lambda_{\mathfrak{q}}(\mathcal{X}(\mathcal{O}_{\mathfrak{q}}))$ directly from Proposition 2.4. Finally, if q ramifies in K, then $\lambda_{\mathfrak{q}}(\mathcal{X}(\mathcal{O}_{\mathfrak{q}}))$ may be deduced from Proposition 2.4, Lemma 2.1(i) and (4). Note that some points in $\mathcal{X}(\mathcal{O}_{\mathfrak{q}})$ may map to nonintegral points in a minimal model at $\mathfrak{q}$ (see also Lemma 6.4).

K	q	splitting	reduction (Tamagawa)	$\lvert\Delta/\Delta_{\min}\rvert_{\mathfrak{q}}$	$\lambda_{\mathfrak{q}}(\mathcal{X}(\mathcal{O}_{\mathfrak{q}}))$
$\mathbb{Q}(\sqrt{-11})$	2	inert	–	1	W_2
	3	split	–	1	W_3
	11	ramified	good	11^{-6}	W_{11}
$\mathbb{Q}(\sqrt{33})$	2	split	–	1	W_2
	3	ramified	nonsplit I_2 (2)	3^{-6}	W_3
	11	ramified	good	11^{-6}	W_{11}
$\mathbb{Q}(\sqrt{-3})$	2	inert	–	1	W_2
	3	ramified	nonsplit I_2 (2)	3^{-6}	W_3
	11	inert	I_0^* (4)	1	W_{11}

5. The rank 1 case

5A. ***Algebraic nonrational points in $\mathcal{X}(\mathbb{Z}_p)'_2$ in rank 1.*** We retain the notation of Theorems 1.6 and 1.7. The set consisting of the torsion points in $\psi(w)$ is equal to $\phi(w)$. Therefore, the results of Section 3 translate into results for $\mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{tors}}$.

Since each $\psi(w)$ is defined by a single p-adic equation, in most cases it is expected that $\mathcal{X}(\mathbb{Z}_p)'_2$ should be strictly larger than $\mathcal{X}(\mathbb{Z})$. The question we investigate in this subsection is which algebraic nontorsion points could arise in $\mathcal{X}(\mathbb{Z}_p)'_2$. The following elementary lemma shows that if a nontorsion point in $\mathcal{X}(\mathbb{Z}_p)'_2$ comes from a quadratic point in the saturation of $E(\mathbb{Q})$, then its belonging to $\mathcal{X}(\mathbb{Z}_p)'_2$ cannot be explained by automorphisms (cf. Section 3A).

Lemma 5.1. *Let E be an elliptic curve over $\mathbb{Q}$ and K a quadratic field with* $\mathrm{Gal}(K/\mathbb{Q}) = \langle\tau\rangle$. *Let $P \in E(K) \setminus E(\mathbb{Q})$ such that $mP \in E(\mathbb{Q})$ for some nonzero integer m. Then, if $\psi(P) = \tau(P)$ for some $\psi \in \mathrm{Aut}(E/\overline{\mathbb{Q}})$, P has finite order.*

Proof. The hypotheses on P imply that $\psi \neq \mathrm{id}$. Since $mP \in E(\mathbb{Q})$, we have $O = mP - \tau(mP) = m(P - \tau(P))$. If $\psi = -\mathrm{id}$, then

$$\begin{cases}\tau(P) = -P \\ m(P - \tau(P)) = O\end{cases} \iff \begin{cases}\tau(P) = -P \\ 2mP = O;\end{cases}$$

thus, P has order dividing $2m$.

For more general ψ, let

$$[\,\cdot\,] : R \simeq \mathrm{End}(E)$$

where $R \subset \mathbb{C}$. Then there exists a root of unity ζ such that $\psi(P) = [\zeta]P$. Therefore,

$$\begin{cases}\tau(P) = [\zeta]P \\ m(P - \tau(P)) = O\end{cases} \iff \begin{cases}\tau(P) = [\zeta]P \\ [m(1-\zeta)]P = O\end{cases}$$

and so $P \in E(K)[mN_{\mathbb{Q}(\zeta)/\mathbb{Q}}(1-\zeta)]$. □

We could try and use noncyclotomic idele class characters to motivate the existence of some algebraic points of infinite order in $\mathcal{X}(\mathbb{Z}_p)'_2$, following the ideas of Section 3B. For example, what we could hope to prove is that at a certain $z \in \mathcal{X}(\overline{\mathbb{Z}})$ satisfying $mz \in E(\mathbb{Q})$ for some nonzero integer m, the quantity $2D_2(z) + C(\mathrm{Log}(z))^2 + \|w\|$, for some $w \in W$, equals the value of *some* p-adic height function at z. If such a p-adic height comes from a character which restricts to the cyclotomic character on $\mathbb{A}_{\mathbb{Q}}^{\times}$ with the right normalisations, then looking at the equation defining $\psi(w)$ we see that this is enough to show $z \in \mathcal{X}(\mathbb{Z}_p)'_2$.

However, our computations (more on this in Section 5B) also recovered some algebraic nontorsion points defined over real quadratic fields and we know that the space of idele class characters of a real quadratic field is one-dimensional. Therefore, in the following proposition we present a sufficient condition for a point defined over a quadratic field to belong to $\mathcal{X}(\mathbb{Z}_p)'_2$, which looks less geometric or algebraic in nature compared to the results of Section 3. However, we then discuss in Remark 5.3 when we expect the hypotheses of the proposition to be satisfied.

Proposition 5.2. *Suppose that E satisfies the assumptions of Theorem 1.7 and that p is an odd prime of good ordinary reduction. Let K be a quadratic field in which p splits. Fix an embedding $\rho : K \hookrightarrow \mathbb{Q}_p$ and let τ be the nontrivial element in* $\mathrm{Gal}(K/\mathbb{Q})$*. Suppose that $z \in \mathcal{X}(\mathcal{O}_K)$ is such that $mz \in E(\mathbb{Q}) \setminus \{O\}$ for some nonzero integer m and that*

$$f_m(z) = \zeta f_m(\tau(z)), \tag{17}$$

for some root of unity ζ. For each rational prime q, let $\mathfrak{q}$ be one (any) prime of K above q and $\lambda_{\mathfrak{q}}$ the local height at $\mathfrak{q}$ with respect to the model $\mathcal{E}$. If

$$\sum_{q \in S} \lambda_{\mathfrak{q}}(z) = \|w\|$$

for some $w \in W$, then $\rho(z) \in \mathcal{X}(\mathbb{Z}_p)_2'$.

Proof. Let $\mathfrak{q} \nmid p$. By quasiquadraticity (Section 2A1(iv)) applied twice and the assumptions on z and m, we have

$$\begin{aligned}\lambda_{\mathfrak{q}}(z) &= \frac{1}{m^2}\big(\lambda_{\mathfrak{q}}(mz) + 2\log|f_m(z)|_{\mathfrak{q}}\big)\\ &= \frac{1}{m^2}\big(\lambda_{\mathfrak{q}}(\tau(mz)) + 2\log|\zeta f_m(\tau(z))|_{\mathfrak{q}}\big)\\ &= \frac{1}{m^2}\big(\lambda_{\mathfrak{q}}(m\tau(z)) + 2\log|f_m(\tau(z))|_{\mathfrak{q}}\big)\\ &= \lambda_{\mathfrak{q}}(\tau(z)) = \lambda_{\tau(\mathfrak{q})}(z).\end{aligned}$$

Similarly, if $p\mathcal{O}_K = \mathfrak{p}_1\mathfrak{p}_2$, then (without loss of generality)

$$m^2\lambda_{\mathfrak{p}_1}(z) = \lambda_p(mz) + 2\log(\rho(f_m(z))) = \lambda_p(mz) + 2\log(\rho(f_m(\tau(z)))) = m^2\lambda_{\mathfrak{p}_2}(z).$$

Therefore,

$$h_p(z) = \lambda_p(\rho(z)) + \|w\|.$$

Since z is in the saturation of $E(\mathbb{Q})$ (i.e., $mz \in E(\mathbb{Q})$), then $z \in \psi(w)$. □

Remark 5.3. Let E, K and p be as in Proposition 5.2. Suppose that $z \in \mathcal{X}(\mathcal{O}_K)$ is such that $mz \in E(\mathbb{Q}) \setminus \{O\}$ and write $x(mz) = n(mz)/(d(mz)^2)$ for some coprime integers $n(mz)$ and $d(mz) > 0$. When can we expect (17) to hold? By [Wuthrich 2004, §2] and our assumptions, we know that

$$\frac{g_m(z)}{f_m(z)^2} = x(mz) = x(m\tau(z)) = \frac{g_m(\tau(z))}{f_m(\tau(z))^2},$$

where $g_m(z)$ can be written as a univariate polynomial in $x(z)$ over $\mathbb{Z}$. For $w \in \{z, \tau(z)\}$, define

$$\delta_m(w) = \frac{f_m(w)}{d(mw)}.$$

By Proposition 1 of [loc. cit.], $\delta_m(w)$ is a unit at all primes at which w has nonsingular reduction. Furthermore, we have

$$\frac{\delta_m(z)}{\tau(\delta_m(\tau(z)))} = \frac{f_m(z)}{\tau(f_m(\tau(z)))} = 1.$$

Thus, if for example z has good reduction at all primes which are split in K, then $\tau(\delta_m(\tau(z))) = \delta_m(\tau(z))$ up to multiplication by elements of $\mathcal{O}_K^\times$: thus, in this case, $f_m(z) = uf_m(\tau(z))$ for some $u \in \mathcal{O}_K^\times$. If K is imaginary then u is a root of unity; otherwise u may or may not be. Note that if K is imaginary, we could have also avoided talking about division polynomials and followed a strategy similar to Proposition 3.12. Conversely, Proposition 5.2 is often not applicable for torsion points as it requires the existence of a *nonzero* multiple of z in $E(\mathbb{Q})$.

5B. ***Computations in rank* 1.** The technique explained in Section 4A to compute $\mathcal{X}(\mathbb{Z}_p)_2$ in the rank 0 case can easily be adapted to compute $\mathcal{X}(\mathbb{Z}_p)_2'$ when the Mordell–Weil group has rank 1. As remarked in Section 5A, the expectation is that $\mathcal{X}(\mathbb{Z}_p)_2'$ should generally be larger than $\mathcal{X}(\mathbb{Z})$.

We ran the code on all the 14783 rank 1 elliptic curves of conductor at most 5000 and let p be the smallest prime greater than or equal to 5 at which the curve has good ordinary reduction. The first observation is that it can happen that there are no points in $\mathcal{X}(\mathbb{Z}_p)_2'$ beside those in $\mathcal{X}(\mathbb{Z})$. For example, the curves of conductor at most 500

- satisfying the assumptions of Theorem 1.7(1) are 254.b1, 430.c1 [LMFDB 2019];
- not satisfying Theorem 1.7(1), but for which $\mathcal{X}(\mathbb{Z}_p)_2' = \mathcal{X}(\mathbb{Z})$ are 297.b1, 325.b1, 325.b2, 467.a1 [LMFDB 2019].

Studying the torsion points of $\mathcal{X}(\mathbb{Z}_p)_2'$ morally provides more data on the extra points that can arise in $\mathcal{X}(\mathbb{Z}_p)_2$ when E has rank 0. No new phenomenon was observed, except that torsion points defined over some degree 4 number fields were also recovered on the two CM elliptic curves 576.e1 and 576.e2 [LMFDB 2019] of j-invariant 54000. The appearance of the latter points can be proved in a similar way to Proposition 3.14.

As far as algebraic nontorsion points are concerned, on 26 curves we identified nontorsion points defined over quadratic extension of $\mathbb{Q}$. All of these were explained by Proposition 5.2 and only on two curves the points were defined over real quadratic fields. We now present an example in which some algebraic torsion and nontorsion points were recovered in $\mathcal{X}(\mathbb{Z}_p)_2' \setminus \mathcal{X}(\mathbb{Z})$. Afterwards, we also include for completeness an example in which the extra algebraic nontorsion points are real.

Example 5.4. Consider the elliptic curve 576.e4 [LMFDB 2019]

$$E : y^2 = x^3 + 8,$$

whose Mordell–Weil group over $\mathbb{Q}$ has rank 1 and is generated, modulo torsion, by the point $z_0 = (1, 3)$. Let $p = 7$, at which E has good ordinary reduction. We have $S = \{2, 3\}$ and $W = W_2 \times W_3$ where

$$W_2 = \{0, -\log 2\} \quad \text{and} \quad W_3 = \left\{0, -\tfrac{1}{2}\log 3\right\}.$$

Write

$$\mathcal{X}(\mathbb{Z}_p)'_2 = \mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{tors}} \cup \mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{nontors}},$$

where the subscripts tors and nontors have the obvious meaning. Let $K = \mathbb{Q}(\sqrt{-3})$ and let τ generate the Galois group of $K/\mathbb{Q}$. Assuming that we have fixed an embedding of $\mathbb{Q}(\sqrt{-3})$ into $\mathbb{Q}_p$, we find that

$$\begin{aligned}\mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{tors}} &= \{(-2, 0), (1 \pm \sqrt{-3}, 0)\},\\ \mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{nontors}} &= \pm\{(1, 3), (2, -4), (46, -312), (-5 \pm \sqrt{-3}, 6 \pm 6\sqrt{-3})\} \cup A^{\mathrm{nonalg?}},\end{aligned}$$

where $A^{\mathrm{nonalg?}}$ denotes the set of points of $\mathcal{X}(\mathbb{Z}_p)'_2$ which have not been recognised as algebraic. Note that $A^{\mathrm{nonalg?}}$ modulo $\pm$ consists of 15 points. Corollary 3.5, together with the observation at the beginning of this section, proves why the two-torsion quadratic points $(1 \pm \sqrt{-3}, 0)$ belong to $\mathcal{X}(\mathbb{Z}_p)'_2$.

Consider now

$$Q \in \{\pm(-5 \pm \sqrt{-3}, 6 \pm 6\sqrt{-3})\}.$$

We show why $Q \in \mathcal{X}(\mathbb{Z}_p)'_{2,\mathrm{nontors}}$. Without loss of generality we assume that $Q = (-5+\sqrt{-3}, 6+6\sqrt{-3})$. As

$$Q = -z_0 + (1 - \sqrt{-3}, 0),$$

$2Q \in E(\mathbb{Q})$. We have

$$f_2(Q) = 2y(Q) = \left(\tfrac{1}{2}(-1 + \sqrt{-3})\right) f_2(\tau(Q)).$$

Therefore, in order to apply Proposition 5.2, it suffices to verify the condition on the local heights at the bad primes. For each prime $\mathfrak{q} \nmid p$, we could use the formula involving $f_2(Q)$ in order to compute $\lambda_{\mathfrak{q}}(Q)$, as in the proof of the proposition. We choose to compute it instead by the quasiparallelogram law

$$\lambda_{\mathfrak{q}}(Q) = \lambda_{\mathfrak{q}}(z_0) + \lambda_{\mathfrak{q}}(1 - \sqrt{-3}, 0) - \log|\sqrt{-3}|_{\mathfrak{q}} = \lambda_{\mathfrak{q}}(z_0) + \lambda_{\mathfrak{q}}(-2, 0) - \log|\sqrt{-3}|_{\mathfrak{q}},$$

which gives

$$\lambda_{\mathfrak{q}}(Q) = \begin{cases} -\frac{1}{2}\log 3 & \text{if } \mathfrak{q} \mid 3,\\ -\log 2 & \text{if } \mathfrak{q} \mid 2,\\ 0 & \text{if } \mathfrak{q} \nmid 2, 3, p.\end{cases}$$

The fact that Q is in $\mathcal{X}(\mathbb{Z}_p)'_2$ then follows from Proposition 5.2.

Example 5.5. Consider the elliptic curve 525.c1 [LMFDB 2019]

$$E : y^2 + xy = x^3 + x^2 - 450x + 3375.$$

Let p be an odd prime of good ordinary reduction split in $\mathbb{Q}(\sqrt{5})$ and fix an embedding $\mathbb{Q}(\sqrt{5}) \hookrightarrow \mathbb{Q}_p$. Then by Proposition 5.2 with $m = 2$, the infinite order points

$$\pm\left(10 \pm 5\sqrt{5}, \tfrac{5}{2}(23 \mp \sqrt{5})\right)$$

belong to $\mathcal{X}(\mathbb{Z}_p)'_2$.

6. Rational points on bielliptic curves

Let C be a smooth projective curve over $\mathbb{Q}$ of genus g and whose Jacobian J has Mordell–Weil rank equal to g. Assume in addition that the Néron–Severi group of J has rank at least equal to 2, that p is an odd prime of good reduction for C and that the p-adic closure of $J(\mathbb{Q})$ has finite index in $J(\mathbb{Q}_p)$. Balakrishnan and Dogra [2018, Theorem 1.2] used the Chabauty–Kim method to explicitly describe a finite set

$$C(\mathbb{Q}_p)_Z \subset C(\mathbb{Q}_p),$$

which depends upon the choice of a correspondence $Z \subset C \times C$ and which contains all the $\mathbb{Q}$-rational points of C. Note that one has $C(\mathbb{Q}_p)_2 \subset C(\mathbb{Q}_p)_Z$. The authors then made Theorem 1.2 algorithmic when C is a bielliptic curve of genus 2, under the extra assumption that J is ordinary at p (we remove this hypothesis here). Let

$$C : y^2 = x^6 + a_4x^4 + a_2x^2 + a_0, \qquad a_i \in \mathbb{Q},$$

be a genus 2 bielliptic curve with a rank-2 Jacobian and consider the associated maps $\varphi_i : C \to (E_i, O_{E_i})$, described affinely by

$$\begin{aligned} E_1 &: y^2 = x^3 + a_4x^2 + a_2x + a_0, &\quad \varphi_1(x, y) &= (x^2, y) \\ E_2 &: y^2 = x^3 + a_2x^2 + a_4a_0x + a_0^2, &\quad \varphi_2(x, y) &= (a_0x^{-2}, a_0yx^{-3}). \end{aligned}$$

Assume that each of E_1 and E_2 has rank 1. Then one may pick Z in such a way that $C(\mathbb{Q}_p)_Z$ can be described in terms of local and global p-adic heights of the images of the points of C in the two elliptic curves (we assume that the given equations for E_1 and E_2 are minimal at p). The superscript ^{E_i} indicates on which curve we are computing these quantities. Let P_i be a point of infinite order in $E_i(\mathbb{Q})$ and write

$$c_i = \frac{h_p^{E_i}(P_i)}{\mathrm{Log}^{E_i}(P_i)^2},$$

where, as usual, $h_p^{E_i}$ is the global p-adic height of Mazur–Tate or Bernardi depending on whether the reduction is ordinary or not. Let $Q_1 = (0, \sqrt{a_0}) \in E_1(\mathbb{Q}(\sqrt{a_0}))$ and $Q_2 = (0, a_0) \in E_2(\mathbb{Q})$. We assume that a_0 is a square in $\mathbb{Q}_p$. Furthermore, let

$$\begin{aligned} C^{(1)}(\mathbb{Q}_p) &= C(\mathbb{Q}_p) \setminus (](0, \sqrt{a_0})[\cup](0, -\sqrt{a_0})[), \\ C^{(2)}(\mathbb{Q}_p) &= C(\mathbb{Q}_p) \setminus (]\infty^+[\cup]\infty^-[), \\ C^{(i)}(\mathbb{Q}) &= C(\mathbb{Q}) \cap C^{(i)}(\mathbb{Q}_p) \quad \text{for } i = 1, 2, \end{aligned}$$

where the inverted square brackets denote the residue disk modulo p around the given point and $\infty^\pm = (1 : \pm 1 : 0) \in C(\mathbb{Q})$.

Theorem 6.1 (Balakrishnan–Dogra[10]). *For each $i \in \{1, 2\}$, the following set is finite*:

$$W^i = \left\{ \sum_{q \neq p} \left(\lambda_q^{E_i}(\varphi_i(z_q) + Q_i) + \lambda_q^{E_i}(\varphi_i(z_q) - Q_i) - 2\lambda_q^{E_{3-i}}(\varphi_{3-i}(z_q)) \right) : (z_q) \in \prod_{q \neq p} C(\mathbb{Q}_q) \setminus \{\varphi_i^{-1}(\pm Q_i)\} \right\}.$$

[10] With a small correction; see Remark 6.2.

Furthermore,

$$C^{(i)}(\mathbb{Q}) \subset \Big\{ z \in C^{(i)}(\mathbb{Q}_p) : 2\lambda_p^{E_{3-i}}(\varphi_{3-i}(z)) - \lambda_p^{E_i}(\varphi_i(z)+Q_i) - \lambda_p^{E_i}(\varphi_i(z)-Q_i) \\ - 2c_{3-i}\operatorname{Log}^{E_{3-i}}(\varphi_{3-i}(z))^2 + 2c_i \operatorname{Log}^{E_i}(\varphi_i(z))^2 + 2h_p^{E_i}(Q_i) \in W^i \Big\}.$$

The second assertion in Theorem 6.1 can be derived from the parallelogram law satisfied by the global p-adic height (as a consequence of Section 2A1(iii) and Section 2A2(iii)) and the fact that there is at most one quadratic function (up to multiplication by a scalar) on each of $E_1(\mathbb{Q})$ and $E_2(\mathbb{Q})$, due to the assumption on their ranks, in analogy to the proof of Theorem 1.7. The set of p-adic points described in the theorem is $C(\mathbb{Q}_p)_Z \cap C^{(i)}(\mathbb{Q}_p)$. We will explicitly determine the sets W^i for Example 6.3. For a general bielliptic curve we will give in Proposition 6.5 a description of a finite set containing W^i, hence giving a proof of finiteness of W^i.

Remark 6.2. If Q_1 is not defined over $\mathbb{Q}$, with the notation $\lambda_q^{E_1}(\varphi_1(z) \pm Q_1)$ in Theorem 6.1 we mean $\lambda_{\mathfrak{q}}^{E_1}(\varphi_1(z) \pm Q_1)$, where $\mathfrak{q}$ is any prime of $\mathbb{Q}(\sqrt{a_0})$ lying above the rational prime q. Indeed, since $\tau(Q_1) = -Q_1$, where $\langle \tau \rangle = \mathrm{Gal}(\mathbb{Q}(\sqrt{a_0})/\mathbb{Q})$, there is no dependence on $\mathfrak{q} \mid q$ and

$$h_p^{E_1}(Q_1) = \sum_q \lambda_q^{E_1}(Q_1).$$

The equations defining the sets $C(\mathbb{Q}_p)_Z \cap C^{(1)}(\mathbb{Q}_p)$ in [Balakrishnan and Dogra 2018, Corollary 8.1] contain a typo in the case when Q_1 is not in the saturation of $E_1(\mathbb{Q})$, which we have corrected in Theorem 6.1. Their formula has the term $2c_1 \operatorname{Log}^{E_1}(Q_1)^2$ in place of $2h_p^{E_1}(Q_1)$ and does not hold unless the two quantities are equal.

One of the advantages of computing rational points using Theorem 6.1 for a bielliptic curve, rather than the more general techniques developed in [Balakrishnan et al. 2019a], is that we do not need to have prior knowledge of any affine point in $C(\mathbb{Q})$.

Balakrishnan, Dogra and Müller [Balakrishnan and Dogra 2018] used Theorem 6.1, combined with the Mordell–Weil sieve, to determine precisely the rational points of two bielliptic curves. As in Section 4A, we suggest here that one can replace the computations of double Coleman integrals with computations involving the p-adic sigma function and division polynomials. We use the resulting algorithm to compute $C(\mathbb{Q}_p)_Z$ for a bielliptic curve whose rational points were already found using different Chabauty-type techniques by Wetherell [1997, Proposition 5.1] and Flynn and Wetherell [1999, Example 3.1]. Our methods lead to an alternative provable determination of $C(\mathbb{Q})$.

Example 6.3. Consider the bielliptic curve

$$C : y^2 = x^6 + x^2 + 1;$$

the associated elliptic curves are 496.a1 and 248.a1 [LMFDB 2019], given by the minimal models

$$E_1 : y^2 = x^3 + x + 1, \qquad E_2 : y^2 = x^3 + x^2 + 1.$$

We have

$$Q_1 = (0, 1) \in E_1(\mathbb{Q}), \qquad Q_2 = (0, 1) \in E_2(\mathbb{Q}).$$

Both elliptic curves have rank 1. Let $p = 3$, which is a prime of good reduction for C. Since E_1 is supersingular at p, we use Bernardi's p-adic height for our calculations.

Claim 1. *If $q \neq p, 2$ and $z \in C(\mathbb{Q}_q) \setminus \{\varphi_i^{-1}(\pm Q_i)\}$, then*

$$w_{q,i}(z) := \lambda_q^{E_i}(\varphi_i(z) + Q_i) + \lambda_q^{E_i}(\varphi_i(z) - Q_i) - 2\lambda_q^{E_{3-i}}(\varphi_{3-i}(z)) = 0.$$

Proof. The curves E_1 and E_2 have everywhere good reduction except at 2 and 31. At 31, the Tamagawa number of each E_i is trivial. Assume first that $\varphi_i(z) \neq O_{E_i}$. Then, for each $q \neq p$, the quasiparallelogram law (3) gives

$$w_{q,i}(z) = 2\bigl(\lambda_q^{E_i}(\varphi_i(z)) + \lambda_q^{E_i}(Q_i) - \log|x(\varphi_i(z))|_q - \lambda_q^{E_{3-i}}(\varphi_{3-i}(z))\bigr);$$

thus, if $q \neq 2$, by Lemma 2.1(i),

$$\begin{aligned} w_{q,i}(z) &= 2(\log(\max\{1, |x(\varphi_i(z))|_q\}) - \log(\max\{1, |x(\varphi_{3-i}(z))|_q\}) - \log|x(\varphi_i(z))|_q) \\ &= 2(\log(\max\{1, |x(\varphi_i(z))|_q\}) - \log(\max\{1, |x(\varphi_i(z))|_q^{-1}\}) - \log|x(\varphi_i(z))|_q) = 0. \end{aligned}$$

It remains to consider the case $\varphi_i(z) = O_{E_i}$. Then $\varphi_{3-i}(z) \in \{\pm Q_{3-i}\}$ and

$$w_{q,i}(z) = 2\lambda_q^{E_i}(Q_i) - 2\lambda_q^{E_{3-i}}(Q_{3-i}) = 0$$

by Proposition 2.2, since Q_i and Q_{3-i} are integral and the Tamagawa numbers at q are equal to 1. □

Claim 2. *We have*

$$W^1 = \{0, \log 2\}, \qquad W^2 = \{-\log 2, -2\log 2\}.$$

Proof. By Claim 1,

$$W^i = \bigl\{w_{2,i}(z) := \lambda_2^{E_i}(\varphi_i(z) + Q_i) + \lambda_2^{E_i}(\varphi_i(z) - Q_i) - 2\lambda_2^{E_{3-i}}(\varphi_{3-i}(z)) : z \in C(\mathbb{Q}_2) \setminus \{\varphi_i^{-1}(\pm Q_i)\}\bigr\}.$$

First note that the curve E_1 has Tamagawa number equal to 1 at 2, whereas E_2 has Tamagawa number equal to 2 and reduction type III. If $|x(z)|_2 \le 1$, then $\varphi_2(z)$ has good reduction at 2 and $\lambda_2^{E_2}(\varphi_2(z)) = \log|x(z)^{-2}|_2$; otherwise $\varphi_2(z)$ reduces to a singular point modulo 2 and, by Proposition 2.4, $\lambda_2^{E_2}(\varphi_2(z)) = -\frac{1}{2}\log 2$. Furthermore $\lambda_2^{E_2}(Q_2) = -\frac{1}{2}\log 2$. Therefore, similarly to the proof of Claim 1, if $\varphi_1(z) \neq O_{E_1}$, then by the quasiparallelogram law we have

$$w_{2,1}(z) = 2\bigl(\lambda_2^{E_1}(\varphi_1(z)) - \lambda_2^{E_2}(\varphi_2(z)) - \log|x(\varphi_1(z))|_2\bigr) = \begin{cases} 0 & \text{if } |x(z)|_2 \le 1, \\ \log 2 & \text{if } |x(z)|_2 > 1; \end{cases}$$

for the remaining points $z = \infty^{\pm}$ we get

$$w_{2,1}(z) = -2\lambda_2^{E_2}(Q_2) = \log 2,$$

thus proving the claim for W^1. The set W^2 is determined in a very similar fashion and we leave the details to the reader. □

We can now compute $C(\mathbb{Q}_p)_Z$ as a union of the two $C(\mathbb{Q}_p)_Z \cap C(\mathbb{Q}_p)^{(i)}$. We find

$$C(\mathbb{Q}_p)_Z = \left\{\infty^{\pm}, (0, \pm 1), \left(\pm\tfrac{1}{2}, \pm\tfrac{9}{8}\right)\right\} \sqcup A \tag{18}$$

where A is a set of size 4. Note, however, that up to the automorphisms $(x, y) \mapsto (-x, y)$, $(x, y) \mapsto (x, -y)$ and their composites, A actually consists of one point:

$$P = (2\cdot 3 + 2\cdot 3^3 + 2\cdot 3^5 + 3^8 + O(3^9), 2 + 2\cdot 3 + 2\cdot 3^3 + 2\cdot 3^5 + 3^6 + 2\cdot 3^8 + O(3^9)).$$

We follow the same strategy to the one of the proof of [Balakrishnan and Dogra 2018, Theorem 8.6] to rule out the possibility that P could be rational. The image of the point P under φ_2 is a point in $E_2(\mathbb{Q}_p)$ whose x-coordinate has valuation $\operatorname{ord}_p(x(\varphi_2(P))) = -2$. On the other hand, the Mordell–Weil group $E_2(\mathbb{Q}) \cong \mathbb{Z}$ is generated by Q_2 and, thus, if $\varphi_2(P) \in E_2(\mathbb{Q})$, there must exist a multiple of Q_2 whose x-coordinate has p-adic valuation equal to -2. As the smallest multiple of Q_2 in the formal group at p is $6Q_2 = \left(\frac{55}{81}, -\frac{971}{729}\right)$, we have reached a contradiction, since the set of points in the formal group whose x-coordinate has valuation at most -4 is a group.

6A. ***Explicit formulae for the sets W^i.*** Theorem 6.1 asserts that the sets W^i, for $i = 1, 2$, are finite, but does not describe them explicitly. In order to obtain an implementation for the computations of the sets $C(\mathbb{Q}_p)_Z \cap C^{(i)}(\mathbb{Q}_p)$ for an arbitrary genus 2 bielliptic curve C, it would be convenient to have a characterisation of W^i that can be made algorithmic, in analogy with that of the sets W_q of Theorems 1.6 and 1.7.

We assume in this section that the coefficients a_0, a_2, a_4 defining C are in $\mathbb{Z}$.

For $i = 1, 2$, let $W_q^{E_{i,\min}}$ be the set W_q from Section 1 for a global minimal model $E_{i,\min}$ of E_i if q is a prime of bad reduction and with nontrivial Tamagawa number. If $E_{i,\min}$ has good reduction at $q \in \{2, 3\} \setminus \{p\}$ and $\bar{E}_{i,\min}(\mathbb{F}_q) = \{O\}$ or $q = 2$ and $E_{i,\min}$ has split multiplicative reduction I_1 at q, let $W_q^{E_{i,\min}} = \varnothing$. For all other $q \neq p$, let $W_q^{E_{i,\min}} = \{0\}$.

Let $W_q^{E_i}$ be the set of values attained by $\lambda_q^{E_i}$ on the points of $E_i(\mathbb{Q}_q)$ of the form (x, y) with $x, y \in \mathbb{Z}_q$. Let

$$V_q^{E_i} = W_q^{E_i} \cup \{0\}.$$

Write

$$\delta^{E_i} = \frac{\Delta^{E_i}}{\Delta^{E_{i,\min}}},$$

where Δ^{E_i} and $\Delta^{E_{i,\min}}$ are the discriminants of E_i and $E_{i,\min}$.

For sets A, B of elements in a field F, we write $A + B$ for their Minkowski addition and $-A$ for the set consisting of the additive inverses of the elements in A. If $A = \{a\}$, we write $a + B$ for $A + B$.

Lemma 6.4.

$$W_q^{E_i} = \tfrac{1}{6}\log|\delta^{E_i}|_q + \left(W_q^{E_{i,\min}} \cup \left\{2k\log q : 1 \le k \le \tfrac{1}{12}\operatorname{ord}_q(\delta^{E_i})\right\}\right).$$

In particular, $W_q^{E_i}$ *is finite*; *it equals* $\{0\}$ *for all but finitely many* q.

Proof. Let x, y and $x_{\min}$, $y_{\min}$ be the Weierstrass coordinates for E_i and $E_{i,\min}$, respectively. Then there exist $u, r, s, t \in \mathbb{Q}$, $u \neq 0$, such that

$$x = u^2 x_{\min} + r, \quad y = u^3 y_{\min} + s u^2 x_{\min} + t.$$

Since $\operatorname{ord}_q(\Delta^{E_{i,\min}}) \leq \operatorname{ord}_q(\Delta^{E_i})$ for each prime q, the scalars u, r, s, t are furthermore all integral (see [Connell 1999, Lemma 5.3.1]).

Now let $P \in E_i(\mathbb{Q}_q)$ such that $x(P), y(P) \in \mathbb{Z}_q$. By (4), we have

$$\lambda_q^{E_i}(P) = \lambda_q^{E_{i,\min}}(P) + \tfrac{1}{6} \log |\delta^{E_i}|_q.$$

If $x_{\min}(P) \in \mathbb{Z}_q$, then $\lambda_q^{E_{i,\min}}(P) \in W_q^{E_{i,\min}}$ by Propositions 2.4, 2.2 and Lemma 2.3. Otherwise, by Lemma 2.1(i), we have $\lambda_q^{E_{i,\min}}(P) = \log |(x(P) - r)/u^2|_q$, where by assumption $|x(P) - r|_q \leq 1$ and the valuation of $(x(P) - r)/u^2$ is an even negative integer. Since $\delta^{E_i} = u^{12}$, this completes the proof of $\subseteq$. To see why the inclusion is actually an equality, notice that the preimage of an integral point in $E_{i,\min}(\mathbb{Q}_q)$ is certainly an integral point on $E_i(\mathbb{Q}_q)$. Furthermore, the points on $E_{i,\min}(\mathbb{Q}_q)$ in the formal group are parametrised by $t \in q\mathbb{Z}_q$ as follows: $t \mapsto (x_{\min}(t), y_{\min}(t))$, where

$$x_{\min}(t) = \frac{1}{t^2} - \frac{a_{1,\min}}{t} - a_{2,\min} - a_{3,\min} t + \cdots \in \mathbb{Z}[a_{1,\min}, \ldots, a_{6,\min}]((t)),$$

where $a_{1,\min}, \ldots, a_{6,\min}$ are the Weierstrass coefficients of $E_{i,\min}$; in particular, we have $\operatorname{ord}_q(x_{\min}(t)) = -2 \operatorname{ord}_q(t)$. Thus for each $1 \leq k \leq \frac{1}{12} \operatorname{ord}_q(\delta^{E_i})$, setting $t = q^k$ gives a point on $E_{i,\min}(\mathbb{Q}_q)$ whose preimage in $E_i(\mathbb{Q}_q)$ has integral x-coordinate. The second assertion of the lemma follows from the explicit description of the sets $W_q^{E_{i,\min}}$. □

Proposition 6.5. *With the notation of Theorem 6.1, suppose that* $\operatorname{ord}_\ell(a_0) \in \{0, 1\}$ *for each prime* ℓ. *For each prime* $q \neq p$, *let*

$$W_q^{1\prime} = \{2v + 2\lambda_q^{E_1}(Q_1) : v \in V_q^{E_1} + (-W_q^{E_2})\} \cup \{2v + 2\lambda_q^{E_1}(Q_1) - 2\log|a_0|_q : v \in W_q^{E_1}\}$$
$$W_q^{2\prime} = \{2v + 2\lambda_q^{E_2}(Q_2) - 2\log|a_0|_q : v \in W_q^{E_2} + (-V_q^{E_1})\} \cup \{2v + 2\lambda_q^{E_2}(Q_2) : v \in -W_q^{E_1}\}.$$

Then W^i *is a subset of the finite set*

$$W^{i\prime} = \left\{\sum_{q \neq p} w'_{q,i} : w'_{q,i} \in W_q^{i\prime}\right\} = \left\{2h_p^{E_i}(Q_i) - 2\lambda_p^{E_i}(Q_i) + \sum_{q \neq p} (w'_{q,i} - 2\lambda_q^{E_i}(Q_i)) : w'_{q,i} \in W_q^{i\prime}\right\}.$$

Proof. By definition,

$$W^i = \left\{\sum_{q \neq p} w_{q,i} : w_{q,i} \in W_q^i\right\},$$

where

$$W_q^i = \{w_{q,i}(z) := \lambda_q^{E_i}(\varphi_i(z) + Q_i) + \lambda_q^{E_i}(\varphi_i(z) - Q_i) - 2\lambda_q^{E_{3-i}}(\varphi_{3-i}(z)) : z \in C(\mathbb{Q}_q) \setminus \{\varphi_i^{-1}(\pm Q_i)\}\}.$$

Let $z \in C(\mathbb{Q}_q) \setminus \{\varphi_i^{-1}(\pm Q_i)\}$. If $\varphi_i(z) = O_{E_i}$, then

$$w_{q,i}(z) = 2\lambda_q^{E_i}(Q_i) - 2\lambda_q^{E_{3-i}}(Q_{3-i}),$$

which belongs to $2\lambda_q^{E_i}(Q_i) - 2W_q^{E_{3-i}}$. When $i = 2$, note that, while it is not always the case that Q_1 is defined over $\mathbb{Q}_q$ (and hence that $\lambda_q^{E_1}(Q_1) \in W_q^{E_1}$), here this follows from the assumption that its preimage under φ_1 is in $C(\mathbb{Q}_q)$.

Otherwise, by the quasiparallelogram law (3), we have

$$\begin{aligned} w_{q,i}(z) &= 2\big(\lambda_q^{E_i}(\varphi_i(z)) + \lambda_q^{E_i}(Q_i) - \log|x(\varphi_i(z))|_q - \lambda_q^{E_{3-i}}(\varphi_{3-i}(z))\big) \\ &= 2\big(\lambda_q^{E_i}(\varphi_i(z)) + \lambda_q^{E_i}(Q_i) + \log|x(\varphi_{3-i}(z))|_q - \log|a_0|_q - \lambda_q^{E_{3-i}}(\varphi_{3-i}(z))\big). \end{aligned}$$

Note that the assumption that $0 \le \operatorname{ord}_q(a_0) \le 1$ implies that, for each $i = 1, 2$,

$$|x(\varphi_i(z))|_q > 1 \Rightarrow |x(\varphi_{3-i}(z))|_q < 1 \tag{19}$$

$$|x(\varphi_{3-i}(z))|_q < 1 \Rightarrow |x(\varphi_i(z))|_q \ge 1. \tag{20}$$

If $|x(\varphi_i(z))|_q > 1$, then $\varphi_i(z)$ reduces to a nonsingular point modulo q, with respect to the Weierstrass equation for E_i. Thus $\lambda_q^{E_i}(\varphi_i(z)) = \log|x(\varphi_i(z))|_q$. Furthermore, by (19), $\varphi_{3-i}(z)$ is integral with respect to the Weierstrass equation defining E_{3-i} and we have $\lambda_q^{E_{3-i}}(\varphi_{3-i}(z)) \in W_q^{E_{3-i}}$. Therefore

$$\frac{w_{q,i}(z)}{2} \in \lambda_q^{E_i}(Q_i) + (-W_q^{E_{3-i}}).$$

Similarly, if $|x(\varphi_{3-i}(z))|_q > 1$, then

$$\frac{w_{q,i}(z)}{2} \in (\lambda_q^{E_i}(Q_i) - \log|a_0|_q) + W_q^{E_i}.$$

It remains to consider the case when $|x(z)|_q = 1$. Then

$$\frac{w_{q,i}(z)}{2} \in W_q^{E_i} + (-W_q^{E_{3-i}}) + \begin{cases} \lambda_q^{E_i}(Q_i) & \text{if } i = 1, \\ \lambda_q^{E_i}(Q_i) - \log|a_0|_q & \text{if } i = 2. \end{cases} \qquad \square$$

Proposition 6.5 and Lemma 6.4 turn Theorem 6.1 into an algorithm for computing a finite set of p-adic points containing $C(\mathbb{Q})$, for an arbitrary genus 2 bielliptic curve C whose associated elliptic curves E_1 and E_2 have Mordell–Weil rank equal to 1. Furthermore, to improve the estimates of the sets W^i provided by Proposition 6.5 we may use the fact that the contributions at primes of potential good reduction for C are trivial (cf. [Balakrishnan and Dogra 2018, Theorem 1.2(i)]). It seems unlikely to the author that the elementary approach of Proposition 6.5 could show the latter for an arbitrary curve, since it is hard to imagine how the proof could be made sensitive to the difference between C being of potential good reduction and its Jacobian only being of potential good reduction. Furthermore, even at primes not of potential good reduction, the sets $W_q^{i\prime}$ might be larger than W_q^i. Nevertheless, having fixed an explicit curve C, the steps of the proof of the proposition should guide the reader through computing W^i precisely.

Here we are instead interested in an algorithm which does not require prior computations of W^i. So let $W^{i''}$ be obtained from $W^{i'}$ of Proposition 6.5 by replacing $W_q^{i'}$ with $\{0\}$ whenever q is a prime of potential good reduction. We implemented in `SageMath` the results of this section and could test them for several bielliptic curves, including the ones of [Balakrishnan and Dogra 2018] and the bielliptic curve

$$C : y^2 = (x^2 + 1)(x^2 + 3)(x^2 + 7), \tag{21}$$

which appears in [Flynn and Wetherell 1999, p. 532] as the only curve amongst 50 bielliptic curves for which the methods of Flynn and Wetherell to find rational points failed.

For instance, in Example 6.3 and for the curve of [Balakrishnan and Dogra 2018, §8.3] we have $W^i = W^{i'} = W^{i''}$ for each $i = 1$ and $i = 2$, but for the curve of [Balakrishnan and Dogra 2018, §8.4], the sets $W^{i''}$ have size 3, whereas W^i has size 1.

We also remark that in Example 6.3, as well as the two examples of [Balakrishnan and Dogra 2018], the elliptic curve E_1 has trivial Tamagawa number at all primes and E_2 has trivial Tamagawa numbers everywhere except for at one prime where it has Tamagawa number equal to two or three. In other words, finding precise expressions for the sets W^i by hand is straightforward. In Example 6.3 as well as [Balakrishnan and Dogra 2018, §8.3] the task is further simplified by the fact that the Weierstrass equations for E_1 and E_2 have minimal discriminant. On the other hand, we cannot expect this for a generic curve. For example, the elliptic curves corresponding to (21) have Tamagawa numbers (2, 1) respectively (4, 2) at the primes not of potential good reduction and analysing what happens at each prime by hand might be rather tedious. In this case, we find $\#W^{1''} = \#W^{2''} = 18$ and in fact this results in many points in $C(\mathbb{Q}_p)$ that are probably not rational (142 when $p = 5$).

Acknowledgements

I am very grateful to Jennifer Balakrishnan for many useful discussions and for providing me with helpful feedback on several drafts. This paper started from questions raised at group meetings organised by Minhyong Kim in which I took part together with Jamie Beacom, Noam Kantor and Alex Saad. I am indebted to all of them and thank Minhyong for encouraging me to write up. I also thank the anonymous referee, as well as Steffen Müller, and my DPhil examiners Victor Flynn and Chris Wuthrich, for several useful suggestions, which led to many improvements to the article. I was supported by EPSRC and by Balliol College through a Balliol Dervorguilla scholarship, and by an NWO Vidi grant.

References

[Balakrishnan 2016] J. S. Balakrishnan, "On 3-adic heights on elliptic curves", *J. Number Theory* **161** (2016), 119–134. MR Zbl

[Balakrishnan and Besser 2015] J. S. Balakrishnan and A. Besser, "Coleman–Gross height pairings and the p-adic sigma function", *J. Reine Angew. Math.* **698** (2015), 89–104. MR Zbl

[Balakrishnan and Dogra 2018] J. S. Balakrishnan and N. Dogra, "Quadratic Chabauty and rational points, I: p-adic heights", *Duke Math. J.* **167**:11 (2018), 1981–2038. With an appendix by J. S. Müller. MR Zbl

[Balakrishnan et al. 2010] J. S. Balakrishnan, R. W. Bradshaw, and K. S. Kedlaya, "Explicit Coleman integration for hyperelliptic curves", pp. 16–31 in *Algorithmic number theory* (Nancy, France, 2010), edited by G. Hanrot et al., Lecture Notes in Comput. Sci. **6197**, Springer, 2010. MR Zbl

[Balakrishnan et al. 2016] J. S. Balakrishnan, A. Besser, and J. S. Müller, "Quadratic Chabauty: p-adic heights and integral points on hyperelliptic curves", *J. Reine Angew. Math.* **720** (2016), 51–79. MR Zbl

[Balakrishnan et al. 2018] J. S. Balakrishnan, I. Dan-Cohen, M. Kim, and S. Wewers, "A non-abelian conjecture of Tate–Shafarevich type for hyperbolic curves", *Math. Ann.* **372**:1-2 (2018), 369–428. MR Zbl

[Balakrishnan et al. 2019a] J. Balakrishnan, N. Dogra, J. S. Müller, J. Tuitman, and J. Vonk, "Explicit Chabauty–Kim for the split Cartan modular curve of level 13", *Ann. of Math.* (2) **189**:3 (2019), 885–944. MR Zbl

[Balakrishnan et al. 2019b] J. S. Balakrishnan, A. Besser, F. Bianchi, and J. S. Müller, "Explicit quadratic Chabauty over number fields", 2019. To appear in *Israel J. Math.* arXiv

[Balakrishnan et al. 2019c] J. S. Balakrishnan, F. Bianchi, V. Cantoral-Farfán, M. Çiperiani, and A. Etropolski, "Chabauty–Coleman experiments for genus 3 hyperelliptic curves", pp. 67–90 in *Research directions in number theory*, edited by J. Balakrishnan et al., Assoc. Women Math. Ser. **19**, Springer, 2019.

[Bernardi 1981] D. Bernardi, "Hauteur p-adique sur les courbes elliptiques", pp. 1–14 in *Séminaire de Théorie des Nombres* (Paris, 1979-1980), edited by M.-J. Bertin, Progr. Math. **12**, Birkhäuser, Boston, 1981. MR Zbl

[Bianchi 2019] F. Bianchi, Code for "Quadratic Chabauty for (bi)elliptic curves and Kim's conjecture", 2019, available at https://github.com/bianchifrancesca/quadratic_chabauty. Sage math code.

[Coates and Kim 2010] J. Coates and M. Kim, "Selmer varieties for curves with CM Jacobians", *Kyoto J. Math.* **50**:4 (2010), 827–852. MR Zbl

[Coleman and Gross 1989] R. F. Coleman and B. H. Gross, "p-adic heights on curves", pp. 73–81 in *Algebraic number theory*, edited by J. Coates et al., Adv. Stud. Pure Math. **17**, Academic Press, Boston, 1989. MR Zbl

[Connell 1999] I. Connell, "Elliptic curve handbook", preprint, 1999, available at http://webs.ucm.es/BUCM/mat/doc8354.pdf.

[Cremona et al. 2006] J. E. Cremona, M. Prickett, and S. Siksek, "Height difference bounds for elliptic curves over number fields", *J. Number Theory* **116**:1 (2006), 42–68. MR Zbl

[Ellenberg and Hast 2017] J. S. Ellenberg and D. R. Hast, "Rational points on solvable curves over $\mathbb{Q}$ via non-abelian Chabauty", preprint, 2017. arXiv

[Flynn and Wetherell 1999] E. V. Flynn and J. L. Wetherell, "Finding rational points on bielliptic genus 2 curves", *Manuscripta Math.* **100**:4 (1999), 519–533. MR Zbl

[Harvey 2008] D. Harvey, "Efficient computation of p-adic heights", *LMS J. Comput. Math.* **11** (2008), 40–59. MR Zbl

[Katz 1976] N. M. Katz, "p-adic interpolation of real analytic Eisenstein series", *Ann. of Math.* (2) **104**:3 (1976), 459–571. MR Zbl

[Kim 2005] M. Kim, "The motivic fundamental group of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ and the theorem of Siegel", *Invent. Math.* **161**:3 (2005), 629–656. MR Zbl

[Kim 2009] M. Kim, "The unipotent Albanese map and Selmer varieties for curves", *Publ. Res. Inst. Math. Sci.* **45**:1 (2009), 89–133. MR Zbl

[Kim 2010a] M. Kim, "Massey products for elliptic curves of rank 1", *J. Amer. Math. Soc.* **23**:3 (2010), 725–747. Appendix and correction by J. S. Balakrishnan, K. S. Kedlaya, and M. Kim in **24**:1 (2011), 281–291. MR Zbl

[Kim 2010b] M. Kim, "p-adic L-functions and Selmer varieties associated to elliptic curves with complex multiplication", *Ann. of Math.* (2) **172**:1 (2010), 751–759. MR Zbl

[Koblitz 1984] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, 2nd ed., Grad. Texts in Math. **58**, Springer, 1984. MR Zbl

[Kolyvagin 1988] V. A. Kolyvagin, "Finiteness of $E(\mathbb{Q})$ and $Ш(E, \mathbb{Q})$ for a subclass of Weil curves", *Izv. Akad. Nauk SSSR Ser. Mat.* **52**:3 (1988), 522–540. In Russian; translated in *Math. USSR-Izv.* **32**:3 (1989), 523–541. MR Zbl

[Lang 1973] S. Lang, *Elliptic functions*, Addison-Wesley, Reading, MA, 1973. MR Zbl

[LMFDB 2019] The LMFDB Collaboration, "The L-functions and modular forms database", electronic reference, 2019, available at http://www.lmfdb.org. Online; accessed 19 March 2019.

[Mazur and Tate 1991] B. Mazur and J. Tate, "The p-adic sigma function", *Duke Math. J.* **62**:3 (1991), 663–688. MR Zbl

[Mazur et al. 2006] B. Mazur, W. Stein, and J. Tate, "Computation of p-adic heights and log convergence", *Doc. Math.* Coates Birthday volume (2006), 577–614. MR Zbl

[Murty and Murty 1997] M. R. Murty and V. K. Murty, *Non-vanishing of L-functions and applications*, Progr. Math. **157**, Birkhäuser, Basel, 1997. MR Zbl

[Pethő et al. 1999] A. Pethő, H. G. Zimmer, J. Gebel, and E. Herrmann, "Computing all S-integral points on elliptic curves", *Math. Proc. Cambridge Philos. Soc.* **127**:3 (1999), 383–402. MR Zbl

[SageMath 2017–2019] W. A. Stein et al., "Sage mathematics software", 2017–2019, available at http://www.sagemath.org. Versions 7.5.1–8.6.

[Silverman 1988] J. H. Silverman, "Computing heights on elliptic curves", *Math. Comp.* **51**:183 (1988), 339–358. MR Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Grad. Texts in Math. **151**, Springer, 1994. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Grad. Texts in Math. **106**, Springer, 2009. MR Zbl

[Smart 1994] N. P. Smart, "S-integral points on elliptic curves", *Math. Proc. Cambridge Philos. Soc.* **116**:3 (1994), 391–399. MR Zbl

[Stein and Wuthrich 2013] W. Stein and C. Wuthrich, "Algorithms for the arithmetic of elliptic curves using Iwasawa theory", *Math. Comp.* **82**:283 (2013), 1757–1792. MR Zbl

[Stoll 2006] M. Stoll, "Finite descent obstructions and rational points on curves", preprint, 2006, available at https://tinyurl.com/stollpreprint. Draft version no. 8.

[Stoll 2007] M. Stoll, "Finite descent obstructions and rational points on curves", *Algebra Number Theory* **1**:4 (2007), 349–391. MR Zbl

[Stroeker and Tzanakis 1994] R. J. Stroeker and N. Tzanakis, "Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms", *Acta Arith.* **67**:2 (1994), 177–196. MR Zbl

[Vignéras 1981] M.-F. Vignéras, "Valeur au centre de symétrie des fonctions L associées aux formes modulaires", pp. 331–356 in *Séminaire de Théorie des Nombres* (Paris, 1979-1980), edited by M.-J. Bertin, Progr. Math. **12**, Birkhäuser, Boston, 1981. Zbl

[Waldspurger 1981] J.-L. Waldspurger, "Sur les coefficients de Fourier des formes modulaires de poids demi-entier", *J. Math. Pures Appl.* (9) **60**:4 (1981), 375–484. MR Zbl

[Wetherell 1997] J. L. Wetherell, *Bounding the number of rational points on certain curves of high rank*, Ph.D. thesis, University of California, Berkeley, 1997, available at https://search.proquest.com/docview/304343505.

[Wuthrich 2004] C. Wuthrich, "On p-adic heights in families of elliptic curves", *J. Lond. Math. Soc.* (2) **70**:1 (2004), 23–40. MR Zbl

Communicated by Kiran S. Kedlaya
Received 2019-04-12 Revised 2020-02-01 Accepted 2020-04-23

francesca.bianchi@rug.nl Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, Netherlands

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2417

The Prasad conjectures for GSp_4 and PGSp_4

Hengfei Lu

We use the theta correspondence between $\mathrm{GSp}_4(E)$ and $\mathrm{GO}(V)$ to study the GSp_4-distinction problems over a quadratic extension E/F of nonarchimedean local fields of characteristic 0. With a similar strategy, we investigate the distinction problem for the pair $(\mathrm{GSp}_4(E), \mathrm{GSp}_{1,1}(F))$, where $\mathrm{GSp}_{1,1}$ is the unique inner form of GSp_4 defined over F. Then we verify the Prasad conjecture for a discrete series representation $\bar{\tau}$ of $\mathrm{PGSp}_4(E)$.

1. Introduction

Let F be a finite field extension over $\mathbb{Q}_p$ and E be a quadratic extension over F with associated Galois group $\mathrm{Gal}(E/F) = \{1, \sigma\}$ and associated quadratic character $\omega_{E/F}$ of $F^\times$. Let W_F be the Weil group of F and WD_F be the Weil–Deligne group. Then $\omega_{E/F}$ is a quadratic character of W_F with kernel W_E. Let $\boldsymbol{G}$ be a connected reductive group defined over F and $\boldsymbol{G}(F)$ (resp. $\boldsymbol{G}(E)$) be the F-rational (resp. E-rational) points. Let $\mathrm{Irr}(\boldsymbol{G}(E))$ denote the set of irreducible smooth representations of $\boldsymbol{G}(E)$. Given a representation $\tau \in \mathrm{Irr}(\boldsymbol{G}(E))$ and a character χ of $\boldsymbol{G}(F)$, we say that τ is $(\boldsymbol{G}(F), \chi)$-distinguished or has a nonzero $(\boldsymbol{G}(F), \chi)$-period if

$$\mathrm{Hom}_{\boldsymbol{G}(F)}(\tau, \chi) \neq 0.$$

If χ is the trivial character, then τ is called $\boldsymbol{G}(F)$-distinguished. There exists a rich literature, such as [Beuzart-Plessis 2018; Flicker 1991; Gan and Raghuram 2013; Lu 2017b; Matringe 2011; Prasad 2015], trying to classify all $\boldsymbol{G}(F)$-distinguished representations of $\boldsymbol{G}(E)$. The method often used to study the distinction problems is the relative trace formula, such as in [Beuzart-Plessis 2018; Flicker and Hakim 1994], which is powerful especially for the global period problems. This paper focuses on the local period

MSC2010: primary 22E50; secondary 11F27.
Keywords: Langlands correspondence, see-saw diagrams, theta lift, quaternionic Hermitian groups, the Prasad conjecture.

problems for $\boldsymbol{G} = \mathrm{GSp}_4, \mathrm{PGSp}_4$ and their inner forms. The main tool in this paper is the local theta correspondence appearing in [Gan and Takeda 2011b; Kudla and Rallis 1992; Yamana 2011].

Let V be the unique nonsplit quaternion algebra D_E with quadratic form N_{D_E} over E, or the split 6-dimensional quadratic space $\mathbb{H}_E^3$ over E. Then

$$\mathrm{GSO}(V) \cong \begin{cases} \mathrm{GSO}_{4,0}(E) = D_E^\times(E) \times D_E^\times(E)/\{(t, t^{-1}) : t \in E^\times\} & \text{if } V = D_E, \\ \mathrm{GSO}_{3,3}(E) = \mathrm{GL}_4(E) \times E^\times/\{(t^{-1}, t^2) : t \in E^\times\} & \text{if } V = \mathbb{H}_E^3, \end{cases}$$

and any irreducible representation of $\mathrm{GSO}(V)$ must be of the form

- $\pi_1 \boxtimes \pi_2$ with $\omega_{\pi_1} = \omega_{\pi_2}$ if $V = D_E$;
- $\Pi \boxtimes \mu$ with $\omega_\Pi = \mu^2$ if $V = \mathbb{H}_E^3$.

Here for each i, π_i is an irreducible representation of $D_E^\times(E)$.

Gan and Takeda [2011b] have studied the explicit theta correspondence between $\mathrm{GSO}(V)$ and $\mathrm{GSp}_4(E)$ and proved that any irreducible representation τ of $\mathrm{GSp}_4(E)$ falls into one of the following two disjoint families of representations:

- $\tau = \theta(\pi_1 \boxtimes \pi_2)$ with $\omega_{\pi_1} = \omega_{\pi_2}$;
- $\tau = \theta(\Pi \boxtimes \mu)$ with $\mu = \omega_\tau$ and $\omega_\Pi = \mu^2$.

The see-saw identity (sometimes called the local Siegel–Weil identity) plays a vital role in the proof of our main theorems. More precisely, suppose that $G \times H$ is a reductive dual pair, with a Weil representation ω_ψ over F. Let $H' \times G'$ be another dual pair contained in the same ambient group, with $G \subset G'$ and $H' \subset H$. Via a so-called see-saw diagram

$$\begin{tikzcd} G' \arrow[d, no head] \arrow[dr, no head] & H \arrow[d, no head] \arrow[dl, no head] \\ G & H' \end{tikzcd}$$

we have

$$\dim \mathrm{Hom}_G(\Theta_\psi(\chi), \pi) = \dim \mathrm{Hom}_{G \times H'}(\omega_\psi, \pi \boxtimes \chi) = \dim \mathrm{Hom}_{H'}(\Theta_\psi(\pi), \chi)$$

for a representation $\pi \in \mathrm{Irr}(G)$ and a character χ of H'. Typically, $\Theta_\psi(\chi)$ is a simpler representation, such as a degenerate principal series representation of G', and the multiplicity $\dim \mathrm{Hom}_G(\Theta_\psi(\chi), \pi)$ has a better chance of being understood; see [Gan 2019]. In order to use the see-saw identity, we need to study the big theta lift $\Theta(\tau)$ to $\mathrm{GO}(V)$ of a generic representation τ of $\mathrm{GSp}_4(E)$. In fact, we have studied the general (almost equal rank) case for the irreducibility of big theta lifts to $\mathrm{GO}_{n+1,n+1}(F)$ of a generic representation of $\mathrm{GSp}_{2n}(F)$ in Section 3C. After computing the big theta lifts following [Gan and Ichino 2014; Gan and Takeda 2011b], we use the local theta correspondences between $\mathrm{GSp}_4(E)$ and $\mathrm{GSO}(V)$ and the see-saw identities to discuss GSp_4-period problems, by transferring the period problem for GSp_4 to various analogous period problems for GL_2, GL_4 and their various forms (not necessarily inner). Then we obtain the following results:

Theorem 1.1 (Theorem 4.4.9). *Suppose that* $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$ *with a central character* ω_τ *and* $\omega_\tau|_{F^\times} = \mathbf{1}$.

(i) *If* $\tau = \theta(\Sigma)$ *is an irreducible representation of* $\mathrm{GSp}_4(E)$, *where* Σ *is an irreducible representation of* $\mathrm{GO}_{4,0}(E)$, *then the representation* τ *is not* $\mathrm{GSp}_4(F)$*-distinguished.*

(ii) *If* $\tau = \theta(\pi_1 \boxtimes \pi_2)$, *where* $\pi_1 \boxtimes \pi_2$ *is a generic representation of* $\mathrm{GSO}_{2,2}(E)$, *then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = \begin{cases} 2 & \text{if } \pi_i \ncong \pi_0 \text{ are both } \mathrm{GL}_2(F)\text{-distinguished}, \\ 1 & \text{if } \pi_1 \ncong \pi_2 \text{ but } \pi_1^\sigma \cong \pi_2^\vee, \\ 1 & \text{if } \pi_1 \cong \pi_2 \text{ is } \mathrm{GL}_2(F)\text{-distinguished but not } (\mathrm{GL}_2(F), \omega_{E/F})\text{-distinguished}, \\ 1 & \text{if } \pi_2 \text{ is } \mathrm{GL}_2(F)\text{-distinguished and } \pi_1 \cong \pi_0, \\ 0 & \text{otherwise}. \end{cases}$$

Here $\pi_0 = \pi(\chi_1, \chi_2)$ *with* $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \mathbf{1}$ *is a principal series representation of* $\mathrm{GL}_2(F)$. *Note that these conditions are mutually exclusive.*

(iii) *Assume that* τ *is not in case* (i) *or* (ii) *and that* $\tau = \theta(\Pi \boxtimes \chi)$ *is generic, where* $\Pi \boxtimes \chi$ *is a representation of* $\mathrm{GSO}_{3,3}(E)$. *Then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = \begin{cases} 1 & \text{if } \Pi \text{ is } \mathrm{GL}_4(F)\text{-distinguished}, \\ 0 & \text{otherwise}. \end{cases}$$

The full local Langlands conjecture for GSp_4 (see Theorem 4.4.7) has been proved by Gan and Takeda [2011a]. Then we can verify the Prasad conjecture for GSp_4 in Section 6C. More precisely, let G_0 be a quasisplit group defined over F (denoted by G^{op} in [Prasad 2015]) such that

$${}^L G_0 = \mathrm{GSp}_4(\mathbb{C}) \rtimes \mathrm{Gal}(E/F),$$

where the nontrivial element $\sigma \in \mathrm{Gal}(E/F)$ acts on $\mathrm{GSp}_4(\mathbb{C})$ by

$$\sigma(g) = \mathrm{sim}(g)^{-1} \cdot g.$$

Here $\mathrm{sim}\colon \mathrm{GSp}_4(\mathbb{C}) \to \mathbb{C}^\times$ is the similitude character. Let ϕ_τ be the Langlands parameter of τ. Define

$$F(\phi_\tau) = \{\tilde{\phi} : WD_F \to {}^L G_0 \mid \tilde{\phi}|_{WD_E} = \phi_\tau\}. \tag{1-1}$$

Theorem 1.2 (the Prasad conjecture for GSp_4). *Let* τ *be an irreducible smooth representation of* $\mathrm{GSp}_4(E)$ *with enhanced Langlands parameter* $(\phi_\tau, \lambda_\tau)$ (*called the Langlands-Vogan parameter*). *Assume that the* L*-packet* Π_{ϕ_τ} *is generic. Then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_{E/F}) = \begin{cases} |F(\phi_\tau)| & \text{if } \tau \text{ is generic, i.e., } \lambda_\tau \text{ is trivial}, \\ 0 & \text{otherwise}, \end{cases}$$

where $F(\phi_\tau)$ *is defined in* (1-1) *and* $|F(\phi_\tau)|$ *denotes the cardinality of the set* $F(\phi_\tau)$.

We will prove analogous results for the inner form in Section 5. Let D be the 4-dimensional quaternion division algebra of F. In a similar way, we study the period problem for the inner form $\mathrm{GU}_2(D)=\mathrm{GSp}_{1,1}$, i.e., try to figure out the multiplicity

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C})$$

for a representation $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$. We will not state the results of the inner form case in the introduction; the precise results can be found in Theorem 5.3.1.

Combining Theorem 1.1 and its analog for inner forms, we can verify the conjecture of Dipendra Prasad [2015, Conjecture 2] for PGSp_4. Given a quasisplit reductive group $\boldsymbol{G}$ defined over F and a quadratic extension E/F, assuming the Langlands–Vogan conjectures for $\boldsymbol{G}$, Prasad [2015] used the recipes from the Galois side to give a formula for the individual multiplicity

$$\dim \mathrm{Hom}_{G_\alpha(F)}(\tau, \chi_{\boldsymbol{G}}),$$

where

- τ is an irreducible discrete series representation of $\boldsymbol{G}(E)$;
- $\chi_{\boldsymbol{G}}$ is a quadratic character of $\boldsymbol{G}(F)$ depending on $\boldsymbol{G}$ and E;
- G_α is any pure inner form of $\boldsymbol{G}$ defined over F satisfying $G_\alpha(E)=\boldsymbol{G}(E)$.

In Section 7, we will focus on the case $\boldsymbol{G}=\mathrm{PGSp}_4$. Then $H^1(F,\boldsymbol{G})=\{\mathrm{PGSp}_4, \mathrm{PGU}_2(D)\}$ and $\chi_{\boldsymbol{G}}=\omega_{E/F}$. The local Langlands correspondences for the quasisplit groups SO_n and Sp_{2n} over a nonarchimedean local field have been verified by Arthur [2013] under certain assumptions which have been removed by Mœglin and Waldspurger [2016a; 2016b; 2018]. We can use the results from the local Langlands correspondence for $\mathrm{SO}_5=\mathrm{PGSp}_4$ freely. More precisely, if $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$ with a trivial central character, then τ corresponds to a representation of $\mathrm{PGSp}_4(E)$, denoted by $\bar{\tau}$. Given a discrete series representation $\bar{\tau}$ of $\mathrm{PGSp}_4(E)$ with the enhanced L-parameter $(\phi_{\bar{\tau}}, \lambda_{\bar{\tau}})$, where $\lambda_{\bar{\tau}}$ is a character of the component group $\pi_0(Z(\phi_{\bar{\tau}}))$, set

$$F(\phi_{\bar{\tau}})=\{\tilde{\phi}: WD_F \to \mathrm{Sp}_4(\mathbb{C}) \mid \tilde{\phi}|_{WD_E}=\phi_{\bar{\tau}}\}.$$

Up to the twisting by the quadratic character $\omega_{E/F}$, there are several orbits in $F(\phi_{\bar{\tau}})$, denoted by $\sqcup_{i=1}^r \mathcal{O}(\tilde{\phi}_i)$. Each orbit $\mathcal{O}(\tilde{\phi}_i)$ corresponds to a unique subset $\mathcal{C}_i$ of $H^1(W_F,\boldsymbol{G})$. (See Section 6A for more details.)

Theorem 1.3. *Let notation be as above. Given a discrete series representation $\bar{\tau}$ of $\mathrm{PGSp}_4(E)$, we have*

$$\dim \mathrm{Hom}_{G_\alpha(F)}(\bar{\tau}, \omega_{E/F})=\sum_{i=1}^r m(\lambda_{\bar{\tau}}, \tilde{\phi}_i)\mathbf{1}_{\mathcal{C}_i}(G_\alpha)/d_0(\tilde{\phi}_i), \tag{1-2}$$

where

- $\mathbf{1}_{\mathcal{C}_i}$ *is the characteristic function of the set* $\mathcal{C}_i$;
- $m(\lambda_{\bar{\tau}}, \tilde{\phi})$ *is the multiplicity for the trivial representation contained in the restricted representation* $\lambda_{\bar{\tau}}|_{\pi_0(Z(\tilde{\phi}))}$;
- $d_0(\tilde{\phi})=|\mathrm{Coker}\{\pi_0(Z(\tilde{\phi})) \to \pi_0(Z(\phi_{\bar{\tau}}))^{\mathrm{Gal}(E/F)}\}|$, *where* $|-|$ *denotes its cardinality.*

Remark 1.4. We would like to highlight the fact that the square-integrable representation $\bar{\tau}$ may be nongeneric and so $\bar{\tau}$ is not $\mathrm{PGSp}_4(F)$-distinguished (see Theorem 5.3.1) but $\bar{\tau}$ contains a nonzero period for the pure inner form $\mathrm{PGSp}_{1,1}(F)$. It is different from the case $\boldsymbol{G} = \mathrm{PGL}_2$ that if a representation $\bar{\pi}$ of $\mathrm{PGL}_2(E)$ is $\mathrm{PD}^\times(F)$-distinguished, then $\bar{\pi}$ must be $\mathrm{PGL}_2(F)$-distinguished (see Lemma 4.4.5).

In fact, we have shown that the equality (1-2) holds for almost all generic representations in Section 7, except that the Langlands parameter $\phi_{\bar{\tau}} = 2\chi_F|_{W_E} \oplus \phi_2$ with ϕ_2 conjugate-symplectic (in the sense of [Gan et al. 2012, §3]) and $\chi_F^2 = \omega_{E/F}$. However, there is a weaker version of the Prasad conjecture which determines the sum of $\dim \mathrm{Hom}_{G_\alpha(F)}(\bar{\tau}, \chi_{\boldsymbol{G}})$ as G_α runs over all pure inner forms of $\boldsymbol{G}$ satisfying $G_\alpha(E) = \boldsymbol{G}(E)$. It involves the degree of the base change map

$$\Phi : \mathrm{Hom}(WD_F, \mathrm{Sp}_4(\mathbb{C})) \to \mathrm{Hom}(WD_E, \mathrm{Sp}_4(\mathbb{C}))$$

for the exception case, i.e., the identity

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}) + \dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = \sum_{\tilde{\phi} \in F(\phi_{\bar{\tau}})} m(\lambda_{\bar{\tau}}, \tilde{\phi}) \frac{\deg \Phi(\tilde{\phi})}{d_0(\tilde{\phi})} \tag{1-3}$$

when the L-packet $\Pi_{\phi_{\bar{\tau}}}$ is generic, which is the original identity formulated by Prasad.

There is a brief introduction to the proof of Theorem 1.3. After introducing the local theta correspondence between quaternionic unitary groups following [Yamana 2011], we use the isomorphism $\mathrm{GU}_2(R) = \mathrm{GSp}_{1,1}(E) \cong \mathrm{GSp}_4(E)$, where $R \cong \mathrm{Mat}_{2,2}(E)$ is the split quaternion algebra over E, to embed the group $\mathrm{GSp}_{1,1}(F)$ into $\mathrm{GSp}_4(E)$. Then one can use the see-saw identity to transfer the inner form $\mathrm{GSp}_{1,1}$-period problem to $\mathrm{GO}^*_{3,0}$ or $\mathrm{GO}^*_{1,1}$ side, which are closely related to GL_n-period problems. But we need to be very careful when we use the see-saw identity for a pair of quaternionic unitary groups. (See Remark 5.2.4.) Once the see-saw identity for the quaternionic unitary groups has been set up, the rest of the proof for the inner form case is similar to the case for GSp_4-period. Then we obtain the results for the distinction problems for the automorphic side. For the Galois side, i.e., the right-hand side of (1-3), it will be checked case by case in Section 7.

Remark 1.5. Raphaël Beuzart-Plessis [2018, Theorem 1] used the local trace formula to deal with the distinction problems for the Galois pair $(G'(E), G'(F))$ for the stable square-integrable representations, where G' is an inner form of $\boldsymbol{G}$ defined over F, which generalizes [Prasad 1992, Theorem C].

The paper is organized as follows. In Section 2, we set up the notation about the local theta correspondence. In Section 3, we will study the irreducibility for the big theta lift of a generic representation in the almost equal rank case, which generalizes the results of [Gan and Ichino 2014, Proposition C.4] for the tempered representations. The detailed computation for the explicit big theta lift $\Theta(\tau)$ to $\mathrm{GO}(V)$ will be given in Section 3E. In Section 4, we will study the distinction problems for GSp_4 over a quadratic extension E/F. The proof of Theorem 1.1 will be given in Section 4D. The analogous results for the inner form $\mathrm{GSp}_{1,1}$ will be given in Section 5. In Section 6A, we will introduce the Prasad conjecture for a reductive quasisplit group $\boldsymbol{G}$ defined over F. Then we will verify the Prasad conjecture for GSp_4 in Section 6C. Finally, the proof of Theorem 1.3 will be given in Section 7.

2. The local theta correspondences for similitudes

In this section, we will briefly recall some results about the local theta correspondence, following [Gan and Takeda 2011b; Kudla 1996; Roberts 2001].

Let F be a nonarchimedean local field of characteristic zero. Consider the dual pair $\mathrm{O}(V) \times \mathrm{Sp}(W)$. For simplicity, we may assume that $\dim V$ is even. Fix a nontrivial additive character ψ of F. Let ω_ψ be the Weil representation for $\mathrm{O}(V) \times \mathrm{Sp}(W)$. If π is an irreducible smooth representation of $\mathrm{O}(V)$ (resp. $\mathrm{Sp}(W)$), the maximal π-isotypic quotient of ω_ψ has the form

$$\pi \boxtimes \Theta_\psi(\pi)$$

for some smooth representation $\Theta_\psi(\pi)$ of $\mathrm{Sp}(W)$ (resp. some smooth representation $\Theta_\psi(\pi)$ of $\mathrm{O}(V)$). We call $\Theta_\psi(\pi)$ or $\Theta_{V,W,\psi}(\pi)$ the big theta lift of π. It is known that $\Theta_\psi(\pi)$ is of finite length and hence is admissible. Let $\theta_\psi(\pi)$ or $\theta_{V,W,\psi}(\pi)$ be the maximal semisimple quotient of $\Theta_\psi(\pi)$, which is called the small theta lift of π.

Theorem 2.1 (Howe duality conjecture [Gan and Takeda 2016a; 2016b]).

- *$\theta_\psi(\pi)$ is irreducible whenever $\Theta_\psi(\pi)$ is nonzero.*
- *The map $\pi \mapsto \theta_\psi(\pi)$ is injective on its domain.*

This has been proved by Waldspurger [1990] when $p \neq 2$.

We extend the Weil representation to the case of similitude groups. Let λ_V and λ_W be the similitude factors of $\mathrm{GO}(V)$ and $\mathrm{GSp}(W)$ respectively. We shall consider the group

$$R = \mathrm{GO}(V) \times \mathrm{GSp}^+(W),$$

where $\mathrm{GSp}^+(W)$ is the subgroup of $\mathrm{GSp}(W)$ consisting of elements g such that $\lambda_W(g)$ lies in the image of λ_V. Define

$$R_0 = \{(h, g) \in R \mid \lambda_V(h)\lambda_W(g) = 1\}$$

to be the subgroup of R. The Weil representation ω_ψ extends naturally to the group R_0 via

$$\omega_\psi(g, h)\phi = |\lambda_V(h)|_F^{-\frac{1}{8}\dim V\cdot\dim W}\,\omega(g_1, 1)(\phi \circ h^{-1}),$$

where $|-|_F$ is the absolute value on F and

$$g_1 = g \begin{pmatrix} \lambda_W(g)^{-1} & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{Sp}(W).$$

Here the central elements $(t, t^{-1}) \in R_0$ acts by the quadratic character $\chi_V(t)^{(\dim W)/2}$, which is slightly different from the normalization used in [Roberts 2001].

Now we consider the compactly induced representation

$$\Omega = \mathrm{ind}_{R_0}^{R}\,\omega_\psi.$$

As a representation of R, Ω depends only on the orbit of ψ under the evident action of $\operatorname{Im}\lambda_V \subset F^\times$. For example, if λ_V is surjective, then Ω is independent of ψ. For any irreducible representation π of $\mathrm{GO}(V)$ (resp. $\mathrm{GSp}^+(W)$), the maximal π-isotropic quotient of Ω has the form

$$\pi \otimes \Theta_\psi(\pi),$$

where $\Theta_\psi(\pi)$ is some smooth representation of $\mathrm{GSp}^+(W)$ (resp. $\mathrm{GO}(V)$). Similarly, we let $\theta_\psi(\pi)$ be the maximal semisimple quotient of $\Theta_\psi(\pi)$. Note that though $\Theta_\psi(\pi)$ may be reducible, it has a central character $\omega_{\Theta_\psi(\pi)}$ given by

$$\omega_{\Theta_\psi(\pi)} = \chi_V^{(\dim W)/2}\omega_\pi.$$

There is an extended Howe conjecture for similitude groups, which says that $\theta_\psi(\pi)$ is irreducible whenever $\Theta_\psi(\pi)$ is nonzero and the map $\pi \mapsto \theta_\psi(\pi)$ is injective on its domain. It was shown by Roberts [1996] that this follows from Theorem 2.1.

If λ_V is surjective, we have $\mathrm{GSp}^+(W) = \mathrm{GSp}(W)$.

Proposition 2.2 [Gan and Takeda 2011a, Proposition 2.3]. *Suppose that π is a supercuspidal representation of* $\mathrm{GO}(V)$ (*resp.* $\mathrm{GSp}(W)$). *Then $\Theta_\psi(\pi)$ is either zero or is an irreducible representation of* $\mathrm{GSp}^+(W)$ (*resp.* $\mathrm{GO}(V)$).

2A. *First occurrence indices for pairs of orthogonal Witt towers.* Let W_n ($n \geq 1$) be the $2n$-dimensional symplectic vector space with associated symplectic group $\mathrm{Sp}(W_n)$ and consider the two towers of orthogonal groups attached to the quadratic spaces with trivial discriminant. More precisely, let $\mathbb{H}$ be the split 2-dimensional quadratic space over F and D be the quaternion division algebra over F. Let

$$V_{2r}^+ = \mathbb{H}^r \quad \text{and} \quad V_{2r}^- = D(F) \oplus \mathbb{H}^r,$$

and denote the orthogonal groups by $\mathrm{O}(V_{2r}^+) = \mathrm{O}_{r,r}$ and $\mathrm{O}(V_{2r}^-) = \mathrm{O}_{r+4,r}$, respectively. For an irreducible representation π of $\mathrm{Sp}(W_n)$, one may consider the theta lifts $\theta_{2r}^+(\pi)$ and $\theta_{2r}^-(\pi)$ to $\mathrm{O}(V_{2r}^+)$ and $\mathrm{O}(V_{2r}^-)$ respectively, with respect to a fixed nontrivial additive character ψ. Set

$$\begin{cases} r^+(\pi) = \inf\{r : \theta_{2r}^+(\pi) \neq 0\}, \\ r^-(\pi) = \inf\{r : \theta_{2r}^-(\pi) \neq 0\}. \end{cases}$$

Then Kudla and Rallis [2005] and Sun and Zhu [2015] showed:

Theorem 2.3 (conservation relation). *For any irreducible representation π of* $\mathrm{Sp}(W_n)$, *we have*

$$r^+(\pi) + r^-(\pi) = 2n = \dim W_n.$$

On the other hand, one may consider the mirror situation, where one fixes an irreducible representation π of $\mathrm{O}(V_{2r})$ and consider its theta lift $\theta_n(\pi)$ to the tower of symplectic groups $\mathrm{Sp}(W_n)$. Then, with $n(\pi)$ defined in the analogous fashion

$$n(\pi) = \inf\{n : \theta_n(\pi) \neq 0\},$$

one has

$$n(\pi) + n(\pi \otimes \det) = 2r = \dim V_{2r}.$$

For similitude groups, this implies that

$$n(\pi)+n(\pi\otimes\nu)=2r,$$

where ν is the nontrivial character of $\mathrm{GO}(V_{2r})/\mathrm{GSO}(V_{2r})$.

3. The irreducibility of the big theta lift

Let τ be an irreducible representation of $\mathrm{Sp}_{2n}(F)$. Gan and Ichino [2014, Proposition C.4] showed that the big theta lift $\Theta^+_{2n+2}(\tau)$ to $\mathrm{O}_{n+1,n+1}(F)$ (called the almost equal rank case) is irreducible if τ is tempered. This includes the case $p=2$ since the Howe duality conjecture has been proved in [Gan and Takeda 2016b]. We will use the generalized standard module [Heiermann 2016, Theorem 3.2] to study the case when Π_{ϕ_τ} is generic (see Theorem 3.2).

In Section 3C, we mainly study the big theta lift to the split group $\mathrm{O}_{n+1,n+1}(F)$ from a representation τ of $\mathrm{Sp}_{2n}(F)$ when the associated L-packet Π_{ϕ_τ} is generic. Then we will focus on the computation for $n=2$.

3A. ***Notation.*** Let us introduce the notation used in this section.

- $|-|_F$ (resp. $|-|_E$) is an absolute value defined on F (resp. E).
- $P_{\vec{n}}$ (resp. $Q_{\vec{n}}$) is a parabolic subgroup of Sp_{2n} (resp. $\mathrm{O}_{n+1,n+1}$) defined over F.
- ϕ_τ is the Langlands parameter or L-parameter of τ and $\phi_\tau^\vee$ is the dual parameter of ϕ_τ.
- $\tau^\vee$ is the contragredient representation of τ.
- Π_{ϕ_τ} is the L-packet containing τ.
- $\mathcal{W}_r$ is the symplectic vector space over E of dimension $2r$.
- Z is a line in $\mathcal{W}_2$ and Y is a maximal isotropic subspace in $\mathcal{W}_2$.
- $Q(Z)$ (resp. $P(Y)$) is the Klingen (resp. Siegel) parabolic subgroup of $\mathrm{GSp}_4(E)=\mathrm{GSp}(\mathcal{W}_2)$.
- B (resp. B_0) is the Borel subgroup of $\mathrm{GSp}_4(E)$ (resp. $\mathrm{GL}_4(E)$).
- P is the parabolic subgroup of $\mathrm{GL}_4(E)$ with Levi component $\mathrm{GL}_2(E)\times\mathrm{GL}_2(E)$.
- $\Theta^+_{2r}(\tau)$ (resp. $\Theta_6(\tau)$) is the big theta lift to $\mathrm{GO}_{r,r}(E)$ (resp. $\mathrm{GSO}_{3,3}(E)$) of τ of $\mathrm{GSp}_4(E)$.
- $\theta^+_6(\tau)$ (resp. $\theta_6(\tau)$) is the small theta lift to $\mathrm{GO}_{3,3}(E)$ (resp. $\mathrm{GSO}_{3,3}(E)$) of τ of $\mathrm{GSp}_4(E)$.

3B. ***The standard module conjecture.*** Let $\boldsymbol{G}$ be a quasisplit reductive group defined over F. Fix a Borel subgroup $\boldsymbol{B}=\boldsymbol{TU}$ of $\boldsymbol{G}$. Let π be an irreducible smooth representation of $\boldsymbol{G}(F)$. If there exists a nondegenerate character ψ_U of $\boldsymbol{U}(F)$ such that $\mathrm{Hom}_{\boldsymbol{U}(F)}(\pi,\psi_U)\neq 0$, then we say π is ψ_U-generic or generic. If the L-packet Π_{ϕ_π} contains a generic representation, then we call Π_{ϕ_π} a generic L-packet. Let $\boldsymbol{P}=\boldsymbol{MN}$ be a standard parabolic subgroup of $\boldsymbol{G}$. Suppose that there exists a generic tempered representation ρ of $\boldsymbol{M}(F)$ such that π is isomorphic to the Langlands quotient $J(\rho,\chi)$, where χ is a character of $\boldsymbol{M}(F)$ and lies in the positive Weyl chamber with respect to $\boldsymbol{P}(F)$. (See [Heiermann and Opdam 2013, p. 777] for more details.)

Theorem 3.1 (the standard module conjecture). *If $\pi = J(\rho, \chi)$ is a generic representation of $\boldsymbol{G}(F)$, then $\mathrm{Ind}_{\boldsymbol{P}(F)}^{\boldsymbol{G}(F)}(\rho \otimes \chi)$ (normalized induction) is irreducible. Moreover, for any irreducible representation ρ' of $\boldsymbol{M}(F)$ lying inside the L-packet Π_{ϕ_ρ}, $\mathrm{Ind}_{\boldsymbol{P}(F)}^{\boldsymbol{G}(F)}(\rho' \otimes \chi)$ is irreducible.*

Heiermann and Opdam [2013] proved the standard module conjecture. Later Heiermann [2016, Theorem 3.2] proved its generalized version i.e., the "moreover" part of Theorem 3.1. The following subsection will focus on the cases $\boldsymbol{G} = \mathrm{Sp}_{2n}$ and $\boldsymbol{G} = \mathrm{O}_{n+1,n+1}$.

3C. ***Theta lift from*** $\mathbf{Sp}_{2n}(F)$ ***to*** $\mathbf{O}_{n+1,n+1}(F)$. Suppose that τ is a generic irreducible admissible representation of $\mathrm{Sp}_{2n}(F)$. Assume that there exists a parabolic subgroup $P_{\vec{n}} = M_{\vec{n}} N_{\vec{n}}$ of Sp_{2n} and an irreducible representation $\pi_1 \otimes \pi_2 \otimes \cdots \otimes \pi_r \otimes \tau_0$ of $M_{\vec{n}}(F) \cong \mathrm{GL}_{n_1}(F) \times \cdots \times \mathrm{GL}_{n_r}(F) \times \mathrm{Sp}_{2n_0}(F)$ (for $n_1 + n_2 + \cdots + n_r + n_0 = n$) such that τ is the unique irreducible quotient of the standard module

$$\mathrm{Ind}_{P_{\vec{n}}(F)}^{\mathrm{Sp}_{2n}(F)}(\pi_1|-|_F^{s_1} \otimes \cdots \otimes \pi_r|-|_F^{s_r} \otimes \tau_0)(\text{normalized induction}), \tag{3-1}$$

where $s_1 > s_2 > \cdots > s_r > 0$, $n \geq n_0$, each π_i is a tempered representation of $\mathrm{GL}_{n_i}(F)$ and τ_0 is a tempered representation of $\mathrm{Sp}_{2n_0}(F)$. Moreover, the Langlands parameter $\phi_\tau : WD_F \to \mathrm{SO}_{2n+1}(\mathbb{C})$ is given by

$$\phi_\tau = \phi_{\pi_1}|-|_F^{s_1} \oplus \cdots \oplus \phi_{\pi_r}|-|_F^{s_r} \oplus \phi_{\tau_0} \oplus \phi_{\pi_r}^\vee|-|_F^{-s_r} \oplus \cdots \oplus \phi_{\pi_1}^\vee|-|_F^{-s_1},$$

where each ϕ_{π_i} is the Langlands parameter of π_i and ϕ_{τ_0} is the Langlands parameter of τ_0. Here we identify the characters of $F^\times$ and the characters of the Weil group W_F by the local class field theory. Due to Theorem 3.1, the generic representation τ is isomorphic to the standard module, i.e., the standard module is irreducible. Thanks to [Gan and Ichino 2014, Proposition C.4], the small theta lift $\theta_{2n+2}^+(\tau)$ is the unique irreducible quotient of the standard module

$$\mathrm{Ind}_{Q_{\vec{n}}(F)}^{\mathrm{O}_{n+1,n+1}(F)}(\pi_1|-|_F^{s_1} \otimes \cdots \otimes \pi_r|-|_F^{s_r} \otimes \Theta_{2n_0+2}^+(\tau_0)), \tag{3-2}$$

where $Q_{\vec{n}}(F)$ is the parabolic subgroup of $\mathrm{O}_{n+1,n+1}(F)$ with Levi component $\mathrm{GL}_{n_1}(F) \times \cdots \times \mathrm{GL}_{n_r}(F) \times \mathrm{O}_{n_0+1,n_0+1}(F)$. We will show that (3-2) equals $\theta_{2n+2}^+(\tau)$ under certain conditions.

Theorem 3.2. *Let $P_{\vec{n}}$ (resp. $Q_{\vec{n}}$) be a parabolic subgroup of Sp_{2n} (resp. $\mathrm{O}_{n+1,n+1}$) defined as above. If the irreducible representation τ is generic and so τ is isomorphic to the standard module* (3-1), *and the standard L-function of τ is regular at $s = 1$, then $\Theta_{2n+2}^+(\tau)$ is irreducible.*

There is another key input in the proof of Theorem 3.2.

Theorem 3.3. *Let $\boldsymbol{G}$ be Sp_{2n} or $\mathrm{SO}_{n+1,n+1}$. Let π be an irreducible representation of $\boldsymbol{G}(F)$. The L-packet Π_{ϕ_π} is generic if and only if the adjoint L-function $L(s, \phi_\pi, \mathrm{Ad})$ is regular at $s = 1$.*

Proof. See [Liu 2011, Theorem 1.2; Jantzen and Liu 2014, Theorem 1.5]. ☐

Proof of Theorem 3.2. We will show that $\Theta_{2n+2}^+(\tau)|_{\mathrm{SO}_{n+1,n+1}(F)}$ is irreducible. If $n = n_0$, then it follows from [Gan and Ichino 2014, Proposition C.4]. Assume that $s_1 > 0$. Then there exists a surjection

$$\mathrm{Ind}_{Q_{\vec{n}}(F)}^{\mathrm{O}_{n+1,n+1}(F)}(\pi_1|-|_F^{s_1} \otimes \cdots \otimes \pi_r|-|_F^{s_r} \otimes \Theta_{2n_0+2}^+(\tau_0)) \twoheadrightarrow \Theta_{2n+2}^+(\tau).$$

Due to [Gan and Ichino 2014, Proposition C.4], if τ_0 is tempered, then $\Theta^+_{2n_0+2}(\tau_0)$ is irreducible and generic. Moreover, if

$$\phi_{\tau_0} : WD_F \to \mathrm{SO}_{2n_0+1}(\mathbb{C})$$

is the Langlands parameter of τ_0, then $\phi_{\theta^+_{2n_0+2}(\tau_0)} = \phi_{\tau_0} \oplus \mathbb{C}$ due to [Gan and Ichino 2014, Theorem C.5]. Assume that $\phi_\tau = \phi_0 \oplus \phi_{\tau_0} \oplus \phi_0^\vee$ with ϕ_{τ_0} tempered and $\phi_0 = \bigoplus_i \phi_{\pi_i}|-|^{s_i}$. Then due to [Gan and Ichino 2014, Proposition C.4], we have $\phi_{\theta^+_{2n+2}(\tau)} = \phi_0 \oplus (\phi_{\tau_0} \oplus \mathbb{C}) \oplus \phi_0^\vee$. Observe that

$$L(s, \mathrm{Ad}_{\mathrm{SO}_{2n+2}} \circ \phi_{\theta^+_{2n+2}(\tau)}) = L(s, \mathrm{Ad}_{\mathrm{SO}_{2n+1}} \circ \phi_\tau) \cdot L(s, \phi_\tau, \mathrm{Std}),$$

where $L(s, \phi_\tau, \mathrm{Std})$ is the standard L-function of τ. By [Liu 2011, Theorem 1.2] and the assumption that τ is generic, we obtain that $L(s, \mathrm{Ad}_{\mathrm{SO}_{2n+1}} \circ \phi_\tau)$ is regular at $s = 1$. So $L(s, \mathrm{Ad}_{\mathrm{SO}_{2n+2}} \circ \phi_{\theta^+_{2n+2}(\tau)})$ is regular at $s = 1$. Thanks to [Jantzen and Liu 2014, Theorem 1.5], the L-packet $\Pi_{\phi_{\theta^+_{2n+2}(\tau)}}$ is generic. By the generalization of the standard module conjecture [Heiermann 2016, Theorem 3.2] that the standard module with a generic quotient is irreducible,

$$\theta^+_{2n+2}(\tau) = \Theta^+_{2n+2}(\tau) = \mathrm{Ind}^{\mathrm{O}_{n+1,n+1}(F)}_{Q_{\bar{n}}(F)}(\pi_1|-|^{s_1}_F \otimes \cdots \otimes \pi_r|-|^{s_r}_F \otimes \Theta^+_{2n_0+2}(\tau_0)),$$

i.e., $\Theta^+_{2n+2}(\tau)$ is irreducible. □

Remark 3.4. Similarly, if Σ is a generic representation of $\mathrm{O}_{n,n}(F)$ and $L(s, \Sigma, \mathrm{Std})$ is regular at $s = 1$, then the big theta lift $\Theta_n(\Sigma)$ to $\mathrm{Sp}_{2n}(F)$ is irreducible. However, if τ is a generic representation of $\mathrm{Sp}_{2n}(F)$ and $L(s, \tau, \mathrm{Std})$ is regular at $s = 1$, the big theta lift to nonsplit group $\mathrm{O}(V_F)$ may be reducible when V_F is a $(2n+2)$-dimensional quadratic space over F with nontrivial discriminant. (See [Lu 2017b, Proposition 3.8(iii)].)

Remark 3.5. There exists an isomorphism between the characters $\lambda_{\theta^+_{2n+2}(\tau)} \cong \lambda_{\theta^+_{2n_0+2}(\tau_0)}$, the latter of which is given in [Atobe and Gan 2017, Theorem 4.3] in terms of the character λ_{τ_0}, conjectured in [Prasad 1993].

Corollary 3.6. *Let Π_{ϕ_τ} be the L-packet of $\mathrm{Sp}_{2n}(F)$ containing τ. Suppose that Π_{ϕ_τ} is generic. If the standard L-function $L(s, \phi_\tau, \mathrm{Std})$ is a factor of the adjoint L-function $L(s, \mathrm{Ad} \circ \phi_\tau)$, then the big theta lift $\Theta^+_{2n+2}(\tau)$ to $\mathrm{O}_{n+1,n+1}(F)$ is irreducible for any $\tau \in \Pi_{\phi_\tau}$.*

For the rest of this section, we will compute the big theta lifts between $\mathrm{GSp}_4(E)$ and $\mathrm{GO}(V)$ explicitly when $\dim_E V = 4$ or 6.

3D. *Representations of* GO(V). Let π_i be an irreducible representations of $\mathrm{GL}_2(E)$ with central character ω_{π_i} and $\omega_{\pi_1} = \omega_{\pi_2}$. Then $\pi_1 \boxtimes \pi_2$ is an irreducible representation of the similitude group

$$\mathrm{GSO}_{2,2}(E) \cong \mathrm{GL}_2(E) \times \mathrm{GL}_2(E)/\{(t, t^{-1}) : t \in E^\times\}.$$

If $\pi_1 \neq \pi_2$, then $\Sigma = \mathrm{Ind}^{\mathrm{GO}_{2,2}(E)}_{\mathrm{GSO}_{2,2}(E)}(\pi_1 \boxtimes \pi_2)$ is an irreducible smooth representation of $\mathrm{GO}_{2,2}(E)$ and $\Sigma \cong \Sigma \otimes \nu$, where $\nu|_{\mathrm{O}_{2,2}(E)} = \det$. If $\pi_1 = \pi_2$, then there are two extensions $(\pi_1 \boxtimes \pi_1)^\pm$ and only one of them participates in the theta lift between $\mathrm{GSp}_4(E)$ and $\mathrm{GO}_{2,2}(E)$, denoted by $(\pi_1 \boxtimes \pi_1)^+ = \Sigma$. Moreover, we have $(\pi_1 \boxtimes \pi_1)^+ \otimes \nu \cong (\pi_1 \boxtimes \pi_1)^-$. (See [Gan and Takeda 2011b, §6].)

Any irreducible representation of

$$\mathrm{GSO}_{3,3}(E) = \mathrm{GL}_4(E) \times \mathrm{GL}_1(E)/\{(t, t^{-2}) : t \in E^\times\}$$

is of the form

$$\Pi \boxtimes \chi,$$

where Π is a representation of $\mathrm{GL}_4(E)$ with central character ω_Π, χ is a character of $E^\times$ and $\chi^2 = \omega_\Pi$.

3E. *Representations of* $\mathbf{GSp_4}(\boldsymbol{E})$. Assume that $\tau = \theta(\pi_1 \boxtimes \pi_2)$ is a representation of $\mathrm{GSp}_4(E)$ and $\pi_1 \boxtimes \pi_2 \in \mathrm{Irr}(\mathrm{GSO}_{2,2}(E))$. Then τ is generic if and only if $\pi_1 \boxtimes \pi_2$ is generic due to [Gan and Takeda 2011b, Corollary 4.2(ii)]. We follow the notation in [Gan and Takeda 2011b] to describe the nondiscrete series representations of $\mathrm{GSp}_4(E)$. Thanks to [Gan and Takeda 2011b, Proposition 5.3], the nondiscrete series representations of $\mathrm{GSp}_4(E)$ fall into the following three families:

- $\tau \hookrightarrow I_{Q(Z)}(\chi|-|_E^{-s}, \pi)$ with χ a unitary character, $s \geq 0$ and π a discrete series representation of $\mathrm{GL}_2(E)$ up to twist;
- $\tau \hookrightarrow I_{P(Y)}(\pi|-|_E^{-s}, \chi)$ with χ an arbitrary character, $s \geq 0$ and π a unitary discrete series representation of $\mathrm{GL}(Y)$;
- $\tau \hookrightarrow I_B(\chi_1|-|_E^{-s_1}, \chi_2|-|_E^{-s_2}; \chi)$, where χ_1, χ_2 are unitary and $s_1 \geq s_2 \geq 0$.

Note that if τ itself is generic and nontempered, then those embeddings are in fact isomorphisms due to the standard module conjecture for GSp_4, except

$$\tau \hookrightarrow I_{Q(Z)}(\mathbf{1}, \pi).$$

For instance, $\tau = J_{P(Y)}(\pi|-|_E^{s}, \chi)$ with $s \geq 0$. If τ is generic, then $I_{P(Y)}(\pi|-|_E^{s}, \chi)$ is irreducible and so

$$\tau = I_{P(Y)}(\pi|-|_E^{s}, \chi) \cong I_{P(Y)}(\pi^\vee|-|_E^{-s}, \chi\omega_\pi|-|_E^{2s})$$

with $s \geq 0$. (See [Gan and Takeda 2011b, Lemma 5.2].)

If the big theta lift $\Theta_6^+(\tau)$ to $\mathrm{GO}_{3,3}(E)$ of τ is irreducible, the restricted representation $\Theta_6^+(\tau)|_{\mathrm{GSO}_{3,3}(E)}$ is irreducible due to [Prasad 1993, §5, p. 282]. We use $\Theta_6(\tau)$ to denote the big theta lift to $\mathrm{GSO}_{3,3}(E)$ of τ.

Proposition 3.7. *Let τ be a generic irreducible representation of $\mathrm{GSp}_4(E)$. Then the big theta lift $\Theta_6(\tau)$ to $\mathrm{GSO}_{3,3}(E)$ of τ is an irreducible representation unless $\tau = I_{Q(Z)}(|-|_E, \pi)$ with π essentially square-integrable. If $\tau = I_{Q(Z)}(|-|_E, \pi)$, then $\Theta_6(\tau) = I_P(\pi|-|_E, \pi) \boxtimes \omega_\pi|-|_E$ is reducible.*

Proof. If τ is a tempered representation, then $\Theta_6^+(\tau)$ is irreducible due to [Gan and Ichino 2014, Proposition C.4] (which holds even for $p = 2$ since the Howe duality conjecture holds) and so $\Theta_6(\tau)$ is irreducible. Assume that the generic representation τ is not essentially tempered. There are 4 cases:

- If $\tau = I_B(\chi_1, \chi_2; \chi)$ is irreducible, then none of the characters $\chi_1, \chi_2, \chi_1/\chi_2, \chi_1\chi_2$ is $|-|_E^{\pm 1}$ and so $I_{B_0}(\mathbf{1}, \chi_2, \chi_1, \chi_1\chi_2)$ has a generic quotient where B_0 is a Borel subgroup of $\mathrm{GL}_4(E)$. Thus $\Theta_6(\tau) = I_{B_0}(\mathbf{1}, \chi_2, \chi_1, \chi_1\chi_2) \cdot \chi \boxtimes \chi^2\chi_1\chi_2$ is irreducible due to the standard module conjecture for GL_4.

- If $\tau = I_{P(Y)}(\pi, \chi)$, then $\Theta_6(\tau)$ is a quotient of

$$I_Q(\mathbf{1}, \pi, \omega_\pi) \cdot \chi \boxtimes \chi^2 \omega_\pi,$$

  where Q is a parabolic subgroup of $\mathrm{GL}_4(E)$ with Levi subgroup $\mathrm{GL}_1(E) \times \mathrm{GL}_2(E) \times \mathrm{GL}_1(E)$. Due to [Gan and Takeda 2011b, Proposition 13.2], the adjoint L-function $L(s, As \circ \phi_\tau)$ is regular at $s = 1$. Since the standard L-function $L(s, \tau, \mathrm{Std})$ is a factor of $L(s, \mathrm{Ad} \circ \phi_\tau)$, we have $L(s, \tau, \mathrm{Std})$ is regular at $s = 1$. Then $I_Q(\mathbf{1}, \pi, \omega_\pi)$ is irreducible and so $\Theta_6(\tau) = I_Q(\mathbf{1}, \pi, \omega_\pi) \cdot \chi \boxtimes \chi^2 \omega_\pi$ is irreducible.

- If $\tau = I_{Q(Z)}(\chi, \pi)$ with $\chi \neq \mathbf{1}$, then there is an epimorphism

$$I_P(\pi \cdot \chi, \pi) \boxtimes \omega_\pi \chi \twoheadrightarrow \Theta_6(\tau)$$

  of $\mathrm{GSO}_{3,3}(E)$-representations, where P is a parabolic subgroup of $\mathrm{GL}_4(E)$ with Levi subgroup $\mathrm{GL}_2(E) \times \mathrm{GL}_2(E)$. Gan and Takeda [2011b, Proposition 13.2] have proved that $I_P(\pi \cdot \chi, \pi)$ is irreducible if $I_{Q(Z)}(\chi, \pi)$ is irreducible and $\chi \neq |-|_E$. If $\chi = |-|_E$ and π is essentially square-integrable, applying [Gan and Takeda 2011b, Corollary 4.4] that τ is generic implies that $\Theta_6(\tau)$ is generic, then $\Theta_6(\tau) = I_P(\pi \cdot \chi, \pi) \boxtimes \omega_\pi \chi$ and $\theta_6(\tau) = J_P(\pi \cdot \chi, \pi) \boxtimes \omega_\pi \chi$ is the Langlands quotient.

- If $\tau \hookrightarrow I_{Q(Z)}(\mathbf{1}, \pi)$, then $\Theta_6(\tau)$ is either zero or $I_P(\pi, \pi) \boxtimes \omega_\pi$, where P is a parabolic subgroup of $\mathrm{GL}_4(E)$ with Levi subgroup $\mathrm{GL}_2(E) \times \mathrm{GL}_2(E)$. In fact, $\Theta_6(\tau) = 0$ only when τ is a nongeneric constituent representation of $I_{Q(Z)}(\mathbf{1}, \pi)$.

This finishes the proof of Proposition 3.7. □

Remark 3.8. Similarly one can prove that if Σ is a generic representation of $\mathrm{GSO}_{2,2}(E)$ and $L(s, \Sigma, \mathrm{Std})$ is regular at $s = 1$, then the big theta lift $\Theta_2(\Sigma)$ to $\mathrm{GSp}_4(E)$ is an irreducible representation.

Let us turn the table around. The rest of this subsection focuses on the computation of local theta lifts to $\mathrm{GO}_{2,2}(E)$ from $\mathrm{GSp}_4(E)$.

Proposition 3.9. *Let τ be a generic irreducible representation of $\mathrm{GSp}_4(E)$. Assume that $\theta_4^+(\tau) \neq 0$.*

(i) *If $\tau = I_{Q(Z)}(\mathbf{1}, \pi(\mu_1, \mu_2))$, then the big theta lift $\Theta_4^+(\tau)$ to $\mathrm{GO}_{2,2}(E)$ of τ is $\mathrm{Ext}^1_{\mathrm{GO}_{2,2}(E)}(\Sigma^+, \Sigma^-)$, where $\Sigma^\pm$ are two distinct extensions of $\pi(\mu_1, \mu_2) \boxtimes \pi(\mu_1, \mu_2)$ from $\mathrm{GSO}_{2,2}(E)$ to $\mathrm{GO}_{2,2}(E)$.*

(ii) *If $\tau \neq I_{Q(Z)}(\mathbf{1}, \pi(\mu_1, \mu_2))$, then $\Theta_4^+(\tau)$ is an irreducible representation of $\mathrm{GO}_{2,2}(E)$.*

Proof. (i) If $\tau = I_{Q(Z)}(\mathbf{1}, \pi(\mu_1, \mu_2))$, then the small theta lift $\theta_4^+(\tau)$ equals Σ^+ by the Howe duality, where Σ^+ is the extension to $\mathrm{GO}_{2,2}(E)$ of $\pi(\mu_1, \mu_2) \boxtimes \pi(\mu_1, \mu_2)$. Let ψ_U be a nondegenerate character of the standard unipotent subgroup U of $\mathrm{GO}_{2,2}(E)$. Then

$$\dim \mathrm{Hom}_U(\Theta_4^+(\tau), \psi_U) = \dim \mathrm{Hom}_{H(\mathcal{W}_1) \rtimes \mathrm{Sp}(\mathcal{W}_1)}(\tau, \omega_\psi) = 2, \tag{3-3}$$

where $\mathcal{W}_2 = Z \oplus \mathcal{W}_1 \oplus Z^*$, $H(\mathcal{W}_1)$ is the Heisenberg group of $\mathcal{W}_1$ equipped with the Weil representation ω_ψ and τ is the representation of $\mathrm{GSp}(\mathcal{W}_2)$. Thus the big theta lift $\Theta_4^+(\tau)$ to $\mathrm{GO}_{2,2}(E)$ is reducible. There is a short exact sequence of $\mathrm{GO}_{2,2}(E)$-representations

$$\Sigma^- \oplus \Sigma^+ \longrightarrow \Theta_4^+(\tau) \longrightarrow \Sigma^+ \longrightarrow 0. \tag{3-4}$$

However, we can not determine $\Theta_4^+(\tau)$ at this moment. Note that

$$\dim \mathrm{Ext}^1_{\mathrm{GSO}_{2,2}(E)}(\pi(\mu_1,\mu_2)\boxtimes\pi(\mu_1,\mu_2), \pi(\mu_1,\mu_2)\boxtimes\pi(\mu_1,\mu_2)) = 1$$

due to [Adler and Prasad 2012, Theorem 1]. Here Ext^1 is the extension functor defined on the category of all smooth representations with a fixed central character. Then $\dim \mathrm{Ext}^1_{\mathrm{GO}_{2,2}(E)}(\Sigma^+, \Sigma^-\oplus\Sigma^+) = 1$ by Frobenius reciprocity, which implies that either $\mathrm{Ext}^1_{\mathrm{GO}_{2,2}(E)}(\Sigma^+,\Sigma^-)$ or $\mathrm{Ext}^1_{\mathrm{GO}_{2,2}(E)}(\Sigma^+,\Sigma^+)$ is zero. Assume that B is the Borel subgroup of $\mathrm{GSO}_{2,2}(E)$. Set $\widetilde{B} = B \rtimes \mu_2$ to be a subgroup of $\mathrm{GO}_{2,2}(E)$ and $\widetilde{B} \cap \mathrm{GSO}_{2,2}(E) = B$. Since

$$\pi(\mu_1,\mu_2)\boxtimes\pi(\mu_1,\mu_2) = \mathrm{Ind}_B^{\mathrm{GSO}_{2,2}(E)}\chi \,(\text{normalized induction}),$$

there are two extensions $\chi^\pm$ to $\widetilde{B}$ of χ of B. We may assume without loss of generality that $\Sigma^+ = \mathrm{Ind}_{\widetilde{B}}^{\mathrm{GO}_{2,2}(E)}\chi^+$ and $\Sigma^- = \mathrm{Ind}_{\widetilde{B}}^{\mathrm{GO}_{2,2}(E)}\chi^-$. Note that $\mathrm{Ext}^1_{\widetilde{B}}(\chi^+,\chi^-)\neq 0$. Then there is a short exact sequence of $\mathrm{GO}_{2,2}(E)$-representations

$$0 \longrightarrow \Sigma^- \longrightarrow \mathrm{Ind}_{\widetilde{B}}^{\mathrm{GO}_{2,2}(E)}(\mathrm{Ext}^1_{\widetilde{B}}(\chi^+,\chi^-)) \longrightarrow \Sigma^+ \longrightarrow 0,$$

which is not split. Hence $\mathrm{Ext}^1_{\mathrm{GO}_{2,2}(E)}(\Sigma^+,\Sigma^-)\neq 0$. Together with (3-3) and (3-4), one can obtain the desired equality $\Theta_4^+(\tau) = \mathrm{Ext}^1_{\mathrm{GO}(2,2)(E)}(\Sigma^+,\Sigma^-)$.

(ii) If τ is a (essentially) discrete series representation, then it follows from [Atobe and Gan 2017, Proposition 5.4].

- If $\tau = I_{Q(Z)}(\mu_0, \pi(\mu_1,\mu_2))$ with $\mu_0 \neq \mathbf{1}$, then there exists only one orbit in the double coset $Q(Z)\backslash \mathrm{GSp}_4(E)/H(\mathcal{W}_1)\rtimes \mathrm{Sp}(\mathcal{W}_1)$ that contributes to the multiplicity

  $$\dim \mathrm{Hom}_{H(W_1)\rtimes \mathrm{Sp}(\mathcal{W}_1)}(\tau, \omega_\psi),$$

  and so $\Theta_4^+(\tau)$ is irreducible.

- If $\tau \subset I_{Q(Z)}(\mathbf{1},\pi)$ with π square-integrable, then τ is tempered. Due to [Atobe and Gan 2017, Proposition 5.5], $\Theta_4^+(\tau)$ is tempered. Note that $\theta_4^+(\tau)$ is a discrete series representation which is projective in the category of the tempered representations. Thus $\Theta_4^+(\tau) = \theta_4^+(\tau)$ is irreducible. Otherwise, it will contradict the Howe duality conjecture (see Theorem 2.1).

- If $\tau = I_{P(Y)}(\pi,\chi)$, then $\dim \mathrm{Hom}_U(\Theta_4^+(\tau), \psi_U) = 1$ and so $\Theta_4^+(\tau)$ is irreducible.

This finishes the proof of Proposition 3.9. □

4. The $\mathrm{GSp}_4(F)$-distinguished representations

This section focuses on the proof of Theorem 1.1. First, we will introduce the see-saw identity in the similitude group in Section 4B. Then we will study the filtrations of various degenerate principal series representations restricted to reductive subgroups in Section 4C, which involves the complicated computation for the double coset decompositions. The proof of Theorem 1.1 will be given in the last subsection.

4A. ***Notation.***

- $\mathbb{C}$ or $\mathbf{1}$ is the trivial representation.
- $\mathbb{H}$ (resp. $\mathbb{H}_E$) is the split 2-dimensional quadratic space over F (resp. E).
- $(-,-)_E$ is the Hilbert symbol on $E^\times \times E^\times$.
- $\mathrm{Res}_{E/F} V$ is a quadratic space over F while V is a quadratic space over E.
- $\mathrm{GSp}(W_n) = \mathrm{GSp}_{2n}(F)$ is the symplectic similitude group.
- $\mathrm{GU}_2(D) = \mathrm{GSp}_{1,1}$ is the unique inner form of GSp_4.
- λ_W (resp. λ_V) is the similitude character of $\mathrm{GSp}_4(E)$ (resp. $\mathrm{GO}(V)$).
- $\mathrm{GSp}_4(E)^\natural = \{g \in \mathrm{GSp}_4(E) \mid \lambda_W(g) \in F^\times\}$ is the subgroup of $\mathrm{GSp}_4(E)$ and similarly for $\mathrm{GO}_{2,2}(E)^\natural$.
- P' (resp. $P^\natural$) is a parabolic (resp. Siegel parabolic) subgroup of $\mathrm{GSp}_4(E)^\natural$ and $Q^\natural$ is the Siegel parabolic subgroup of $\mathrm{GO}_{2,2}(E)^\natural$. And $R_{\bar{P}'}$ (resp. $R_{\bar{P}^\natural}$) is the Jacquet functor with respect to the parabolic subgroup opposite to P' (resp. $P^\natural$).
- ind denotes the compact induction.
- $R_r(\mathbf{1})$ is the big theta lift to $\mathrm{GO}_{4,4}(F)$ of the trivial representation of $\mathrm{GSp}(W_r)$.
- $R^{m,n}(\mathbf{1})$ is the big theta lift to $\mathrm{GSp}_8(F)$ of the trivial representation of $\mathrm{GO}_{m,n}(F)$.
- Σ is a generic representation of $\mathrm{GO}(V)$.
- Q_r is the Siegel parabolic subgroup of $H_r = \mathrm{GO}_{r,r}(F)$.
- $I_{Q_r}^{H_r}(s)$ is the degenerate Siegel principal series of H_r.
- $X_4 = Q_4 \backslash H_4$ is the projective variety.
- $\mathcal{I}(s)$ is the degenerate Siegel principal series of $\mathrm{GSp}_8(F)$.
- $\mathrm{Mat}_{m,n}(F)$ is the matrix space over F consisting of all $m \times n$ matrices.

4B. ***See-saw identity for orthogonal-symplectic dual pairs.*** Following the notation in [Prasad 1996], for a quadratic space (V, q) of even dimension over E, let $\mathrm{Res}_{E/F} V$ be the same space V but now thought of as a vector space over F with a quadratic form

$$q_F(v) = \tfrac{1}{2} \operatorname{tr}_{E/F} q(v).$$

If W_0 is a symplectic vector space over F, then $W_0 \otimes_F E$ is a symplectic vector space over E. Then we have the following isomorphism of symplectic spaces over F:

$$\mathrm{Res}_{E/F}[(W_0 \otimes_F E) \otimes_E V] \cong W_0 \otimes_F \mathrm{Res}_{E/F} V =: \boldsymbol{W}.$$

There is a pair

$$\big(\mathrm{GSp}(W_0), \mathrm{GO}(\mathrm{Res}_{E/F} V)\big) \quad \text{and} \quad \big(\mathrm{GSp}(W_0 \otimes_F E), \mathrm{GO}(V)\big)$$

of similitude dual reductive pairs in the symplectic similitude group $\mathrm{GSp}(\boldsymbol{W})$. A pair (G, H) and (G', H') of dual reductive pairs in a symplectic similitude group is called a see-saw pair if $H \subset G'$ and $H' \subset G$. The following lemma is quite useful in this section. See [Prasad 1996, Lemma, p. 6].

Lemma 4.2.1. *For a see-saw pair of dual reductive pairs (G, H) and (G', H'), let π be an irreducible representation of H and π' of H'. Then we have the following isomorphism*:

$$\mathrm{Hom}_H(\Theta_\psi(\pi'), \pi) \cong \mathrm{Hom}_{H'}(\Theta_\psi(\pi), \pi').$$

Let $\mathrm{GSp}(W_0 \otimes_F E)^\natural$ be the subgroup of $\mathrm{GSp}(W_0 \otimes_F E)$ where the similitude factor takes values in $F^\times$. Similarly we define

$$\mathrm{GO}(V)^\natural = \{h \in \mathrm{GO}(V) \mid \lambda_V(h) \in F^\times\}.$$

Then we have a see-saw diagram

$$\begin{tikzcd} \mathrm{GSp}(W_0 \otimes_F E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GO}(\mathrm{Res}_{E/F} V) \arrow[d, dash] \arrow[dl, dash] \\ \mathrm{GSp}(W_0) & \mathrm{GO}(V)^\natural \end{tikzcd}$$

Replace W_0 by a 4-dimensional symplectic space W_2 over F with a symplectic similitude group $\mathrm{GSp}_4(F)$. Then there is a see-saw pair

$$\big(\mathrm{GSp}_4(E)^\natural, \mathrm{GO}(V)^\natural\big) \quad \text{and} \quad \big(\mathrm{GSp}_4(F), \mathrm{GO}(\mathrm{Res}_{E/F} V)\big)$$

in the similitude symplectic group $\mathrm{GSp}(\boldsymbol{W})$, where $\boldsymbol{W} = \mathrm{Res}_{E/F}((W_2 \otimes_F E) \otimes_E V)$ and

$$\mathrm{GSp}_4(E)^\natural = \{g \in \mathrm{GSp}_4(E) | \lambda_W(g) \in F^\times\}.$$

Remark 4.2.2. Let V_F be a quadratic space over F. If the image of the similitude character λ_{V_F} is not surjective, then we need to consider the dual pair $R = \mathrm{GSp}_{4n}(F)^+ \times \mathrm{GO}(V_F)$. Moreover, $\mathrm{GSp}_{4n}(F) \times \mathrm{GO}(V_F)$ is not a dual pair in the usual sense. However, for our purpose (see Lemma 4.4.1), we will consider the induction in stages (see [Gan 2011, §9.7])

$$\mathrm{Ind}_R^{\mathrm{GSp}_{4n}(F)\times\mathrm{GO}(V_F)} \Omega_\psi = \mathrm{Ind}_R^{\mathrm{GSp}_{4n}(F)\times\mathrm{GO}(V_F)} \mathrm{ind}_{R_0}^R \, \omega_\psi,$$

where Ω_ψ (resp. ω_ψ) is the Weil representation of R (resp. R_0) defined in Section 2. Suppose that $V_F \otimes_E E$ is a split quadratic space over E. Then

$$\begin{aligned} \mathrm{Hom}_{\mathrm{GO}(V_F)}(\Theta_\psi(\tau), \chi) &= \mathrm{Hom}_{\mathrm{GSp}_{2n}(E)^\natural \times \mathrm{GO}(V_F)}(\mathrm{Ind}_R^{\mathrm{GSp}_{4n}(F)\times\mathrm{GO}(V_F)} \Omega_\psi, \tau \boxtimes \chi) \\ &= \mathrm{Hom}_{\mathrm{GSp}_{2n}(E)^\natural}(\mathrm{Ind}_{\mathrm{GSp}_{4n}(F)^+}^{\mathrm{GSp}_{4n}(F)} \Theta_\psi(\chi), \tau) \end{aligned}$$

for a representation $\tau \in \mathrm{Irr}(\mathrm{GSp}_{2n}(E)^\natural)$ and a character χ of $\mathrm{GO}(V_F)$.

In order to use Lemma 4.2.1, we need to figure out the discriminant and Hasse invariant of the quadratic space $\mathrm{Res}_{E/F} V$ over F.

Assume that $E = F(\sqrt{d})$ is a quadratic field extension of F, where $d \in F^\times \setminus F^{\times 2}$. Let D_E be the nonsplit quaternion algebra with involution $*$ defined over E with a norm map N_{D_E}, which is a 4-dimensional quadratic space V over E. More precisely, D_E is a noncommutative E-algebra generated by $1, i$ and j, denoted by $\left(\frac{a,b}{E}\right)$, where $i^2 = a$, $j^2 = b$, $ij = -ji$, $a, b \in E^\times$ and $(a,b)_E = -1$. Here $(-,-)_E$ is the Hilbert symbol defined on $E^\times \times E^\times$. Then there is an isomorphism for the vector space $\mathrm{Res}_{E/F} V$,

$$\mathrm{Res}_{E/F} D_E \cong \mathrm{Span}_F\left\{1, \sqrt{d}, i, \sqrt{d}i, j, \sqrt{d}j, ij, \sqrt{d}ij\right\}$$

as F-vector spaces. Given a vector $v \in V$, set

$$q_F(v) = \tfrac{1}{2}\,\mathrm{tr}_{E/F} \circ N_{D_E}(v) \quad \text{and} \quad (v_i, v_j) = q_F(v_i + v_j) - q_F(v_i) - q_F(v_j).$$

Lemma 4.2.3. *The quadratic space* $\mathrm{Res}_{E/F} D_E$ *with quadratic form* $\frac{1}{2}\,\mathrm{tr}_{E/F} \circ N_{D_E}$ *over F has dimension* 8, *discriminant* 1 *and Hasse-invariant* -1.

Proof. The nonsplit quaternion algebra over a nonarchimedean local field is unique. We may assume that

$$i^2 = a \in F^\times$$

and $j^2 = b = b_1 + b_2\sqrt{d}$, $N_{E/F}(b) = b_1^2 - b_2^2 d$, $b_i \in F$.

For an element $v = x_1 + x_2 i + x_3 j + x_4 ij$ in D_E with $x_i \in E$, we have

$$\tfrac{1}{2}(v,v) = N_{D_E}(v) = vv^* = x_1^2 - a x_2^2 - b x_3^2 + ab x_4^2$$

and the corresponding matrix for the quadratic space $(\mathrm{Res}_{E/F} D_E, q_F)$ is

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2d & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2ad & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2b_1 & -2b_2 d & 0 & 0 \\ 0 & 0 & 0 & 0 & -2b_2 d & -2b_1 d & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2ab_1 & 2dab_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2dab_2 & 2dab_1 \end{pmatrix}.$$

The discriminant algebra of $\mathrm{Res}_{E/F} D_E$ is trivial in $F^\times / F^{\times 2}$. If $b_1 = 0$, then the Hasse-invariant is

$$(-d, a) = -1$$

since $(b_2\sqrt{d}, a)_E = -1$, where $(-,-)$ is the Hilbert symbol defined on $F^\times \times F^\times$. If $b_1 \neq 0$, then the Hasse-invariant is

$$(d,d)(-a,-ad)\left(-b_1, \frac{N_{E/F}(b)d}{-b_1}\right)(N_{E/F}(b)d, -1)\left(ab_1, \frac{N_{E/F}(b)d}{ab_1}\right) = (a, N_{E/F}(b)) = (a,b)_E = -1,$$

because $(a,b)_E = (a, N_{E/F}(b))$ for all $a \in F^\times$ and $b \in E^\times$. $\square$

Now let V be the split $2n$-dimensional quadratic space $\mathbb{H}_E^n$ over E. There is a basis $\{e_i, e'_j\}_{1\le i,j\le n}$ for the quadratic space V satisfying $\langle e_i, e'_j\rangle = \delta_{ij}$ and the other inner products are zero. Then we fix the basis

$$\{e_i, \sqrt{d}e_i, e'_j, e'_j/\sqrt{d}\}_{1\le i,j\le n}$$

for $\mathrm{Res}_{E/F} V$. It is straightforward to check that the vector space $\mathrm{Res}_{E/F} V$ is isomorphic to the split $4n$-dimensional quadratic space $\mathbb{H}^{2n}$ over F.

4C. *The structure of degenerate principal series.* In this subsection, we follow the notation in [Gan and Ichino 2011; Kudla 1996]. Let $H_n = \mathrm{GO}(\mathbb{H}^n)$ be the orthogonal similitude group. Define the quadratic character ν to be

$$\nu(h) = \det(h)\cdot\lambda_V^{-n}(h) \text{ for } h\in \mathrm{GO}(\mathbb{H}^n)$$

so that $\nu|_{\mathrm{O}(\mathbb{H}^n)} = \det$. Define

$$\mathrm{GSO}(\mathbb{H}^n) = \ker\nu = \{h\in\mathrm{GO}(\mathbb{H}^n) \mid \lambda(h)^n = \det(h)\}.$$

Assume that Q_n is the standard Siegel parabolic subgroup of H_n, i.e.,

$$Q_n = \left\{ \begin{pmatrix} A^{-1} & \\ & \lambda A^t\end{pmatrix}\begin{pmatrix} I & X\\ & I\end{pmatrix} \;\middle|\; A\in \mathrm{GL}_n(F),\, X\in \mathrm{Mat}_{n,n}(F) \text{ and } X+X^t=0\right\}$$

with modular character $|\det A|_F^{1-n}|\lambda|_F^{-n(n-1)/2}$. Then $Q_n\backslash H_n$ is a projective variety and a homogenous space equipped with H_n-action. Each point on $Q_n\backslash H_n$ corresponds to an isotropic subspace in $\mathbb{H}^n$ of dimension n. Set the degenerate principal series representation $I_{Q_n}^{H_n}(s)$ as

$$I_{Q_n}^{H_n}(s) = \bigl\{ f : H_n\to\mathbb{C} \mid f(xg) = \delta_{Q_n}(x)^{1/2+s/(n-1)} f(g) \text{ for } x\in Q_n,\, g\in H_n\bigr\}.$$

Let W_r be the symplectic space with a symplectic similitude group $\mathrm{GSp}(W_r)$. Set $\mathbf{1}_W$ to be the trivial representation of $\mathrm{GSp}(W_r)$. Then the big theta lift $\Theta_r(\mathbf{1}_W)$ to H_n of the trivial representation $\mathbf{1}_W$ is isomorphic to a subrepresentation of $I_{Q_n}^{H_n}(s_0)$, where

$$s_0 = r - \tfrac{1}{2}(n-1).$$

The image of $\Theta_r(\mathbf{1}_W)$ in $I_{Q_n}^{H_n}(s_0)$ is denoted by $R_r(\mathbf{1})$, i.e.,

$$\Theta_r(\mathbf{1}_W) = R_r(\mathbf{1}) \subset I_{Q_n}^{H_n}(s_0).$$

Let us come back to the GSp_4-cases. Assume that $r=2$ and $n=4$.

Proposition 4.3.1. *There is an exact sequence of H_4-modules*

$$0\longrightarrow R_2(\mathbf{1}) \longrightarrow I_{Q_4}^{H_4}\bigl(\tfrac{1}{2}\bigr) \longrightarrow R_1(\mathbf{1})\otimes\nu \longrightarrow 0.$$

Proof. Note that $R_2(\mathbf{1})|_{\mathrm{O}_{4,4}(F)}$ is isomorphic to the big theta lift of the trivial representation $\mathbf{1}_W$ from $\mathrm{Sp}_4(F)$ to $\mathrm{O}_{4,4}(F)$, and similarly for the big theta lift $R_1(\mathbf{1})$. There is only one orbit for the double coset

$$Q_4\backslash H_4/\mathrm{O}_{4,4}(F) = (Q_4\cap \mathrm{O}_{4,4}(F))\backslash \mathrm{O}_{4,4}(F)/\mathrm{O}_{4,4}(F).$$

Applying Mackey theory, we have $I_{Q_4}^{H_4}\big(\frac{1}{2}\big)|_{\mathrm{O}_{4,4}(F)} \cong I_{Q_4\cap \mathrm{O}_{4,4}(F)}^{\mathrm{O}_{4,4}(F)}\big(\frac{1}{2}\big)$. Then the sequence is still the same when restricted to the orthogonal group $\mathrm{O}_{4,4}(F)$. The sequence is exact when restricted to the orthogonal group $\mathrm{O}_{4,4}(F)$ due to the structure of degenerate principal series (see [Gan and Ichino 2014, Proposition 7.2]). By the construction of the extended Weil representation, the sequence is exact as H-modules. □

Similarly, let $P_4 = M_4N_4$ be the Siegel parabolic subgroup of $\mathrm{GSp}(W_4) = \mathrm{GSp}_8(F)$ where $M_4 \cong \mathrm{GL}_1(F) \times \mathrm{GL}_4(F)$ is the Levi part of the parabolic subgroup. Let $\mathcal{I}(s)$ be the degenerate normalized induced representation of $\mathrm{GSp}_8(F)$ associated to P_4, i.e.,

$$\mathcal{I}(s) = \big\{f : \mathrm{GSp}_8(F) \to \mathbb{C} \mid f(pg) = \delta_{P_4}(p)^{(1/2)+(s/5)} f(g) \text{ for } p \in P_4,\ g \in \mathrm{GSp}_8(F)\big\}.$$

Then we have:

Proposition 4.3.2. *There is an exact sequence of* $\mathrm{GSp}_8(F)$*-modules*

$$0 \longrightarrow R^{3,3}(\mathbf{1}) \longrightarrow \mathcal{I}\big(\tfrac{1}{2}\big) \longrightarrow R^{4,0}(\mathbf{1}) \longrightarrow 0,$$

where $\mathcal{I}(s)$ is the degenerate normalized induced representation of $\mathrm{GSp}_8(F)$ *and* $R^{3,3}(\mathbf{1})$ (*resp.* $R^{4,0}(\mathbf{1})$) *is the big theta lift to* $\mathrm{GSp}_8(F)$ *of the trivial representation of* $\mathrm{GO}_{3,3}(F)$ (*resp.* $\mathrm{GO}_{4,0}(F)$).

Now we use Mackey theory to study $I_{Q_4}^{H_4}\big(\frac{1}{2}\big)|_{\mathrm{GO}_{2,2}(E)^\natural}$ which involves the computation for the double coset $Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural$. Denote $X_4 = Q_4\backslash H_4$ as the projective variety.

4C1. *Double cosets.* Now let us consider the double coset

$$Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural.$$

Assume that $V = \mathbb{H}_E^2$ with basis $\{e_i, e_j'\}_{1\le i,j\le 2}$ and $\langle e_i, e_j'\rangle = \delta_{ij}$. Fix the basis

$$\{e_1, \sqrt{d}e_1, e_2, \sqrt{d}e_2, e_1', e_1'/\sqrt{d}, e_2', e_2'/\sqrt{d}\}$$

for $V_F = \mathrm{Res}_{E/F} V$. The inner product $\langle\langle -, -\rangle\rangle$ on V_F is given by

$$\langle\langle x, y\rangle\rangle := \tfrac{1}{2}\,\mathrm{tr}_{E/F}(\langle x, y\rangle)$$

for $x, y \in V$. Let us fix an embedding $i : \mathrm{GO}_{2,2}(E)^\natural \to \mathrm{GSO}_{4,4}(F)$.

The double coset decomposition for the case at hand can be obtained from more general case. Assume that $\boldsymbol{V}$ is a symplectic space or a split quadratic space over E of dimension $2n$, with a nondegenerate bilinear form $B : \boldsymbol{V} \times \boldsymbol{V} \to E$. Let $U(\boldsymbol{V})$ be the isometry group, i.e.,

$$U(\boldsymbol{V}) = \big\{g \in \mathrm{GL}(\boldsymbol{V}) \mid B(gx, gy) = B(x, y) \text{ for all } x, y \in \boldsymbol{V}\big\}$$

which is a symplectic group or an orthogonal group. Then $\mathrm{Res}_{E/F}\boldsymbol{V}$ is a vector space over F of dimension $4n$ with a nondegenerate bilinear form $\frac{1}{2}\,\mathrm{tr}_{E/F} \circ B$.

Lemma 4.3.3. *Let P be a Siegel parabolic subgroup of $U(\mathrm{Res}_{E/F}\boldsymbol{V})$. Then each point in the homogeneous space $X = P\backslash U(\mathrm{Res}_{E/F}\boldsymbol{V})$ corresponds to a $2n$-dimensional maximal isotropic subspace in $\mathrm{Res}_{E/F}\boldsymbol{V}$ and the finite double cosets $X/U(\boldsymbol{V})$ can be parametrized by a pair*

$$(\dim_E E\cdot L,\, B_L),$$

where $L \subset \mathrm{Res}_{E/F}\boldsymbol{V}$ is a maximal isotropic subspace with respect to the inner product $\langle\langle -,-\rangle\rangle$ over F,

$$E\cdot L := \{e\cdot x \mid e\in E,\, x\in L\}$$

is a linear E-subspace in $\boldsymbol{V}$ and

$$B_L : L/L_0 \times L/L_0 \to \sqrt{d}\cdot F \tag{4-1}$$

is a nondegenerate bilinear form inherited from V, where

$$L_0 = \{x\in L : B(x,y) = 0 \text{ for all } y\in L\}.$$

Moreover, if $L = L_0$, then L lies in the closed orbit. If $L_0 = 0$, then L lies in the open orbit.

Proof. Under a suitable basis for L, the bilinear form for $B|_L$ corresponds to a matrix $\sqrt{d}\cdot T$, where $T\in M_{2n}(F)$. Moreover, we can choose T such that it is a diagonal (resp. an anti-diagonal) matrix if $B(x,y) = B(y,x)$ (resp. $B(y,x) = -B(x,y)$). Then

$$\dim_E E\cdot L = n + \tfrac{1}{2}\cdot \mathrm{rank}(T),$$

which is invariant under $U(\boldsymbol{V})$-action. The bilinear form B_L corresponds to a matrix $\sqrt{d}\cdot T'$, i.e.,

$$T = \begin{pmatrix} 0 & 0 & 0\\ 0 & T' & 0\\ 0 & 0 & 0\end{pmatrix}$$

where T' is invertible and $\mathrm{rank}(T) = \mathrm{rank}(T')$.

Assume that there are two isotropic subspaces L_1 and L_2 satisfying

$$\dim_E E\cdot L_1 = \dim_E E\cdot L_2 = l \quad\text{and}\quad B_{L_1}\cong B_{L_2}.$$

This means that there exists $g\in \mathrm{GL}_l(E)$ such that $g : E\cdot L_1 \to E\cdot L_2$ satisfying

$$B_{L_1}(x,y) = B_{L_2}(gx, gy).$$

It is easy to lift g to $g_E\in U(\boldsymbol{V})$ such that $g_E L_1 = L_2$.

In fact, $g = \left(\begin{smallmatrix} g_1 & 0\\ 0 & g_2\end{smallmatrix}\right)$ lies in a subgroup of $\mathrm{GL}_l(E)$, which can be regarded as a Levi subgroup of $U(\boldsymbol{V})$, and

$$B_L(gx, gy) = B_L(g_2x', g_2y')$$

when $x - x', y - y'\in L_0$. Then $g_E = \begin{pmatrix} g_1 & g_2 & \\ & & g_1^*\end{pmatrix}\in U(\boldsymbol{V})$, where g_1^* depends on g_1 and $\boldsymbol{V}$. □

Remark 4.3.4. There is only one closed orbit in the double coset $P\backslash U(\mathrm{Res}_{E/F}\boldsymbol{V})/U(\boldsymbol{V})$. When $T = 0$, the subspace $E\cdot L$ is the maximal isotropic subspace of $\boldsymbol{V}$ and so $U(\boldsymbol{V})$ acts on the subvariety $\{L : L = L_0\}\subset X$ transitively.

Consider the double coset decomposition of

$$Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural.$$

There are several F-rational orbits in $Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural$. By Lemma 4.3.3, there are two invariants for the orbit $\mathrm{GO}_{2,2}(E)^\natural \cdot L$:

- the dimension $\dim_E(E \cdot L)$, and
- the bilinear form B_L (defined in (4-1)) up to scaling in $F^\times$.

By the classification of 4-dimensional quadratic spaces over F, there are 4 elements lying in the kernel

$$\ker\{H^1(F, \mathrm{O}_4) \to H^1(E, \mathrm{O}_4)\},$$

which are

- the split quaternion algebra $\mathrm{Mat}_{2,2}(F)$ with $q(v) = \det(v)$ for $v \in \mathrm{Mat}_{2,2}(F)$,
- the quaternion division algebra $D(F)$ with the norm map $N_{D/F}$,
- the nonsplit 4-dimensional quadratic space $V_3 = E \oplus \mathbb{H}$ with $q(e, x, y) = N_{E/F}(e) - xy$, and
- $V_4 = \epsilon V_3$ with $\epsilon \in F^\times \setminus N_{E/F}(E^\times)$.

However, we consider the double coset

$$Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural$$

for the similitude groups and observe that V_3 and V_4 are in the same orbit in $Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural$. More precisely, $\mathrm{Mat}_{2,2}(F)$, $D(F)$ and $E \oplus \mathbb{H}$ are three representatives in the union of the open orbits $\mathrm{GO}_{2,2}(E)^\natural \cdot L$ in $X_4/\mathrm{GO}_{2,2}(E)^\natural$.

Proposition 4.3.5. *Pick a point $L \in X_4/\mathrm{GO}_{2,2}(E)^\natural$ lying in an open orbit. Then the stabilizer of L in $\mathrm{GO}_{2,2}(E)^\natural$ is isomorphic to the similitude group $\mathrm{GO}(L)$.*

Proof. For $g \in \mathrm{GO}_{2,2}(E)^\natural$ with $g(L) = L$, we have

$$\langle gl_1, gl_2\rangle = \lambda(g) \cdot \langle l_1, l_2\rangle$$

and so $\langle\langle gl_1, gl_2\rangle\rangle = \lambda(g) \cdot \langle\langle l_1, l_2\rangle\rangle$. This means $g \in \mathrm{GO}(L)$. Conversely, if $h \in \mathrm{GO}(L, (1/\sqrt{d})q_E|_L)$, set

$$h_E : x \otimes e \mapsto h(x) \otimes e$$

for $x \otimes e \in L \otimes E \cong L \cdot E = V$. Then $h_E(L) = L$ and

$$\langle h_E(x_1 \otimes e_1), h_E(x_2 \otimes e_2)\rangle = e_1 e_2 \lambda(h)\langle\langle x_1, x_2\rangle\rangle = \lambda(h)\langle x_1 \otimes e_1, x_2 \otimes e_2\rangle,$$

i.e., $h_E \in \mathrm{GO}_{2,2}(E)^\natural$. Then we get a bijection between the similitude orthogonal group $\mathrm{GO}(L)$ and the stabilizer of L in $\mathrm{GO}_{2,2}(E)^\natural$. Observe that the map $h \mapsto h_E$ is a group homomorphism. Then $\mathrm{GO}(L)$ is isomorphic to the stabilizer of L via the map $h \mapsto h_E$. $\square$

There are three F-rational open orbits $\mathrm{GO}_{2,2}(E)^\natural \cdot L$ where L represents one of $\mathrm{Mat}_{2,2}(F)$, $D(F)$ or $E \oplus \mathbb{H}$, whose stabilizers are $\mathrm{GO}_{2,2}(F)$, $\mathrm{GO}_{4,0}(F)$ and $\mathrm{GO}_{3,1}(F)$ respectively. There is one closed orbit $\mathrm{GO}_{2,2}(E)^\natural \cdot L$ which has stabilizer

$$\mathrm{GO}_{2,2}(E)^\natural \cap Q_4 =: Q^\natural \cong \left\{ \begin{pmatrix} A^{-1} & * \\ 0 & \lambda A^t \end{pmatrix} \,\middle|\, A \in \mathrm{GL}_2(E), \lambda \in F^\times \right\}.$$

There are two intermediate orbits with representatives L_1, L_2 and $\dim_E(E \cdot L_i) = 3$. The stabilizers are isomorphic to

$$(\mathrm{GL}_1(E) \times \mathrm{GO}_{1,1}(F)) \cdot \mathrm{Mat}_{2,2}(F) \quad \text{and} \quad (\mathrm{GL}_1(E) \times \mathrm{GO}(\mathcal{V}_E)) \cdot \mathrm{Mat}_{2,2}(F),$$

where $\mathcal{V}_E$ is the 2-dimensional quadratic space over F whose discriminant algebra is E.

Remark 4.3.6. For $(g, t) \in \mathrm{GL}_2(E) \times F^\times$, we set

$$\beta((g, t)) = (g, \sigma(g) \cdot t) \in \mathrm{GL}_2(E) \times \mathrm{GL}_2(E).$$

Then $\beta : \mathrm{GSO}_{3,1}(F) \to \mathrm{GSO}_{2,2}(E)^\natural$ is an embedding due to the exact sequences

$$\begin{tikzcd} 1 \arrow[r] & E^\times \arrow[r, "i_1"] \arrow[d, equal] & \mathrm{GL}_2(E) \times F^\times \arrow[r] \arrow[d, "\beta"] & \mathrm{GSO}_{3,1}(F) \arrow[r] \arrow[d] & 1 \\ 1 \arrow[r] & E^\times \arrow[r, "i_2"] & \mathrm{GL}_2(E) \times \mathrm{GL}_2(E) \arrow[r] & \mathrm{GSO}_{2,2}(E) \arrow[r] & 1 \end{tikzcd}$$

where $i_1(e) = (e, N_{E/F}(e)^{-1})$ and $i_2(e) = (e, e^{-1})$ for $e \in E^\times$.

There are several orbits for $X_4/\mathrm{GO}_{2,2}(E)^\natural$. By Mackey theory, there is a decreasing filtration of $\mathrm{GO}_{2,2}(E)^\natural$-modules for $I_{Q_4}^{H_4}(s)|_{\mathrm{GO}_{2,2}(E)^\natural}$.

4C2. *Filtration.* Consider the filtration

$$I_{Q_4}^{H_4}(s) = I_2(s) \supset I_1(s) \supset I_0(s) \supset 0$$

of $I_{Q_4}^{H_4}(s)|_{\mathrm{GO}_{2,2}(E)^\natural}$ with a sequence of subquotients

$$\begin{aligned} I_0(s) &= \mathrm{ind}_{\mathrm{GO}_{2,2}(F)}^{\mathrm{GO}_{2,2}(E)^\natural} \mathbb{C} \oplus \mathrm{ind}_{\mathrm{GO}_{4,0}(F)}^{\mathrm{GO}_{2,2}(E)^\natural} \mathbb{C} \oplus \mathrm{ind}_{\mathrm{GO}_{3,1}(F)}^{\mathrm{GO}_{2,2}(E)^\natural} \mathbb{C}, \\ I_2(s)/I_1(s) &\cong \mathrm{ind}_{Q^\natural}^{\mathrm{GO}_{2,2}(E)^\natural} \delta_{Q^\natural}^{s+1}, \end{aligned}$$

where $Q^\natural$ is the Siegel parabolic subgroup of $\mathrm{GO}_{2,2}(E)^\natural$ with modular character $\delta_{Q^\natural}$ and

$$I_1(s)/I_0(s) \cong \mathrm{ind}_{(\mathrm{GL}_1(E) \times \mathrm{GO}_{1,1}(F)) \cdot N}^{\mathrm{GO}_{2,2}(E)^\natural} \delta_Q^{\frac{1}{2}+\frac{s}{3}} \delta_1^{-\frac{1}{2}} \oplus \mathrm{ind}_{Q'}^{\mathrm{GO}_{2,2}(E)^\natural} \delta_Q^{\frac{1}{2}+\frac{s}{3}} \delta_2^{-\frac{1}{2}}$$

where $Q' = (\mathrm{GL}_1(E) \times \mathrm{GO}(\mathcal{V}_E)) \cdot N$, $N \cong \mathrm{Mat}_{2,2}(F)$ and

$$\delta_i(t, h) = |N_{E/F}(t^2) \cdot \lambda_V(h)^{-2}|_F$$

for $t \in \mathrm{GL}_1(E)$ and $h \in \mathrm{GO}_{1,1}(F)$ or $\mathrm{GO}(\mathcal{V}_E)$, where $\mathcal{V}_E$ is the nonsplit 2-dimensional quadratic space.

Remark 4.3.7. We would like to highlight the fact that on the open orbits related to $I_0(s)$, the group embedding $\mathrm{GO}_{2,2}(F) \hookrightarrow \mathrm{GO}_{2,2}(E)^\natural$ (and similarly for the other two group embeddings) is not induced from the geometric embedding $i : \mathrm{GO}(L) \hookrightarrow \mathrm{GO}(L \otimes_F E)$, but the composite map $\mathrm{Ad}_{h^\delta} \circ i$ of the adjoint map Ad_{h^δ} and the geometric embedding i where

$$h^\delta = \begin{pmatrix} \sqrt{d} & \\ & 1 \end{pmatrix} \in \mathrm{GO}(2,2)(E).$$

However, it does not affect the results when we consider the distinction problems for the similitude groups. In Section 4D, we will show that the results on the open orbits determine the distinction problems $\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}\big(\tfrac{1}{2}\big), \Sigma\big)$ when Σ is a generic representation.

Recall that

$$\mathrm{GSp}_4(E)^\natural = \{g \in \mathrm{GSp}_4(E) \mid \lambda_W(g) \in F^\times\}.$$

When we deal with the case

$$\mathrm{Ind}_{P_4}^{\mathrm{GSp}_8(F)} \delta_{P_4}^{s/5}|_{\mathrm{GSp}_4(E)^\natural},$$

where P_4 is the Siegel parabolic subgroup of $\mathrm{GSp}_8(F)$ with modular character δ_{P_4}, the above results still hold. More precisely, set

$$\mathcal{I}(s) = \big\{f : \mathrm{GSp}_8(F) \to \mathbb{C} \mid f(xg) = \delta_{P_4}(x)^{(1/2)+(s/5)} f(g) \text{ for } x \in P_4,\, g \in \mathrm{GSp}_8(F)\big\}.$$

There is a filtration

$$\mathcal{I}_0(s) \subset \mathcal{I}_1(s) \subset \mathcal{I}_2(s) = \mathcal{I}(s)|_{\mathrm{GSp}_4(E)^\natural}$$

of $\mathcal{I}(s)|_{\mathrm{GSp}_4(E)^\natural}$ such that

- $\mathcal{I}_0(s) \cong \mathrm{ind}_{\mathrm{GSp}_4(F)}^{\mathrm{GSp}_4(E)^\natural} \mathbb{C}$,
- $\mathcal{I}_1(s)/\mathcal{I}_0(s) \cong \mathrm{ind}_{M'N'}^{\mathrm{GSp}_4(E)^\natural} \delta_{P_4}^{(1/2)+(s/5)} \delta_{M'N'}^{-1/2}$ and
- $\mathcal{I}_2(s)/\mathcal{I}_1(s) \cong \mathrm{ind}_{P^\natural}^{\mathrm{GSp}_4(E)^\natural} \delta_{P^\natural}^{(s+1)/3}$,

where $P^\natural$ is the Siegel parabolic subgroup of $\mathrm{GSp}_4(E)^\natural$,

$$M' \cong \mathrm{GL}_1(E) \times \mathrm{GL}_2(F), \quad N' \cong \mathrm{Mat}_{1,1}(E) \oplus \mathrm{Mat}_{2,2}(F)$$

and

$$\delta_{M'N'}(t, g) = |N_{E/F}(t)^4 \cdot \lambda_W(g)^{-4}|_F$$

for $(t, g) \in \mathrm{GL}_1(E) \times \mathrm{GL}_2(F)$. Here the group embedding $\mathrm{GSp}_4(F) \hookrightarrow \mathrm{GSp}_4(E)^\natural$ in $\mathcal{I}_0(s)$ is the composition map $\mathrm{Ad}_{g^\delta} \circ i'$ where $i' : \mathrm{GSp}(W_2) \hookrightarrow \mathrm{GSp}(W_2 \otimes_F E)$ is the geometric embedding and

$$g^\delta = \begin{pmatrix} \sqrt{d} & \\ & 1 \end{pmatrix} \in \mathrm{GSp}_4(E).$$

4D. ***The distinction problem for*** $\mathbf{GSp_4}$**.** Let us recall what we have obtained. Let $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$. Since $\tau|_{\mathrm{Sp}_4(E)}$ is multiplicity-free due to [Adler and Prasad 2006, Theorem 1.4], $\tau|_{\mathrm{GSp}_4(E)^\natural}$ is multiplicity-free. Assume that $\tau = \theta(\pi_1 \boxtimes \pi_2)$ participates in the theta correspondence with $\mathrm{GSO}_{2,2}(E)$. Then the see-saw identity implies that

$$\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) \subset \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\Theta_2(\Sigma), \mathbb{C}) \cong \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_2(\mathbf{1}), \Sigma),$$

where $R_2(\mathbf{1})$ is the image of the big theta lift to H_4 of the trivial representation of $\mathrm{GSp}_4(F)$ in $I_{Q_4}^{H_4}\left(\frac{1}{2}\right)$ and Σ is the irreducible representation of $\mathrm{GO}_{2,2}(E)$ such that $\tau = \theta(\Sigma)$. In fact, if $\pi_1 \ncong \pi_2$, then $\Sigma = \mathrm{Ind}_{\mathrm{GSO}_{2,2}(E)}^{\mathrm{GO}_{2,2}(E)}(\pi_1 \boxtimes \pi_2)$. If $\pi_1 \cong \pi_2$, then there are two extensions to $\mathrm{GO}_{2,2}(E)$ of $\pi_1 \boxtimes \pi_2$. The representation Σ is the unique extension of $\pi_1 \boxtimes \pi_1$ which participates into the theta correspondence with $\mathrm{GSp}_4(E)$, denoted by $(\pi_1 \boxtimes \pi_1)^+$.

Lemma 4.4.1. *Assume that $\pi_1 \boxtimes \pi_2 \in \mathrm{Irr}(\mathrm{GSO}_{2,2}(E))$. Let $\Sigma \in \mathrm{Irr}(\mathrm{GO}_{2,2}(E))$ such that $\Sigma|_{\mathrm{GSO}_{2,2}(E)} \supset \pi_1 \boxtimes \pi_2$ and Σ has a nonzero theta lift to $\mathrm{GSp}_4(E)$. Then*

$$\dim \mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSO}(L)}(\pi_1 \boxtimes \pi_2, \mathbb{C}),$$

where $\mathrm{GO}(L) \hookrightarrow \mathrm{GO}(L \otimes_F E) = \mathrm{GO}_{2,2}(E)$ and the 4-dimensional quadratic space L is one of $\mathrm{Mat}_{2,2}(F)$, $D(F)$ or $E \oplus \mathbb{H}_F$.

Proof. If $\pi_1 \neq \pi_2$, then it follows from Frobenius reciprocity. If $\pi_1 = \pi_2$ and L is either $\mathrm{Mat}_{2,2}(F)$ or $D(F)$, then we consider the see-saw diagram

$$\begin{tikzcd}
\mathrm{GO}_{2,2}(E)^\natural \arrow[d, dash] \arrow[drr, dash] & & \mathrm{GSp}_4(F) \arrow[d, dash] \arrow[dll, dash] \\
\mathrm{GO}(L) & & \mathrm{GSp}_2(E)^\natural
\end{tikzcd}$$

where $\mathrm{GSp}_2(E)^\natural = \{g \in \mathrm{GSp}_2(E) | \lambda_W(g) \in F^\times\}$. We have

$$\mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma \otimes \nu, \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma, \nu) = \mathrm{Hom}_{\mathrm{GSp}_2(E)^\natural}(\Theta_2(\nu), \pi_1) = 0,$$

because the big theta lift $\Theta_2(\nu)$ to $\mathrm{GSp}_4(F)$ is zero by the conservation relation. If $\pi_1 = \pi_2$ and L is $E \oplus \mathbb{H}_F$, then

$$\mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma, \nu) = \mathrm{Hom}_{\mathrm{GSp}_2(E)^\natural}(\mathrm{Ind}_{\mathrm{GSp}_4(F)^+}^{\mathrm{GSp}_4(F)} \Theta_2(\nu), \mathbb{C}) = 0.$$

(See Remark 4.2.2.) Hence

$$\mathrm{Hom}_{\mathrm{GSO}(L)}(\pi_1 \boxtimes \pi_2, \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma \oplus (\Sigma \otimes \nu), \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}(L)}(\Sigma, \mathbb{C}).$$

This finishes the proof. □

Lemma 4.4.2. *Given a representation $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$ with $\omega_\tau|_{F^\times} = \mathbf{1}$, we have*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau^g, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau^\vee, \mathbb{C}),$$

where $\tau^g(x) = \tau(gxg^{-1})$ for $g \in \mathrm{GSp}_4(E)$.

Proof. Note that $\tau^g \cong \tau$ and so $\dim \operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau^g, \mathbb{C}) = \dim \operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C})$. Since $\omega_\tau|_{F^\times}$ is trivial and $\tau^\vee \cong \tau \otimes \omega_\tau^{-1}$, we have

$$\operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau^\vee, \mathbb{C}) = \operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau \otimes \omega_\tau^{-1}, \mathbb{C}) = \operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_\tau|_{F^\times}) = \operatorname{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}). \qquad \square$$

Remark 4.4.3. We have a similar statement for the group $\mathrm{GO}(V)$ when V is a 4-dimensional split quadratic space over E.

There is another key input for the GL_4-distinction problems in our proof of Theorem 1.1.

Theorem 4.4.4 [Matringe 2011, Theorem 5.2]. *Given a generic representation π of $\mathrm{GL}_n(E)$ with a Langlands parameter $\phi_\pi = \triangle_1 \oplus \triangle_2 \oplus \cdots \oplus \triangle_t$ with $\triangle_i : WD_E \to \mathrm{GL}_{n_i}(\mathbb{C})$ irreducible and $\sum_{i=1}^t n_i = n$, then π is $\mathrm{GL}_n(F)$-distinguished if and only if there is a reordering of $\triangle_i'$s and an integer r between 1 and $\frac{1}{2}t$ such that $\triangle_{i+1}^\sigma = \triangle_i^\vee$ for $i = 1, 3, \ldots, 2r-1$ and $\triangle_i$ is conjugate-orthogonal for $i > 2r$.*

Lemma 4.4.5. *Let π be a square-integrable representation of $\mathrm{GL}_2(E)$. Then π is $\mathrm{GL}_2(F)$-distinguished if and only if π is $D^\times(F)$-distinguished. If $\pi = \pi(\chi^{-1}, \chi^\sigma)$, then π is both $\mathrm{GL}_2(F)$-distinguished and $D^\times(F)$-distinguished. Let $\pi_0 = \pi(\chi_1, \chi_2)$ with $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \mathbf{1}$ be an irreducible smooth representation of $\mathrm{GL}_2(E)$. Then π_0 is $\mathrm{GL}_2(F)$-distinguished but not $D^\times(F)$-distinguished. These exhaust all generic $\mathrm{GL}_2(F)$-distinguished representations of $\mathrm{GL}_2(E)$.*

Proof. If π is square-integrable, then it follows from [Prasad 1992, Theorem C]. Let $\pi_0 = \pi(\chi_1, \chi_2)$. By Mackey theory, we know that

$$\dim \operatorname{Hom}_{D^\times(F)}(\pi_0, \mathbb{C}) = \dim \operatorname{Hom}_{E^\times}(\chi_1 \chi_2^\sigma, \mathbb{C}) = \begin{cases} 1 & \text{if } \chi_1\chi_2^\sigma = \mathbf{1}, \\ 0 & \text{otherwise.} \end{cases}$$

If $\chi_1 \neq \chi_2$ and $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \mathbf{1}$, then $\chi_1\chi_2^\sigma \neq \mathbf{1}$. Thus π_0 is not $D^\times(F)$-distinguished. Since the Langlands parameter $\phi_\pi = \chi^{-1} \oplus \chi^\sigma$ (resp. ϕ_{π_0}) is conjugate-orthogonal in the sense of [Gan et al. 2012, §3], π (resp. π_0) is $\mathrm{GL}_2(F)$-distinguished due to [Gan and Raghuram 2013, Theorem 6.2] or Theorem 4.4.4. The last claim follows from Theorem 4.4.4. $\square$

Lemma 4.4.6. *Let π be an essentially discrete series representation of $\mathrm{GL}_2(E)$. Let $\Pi = J_P(\pi|-|_E, \pi)$ be the nongeneric representation of $\mathrm{GL}_4(E)$. Then the following statements are equivalent*:

(i) *Π is either $\mathrm{GL}_4(F)$-distinguished or $(\mathrm{GL}_4(F), \omega_{E/F})$-distinguished.*

(ii) $\Pi^\vee \cong \Pi^\sigma$.

(iii) *$I_P(\pi|-|_E, \pi)$ is both $\mathrm{GL}_4(F)$-distinguished and $(\mathrm{GL}_4(F), \omega)$-distinguished.*

Proof. See [Gurevich et al. 2018, Theorem 6.5]. $\square$

4D1. *The Langlands correspondence for* GSp_4. In this part, we will recall the Langlands correspondence for GSp_4 which has been set up in [Gan and Takeda 2011a].

Let $\Pi(\mathrm{GSp}_4)$ be the set of (equivalence classes of) irreducible smooth representation of $\mathrm{GSp}_4(F)$. Let $\operatorname{Hom}(WD_F, \mathrm{GSp}_4(\mathbb{C}))$ be the set of (equivalence classes of) admissible homomorphisms

$$WD_F \to \mathrm{GSp}_4(\mathbb{C}).$$

Theorem 4.4.7 (Gan–Takeda). *There is a surjective finite to one map*

$$L:\Pi(\mathrm{GSp}_4)\to \mathrm{Hom}(WD_F,\mathrm{GSp}_4(\mathbb{C}))$$

with the following properties:

(i) *τ is a (essentially) discrete series representation of $\mathrm{GSp}_4(F)$ if and only if its L-parameter $\phi_\tau=L(\tau)$ does not factor through any proper Levi subgroup of $\mathrm{GSp}_4(\mathbb{C})$.*

(ii) *For an L-parameter $\phi\in\mathrm{Hom}(WD_F,\mathrm{GSp}_4(\mathbb{C}))$, its fiber Π_ϕ can be naturally parametrized by the set of irreducible characters of the component group*

$$\pi_0(Z(\mathrm{Im}(\phi))/Z_{\mathrm{GSp}_4(\mathbb{C})}).$$

This component group is either trivial or equal to $\mathbb{Z}/2\mathbb{Z}$. When it is $\mathbb{Z}/2\mathbb{Z}$, exactly one of the two representations in Π_ϕ is generic and it is the one indexed by the trivial character of $\pi_0(Z(\mathrm{Im}(\phi))/Z_{\mathrm{GSp}_4(\mathbb{C})})$.

(iii) *The similitude character $\mathrm{sim}(\phi_\tau)$ of ϕ_τ equals the central character ω_τ of τ. Here $\mathrm{sim}:\mathrm{GSp}_4(\mathbb{C})\to\mathbb{C}^\times$ is the similitude character of $\mathrm{GSp}_4(\mathbb{C})$.*

(iv) *The L-parameter of $\tau\otimes(\chi\circ\lambda_W)$ is equal to $\phi_\tau\otimes\chi$. Here $\lambda_W:\mathrm{GSp}_4(F)\to F^\times$ is the similitude character of $\mathrm{GSp}_4(F)$, and we have regarded χ as both a character of $F^\times$ and a character W_F by local class field theory.*

Definition 4.4.8. An irreducible representation τ of $\mathrm{GSp}_4(E)^\natural$ occurs on the boundary of $\mathcal{I}(s)$ if

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\mathcal{I}_{i+1}(s)/\mathcal{I}_i(s),\tau)\neq 0\quad \text{for } i=0 \text{ or } 1.$$

In [Lu 2017a], we have verified the Prasad conjecture for GSp_4 when τ is a tempered representation by showing that τ does not occur on the boundary of $\mathcal{I}\left(\frac{1}{2}\right)$. After discussing with Dmitry Gourevitch, we realized that [Gourevitch et al. 2019, Proposition 4.9] can imply the Prasad conjecture for GSp_4 when the L-packet Π_{ϕ_τ} is generic. Thus we will give a slightly different proof of Theorem 1.1 from the one in [Lu 2017a].

We repeat the statements of Theorem 1.1 as below.

Theorem 4.4.9. *Assume that $\tau\in\mathrm{Irr}(\mathrm{GSp}_4(E))$ with a central character ω_τ satisfying $\omega_\tau|_{F^\times}=\mathbf{1}$.*

(i) *If $\tau=\theta(\Sigma)$ is an irreducible representation of $\mathrm{GSp}_4(E)$, where Σ is an irreducible representation of $\mathrm{GO}_{4,0}(E)$, then τ is not $\mathrm{GSp}_4(F)$-distinguished.*

(ii) *Suppose $\Sigma=(\pi_1\boxtimes\pi_1)^+$ is an irreducible representation of $\mathrm{GO}_{2,2}(E)$ and $\Sigma=\mathrm{Ind}_{\mathrm{GSO}_{2,2}(E)}^{\mathrm{GO}_{2,2}(E)}(\pi_1\boxtimes\pi_2)$ if $\pi_1\neq\pi_2$. If $\tau=\theta(\Sigma)$ is generic, then*

$$\dim\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})=\begin{cases}2 & \text{if } \pi_i\ncong\pi_0 \text{ are both } \mathrm{GL}_2(F)\text{-distinguished},\\ 1 & \text{if } \pi_1\ncong\pi_2 \text{ but } \pi_1^\sigma\cong\pi_2^\vee,\\ 1 & \text{if } \pi_1\cong\pi_2 \text{ is } \mathrm{GL}_2(F)\text{-distinguished but not } (\mathrm{GL}_2(F),\omega_{E/F})\text{-distinguished},\\ 1 & \text{if } \pi_2 \text{ is } \mathrm{GL}_2(F)\text{-distinguished and } \pi_1\cong\pi_0,\\ 0 & \text{otherwise.}\end{cases}$$

Here $\pi_0 = \pi(\chi_1, \chi_2)$ *with* $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = 1$.

(iii) *Assume that* τ *is not in case* (i) *or* (ii), *so that* $\tau = \theta(\Pi \boxtimes \chi)$, *where* $\Pi \boxtimes \chi$ *is a representation of* $\mathrm{GSO}_{3,3}(E)$. *If* τ *is generic, then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = \begin{cases} 1 & \text{if } \phi_\Pi \text{ is conjugate-orthogonal,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (i) If Σ is a representation of $\mathrm{GO}_{4,0}(E)$, then $\tau = \theta(\Sigma) = \Theta(\Sigma)$ and

$$\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\Theta(\Sigma), \mathbb{C}) \cong \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}(\Theta_{W,D',\psi}(\mathbf{1}), \Sigma^+),$$

where $D' = \mathrm{Res}_{E/F} D_E = D(F) \oplus \mathbb{H}^2$ is the 8-dimensional quadratic vector space over F with determinant 1 and Hasse invariant -1 due to Lemma 4.2.3 and $\Theta_{W,D',\psi}(\mathbf{1})$ is the big theta lift to $\mathrm{GO}(V')$ of the trivial representation $\mathbf{1}$. Note that the first occurrence of the trivial representation is $\dim_F W = 4$ in the Witt tower $D \oplus \mathbb{H}^r$, which is bigger than 2. Thus $\Theta_{W,D',\psi}(\mathbf{1}) = 0$. Hence

$$\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\Theta(\Sigma), \mathbb{C}) = 0$$

and so $\tau = \theta(\Sigma)$ is not $\mathrm{GSp}_4(F)$-distinguished.

(ii) By Proposition 4.3.1, there is an exact sequence

$$0 \longrightarrow R_2(\mathbf{1}) \longrightarrow I_{Q_4}^{H_4}\left(\tfrac{1}{2}\right) \longrightarrow \nu \otimes R_1(\mathbf{1}) \longrightarrow 0 \tag{4-2}$$

of H_4-representations, where $R_i(\mathbf{1})$ is the big theta lift to H_4 of the trivial representation $\mathbf{1}$ of $\mathrm{GSp}_{2i}(F)$. We take the right exact contravariant functor $\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(-, \Sigma)$ with respect to (4-2) and get a short exact sequence

$$0 \to \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_1(\mathbf{1})) \otimes \nu, \Sigma) \to \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\left(I_{Q_4}^{H_4}\left(\tfrac{1}{2}\right), \Sigma\right) \to \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_2(\mathbf{1}), \Sigma). \tag{4-3}$$

Consider the following double see-saw diagrams:

$$\begin{array}{ccccc} \mathrm{GSp}_4(E)^\natural & & H_4 & & \mathrm{GSp}_2(E)^\natural \\ | & \times & | & \times & | \\ \mathrm{GSp}_4(F) & & \mathrm{GO}_{2,2}(E)^\natural & & \mathrm{GL}_2(F) \end{array}$$

Note that $\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_2(\mathbf{1}), \Sigma) \cong \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C})$. There is a key observation due to Wee Teck Gan that $\mathrm{GO}_{2,2}(E)^\natural$ is a subgroup of $\mathrm{GSO}_{4,4}(F)$. One has

$$\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_1(\mathbf{1}) \otimes \nu, \Sigma) = \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_1(\mathbf{1}), \Sigma) \cong \mathrm{Hom}_{\mathrm{GSp}_2(F)}(\Theta_1(\Sigma), \mathbb{C}).$$

Here $\Theta_1(\Sigma)$ is the big theta lift to $\mathrm{GSp}_2(E)$ of Σ, which is zero unless $\pi_1 = \pi_2$. Then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) + \dim \mathrm{Hom}_{\mathrm{GSp}_2(F)}(\Theta_2(\Sigma), \mathbb{C}) \geq \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\left(I_{Q_4}^{H_4}\left(\tfrac{1}{2}\right), \Sigma\right). \tag{4-4}$$

Observe that $\mathrm{GO}_{2,2}(E)^\natural$ is the fixed point of a involution on H_4, which is given by the scalar matrix

$$h=\sqrt{d}\in \mathrm{GO}_{2,2}(E)^\natural\subset H_4$$

acting on H_4 by conjugation. Due to [Ólafsson 1987, Theorem 2.5], there exists a polynomial f on H_4 such that the complements of the open orbits in the double coset $Q_4\backslash H_4/\mathrm{GO}_{2,2}(E)^\natural$ is the zero set of f. Thanks to [Gourevitch et al. 2019, Proposition 4.9], the multiplicity $\dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}(\tfrac12),\Sigma\big)$ is at least $\dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_0(\tfrac12),\Sigma\big)$ where the submodule I_0 corresponds to the open orbits. More precisely,

$$I_0(\tfrac12)\cong \mathrm{ind}_{\mathrm{GO}_{4,0}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{2,2}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{3,1}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}$$

and

$$\begin{aligned}\dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}&\big(I_{Q_4}^{H_4}(\tfrac12),\Sigma\big)\\ &\geq \dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(\mathrm{ind}_{\mathrm{GO}_{4,0}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{2,2}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{3,1}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C},\Sigma\big).\end{aligned}\tag{4-5}$$

Together with (4-4), we have

$$\begin{aligned}&\dim\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GSp}_2(F)}(\Theta_2(\Sigma),\mathbb{C})\\ &\ \geq \dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}(\tfrac12),\Sigma\big)\\ &\ \geq \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(\mathrm{ind}_{\mathrm{GO}_{4,0}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{2,2}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C}\oplus\mathrm{ind}_{\mathrm{GO}_{3,1}(F)}^{\mathrm{GO}_{2,2}(E)^\natural}\mathbb{C},\Sigma\big)\\ &\ =\dim\mathrm{Hom}_{\mathrm{GO}_{4,0}(F)}(\Sigma,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(F)}(\Sigma,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GO}_{3,1}(F)}(\Sigma,\mathbb{C})\\ &\ =\dim\mathrm{Hom}_{\mathrm{GSO}_{4,0}(F)}(\pi_1\boxtimes\pi_2,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GSO}_{2,2}(F)}(\pi_1\boxtimes\pi_2,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GSO}_{3,1}(F)}(\pi_1\boxtimes\pi_2,\mathbb{C}).\end{aligned}\tag{4-6}$$

The last equality of (4-6) holds due to Lemma 4.4.1, which also equals

$$\begin{aligned}\dim\mathrm{Hom}_{D^\times(F)}(\pi_1,\mathbb{C})\dim\mathrm{Hom}_{D^\times(F)}(\pi_2,\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1,\mathbb{C})\dim\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_2,\mathbb{C})&\\ +\dim\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi_1^\sigma,\pi_2^\vee).&\end{aligned}$$

In order to get the upper bound for the multiplicity $\dim\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})$, let us turn the table around. There is an exact sequence

$$0\longrightarrow R^{3,3}(\mathbf{1})\longrightarrow\mathcal{I}(\tfrac12)\longrightarrow R^{4,0}(\mathbf{1})\longrightarrow 0$$

of $\mathrm{GSp}_8(F)$-representations, where $\mathcal{I}(s)$ is the degenerate principal series of $\mathrm{GSp}_8(F)$ and $R^{m,n}(\mathbf{1})$ is the big theta lift to $\mathrm{GSp}_8(F)$ of the trivial representation $\mathbf{1}$ of $\mathrm{GO}_{m,n}(F)$. There is only one open orbit in the double coset decomposition $P_4\backslash\mathrm{GSp}_8(F)/\mathrm{GSp}_4(E)^\natural$. In a similar way, by Lemma 4.4.2, [Ólafsson 1987, Theorem 2.5] and [Gourevitch et al. 2019, Proposition 4.9],

$$\begin{aligned}\dim\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})&=\dim\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_0(\tfrac12),\tau\big)\leq\dim\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}(\tfrac12),\tau\big)\\ &\leq\dim\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{3,3}(\mathbf{1}),\tau)+\dim\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{4,0}(\mathbf{1}),\tau)\\ &=\dim\mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}(\Theta_6^+(\tau),\mathbb{C})+\dim\mathrm{Hom}_{\mathrm{GO}_{4,0}(F)}(\Theta_4^+(\tau),\mathbb{C}).\end{aligned}\tag{4-7}$$

Now we separate them into two cases: $\pi_1 \not\cong \pi_2$ and $\pi_1 \cong \pi_2$.

(A) If $\pi_1 \not\cong \pi_2$, then the theta lift $\Theta_1(\Sigma)$ to $\mathrm{GSp}_2(E)$ of Σ is zero,

$$\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(R_1(\mathbf{1}) \otimes \nu, \Sigma) = \mathrm{Hom}_{\mathrm{GSp}_2(F)}(\Theta_1(\Sigma), \mathbb{C}) = 0$$

and $\Sigma = \mathrm{Ind}_{\mathrm{GSO}(2,2)(E)}^{\mathrm{GO}(2,2)(E)}(\pi_1 \boxtimes \pi_2)$. There are several subcases:

(A1) If $\pi_i (i = 1, 2)$ are both $D^\times(F)$-distinguished, which implies that ϕ_{π_i} are conjugate-orthogonal and so that π_i are both $\mathrm{GL}_2(F)$-distinguished due to Lemma 4.4.5, then $\pi_1^\vee \not\cong \pi_2^\sigma$. Otherwise, $\pi_1^\sigma \cong \pi_1^\vee \cong \pi_2^\sigma$, which contradicts the assumption $\pi_1 \not\cong \pi_2$. Then the inequality (4-6) can be rewritten as

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) \geq \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}\big(\tfrac{1}{2}\big), \Sigma\big) \geq 2. \tag{4-8}$$

Flicker [1991] proved that $(\mathrm{GL}_n(E), \mathrm{GL}_n(F))$ is a Gelfand pair, which implies that

$$1 \geq \mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C}).$$

Thus

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) \leq 1 + 1 \tag{4-9}$$

due to the upper bound (4-7). Then (4-8) and (4-9) imply

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 2.$$

(A2) If $\pi_1 = \pi(\chi_1, \chi_2)$, $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \mathbf{1}$ and π_2 is $\mathrm{GL}_2(F)$-distinguished, then Lemma 4.4.5 implies that both ϕ_{π_1} and ϕ_{π_2} are conjugate-orthogonal, $\pi_1^\vee \not\cong \pi_2^\sigma$ and

$$\mathrm{Hom}_{\mathrm{GO}_{4,0}(F)}(\Sigma, \mathbb{C}) = 0 = \mathrm{Hom}_{\mathrm{GO}_{3,1}(F)}(\Sigma, \mathbb{C}).$$

Moreover, $\mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C}) \neq 0$. Since

$$\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}\big(\tfrac{1}{2}\big), \Sigma\big) \geq \dim \mathrm{Hom}_{\mathrm{GO}(2,2)(F)}(\Sigma, \mathbb{C}) + 0 = 1,$$

the desired equality $\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 1$ follows from (4-6) and (4-7).

(A3) If $\pi_1^\sigma \cong \pi_2^\vee$, then Lemma 4.4.1 implies

$$\dim \mathrm{Hom}_{\mathrm{GO}_{3,1}(F)}(\Sigma, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSO}_{3,1}(F)}(\pi_1 \boxtimes \pi_2, \mathbb{C}) = 1.$$

By the previous arguments, we know that $\mathrm{Hom}_{\mathrm{GO}_{2,2}(F)}(\Sigma, \mathbb{C}) = 0$ in this case. Therefore

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 1.$$

In other cases, if $\pi_1^\sigma \not\cong \pi_2^\vee$ and either ϕ_{π_1} or ϕ_{π_2} is not conjugate-orthogonal, then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 0.$$

If not, then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C}) = 1.$$

Set $\Pi\boxtimes\chi=\Theta_6^+(\tau)|_{\mathrm{GSO}_{3,3}(E)}$ as a representation of $\mathrm{GSO}_{3,3}(E)$, which is irreducible due to Proposition 3.7. Then Π is $\mathrm{GL}_4(F)$-distinguished and so ϕ_Π is conjugate-orthogonal.

We consider the following cases:

- If ϕ_{π_1} is conjugate-orthogonal, then ϕ_{π_2} is conjugate-orthogonal by Theorem 4.4.4.
- If ϕ_{π_1} is irreducible, by the assumption $\pi_1^\sigma\ncong\pi_2^\vee$ and Theorem 4.4.4, then ϕ_{π_1} is conjugate-orthogonal, which will imply that ϕ_{π_2} is conjugate-orthogonal as well.
- Now suppose that both ϕ_{π_1} and ϕ_{π_2} are reducible and that neither ϕ_{π_1} nor ϕ_{π_2} is conjugate-orthogonal. Assume that $\phi_{\pi_i}=\chi_{i1}+\chi_{i2}$ $(i=1,2)$. Then
$$\phi_\Pi=\chi_{11}+\chi_{12}+\chi_{21}+\chi_{22},\qquad \chi_{11}\chi_{12}=\chi_{21}\chi_{22}:E^\times/F^\times\to\mathbb{C}^\times.$$
Thanks to Theorem 4.4.4, $\chi_{11}\chi_{21}^\sigma=\mathbf{1}$ and $\chi_{12}\neq\chi_{22}$ but $\chi_{12}|_{F^\times}=\mathbf{1}=\chi_{22}|_{F^\times}$. Furthermore, $\chi_{21}\chi_{22}\cdot(\chi_{21}\chi_{22})^\sigma=\mathbf{1}$ implies
$$\chi_{21}^\sigma\chi_{21}=\mathbf{1}.$$
Similarly $\chi_{11}^\sigma\chi_{11}=\mathbf{1}$. Thus, $\chi_{21}^\sigma=\chi_{21}^{-1}$ and $\chi_{11}=\chi_{21}$. This implies that $\chi_{12}=\chi_{22}$ which contradicts the condition $\chi_{12}\neq\chi_{22}$.

Hence the Langlands parameter ϕ_Π can not be conjugate-orthogonal. Thus $\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})=0$ if $\pi_1^\sigma\ncong\pi_2^\vee$ and either ϕ_{π_1} or ϕ_{π_2} is not conjugate-orthogonal.

(B) If $\pi_1=\pi_2$ is a discrete series representation, then $\Theta_1(\Sigma)=\pi_1$ due to [Atobe and Gan 2017, Proposition 5.4]. If $\pi_1=\pi_2$ is an irreducible principal series representation, applying the functor $\mathrm{Hom}_{\mathrm{GO}_4(E)}(-,\Sigma)$ on the Kudla filtration (see [Gan and Takeda 2011b, Theorem A1]), we have
$$\Theta_1(\Sigma)=\pi_1$$
except for $\pi_1=\pi(\chi,\chi)$. If $\pi_1=\pi(\chi,\chi)$, then there is an exact sequence
$$\pi_1\longrightarrow\Theta_1(\pi_1\boxtimes\pi_1)\longrightarrow\pi_1\longrightarrow 0$$
of $\mathrm{GL}_2(E)$-representations, where we can not deduce $\Theta_1(\pi_1\boxtimes\pi_1)$ directly. There are two choices that $\Theta_1(\pi_1\boxtimes\pi_1)$ is either π_1 or $\mathrm{Ext}_{\mathrm{GL}_2(E)}(\pi_1,\pi_1)$. We will show that $\Theta_1(\pi_1\boxtimes\pi_1)$ has a unique Whittaker model which can imply that $\Theta_1(\pi_1\boxtimes\pi_1)=\pi_1$. Let $N=\left\{\left(\begin{smallmatrix}1&n\\0&1\end{smallmatrix}\right)\mid n\in E\right\}$ be the subgroup of $\mathrm{GSp}_2(E)$. Let ψ_N be a nontrivial character of N. Consider the Whittaker model of $\Theta_1(\pi_1\boxtimes\pi_1)$,
$$\dim\mathrm{Hom}_N(\Theta_1(\pi_1\boxtimes\pi_1),\psi_N)=\dim\mathrm{Hom}_{\mathrm{PGL}_2(E)}(\pi_1\boxtimes\pi_1,\mathbb{C})\leq 1$$
due to [Lu 2017b, Proposition 3.4], which implies that $\Theta_1(\Sigma)=\pi_1$. Therefore the exact sequence (4-3) implies the inequality
$$\dim\mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\mathbb{C})\geq\dim\mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}\big(\tfrac12\big),\Sigma\big)-\dim\mathrm{Hom}_{\mathrm{GSp}_2(F)}(\pi_1,\mathbb{C}).\tag{4-10}$$

We separate them into the following cases:

(B1) If π_1 is $D^\times(F)$-distinguished, then $\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_0\big(\frac{1}{2}\big), \Sigma\big) = 3$. Again, we consider the upper bound (4-7) and the lower bound (4-10) to obtain the equality

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(\mathbb{C})}(\tau, \mathbb{C}) = 2.$$

(B2) If $\pi_1 \cong \pi_0 = \pi(\chi_1, \chi_2)$ with $\chi_1 \neq \chi_2$ and $\chi_1|_{F^\times} = \chi_2|_{F^\times} = 1$, then

$$\dim \mathrm{Hom}_{\mathrm{GO}_{4,0}(F)}(\Sigma, \mathbb{C}) = 0.$$

In a similar way, we can get $\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 1$.

(B3) If π_1 is not $\mathrm{GL}_2(F)$-distinguished but $(\mathrm{GL}_2(F), \omega_{E/F})$-distinguished, then

$$\mathrm{Hom}_{\mathrm{GSp}_2(F)}(\pi_1, \mathbb{C}) = 0 \text{ and } \mathrm{Hom}_{\mathrm{GO}_{3,1}(F)}(\Sigma, \mathbb{C}) \neq 0,$$

which implies that $\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I_{Q_4}^{H_4}\big(\frac{1}{2}\big), \Sigma\big) \geq 1 = \dim \mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C})$. Thus we can deduce that $\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) = 1$.

(iii) If τ is not in case (i) or (ii), then the first occurrence index of τ of $\mathrm{GSp}_4(E)$ in the Witt tower $\mathbb{H}_E^r$ is 3. Observe that $\Theta_6^+(\tau)|_{\mathrm{GSO}_{3,3}(E)}$ is irreducible unless $\tau = \mathrm{Ind}_{Q(Z)}^{\mathrm{GSp}_4(E)}(\chi, \pi)$ with $\chi = |-|_E$.

Suppose that $\tau \neq \mathrm{Ind}_{Q(Z)}^{\mathrm{GSp}_4(E)}(|-|_E, \pi)$. Consider the double see-saw diagrams

$$\begin{tikzcd}
\mathrm{GO}_{2,2}(E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GSp}_8(F) \arrow[d, dash] \arrow[dl, dash] \arrow[dr, dash] & \mathrm{GO}_{3,3}(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\
\mathrm{GO}_{4,0}(F) & \mathrm{GSp}_4(E)^\natural & \mathrm{GO}_{3,3}(F)
\end{tikzcd}$$

By [Kudla and Rallis 1992, p. 211] and Proposition 4.3.1, there are two exact sequences

$$0 \longrightarrow R^{3,3}(\mathbf{1}) \longrightarrow \mathcal{I}\big(\tfrac{1}{2}\big) \longrightarrow R^{4,0}(\mathbf{1}) \longrightarrow 0$$

and

$$0 \longrightarrow R^{4,0}(\mathbf{1}) \oplus R^{2,2}(\mathbf{1}) \longrightarrow \mathcal{I}\big(-\tfrac{1}{2}\big) \longrightarrow R^{5,1}(\mathbf{1}) \cap R^{3,3}(\mathbf{1}) \longrightarrow 0$$

of $\mathrm{GSp}_8(F)$-modules, where $\mathcal{I}(s)$ is the degenerate principal series of $\mathrm{GSp}_8(F)$ and $R^{m,n}(\mathbf{1})$ is the big theta lift to $\mathrm{GSp}_8(F)$ of the trivial representation $\mathbf{1}$ of $\mathrm{GO}_{m,n}(F)$. Assume that τ is generic and its theta lift to $\mathrm{GO}_{2,2}(E)$ is zero. Then

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{4,0}(\mathbf{1}), \tau) = \mathrm{Hom}_{\mathrm{GO}_{4,0}(F)}(\Theta_4^+(\tau), \mathbb{C}) = 0,$$

so that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}\big(-\tfrac{1}{2}\big), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{5,1}(\mathbf{1}) \cap R^{3,3}(\mathbf{1}), \tau).$$

Thus applying Lemma 4.4.2,

$$\begin{aligned}\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) &= \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_0\big(\tfrac{1}{2}\big), \tau\big) \\ &\le \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}\big(\tfrac{1}{2}\big), \tau\big) \\ &\le \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{3,3}(\mathbf{1}), \tau) \\ &= \dim \mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C}) \\ &= \dim \mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}((\Pi \boxtimes \chi)^+, \mathbb{C}) \end{aligned} \tag{4-11}$$

where $(\Pi \boxtimes \chi)^\pm$ are two extensions to $\mathrm{GO}_{3,3}(E)$ of $\Pi \boxtimes \chi$. On the other hand, one has

$$\mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}((\Pi \boxtimes \chi)^-, \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}_{3,3}(F)}(\Theta_6^+(\tau) \otimes \nu, \mathbb{C}) \cong \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\Theta(\nu), \tau) = 0.$$

Then we have an inequality

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) \le \dim \mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Pi \boxtimes \chi, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GL}_4(F)}(\Pi, \mathbb{C}). \tag{4-12}$$

Now we want to obtain the reverse inequality. Note that

$$1 \longrightarrow R^{5,1}(\mathbf{1}) \cap R^{3,3}(\mathbf{1}) \longrightarrow R^{3,3}(\mathbf{1}) \longrightarrow R^{2,2}(\mathbf{1}) \longrightarrow 1$$

is exact (see [Gan and Ichino 2014, Proposition 7.2]). There is an injection

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{3,3}(\mathbf{1}), \tau) \hookrightarrow \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(R^{5,1}(\mathbf{1}) \cap R^{3,3}(\mathbf{1}), \tau) = \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}\big(-\tfrac{1}{2}\big), \tau\big) \tag{4-13}$$

since the theta lifts to $\mathrm{GO}_{2,2}(E)$ and $\mathrm{GO}_{4,0}(E)$ of τ are both zero by the assumption.

We will show that τ does not occur on the boundary of $\mathcal{I}\big(-\frac{1}{2}\big)$ under the assumptions. If τ is nondiscrete, then $\tau = J_{Q(Z)}(\chi, \pi)$, $\chi \ne \mathbf{1}$, due to [Gan and Takeda 2011b, Table 1]. Note that

$$\mathcal{I}_1(s)/\mathcal{I}_0(s) = \mathrm{ind}_{(E^\times \times \mathrm{GSp}_2(F))N'}^{\mathrm{GSp}_4(E)^\natural} \chi',$$

where $N' \cong E \oplus \mathrm{Mat}_{2,2}(F)$ and $\chi'(t, g) = |N_{E/F}(t)^{s+\frac{1}{2}} \cdot \lambda(g)^{-2s-3}|_F$. Set

$$P' = (\mathrm{GL}_1(E) \times \mathrm{GSp}_2(E)^\natural) \cdot N'.$$

Thanks to the second adjoint theorem due to Bernstein, we have

$$\mathrm{Hom}\big(\mathcal{I}_1\big(-\tfrac{1}{2}\big)/\mathcal{I}_0\big(-\tfrac{1}{2}\big), \tau\big) = \mathrm{Hom}_{E^\times \times \mathrm{Sp}_2(E) \times F^\times}\big(\mathbf{1} \otimes \mathrm{ind}_{\mathrm{Sp}_2(F)}^{\mathrm{Sp}_2(E)} \mathbb{C} \otimes |-|_F^{-2}, R_{\bar{P}'}(J_{Q(Z)}(\chi, \pi))\big) = 0,$$

because $R_{\bar{P}'}(J(\chi, \pi)) = \chi \otimes \pi + \chi^{-1} \otimes \pi\chi$ and $\chi \ne \mathbf{1}$. Moreover, the cuspidal supports of $J_{Q(Z)}(\chi, \pi)$ and $\mathcal{I}_2\big(-\frac{1}{2}\big)/\mathcal{I}_1\big(-\frac{1}{2}\big)$ are disjoint. Therefore $\tau = J_{Q(Z)}(\chi, \pi)$ does not occur on the boundary of $\mathcal{I}\big(-\frac{1}{2}\big)$ and so

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}\big(-\tfrac{1}{2}\big), \tau\big) \le \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_0\big(-\tfrac{1}{2}\big), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}).$$

Note that if τ is a discrete series representation, then we have

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_{i+1}\big(-\tfrac{1}{2}\big)/\mathcal{I}_i\big(-\tfrac{1}{2}\big), \tau\big) = 0$$

for $i = 0, 1$. If not, then we will get a contradiction. Suppose that

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_1\big(-\tfrac{1}{2}\big)/\mathcal{I}_0\big(-\tfrac{1}{2}\big), \tau\big) \neq 0.$$

Then $\mathrm{Hom}_{\mathrm{GL}_1(E)}(\mathbf{1}, R_{\bar{P}'}(\tau)) \neq 0$, which contradicts Casselman's criterion [Casselman and Miličić 1982] for the discrete series representation that

$$\mathrm{Hom}_{\mathrm{GL}_1(E)}(|-|_E^s, R_{\bar{P}'}(\tau)) \neq 0$$

implies $s < 0$. Similarly,

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_2\big(-\tfrac{1}{2}\big)/\mathcal{I}_1\big(-\tfrac{1}{2}\big), \tau\big) = \mathrm{Hom}_{\mathrm{GL}_2(E)\times F^\times}(\delta_{P^\natural}^{1/6}, R_{\bar{P}^\natural}(\tau)) = 0$$

and so

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}\big(-\tfrac{1}{2}\big), \tau\big) \leq \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_0\big(-\tfrac{1}{2}\big), \tau\big). \tag{4-14}$$

Therefore one can combine (4-12)–(4-14) to obtain that

$$\begin{aligned}\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \mathbb{C}) &= \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathcal{I}_0\big(-\tfrac{1}{2}\big), \tau\big)\\ &= \dim \mathrm{Hom}_{\mathrm{GSO}_{3,3}(F)}(\Theta_6^+(\tau), \mathbb{C})\\ &= \dim \mathrm{Hom}_{\mathrm{GL}_4(F)}(\Pi, \mathbb{C}).\end{aligned} \tag{4-15}$$

Thus the left-hand side is 1 if and only if Π is $\mathrm{GL}_4(F)$-distinguished.

If $\tau = \mathrm{Ind}_{Q(Z)}^{\mathrm{GSp}_4(E)}(|-|_E, \pi)$ is irreducible, then $\theta_6(\tau) = J_P(\pi|-|_E, \pi) \boxtimes \omega_\pi|-|_E$. It suffices to show that $I_P(\pi|-|_E, \pi)$ is $\mathrm{GL}_4(F)$-distinguished if and only if ϕ_Π is conjugate-self-dual. This follows from Lemma 4.4.6.

Hence we have finished the proof. □

Remark 4.4.10. We can also show that if $\tau = \theta(\pi_1 \boxtimes \pi_2)$ with $\pi_1^\vee \cong \pi_2^\sigma$ is generic, then $\phi_\Pi = \phi_{\pi_1} \oplus \phi_{\pi_2}$ is not only conjugate-orthogonal but also conjugate-symplectic. Keeping this fact in mind will be helpful when we verify the Prasad conjecture for GSp_4 in Section 6C.

Corollary 4.4.11. *The pair* $(\mathrm{GSp}_4(E)^\natural, \mathrm{GSp}_4(F))$ *is not a Gelfand pair.*

For a generic representation τ of $\mathrm{GSp}_4(E)$ with $\omega_\tau|_{F^\times} = \chi_F^2$, we may consider the multiplicity

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\pi, \chi_F)$$

which is equal to $\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\pi \otimes \chi_E^{-1}, \mathbb{C})$, where χ_E is a character of $E^\times$ and $\chi_F = \chi_E|_{F^\times}$. We will focus on the case $\chi_F = \omega_{E/F}$ when we verify the Prasad conjecture for GSp_4 in Section 6C.

5. The $\mathrm{GSp}_{1,1}(F)$-distinguished representations

5A. *Notation.*

- D (resp. D_E) is a quaternion division algebra over F (resp. E) with a standard involution $*$.
- π^{D_E} is the Jacquet–Langlands lift to $D_E^\times(E)$ of π and $\pi^{D_E}\boxtimes\pi^{D_E}$ is a representation of $\mathrm{GSO}_{4,0}(E)$.
- $\mathfrak{W}$ (resp. $\mathfrak{V}$) is a right skew-Hermitian (resp. left Hermitian) D-vector space with isometry group $U(\mathfrak{W})$ (resp. $U(\mathfrak{V})$).
- $\mathfrak{U}^*$ is the dual D-vector space of $\mathfrak{U}$ in $\mathrm{Res}_{R/D}V_R$.
- $\mathfrak{W}\otimes_D\mathfrak{V}$ is a symplectic F-vector space.
- $\mathrm{GO}^*_{3,0}=\mathrm{GL}_1(D_4)\times\mathbb{G}_m/\{(t^{-1},t^2)\}$ (resp. $\mathrm{GO}^*_{r,r}$) is the inner form of $\mathrm{GO}_{3,3}$ (resp. $\mathrm{GO}_{2r,2r}$) defined over F and D_4 is the division F-algebra of degree 4.
- $\mathfrak{I}(s)$ (resp. $I(s)$) is the degenerate principal series of $\mathrm{GSp}_{2,2}(F)$ (resp. $\mathrm{GO}^*_{2,2}(F)$).
- $\mathrm{GSO}^*_{2,0}$ is the inner form of $\mathrm{GSO}_{3,1}$ defined over F.
- $\mathrm{GO}_{5,1}=\mathrm{GL}_2(D_E)\times\mathbb{G}_m/\{(t^{-1},t^2)\}$ is the pure inner form of $\mathrm{GO}_{3,3}$ defined over E and $\Pi^D\boxtimes\chi$ is a representation of $\mathrm{GSO}_{5,1}(E)$.
- B_1 is the minimal parabolic subgroup of $\mathrm{GL}_2(D_E)(E)$.
- $\mathrm{GSp}_{1,0}=D^\times$ (resp. $\mathrm{Sp}_{1,0}$) is the inner form of GL_2 (resp. SL_2).
- $P(Y_D)$ (resp. $\mathfrak{Q}$) is the Siegel parabolic subgroup of $\mathrm{GU}(\mathfrak{V})$ (resp. $\mathrm{GO}^*_{2,2}(F)$).
- $\mathfrak{R}^3(\mathbf{1})$ (resp. $\mathfrak{R}^2(\mathbf{1})$) is the big theta lift to $\mathrm{GSp}_{2,2}(F)$ of the trivial representation of $\mathrm{GO}^*_{3,0}(F)$ (resp. $\mathrm{GO}^*_{1,1}(F)$) and $\mathfrak{R}^{1,j}(\mathbf{1})$ is the big theta lift to $\mathrm{GO}^*_{2,2}(F)$ from $\mathrm{GSp}_{1,j}(F)$.
- $\theta_2^-(\tau)$ (resp. $\Theta_2^-(\tau)$) is the small (resp. big) theta lift to $\mathrm{GO}_{5,1}(E)$ of τ of $\mathrm{GSp}_4(E)$.
- $\Theta_{\mathfrak{W},\mathfrak{V},\psi}(\pi)$ is the big theta lift to $\mathrm{GU}(\mathfrak{V})$ of π of $\mathrm{GU}(\mathfrak{W})$.
- γ_F is the Weil index and $\gamma_F(\psi\circ q)\in\mu_8$ for the character of second degree $x\mapsto\psi(q(x,x))$, where q is a nondegenerate symmetric F-bilinear form.

5B. ***Theta lifts for quaternionic unitary groups.*** In order to study the $\mathrm{GSp}_{1,1}$-distinction problems, we need to introduce the local theta lift for quaternionic unitary groups, following [Gan and Tantono 2014; Gurevich and Szpruch 2015; Yamana 2011].

5B1. *Morita equivalence.* Let $R=\mathrm{Mat}_{2,2}(E)$ be the split quaternion algebra over E. Any left Hermitian (resp. right skew-Hermitian) free R-module (W_R,h_R) corresponds to a symplectic (resp. orthogonal) space (W_E,h_E) over E and

$$\dim_E W_E=2\cdot\dim_R W_R,\ \mathrm{Aut}(W_R,h_R)=\mathrm{Aut}(W_E,h_E).$$

See [Gurevich and Szpruch 2015, §2.1] for more details.

5B2. *Dual pairs.* Let D be the unique nonsplit quaternion algebra over F, with a standard involution $*$. Then $D \otimes_F E \cong R$. There is a D-linear map

$$\mathrm{tr}_{R/D} : R \to D$$

such that $\mathrm{tr}_{R/D}(d) = 2d$ for $d \in D$. Given a 4-dimensional symplectic space $(\mathcal{W}_2, h_E)$ over E, corresponding to a 2-dimensional left Hermitian space (W_R, h_R), we set

$$h_D(x, y) = \tfrac{1}{2}\,\mathrm{tr}_{R/D}(h_R(x, y)) \in D$$

for all $x, y \in W_R$. Then h_D is a nondegenerate Hermitian form on $\mathfrak{V} = \mathrm{Res}_{R/D} W_R$ and $\dim_D \mathfrak{V} = 4$.

Given a left Hermitian space $(\mathfrak{V}, h_D)$ and a right skew-Hermitian space $(\mathfrak{W}, s_D)$, the tensor product space $\mathfrak{W} \otimes_D \mathfrak{V}$ admits a symplectic form defined by

$$\langle w \otimes v, w' \otimes v' \rangle := \tfrac{1}{2}\,\mathrm{tr}_{D/F}((w, w') \cdot (v, v')^*).$$

This gives an embedding of F-groups

$$U(\mathfrak{W}) \times U(\mathfrak{V}) \to \mathrm{Sp}(\mathfrak{W} \otimes_D \mathfrak{V}).$$

Then we can define the Weil representation ω_ψ on $U(\mathfrak{W}) \times U(\mathfrak{V})$, using the complete polarization $\mathfrak{V} = Y_D + Y_D^*$ of $\mathfrak{V}$.

Theorem 5.2.1 [Gan and Sun 2017, Theorem 1.2]. *The Howe duality conjecture holds for the dual pair* $U(\mathfrak{W}) \times U(\mathfrak{V})$.

We can extend it to the similitude group $\mathrm{GU}(\mathfrak{W}) \times \mathrm{GU}(\mathfrak{V})$ following Roberts. (See [Gan and Tantono 2014, §3].)

5B3. *The see-saw diagram.* Let us fix the polarization $W_R = Y_R + Y_R^*$. Then

$$\mathfrak{V} = \mathrm{Res}_{R/D} W_R = Y_D + Y_D^*.$$

Consider the following see-saw diagram:

$$\begin{tikzcd}
\mathrm{GU}(\mathfrak{V}) \arrow[d, dash] \arrow[dr, dash] & \mathrm{GO}_{2,2}(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\
\mathrm{GU}(W_R)^\natural & \mathrm{GO}_{1,1}^*(F)
\end{tikzcd}$$

Here $\mathrm{GU}(W_R)^\natural = \mathrm{GSp}_4(E)^\natural$.

Proposition 5.2.2 [Gurevich and Szpruch 2015, Theorem 8.2]. *Let τ be an irreducible representation of* $\mathrm{GSp}(\mathcal{W}_2) \cong \mathrm{GU}(W_R)$. *Assume that π is an irreducible representation of* $\mathrm{GO}_{1,1}^*(F)$. *Then*

$$\mathrm{Hom}_{\mathrm{GU}(W_R)^\natural}(\Theta_{\mathfrak{W},\mathfrak{V},\psi}(\pi), \tau) = \mathrm{Hom}_{\mathrm{GO}_{1,1}^*(F)}(\Theta_4^+(\tau), \pi).$$

Assume that V_R is a skew-Hermitian free module over R of rank 2, associated to the anisotropic 4-dimensional quadratic space over E given by (D_E, N_{D_E}) such that

$$\mathrm{GU}(V_R) \cong \mathrm{GO}_{4,0}(E).$$

Then $\mathrm{Res}_{R/D} V_R$ is a 4-dimensional skew-Hermitian D-vector space with trivial discriminant. There is a natural embedding

$$\mathrm{SU}(V_R) \cong \mathrm{SO}_{4,0}(E) \hookrightarrow \mathrm{SO}^*_{2,2}(F) = \mathrm{SU}(\mathrm{Res}_{R/D} V_R).$$

Given a 1-dimensional Hermitian vector space $\mathfrak{V}_1$ over D, we consider the theta lift from $\mathrm{GU}(\mathfrak{V}_1) = \mathrm{GSp}_{1,0}(F)$ to $\mathrm{GO}^*_{2,2}(F)$ and the theta lift from $\mathrm{GSO}_{4,0}(E)$ to $\mathrm{GU}(R \otimes_D \mathfrak{V}_1) = \mathrm{GL}_2(E)$. Consider the see-saw diagram

$$\begin{tikzcd}
\mathrm{GU}(\mathrm{Res}_{R/D} V) \arrow[d, dash] \arrow[dr, dash] & \mathrm{GL}_2(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\
\mathrm{GSO}_{4,0}(E)^\natural & \mathrm{GSp}_{1,0}(F)
\end{tikzcd}$$

which is different from the situation in [Gurevich and Szpruch 2015, Theorem 8.2], since there does not exist a natural polarization in the symplectic F-vector space $\mathbb{V} = (\mathrm{Res}_{R/D} V_R) \otimes_D \mathfrak{V}_1$.

Assume that $\mathbb{V} = \mathbb{X} \oplus \mathbb{Y}$ is a polarization. Set the group

$$\mathrm{Mp}(\mathbb{V})_{\mathbb{Y}} = \mathrm{Sp}(\mathbb{V}) \times \mathbb{C}^\times$$

with group law

$$(g_1, z_1)(g_2, z_2) = (g_1 g_2, z_1 z_2 \cdot z_{\mathbb{Y}}(g_1, g_2)),$$

where $z_{\mathbb{Y}}(g_1, g_2) = \gamma_F\big(\tfrac{1}{2}\psi \circ q(\mathbb{Y}, g_2^{-1}\mathbb{Y}, g_1\mathbb{Y})\big)$ is a 2-cocycle (called Rao cocyle) associated to $\mathbb{Y}$ and $q(\mathbb{Y}, g_2^{-1}\mathbb{Y}, g_1\mathbb{Y})$ is the Leray invariant. (See [Kudla 1996, §I.3].)

Suppose that $\mathbb{V} = \mathbb{X}' \oplus \mathbb{Y}'$ is another polarization of $\mathbb{V}$. There is an isomorphism

$$\mathcal{S}(\mathbb{X}) \cong \mathcal{S}(\mathbb{X}').$$

Given $\varphi \in \mathcal{S}(\mathbb{X})$ and $\varphi' \in \mathcal{S}(\mathbb{X}')$, due to [Ichino and Prasanna 2016, Lemma 3.3], we have

$$\varphi(x) = \int_{\mathbb{Y} \cap \mathbb{Y}' \backslash \mathbb{Y}} \psi\big(\tfrac{1}{2}\langle x', y'\rangle - \tfrac{1}{2}\langle x, y\rangle\big)\varphi'(x')\, dy$$

where $x' \in \mathbb{X}'$ and $y' \in \mathbb{Y}'$ are given by $x' + y' = x + y \in \mathbb{V}$.

Lemma 5.2.3 (local Siegel–Weil identity). *Assume that π is an irreducible discrete series representation of $\mathrm{GL}_2(E)$ so that the big theta lift $\Theta(\pi)$ to $\mathrm{GSO}_{4,0}(E)$ is isomorphic to $\pi^{D_E} \boxtimes \pi^{D_E}$, where π^{D_E} is the Jacquet–Langlands lift to $D_E^\times(E)$ of π. Let ϱ be an irreducible representation of $\mathrm{GSp}_{1,0}(F)$. Then*

$$\dim \mathrm{Hom}_{\mathrm{GSO}_{4,0}(E)^\natural}(\Theta(\varrho), \pi^{D_E} \boxtimes \pi^{D_E}) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,0}(F)}(\pi, \varrho),$$

where $\Theta(\varrho)$ is the big theta lift to $\mathrm{GO}^*_{2,2}(F)$ of ϱ.

Proof. It suffices to show that two splittings of $\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_{1,0}(F)$ in $\mathrm{Mp}(\mathbb{V})$ are compatible. Let us fix two polarizations $\mathrm{Res}_{R/D}V_R=\mathfrak{U}\oplus\mathfrak{U}^*$ and $R\otimes_D\mathfrak{V}_1=X\oplus Y$. Then

$$\mathbb{V}=\mathbb{X}\oplus\mathbb{Y}=(\mathfrak{U}\otimes_D\mathfrak{V}_1)\oplus(\mathfrak{U}^*\otimes_D\mathfrak{V}_1)\quad\text{and}\quad\mathbb{V}=\mathbb{X}'\oplus\mathbb{Y}'=(D_E\otimes_E X)\oplus(D_E\otimes_E Y).$$

Choose a fixed element $h_0\in\mathrm{Sp}(\mathbb{V})$ such that

$$\mathbb{X}'=h_0\mathbb{X}\quad\text{and}\quad\mathbb{Y}'=h_0\mathbb{Y}.$$

By [Ichino and Prasanna 2016, Appendix B.4], there is an isomorphism $\alpha_0:\mathrm{Mp}(\mathbb{V})_{\mathbb{Y}'}\to\mathrm{Mp}(\mathbb{V})_{\mathbb{Y}}$ via

$$(h,z)\mapsto(\alpha_0(h),z),$$

where $\alpha_0(h)=h^{-1}\cdot g\cdot h$ for all $h\in\mathrm{Sp}(\mathbb{V})$. Moreover,

$$z_{\mathbb{Y}'}(h_1,h_2)=z_{\mathbb{Y}}(\alpha_0(h_1),\alpha_0(h_2)).$$

Now we fix the splitting $i_{\mathbb{Y}}:\mathrm{O}^*_{2,2}(F)\times\mathrm{Sp}_{1,0}(F)\hookrightarrow\mathrm{Mp}(\mathbb{V})_{\mathbb{Y}}$ and

$$i_{\mathbb{Y}'}:\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_2(E)\hookrightarrow\mathrm{Mp}(\mathbb{V})_{\mathbb{Y}'},$$

where the splitting $i_{\mathbb{Y}}(y,z)=((y,z),\beta_{\mathbb{Y}}(z))$ is defined in [Kudla 1994, Theorem 3.1].

We will show that $i_{\mathbb{Y}}(h)=\alpha_0\circ i_{\mathbb{Y}'}(h)$ for all $h=(y,z)\in\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_{1,0}(F)$. Consider

$$\begin{tikzcd}
\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_{1,0}(F) \arrow[r, hook] \arrow[d, equal] & \mathrm{O}^*_{2,2}(F)\times\mathrm{Sp}_{1,0}(F) \arrow[r, "i_{\mathbb{Y}}"] & \mathrm{Mp}(\mathbb{V})_{\mathbb{Y}} \\
\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_{1,0}(F) \arrow[r, hook] & \mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_2(E) \arrow[r, "i_{\mathbb{Y}'}"] & \mathrm{Mp}(\mathbb{V})_{\mathbb{Y}'} \arrow[u, "\alpha_0"']
\end{tikzcd}$$

Set $i_{\mathbb{Y}}(h)=(h,\beta_{\mathbb{Y}}(h))$. Then $\beta_{\mathbb{Y}}(z)=1$ for all $z\in\mathrm{Sp}_{1,0}(F)$. Similarly, we have

$$\beta_{\mathbb{Y}'}(y)=1$$

for all $y\in\mathrm{SO}_{4,0}(E)$. In order to show that

$$\beta_{\mathbb{Y}}(h)=\beta_{\mathbb{Y}'}(h)$$

for all $h=(y,z)\in\mathrm{SO}_{4,0}(E)\times\mathrm{Sp}_{1,0}(F)$, we will show that $\beta_{\mathbb{Y}}(y)=1=\beta_{\mathbb{Y}'}(z)$.

- If $y\in\mathrm{SO}_{4,0}(E)\subset\mathrm{O}^*_{2,2}(F)=\bigsqcup_{i=0}^2\mathfrak{P}\omega_i\mathfrak{P}$, say $y\in\mathfrak{P}\omega_i\mathfrak{P}$, where $\mathfrak{P}$ is the Siegel parabolic subgroup of $\mathrm{O}^*_{2,2}(F)$, $\omega_0=\mathbf{1}_4$ (the identity matrix in $\mathrm{O}^*_{2,2}(F)$),

$$\omega_1=\begin{pmatrix} & & 1 & \\ & 1 & & \\ 1 & & & \\ & & & 1\end{pmatrix}\quad\text{and}\quad\omega_2=\begin{pmatrix} & & 1 & \\ & & & 1\\ 1 & & & \\ & 1 & & \end{pmatrix},$$

  then $\beta_{\mathbb{Y}}(y)=(-1)^i$. Since ω_1 switches a pair of vectors e_1 and e_1' in a basis $\{e_1,e_2,e_1',e_2'\}$, which corresponds to an element $h\in\mathrm{O}_{4,0}(E)$ with determinant -1, where $\mathfrak{P}$ stabilizes the maximal isotropic subspace $\{e_1,e_2\}$, it follows that

$$\mathrm{SO}_{4,0}(E)\cap\mathfrak{P}\omega_1\mathfrak{P}=\varnothing,$$

  i.e., $\beta_{\mathbb{Y}}(y)=1$.

- If $z \in \mathrm{Sp}_{1,0}(F)$ and so $z = g \in \mathrm{SL}_2(E)$, then $\beta_{\mathbb{Y}'}(z) = \gamma_F\big(x(g), \frac{1}{2}\psi\big)^4 \cdot \gamma_F\big(\frac{1}{2}\psi \circ N_{D_E}\big)^4 = 1$, where

$$x(g) = \begin{cases} N_{E/F}(a_{21}) \pmod{F^{\times 2}} & \text{if } g = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right) \text{ with } a_{21} \neq 0, \\ N_{E/F}(a_{22}) \pmod{F^{\times 2}} & \text{otherwise.} \end{cases}$$

Therefore we have finished the proof. □

Remark 5.2.4. From the proof above, we can see that the see-saw identity does not hold if one replaces $\mathrm{SO}_{4,0}(E)$ by $\mathrm{O}_{4,0}(E)$ in this case.

Let V be a free R-module of rank 2 corresponding to the quadratic space $\mathbb{H}_E^2$ by the Morita equivalence. Then $\mathrm{Res}_{R/D} V$ is a skew-Hermitian D-vector space of dimension 4.

Lemma 5.2.5. *Let Σ be an irreducible representation of $\mathrm{GO}_{2,2}(E)$. Let ϱ be an irreducible representation of $\mathrm{GSp}_{1,j}(F)$ for $j = 0$ or 1. Then*

$$\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(\Theta(\varrho), \Sigma) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,j}(F)}(\Theta_{1+j}(\Sigma \cdot \nu^{1+j}), \varrho),$$

where ν is the nontrivial character of $\mathrm{GO}_{2,2}(E)/\mathrm{GSO}_{2,2}(E)$ and $\nu|_{\mathrm{O}_{2,2}(E)} = \det$.

Proof. Consider the see-saw diagram

$$\begin{tikzcd} \mathrm{GO}^*_{2,2}(F) \arrow[d, dash] \arrow[drr, dash] & & \mathrm{GSp}_{2+2j}(E)^\natural \arrow[d, dash] \arrow[dll, dash] \\ \mathrm{GO}_{2,2}(E)^\natural & & \mathrm{GSp}_{1,j}(F) \end{tikzcd}$$

Assume that $\mathfrak{W} = \mathrm{Res}_{R/D} V$. Let us fix the polarization $\mathfrak{W} = \mathfrak{U} + \mathfrak{U}^*$ and $\mathbb{H}_E^2 = Y + Y^*$, where Y^* is the dual space of Y. Let $\mathfrak{V}$ be a Hermitian D-vector space with isometric group $\mathrm{GSp}_{1,j}(F)$. Then there exists a natural polarization

$$\mathfrak{W} \otimes_D \mathfrak{V} = \mathfrak{U} \otimes_D \mathfrak{V} + \mathfrak{U}^* \otimes_D \mathfrak{V}.$$

Similarly, $\mathbb{H}_E^2 \otimes_E \mathcal{W}_{1+j} = Y \otimes_E \mathcal{W}_{1+j} + Y^* \otimes_E \mathcal{W}_{1+j}$, where $\mathcal{W}_r$ is the symplectic vector space over E of dimension $2r$. Set $\mathbb{Y} = \mathfrak{U}^* \otimes_D \mathfrak{V}$ and $\mathbb{Y}' = Y^* \otimes_E \mathcal{W}_{1+j}$. Then we have the splitting $i_{\mathbb{Y}}$ and $i_{\mathbb{Y}'}$ defined in [Kudla 1994, Theorem 3.1]. For instance, $i_{\mathbb{Y}'}(y, z) = ((y, z), \beta_{\mathbb{Y}'}(y))$ for $(y, z) \in \mathrm{O}_{2,2}(E) \times \mathrm{Sp}_{2+2j}(E)$ and

$$i_{\mathbb{Y}}(y, z) = ((y, z), \beta_{\mathbb{Y}}(y)) \in \mathrm{Mp}(\mathfrak{W} \otimes_D \mathfrak{V})_{\mathbb{Y}}$$

for $y \in \mathrm{O}^*_{2,2}(F)$ and $z \in \mathrm{Sp}_{1,j}(F)$. Note that $\beta_{\mathbb{Y}'}(y) = 1$ for $y \in \mathrm{O}_{2,2}(E)$ and

$$\beta_{\mathbb{Y}}(y) = (-1)^{(1+j)i}$$

if $y \in \mathfrak{P}\omega_i\mathfrak{P}$, where $\mathrm{O}^*_{2,2}(F) = \bigcup_i \mathfrak{P}\omega_i\mathfrak{P}$ and $\mathfrak{P}$ is the Siegel parabolic subgroup of $\mathrm{O}^*_{2,2}(F)$. Thus

$$\beta_{\mathbb{Y}}(h) = \beta_{\mathbb{Y}'}(h) \cdot (\nu(h))^{1+j}$$

for $h \in \mathrm{O}_{2,2}(E)$. Hence

$$\begin{aligned}\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(\Theta(\varrho), \Sigma) &= \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural \times \mathrm{GSp}_{1,j}(F)}(\omega_{\psi,\mathbb{Y}}, \Sigma \otimes \varrho)\\ &= \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural \times \mathrm{GSp}_{1,j}(F)}(\omega_{\psi,\mathbb{Y}'}, \Sigma \cdot \nu^{1+j} \otimes \varrho)\\ &= \dim \mathrm{Hom}_{\mathrm{GSp}_{1,j}(F)}(\Theta_{1+j}(\Sigma \cdot \nu^{1+j}), \varrho),\end{aligned}$$

where $\omega_{\psi,\mathbb{Y}}$ (resp. $\omega_{\psi,\mathbb{Y}'}$) is the Weil representation on $\mathrm{Mp}(\mathfrak{W} \otimes_D \mathfrak{V})$ emphasizing the splitting $\mathbb{Y} + \mathbb{Y}^*$ (resp. $\mathbb{Y}' + \mathbb{Y}'^*$). This finishes the proof. □

5B4. *Degenerate principal series.* Let us fix the complete polarization

$$\mathfrak{V} = Y_D + Y_D^*.$$

Suppose $\dim_D \mathfrak{V} = 4$. Assume that $\mathfrak{I}(s)$ is the degenerate principal series of $\mathrm{GU}(\mathfrak{V}) = \mathrm{GSp}_{2,2}(F)$ associated to a Siegel parabolic subgroup $P(Y_D)$, i.e.,

$$\mathfrak{I}(s) = \left\{ f : \mathrm{GU}(\mathfrak{V}) \to \mathbb{C} \mid f(pg) = \delta_{P(Y_D)}(p)^{(1/2)+(s/5)} f(g) \text{ for all } p \in P(Y_D),\, g \in \mathrm{GU}(\mathfrak{V}) \right\},$$

where $\delta_{P(Y_D)}$ is the modular character. Similar to Proposition 4.3.1, we have

Lemma 5.2.6. *Assume that $\mathfrak{R}^3(\mathbf{1})$ is the big theta lift to $\mathrm{GU}(\mathfrak{V})$ of the trivial representation of $\mathrm{GO}^*_{3,0}(F)$. Then there is an exact sequence*

$$0 \longrightarrow \mathfrak{R}^3(\mathbf{1}) \longrightarrow \mathfrak{I}\left(\tfrac{1}{2}\right) \longrightarrow \mathfrak{R}^2(\mathbf{1}) \longrightarrow 0,$$

where $\mathfrak{R}^2(\mathbf{1})$ is the big theta lift to $\mathrm{GU}(\mathfrak{V})$ of the trivial representation of $\mathrm{GO}^*_{1,1}(F)$.

Proof. By [Yamana 2011, Theorem 1.4], we may give a similar proof as in Proposition 4.3.1. So we omit it here. □

5B5. *Double cosets.* Assume that $P(Y_D)$ is the Siegel parabolic subgroup of $\mathrm{GU}(\mathfrak{V}) = \mathrm{GSp}_{2,2}(F)$. Then the homogeneous space $X_D = P(Y_D)\backslash \mathrm{GSp}_{2,2}(F)$ corresponds to the set of maximal isotropic subspaces in $\mathfrak{V}$. We consider the double coset $X_D/\mathrm{GU}(W_R)^\natural = X_D/\mathrm{GSp}_4(E)^\natural$, similar to Lemma 4.3.3.

Proposition 5.2.7. *In the double cosets $X_D/\mathrm{GSp}_4(E)^\natural$, there are*

- *one closed orbit with stabilizer $P(Y_D) \cap \mathrm{GSp}_4(E)^\natural$,*
- *one open orbit with stabilizer $\mathrm{GU}_2(D)(F) = \mathrm{GSp}_{1,1}(F) \subset \mathrm{GSp}_4(E)^\natural$ and*
- *one intermediate orbit with a representative*

$$L = Dr(\sqrt{d}e + f) + D\left(e - \frac{1}{\sqrt{d}} f\right) \in X_D,$$

which is a nonfree R-module with stabilizer $(\mathrm{GL}_1(E) \times \mathrm{GSp}_{1,0}(F)) \cdot N$, $N \cong E \oplus D$, where $r = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) = r^2 \in R$ and $W_R = Re + Rf$ with $h_R(e, f) = 1$.

Lemma 5.2.8. *Let τ be an irreducible representation of* $\mathrm{GU}(W_R)^\natural = \mathrm{GSp}_4(E)^\natural$ *and* $\mathrm{GSp}_4(E)^\natural \hookrightarrow \mathrm{GSp}_{2,2}(F)$ *be a natural embedding. Then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}\big(\tfrac{1}{2}\big), \tau\big) \geq \dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau^\vee, \mathbb{C}).$$

Proof. Note that there are three orbits for $P(Y_D)\backslash \mathrm{GSp}_{2,2}(F)/\mathrm{GSp}_4(E)^\natural$. There is a filtration for $\mathfrak{I}\big(\frac{1}{2}\big)\big|_{\mathrm{GSp}_4(E)^\natural}$ as follows:

$$\mathrm{ind}_{\mathrm{GSp}_{1,1}(F)}^{\mathrm{GSp}_4(E)^\natural} \mathbb{C} = \mathfrak{I}_0\big(\tfrac{1}{2}\big) \subset \mathfrak{I}_1\big(\tfrac{1}{2}\big) \subset \mathfrak{I}_2\big(\tfrac{1}{2}\big) = \mathfrak{I}\big(\tfrac{1}{2}\big)\big|_{\mathrm{GSp}_4(E)^\natural},$$

where $\mathfrak{I}_2\big(\frac{1}{2}\big)/\mathfrak{I}_1\big(\frac{1}{2}\big) \cong \mathrm{ind}_{P^\natural}^{\mathrm{GSp}_4(E)^\natural} \delta_{P^\natural}^{1/2}$ and $\mathfrak{I}_1\big(\frac{1}{2}\big)/\mathfrak{I}_0\big(\frac{1}{2}\big) \cong \mathrm{ind}_{MN}^{\mathrm{GSp}_4(E)^\natural} \delta_{P(Y_D)}^{3/5}\delta_3^{-1/2}$, where

$$M \cong \mathrm{GL}_1(E) \times \mathrm{GSp}_{1,0}(F), \quad N \cong D \oplus E \quad \text{and} \quad \delta_3(t,x) = |N_{E/F}(t)^4 \cdot \lambda(x)^{-4}|_F$$

for $(t,d) \in M$. There exists an involution on $\mathrm{GSp}_{2,2}(F)$ such that the fixed points coincides with $\mathrm{GSp}_4(E)^\natural$. Applying [Ólafsson 1987, Theorem 2.5; Gourevitch et al. 2019, Proposition 4.9], we obtain the inequality

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}\big(\tfrac{1}{2}\big), \tau\big) \geq \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_0\big(\tfrac{1}{2}\big), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau^\vee, \mathbb{C}).$$

This finishes the proof. □

5C. *The distinction problem for* $\mathbf{GSp}_{1,1}$. Let $\mathrm{GU}_2(D) = \mathrm{GSp}_{1,1}$ be the inner form of GSp_4 defined over F, whose E-points coincide with $\mathrm{GSp}_4(E)$. Assume that $\tau \in \mathrm{Irr}(\mathrm{GSp}_4(E))$ with $\omega_\tau|_{F^\times} = \mathbf{1}$. In this subsection, we will study the multiplicity

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}).$$

Theorem 5.3.1. *Let τ be an irreducible representation of* $\mathrm{GSp}_4(E)$ *such that* Π_{ϕ_τ} *is generic.*

(i) *If $\tau = \theta(\pi_1 \boxtimes \pi_2)$ is a nongeneric tempered representation of* $\mathrm{GSp}_4(E)$, *where $\pi_1 \boxtimes \pi_2$ is an irreducible smooth representation of* $\mathrm{GSO}_{4,0}(E)$, *then* $\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = 1$ *if and only if one of the following holds*:

- $\pi_1 \ncong \pi_2$ *but* $\pi_1^\vee \cong \pi_2^\sigma$;
- $\pi_1 \cong \pi_2$ *are both* $(D^\times(F), \omega_{E/F})$*-distinguished.*

(ii) *If $\tau = \theta(\pi_1 \boxtimes \pi_2) = \theta(\pi_2 \boxtimes \pi_1)$ is generic, then*

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \begin{cases} 2 & \text{if } \pi_1 = \pi_2 = \pi(\chi^{-1}, \chi^\sigma), \\ 1 & \text{if } \pi_1 = \pi_2 \text{ are square-integrable and } D^\times(F)\text{-distinguished}, \\ 1 & \text{if } \pi_1 \text{ is } D^\times(F)\text{-distinguished and } \pi_2 = \pi_0, \\ 2 & \text{if } \pi_1 \neq \pi_2 \text{ are both } D^\times(F)\text{-distinguished}, \\ 0 & \text{otherwise}. \end{cases}$$

Here $\pi_0 = \pi(\chi_1, \chi_2)$ with $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \mathbf{1}$. Note that these conditions are mutually exclusive.

(iii) *Assume that* τ *is not as in case* (i) *or* (ii), *so that* $\tau = \theta(\Pi^D \boxtimes \chi)$ *is generic, where* $\Pi^D \boxtimes \chi$ *is an irreducible representation of* $\mathrm{GSO}_{5,1}(E)$. *Then* $\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = 1$ *if and only if one of the following holds*:

- ϕ_Π *is irreducible and conjugate-orthogonal or*
- $\phi_\Pi = \phi_\rho + \phi_\rho \mu$ *with* $\rho^\sigma \cong \rho^\vee \mu^{-1}$,

where $\Pi = JL(\Pi^D)$ *is the Jacquet–Langlands lift to* $\mathrm{GL}_4(E)$ *of* Π^D.

Proof. The proof is very similar to the proof of Theorem 4.4.9.

(i) Assume that V_R is a skew-Hermitian free module over R of rank 2, corresponding to D_E by the Morita equivalence. Then $\mathrm{Res}_{R/D} V_R$ is a 4-dimensional skew-Hermitian vector space over D with trivial discriminant. Fix a polarization $\mathrm{Res}_{R/D} V = \mathfrak{U} \oplus \mathfrak{U}^*$. Consider the diagram

$$\begin{tikzcd} \mathrm{GSp}_4(E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GO}^*_{2,2}(F) \arrow[d, dash] \arrow[dl, dash] \arrow[dr, dash] & \mathrm{GL}_2(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\ \mathrm{GSp}_{1,1}(F) & \mathrm{GO}_{4,0}(E)^\natural & \mathrm{GSp}_{1,0}(F) \end{tikzcd}$$

There is an exact sequence of $\mathrm{GO}^*_{2,2}(F)$-representations

$$0 \longrightarrow \mathfrak{R}^{1,1}(\mathbf{1}) \longrightarrow I\left(\tfrac{1}{2}\right) \longrightarrow \mathfrak{R}^{1,0}(\mathbf{1}) \longrightarrow 0,$$

where $I(s)$ is the degenerate principal series of $\mathrm{GO}^*_{2,2}(F)$ and $\mathfrak{R}^{1,j}(\mathbf{1})$ is the theta lift to $\mathrm{GO}^*_{2,2}(F)$ the trivial representation of $\mathrm{GSp}_{1,j}(F)$. Set $\tau = \Theta_2(\Sigma)$, where

$$\Sigma = \begin{cases} \mathrm{Ind}_{\mathrm{GSO}_{4,0}(E)}^{\mathrm{GO}_{4,0}(E)}(\pi_1 \boxtimes \pi_2) & \text{if } \pi_1 \ncong \pi_2, \\ (\pi_1 \boxtimes \pi_1)^+ & \text{if } \pi_1 \cong \pi_2. \end{cases}$$

Note that $\mathrm{GO}_{4,0}(E)$ is an anisotropic group. Using the contravariant exact functor

$$\mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}(-, \Sigma),$$

we obtain a short exact sequence

$$0 \to \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}(\mathfrak{R}^{1,0}(\mathbf{1}), \Sigma) \to \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}\big(I\left(\tfrac{1}{2}\right), \Sigma\big) \to \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}(\mathfrak{R}^{1,1}(\mathbf{1}), \Sigma) \to 0.$$

Applying Lemma 5.2.5, we have

$$0 \to \mathrm{Hom}_{\mathrm{GSp}_{1,0}(F)}(\Theta_1(\Sigma \otimes \nu), \mathbb{C}) \to \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}\big(I\left(\tfrac{1}{2}\right), \Sigma\big) \to \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) \to 0, \tag{5-1}$$

where $\Theta_1(\Sigma \otimes \nu)$ is the big theta lift to $\mathrm{GL}_2(E)$ of $\Sigma \otimes \nu$. There are no F-rational points on the nonidentity connected component of $\mathrm{GO}^*_{2,2}$ (see [Mœglin et al. 1987, pp. 21–22]), so that

$$\mathrm{GO}^*_{2,2}(F) = \mathrm{GSO}^*_{2,2}(F) = \mathfrak{Q} \cdot \mathrm{GO}_{4,0}(E)^\natural,$$

where $\mathfrak{Q}$ is the Siegel parabolic subgroup of $\mathrm{GO}^*_{2,2}(F)$. Then

$$\mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}\big(I\left(\tfrac{1}{2}\right), \Sigma\big) = \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}(\mathrm{ind}_{\mathrm{GO}^*_{2,0}(F)}^{\mathrm{GO}_{4,0}(E)^\natural} \mathbb{C}, \Sigma) = \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma, \mathbb{C}). \tag{5-2}$$

Here $\mathrm{GSO}^*_{2,0}(F)$ sits in the exact sequence

$$\begin{tikzcd} 1 \arrow[r] & E^\times \arrow[r, "i"] \arrow[d, equal] & D_E^\times(E) \times F^\times \arrow[r] \arrow[d] & \mathrm{GSO}^*_{2,0}(F) \arrow[r] \arrow[d, hook] & 1 \\ 1 \arrow[r] & E^\times \arrow[r] & D_E^\times(E) \times D_E^\times(E) \arrow[r] & \mathrm{GSO}_{4,0}(E) \arrow[r] & 1 \end{tikzcd}$$

where $i(e) = (e, N_{E/F}(e)^{-1})$ and the embedding $\mathrm{GSO}^*_{2,0}(F) \hookrightarrow \mathrm{GSO}_{4,0}(E)$ is given by

$$(x, t) \mapsto (x, t \cdot x^\sigma)$$

for $x \in D_E^\times(E)$ and $t \in F^\times$. The σ-action on $D_E^\times(E)$ is induced from the isomorphism $D_E(E) \cong D_E(E) \otimes_E (E, \sigma)$. There are two subcases:

- If $\pi_1 \ncong \pi_2$, then $\pi_1 \boxtimes \pi_2$ does not participate in theta correspondence with $\mathrm{GL}_2(E)$. The short exact sequence (5-1) implies that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}\big(I\big(\tfrac{1}{2}\big), \Sigma\big) = \dim \mathrm{Hom}_{\mathrm{GSO}^*_{2,0}(F)}(\pi_1 \boxtimes \pi_2, \mathbb{C}). \tag{5-3}$$

  Hence one can get

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{D_E^\times(E)}(\pi_2^\vee, \pi_1^\sigma),$$

  where $\pi_1^\sigma = JL^{-1}(JL(\pi_1)^\sigma)$.

- If $\pi_1 = \pi_2$, then the short exact sequence (5-1) implies that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}_{4,0}(E)^\natural}\big(I\big(\tfrac{1}{2}\big), \Sigma\big) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma, \mathbb{C})$$

  because $\Theta_1(\Sigma \otimes \nu) = 0$. Note that

$$\dim \mathrm{Hom}_{\mathrm{GSO}^*_{2,0}(F)}(\pi_1 \boxtimes \pi_1, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma, \mathbb{C}) + \dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma \otimes \nu, \mathbb{C}).$$

  In a similar way, $\dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma \otimes \nu, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,0}(F)}(JL(\pi_1), \mathbb{C})$. Therefore, if $JL(\pi_1)$ is $D^\times(F)$-distinguished, then $\pi_1^\sigma \cong \pi_1^\vee$ and so

$$\dim \mathrm{Hom}_{\mathrm{GSO}^*_{2,0}(F)}(\pi_1 \boxtimes \pi_1, \mathbb{C}) = 1 = \dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma \otimes \nu, \mathbb{C}).$$

  Then $\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{2,0}(F)}(\Sigma, \mathbb{C}) = 0$ if $JL(\pi_1)$ is $D^\times(F)$-distinguished. Also, τ is $\mathrm{GSp}_{1,1}(F)$-distinguished if and only if $JL(\pi_1)^\vee \cong JL(\pi_1)^\sigma$ which is not $D^\times(F)$-distinguished. Thus τ is $\mathrm{GSp}_{1,1}(F)$-distinguished if and only if $JL(\pi_1)$ is $(D^\times(F), \omega_{E/F})$-distinguished, in which case ϕ_{π_1} is conjugate-symplectic.

(Similarly, one can show that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \omega_{E/F}) = \dim \mathrm{Hom}_{D_E^\times(E)}(\pi_2^\vee, \pi_1^\sigma) - \dim \mathrm{Hom}_{D^\times(F)}(\Theta_1(\Sigma \otimes \nu), \omega_{E/F}).$$

Here we use the fact

$$\omega_{E/F} \circ \lambda_V|_{\mathrm{GO}^*_{2,0}(F)} = \mathbf{1}.$$

Hence $\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \omega_{E/F}) = 1$ if and only if either $JL(\pi_1) = JL(\pi_2)$ are both $D^\times(F)$-distinguished or $\pi_1 \ncong \pi_2$ but $\pi_1^\vee = \pi_2^\sigma$. It will be useful when we verify the Prasad conjecture for PGSp_4 in Section 7.)

(ii) We will use a similar argument. Assume that V_R corresponds to $\mathbb{H}_E^2$ by the Morita equivalence. By the conversation relation, we have $\theta_2^-(\tau) = 0$. Via the see-saw diagrams

$$\begin{tikzcd}
\mathrm{GO}_{5,1}(E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GSp}_{2,2}(F) \arrow[d, dash] \arrow[dl, dash] \arrow[dr, dash] & \mathrm{GO}_{2,2}(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\
\mathrm{GO}^*_{3,0}(F) & \mathrm{GSp}_4(E)^\natural & \mathrm{GO}^*_{1,1}(F)
\end{tikzcd}$$

applying Lemma 5.2.6 and Proposition 5.2.2, we have

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}(\tfrac{1}{2}), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\mathfrak{R}^2(\mathbf{1}), \tau) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Theta_4^+(\tau), \mathbb{C}),$$

where $\mathfrak{I}(s)$ is the degenerate principal series of $\mathrm{GSp}_{2,2}(F)$. Due to Lemma 5.2.8,

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) \le \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}(\tfrac{1}{2}), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Theta_4^+(\tau), \mathbb{C}).$$

We want to get the reverse inequality. Consider the diagrams

$$\begin{tikzcd}
\mathrm{GSp}_4(E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GO}^*_{2,2}(F) \arrow[d, dash] \arrow[dl, dash] \arrow[dr, dash] & \mathrm{GL}_2(E)^\natural \arrow[d, dash] \arrow[dl, dash] \\
\mathrm{GSp}_{1,1}(F) & \mathrm{GO}_{2,2}(E)^\natural & \mathrm{GSp}_{1,0}(F)
\end{tikzcd}$$

There is an exact sequence of $\mathrm{GO}^*_{2,2}(F)$-representations

$$0 \longrightarrow \mathfrak{R}^{1,0}(\mathbf{1}) \longrightarrow I\big(-\tfrac{1}{2}\big) \longrightarrow \mathfrak{R}^{1,1}(\mathbf{1}) \longrightarrow 0.$$

Note that $\dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(\mathfrak{R}^{1,0}(\mathbf{1}), \Sigma) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,0}(F)}(\Theta_1(\Sigma \otimes \nu), \mathbb{C}) = 0$. Thanks to [Ólafsson 1987, Theorem 2.5; Gourevitch et al. 2019, Proposition 4.9], we have

$$\begin{aligned}
\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) &= \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(\mathfrak{R}^{1,1}(\mathbf{1}), \Sigma) \\
&= \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}\big(I\big(-\tfrac{1}{2}\big), \Sigma\big) \\
&\ge \dim \mathrm{Hom}_{\mathrm{GO}_{2,2}(E)^\natural}(\mathrm{ind}_{\mathrm{GO}^*_{1,1}(F)}^{\mathrm{GO}_{2,2}(E)^\natural} \mathbb{C}, \Sigma) \\
&= \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Sigma, \mathbb{C}).
\end{aligned}$$

Therefore $\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Theta_4^+(\tau), \mathbb{C})$ unless $\Theta_4^+(\tau)$ is reducible. There is no F-rational points on the nonidentity connected component of $\mathrm{GO}^*_{1,1}$, so that

$$\mathrm{GO}^*_{1,1}(F) = \mathrm{GSO}^*_{1,1}(F).$$

There are two cases: $\pi_1 \ncong \pi_2$ and $\pi_1 = \pi_2$.

Assume that $\pi_1 \ncong \pi_2$. Since

$$\mathrm{GO}^*_{1,1}(F) = \mathrm{GSO}^*_{1,1}(F) \cong \mathrm{GL}_2(F) \times D^\times(F)/\{(t, t^{-1}) : t \in F^\times\},$$

for $\pi_1 \neq \pi_2$ one can obtain that $\Theta_4^+(\tau) = \mathrm{Ind}^{\mathrm{GO}(2,2)(E)}_{\mathrm{GSO}(2,2)(E)}(\pi_1 \boxtimes \pi_2)$ and

$$\mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Sigma, \mathbb{C}) = \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\pi_1 \boxtimes \pi_2, \mathbb{C}) \oplus \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\pi_2 \boxtimes \pi_1, \mathbb{C}). \tag{5-4}$$

There are two subcases:

- If π_i $(i = 1, 2)$ are both $D^\times(F)$-distinguished, then (5-4) implies that
$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Sigma, \mathbb{C}) = 2.$$
- If π_1 is $D^\times(F)$-distinguished and $\pi_2 = \pi(\chi_1, \chi_2)$ with $\chi_1 \neq \chi_2$, $\chi_1|_{F^\times} = \chi_2|_{F^\times} = 1$, then π_2 is $\mathrm{GL}_2(F)$-distinguished but not $D^\times(F)$-distinguished (see Lemma 4.4.5). So (5-4) implies that
$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = 1.$$

If $\pi_1 = \pi_2$ are both square-integrable representations, then

$$\mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Sigma, \mathbb{C}) = \mathrm{Hom}_{\mathrm{GSO}^*_{1,1}(F)}(\pi_1 \boxtimes \pi_1, \mathbb{C}) = \begin{cases} 1 & \text{if } \pi_1 \text{ is } D^\times(F)\text{-distinguished,} \\ 0 & \text{otherwise.} \end{cases}$$

If $\pi_1 = \pi_2 = \pi(\chi^{-1}, \chi^\sigma)$, then $\Theta_4^+(\tau)$ is reducible. We will show that $\tau = I_{Q(Z)}(\mathbf{1}, \pi_1)$ does not occur on the boundary of $\mathfrak{I}\big(\tfrac12\big)$ and hence that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{1,1}(F)}(\Theta_4^+(\tau), \mathbb{C}).$$

There is a filtration

$$\mathrm{ind}^{\mathrm{GSp}_4(E)^\natural}_{\mathrm{GSp}_{1,1}(F)} \mathbb{C} = \mathfrak{I}_0(s) \subset \mathfrak{I}_1(s) \subset \mathfrak{I}_2(s) = \mathfrak{I}(s)|_{\mathrm{GSp}_4(E)^\natural}$$

of $\mathfrak{I}(s)|_{\mathrm{GSp}_4(E)^\natural}$ such that $\mathfrak{I}_2(s)/\mathfrak{I}_1(s) = \mathrm{ind}^{\mathrm{GSp}_4(E)^\natural}_{P^\natural}\, \delta_{P^\natural}^{(s+1)/3}$ and

$$\mathfrak{I}_1(s)/\mathfrak{I}_0(s) = \mathrm{ind}^{\mathrm{GSp}_4(E)^\natural}_{MN}\, \delta_{P(Y_D)}^{(1/2)+(s/5)} \delta_3^{-1/2},$$

where $\delta_3(t, x) = |N_{E/F}(t)^4 \lambda(d)^{-4}|_F$ for $(t, x) \in M = \mathrm{GL}_1(E) \times \mathrm{GSp}_{1,0}(F)$. If

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_1\big(\tfrac12\big)/\mathfrak{I}_0\big(\tfrac12\big), \tau\big) \neq 0,$$

then

$$\mathrm{Hom}_{\mathrm{GL}_1(E)}(|-|_E, R_{\bar{P}''}(I_{Q(Z)}(\mathbf{1}, \pi_1))) \neq 0,$$

which is impossible, where $P'' = (\mathrm{GL}_1(E) \times \mathrm{GL}_2(E)^\natural) \ltimes N$ is a parabolic subgroup of $\mathrm{GSp}_4(E)^\natural$ and $R_{\bar{P}''}$ denotes the Jacquet functor associate to the parabolic opposite to P''. So

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_1\big(\tfrac12\big)/\mathfrak{I}_0\big(\tfrac12\big), \tau\big) = 0.$$

It is quite straightforward to see that

$$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\mathrm{ind}^{\mathrm{GSp}_4(E)^\natural}_{P^\natural}\, \delta^{1/2}_{P^\natural}, I_{Q(Z)}(\mathbf{1}, \pi_1)) = 0$$

by applying the Jacquet functor. Hence $\tau = I_{Q(Z)}(\mathbf{1}, \pi_1)$ does not occur on the boundary of $\mathfrak{I}\big(\tfrac12\big)$.

The big theta lift to $\mathrm{GSO}_{2,2}(E)$ of $\tau = I_{Q(Z)}(\mathbf{1}, \pi_1)$ of $\mathrm{GSp}_4(E)$ is

$$\mathrm{Ext}^1_{\mathrm{GSO}(2,2)(E)}(\pi_1 \boxtimes \pi_1, \pi_1 \boxtimes \pi_1).$$

From the see-saw pairs diagram

$$\begin{tikzcd}[column sep=small]
\mathrm{GSO}_{5,1}(E)^\natural \arrow[d, dash] \arrow[dr, dash] & \mathrm{GSp}_{2,2}(F) \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \mathrm{GSO}_{2,2}(E)^\natural \arrow[dl, dash] \arrow[d, dash] \\
\mathrm{GO}^*_{3,0}(F) & \mathrm{GSp}_4(E)^\natural & \mathrm{GO}^*_{1,1}(F)
\end{tikzcd}$$

one can use the fact $\theta_2^-(\tau) = 0$ to obtain that

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSO}^*_{1,1}(F)}(\mathrm{Ext}^1_{\mathrm{GSO}_{2,2}(E)}(\pi_1 \boxtimes \pi_1, \pi_1 \boxtimes \pi_1), \mathbb{C}) = 2.$$

(iii) Assume that $\theta_4^+(\tau) = 0$. Note that $0 \to \mathfrak{R}^2(\mathbf{1}) \to \mathfrak{I}\big(-\frac{1}{2}\big) \to \mathfrak{R}^3(\mathbf{1}) \to 0$ is exact. Then we can use the same method appearing in (ii) to show that

$$\begin{aligned}\dim \mathrm{Hom}_{\mathrm{GO}^*_{3,0}(F)}(\Theta_2^-(\tau), \mathbb{C}) &= \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\mathfrak{R}^3(\mathbf{1}), \tau) \\ &= \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}\big(-\tfrac{1}{2}\big), \tau\big) \geq \dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}).\end{aligned}$$

We will show that τ does not occur on the boundary of $\mathfrak{I}\big(-\frac{1}{2}\big)$ in this case. Then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}\big(-\tfrac{1}{2}\big), \tau\big) \leq \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_0\big(-\tfrac{1}{2}\big), \tau\big) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C})$$

and so

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}\big(-\tfrac{1}{2}\big), \tau\big).$$

In order to show that τ does not occur on the boundary of $\mathfrak{I}\big(-\frac{1}{2}\big)$, we separate them into two cases.

- If $\tau = I_{Q(Z)}(\chi, \pi)$ with $\chi \neq \mathbf{1}$, then

  $$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_2\big(-\tfrac{1}{2}\big)/\mathfrak{I}_1\big(-\tfrac{1}{2}\big), \tau\big) = \mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}(\mathrm{ind}_{P^\natural}^{\mathrm{GSp}_4(E)^\natural} \delta_{P^\natural}^{1/6}, \tau) = 0.$$

  If $\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_1\big(-\frac{1}{2}\big)/\mathfrak{I}_0\big(-\frac{1}{2}\big), \tau\big) \neq 0$, then $\mathrm{Hom}_{\mathrm{GL}_1(E)}(\mathbf{1}, R_{\overline{P}''}(\tau)) \neq 0$ which is impossible since $R_{\overline{P}''}(\tau) = \chi \otimes \pi \oplus \chi^{-1} \otimes \pi\chi$ and $\chi \neq \mathbf{1}$, where $P'' = (\mathrm{GL}_1(E) \times \mathrm{GL}_2(E)^\natural) \rtimes N$.
- If τ is square-integrable, then $\mathrm{Hom}_{\mathrm{GL}_1(E)}(\mathbf{1}, R_{\overline{P}''}(\tau)) = 0$ due to the Casselman criterion in [Casselman and Miličić 1982] for a discrete series representation that $\mathrm{Hom}_{\mathrm{GL}_1(E)}(|-|_E^s, R_{\overline{P}''}(\tau)) \neq 0$ implies that $s < 0$. Hence $\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_1\big(-\frac{1}{2}\big)/\mathfrak{I}_0\big(-\frac{1}{2}\big), \tau\big) = 0$. In a similar way,

  $$\mathrm{Hom}_{\mathrm{GSp}_4(E)^\natural}\big(\mathfrak{I}_2\big(-\tfrac{1}{2}\big)/\mathfrak{I}_1\big(-\tfrac{1}{2}\big), \tau\big) = \mathrm{Hom}_{\mathrm{GL}_2(E) \times F^\times}(\delta_{P^\natural}^{1/6}, R_{\overline{P}^\natural}(\tau)) = 0.$$

Hence τ does not occur on the boundary of $\mathfrak{I}\big(-\frac{1}{2}\big)$. Moreover, if $\tau \neq I_{Q(Z)}(|-|_E, \rho)$, then $\Theta_2^-(\tau) = \Pi^D \boxtimes \chi$ is irreducible. Then there exists an identity

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{3,0}(F)}(\Pi^D \boxtimes \chi, \mathbb{C}) = \dim \mathrm{Hom}_{D_4^\times(F)}(\Pi^D, \mathbb{C}),$$

where D_4 is the division algebra over F of degree 4.

- If $\Pi = JL(\Pi^D)$ is a square-integrable representation of $\mathrm{GL}_4(E)$, then [Beuzart-Plessis 2018, Theorem 1] and Theorem 4.4.4 imply that

$$\dim \mathrm{Hom}_{\mathrm{GL}_4(F)}(\Pi, \omega_{E/F}) = \dim \mathrm{Hom}_{D_4^\times(F)}(\Pi^D, \omega_{E/F}) = \begin{cases} 1 & \text{if } \phi_\Pi \text{ is conjugate-symplectic,} \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\dim \mathrm{Hom}_{D_4^\times(F)}(\pi^D, \mathbb{C}) = 1$ if and only if ϕ_Π is conjugate-orthogonal.

- If Π^D is an induced representation $\pi(\rho_D, (\rho_D)^\vee \otimes \mu)$ with $\mu \neq \omega_{\rho_D}$, then we use the orbit decomposition $B_1 \backslash \mathrm{GL}_2(D_E)(E)/\mathrm{GL}_1(D_4)(F)$ and Mackey theory to get that

$$\begin{aligned} \dim \mathrm{Hom}_{D_4^\times(F)}(\Pi^D, \mathbb{C}) &= \dim \mathrm{Hom}_{D_E^\times(E)}(\rho_D^\sigma \otimes \rho_D^\vee \cdot \mu, \mathbb{C}) = \dim \mathrm{Hom}_{D_E^\times(E)}(\rho_D^\sigma, \rho_D \cdot \mu^{-1}) \\ &= \begin{cases} 1 & \text{if } \rho_D^\sigma \cong \rho_D \mu^{-1}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{5-5}$$

In this case, $\rho^\sigma = \rho\mu^{-1}$ where $\rho = JL(\rho_D)$ is the Jacquet–Langlands lift to $\mathrm{GL}_2(E)$ and $\phi_\Pi = \phi_\rho \oplus \phi_\rho^\vee \cdot \mu$, which is conjugate-orthogonal due to Theorem 4.4.4.

- If $\Pi^D = \mathrm{Sp}(\rho_D|-|_E^{1/2})$ is a generalized Speh representation and $\tau = I_{Q(Z)}(|-|_E, \rho)$, then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \mathbb{C}) = \dim \mathrm{Hom}_{\mathrm{GO}^*_{3,0}(F)}(\Theta_2^-(\tau), \mathbb{C}) = \begin{cases} 1 & \text{if } \rho^\sigma \cong \rho^\vee |-|_E^{-1}, \\ 0 & \text{otherwise.} \end{cases} \qquad \square$$

6. The Prasad conjecture for GSp_4

6A. ***The Prasad conjecture.*** In this subsection, we give a brief introduction to the Prasad conjecture [2015, Conjecture 2]. One may refer to [Prasad 2015, §13] for more details.

Let $\boldsymbol{G}$ be a quasisplit reductive group defined over a local field F with characteristic zero. Let W_F be the Weil group of F and WD_F be the Weil–Deligne group of F. Let E be a quadratic extension over F. A quadratic character $\chi_{\boldsymbol{G}}$ is introduced in [Prasad 2015, §8] and another quasisplit reductive group G^{op} defined over F is introduced in [Prasad 2015, §7]. Then there is a relation between the fibers of the base change map

$$\Phi : \mathrm{Hom}(WD_F, {}^L G^{op}) \to \mathrm{Hom}(WD_E, {}^L G^{op})$$

from the Galois side and the $\chi_{\boldsymbol{G}}$-distinction problems for $\boldsymbol{G}(E)/\boldsymbol{G}(F)$ from the automorphic side.

More precisely, assume the Langlands–Vogan conjecture in [Vogan 1993]. Given an irreducible representation π of $\boldsymbol{G}(E)$ with an enhanced L-parameter (ϕ_π, λ), where λ is an irreducible representation of the component group $\pi_0(Z(\phi_\pi))$ and the L-packet Π_{ϕ_π} is generic, we have

$$\sum_\alpha \dim \mathrm{Hom}_{G_\alpha(F)}(\pi, \chi_{\boldsymbol{G}}) = \sum_i m(\lambda, \tilde{\phi}_i) \deg \Phi(\tilde{\phi}_i)/d_0(\tilde{\phi}_i),$$

where

- $\alpha \in H^1(W_F, \boldsymbol{G})$ runs over all pure inner forms of $\boldsymbol{G}$ satisfying $G_\alpha(E) = \boldsymbol{G}(E)$;
- $\tilde{\phi}_i \in \mathrm{Hom}(WD_F, {}^L G^{op})$ runs over all parameters of ${}^L G^{op}$ satisfying $\tilde{\phi}_i|_{WD_E} = \phi_\pi$;
- $m(\lambda, \tilde{\phi}) = \dim \mathrm{Hom}_{\pi_0(Z(\tilde{\phi}))}(\mathbf{1}, \lambda)$ is the multiplicity of the trivial representation contained in the restricted representation $\lambda|_{\pi_0(Z(\tilde{\phi}))}$;
- $d_0(\tilde{\phi}) = |\mathrm{Coker}\{\pi_0(Z(\tilde{\phi})) \to \pi_0(Z(\phi_\pi))^{\mathrm{Gal}(E/F)}\}|$.

Remark 6.1.1. If $H^1(F, \boldsymbol{G})$ is trivial such as $\boldsymbol{G} = \mathrm{GSp}_{2n}$, then the automorphic side contains only one term. The Prasad conjecture gives a precise formula for the multiplicity

$$\dim \mathrm{Hom}_{\boldsymbol{G}(F)}(\pi, \chi_{\boldsymbol{G}}).$$

Remark 6.1.2. There exists a counterexample even for GL_2 when Π_{ϕ_π} is not generic. Let $\boldsymbol{G} = \mathrm{GL}_2$, $\chi_{\boldsymbol{G}} = \omega_{E/F}$ and $\pi = \mathbf{1}$ be the trivial representation. Then the automorphic side is zero however the Galois side is nonzero.

Remark 6.1.3. If $\tilde{\phi}$ comes from a square-integrable representation, then $\deg \Phi(\tilde{\phi}) = 1$. The reason, due to Prasad, is that $\tilde{\phi}$ represents a singleton in $\mathrm{Hom}(WD_F, {}^L G^{op})$.

If π is square-integrable, then we have a refined version, i.e., the formula for each dimension

$$\dim \mathrm{Hom}_{G_\alpha(F)}(\pi, \chi_{\boldsymbol{G}}).$$

Let $Z(\widehat{G}^{op})$ be the center of the dual group $\widehat{G}^{op}$. There is a perfect pairing

$$H^1(\mathrm{Gal}(E/F), Z(\widehat{G}^{op})^{W_E}) \times H^1(\mathrm{Gal}(E/F), \boldsymbol{G}(E)) \to \mathbb{Q}/\mathbb{Z}$$

for Prasad's studies [2015, §11] of the character twists. Set $\Omega_{\boldsymbol{G}}(E) = H^1(\mathrm{Gal}(E/F), Z(\widehat{G}^{op})^{W_E})$. Given a parameter $\tilde{\phi} \in H^1(W_F, \widehat{G}^{op})$, we consider the stabilizer $\Omega_{\boldsymbol{G}}(\tilde{\phi}, E) \subset \Omega_{\boldsymbol{G}}(E)$ under the pairing

$$H^1(W_F, Z(\widehat{G}^{op})) \times H^1(W_F, \widehat{G}^{op}) \to H^1(W_F, \widehat{G}^{op}).$$

Set

$$A_{\boldsymbol{G}}(\tilde{\phi}) \subset H^1(\mathrm{Gal}(E/F), \boldsymbol{G}(E)) \cong \Omega_{\boldsymbol{G}}(E)^\vee$$

to be the annihilator of the stabilizer $\Omega_{\boldsymbol{G}}(\tilde{\phi}, E)$. Then there is another perfect pairing

$$\Omega_{\boldsymbol{G}}(E)/\Omega_{\boldsymbol{G}}(\tilde{\phi}, E) \times A_{\boldsymbol{G}}(\tilde{\phi}) \to \mathbb{Q}/\mathbb{Z},$$

meaning that in the orbit $\Omega_{\boldsymbol{G}}(E)/\Omega_{\boldsymbol{G}}(\tilde{\phi}, E)$ of character twists of $\tilde{\phi}$ (which go to a particular parameter under the base change to E) there are exactly as many parameters as there are certain pure inner forms of $\boldsymbol{G}$ over F which trivialize after base change to E.

Consider

$$F(\phi_\pi) = \{\tilde{\phi} : WD_F \to {}^L G^{op} \mid \tilde{\phi}|_{WD_E} = \phi_\pi\} = \sqcup_{i=1}^r \mathcal{O}(\tilde{\phi}_i).$$

Each orbit $\mathcal{O}(\tilde{\phi}_i)$ of $\Omega_{\boldsymbol{G}}(E)$-action on $F(\phi_\pi)$ is associated to a coset $\mathcal{C}_i$ of $A_{\boldsymbol{G}}(\tilde{\phi}_i)$ in $H^1(\mathrm{Gal}(E/F), \boldsymbol{G}(E))$ defining a set of certain pure inner forms G_α of $\boldsymbol{G}$ over F such that $G_\alpha(E) = \boldsymbol{G}(E)$. Then

$$\dim \mathrm{Hom}_{G_{\alpha(F)}}(\pi, \omega_{\boldsymbol{G}}) = \sum_{i=1}^r m(\lambda, \tilde{\phi}_i) \cdot 1_{\mathcal{C}_i}(G_\alpha)/d_0(\tilde{\phi}_i),$$

where

- $1_{\mathcal{C}_i}$ is the characteristic function of the coset $\mathcal{C}_i$;
- $m(\lambda, \tilde{\phi})$ is the multiplicity for the trivial representation contained in the restricted representation $\lambda|_{\pi_0(Z(\tilde{\phi}))}$, which may be zero;
- $d_0(\tilde{\phi}) = |\,\mathrm{Coker}\{\pi_0(Z(\tilde{\phi})) \to \pi_0(Z(\phi_\pi))^{\mathrm{Gal}(E/F)}\}|$.

6B. ***The Prasad conjecture for*** $\mathbf{GL_2}$**.** Before we give the proof of Theorem 1.2, let us recall the Prasad conjecture for $\boldsymbol{G}=\mathrm{GL}_2=\mathrm{GSp}_2$. Set $\boldsymbol{G}=\mathrm{GL}_2$. Then $\chi_{\boldsymbol{G}}=\omega_{E/F}$ and $G^{op}=\mathrm{U}(2,E/F)$ is the quasisplit unitary group, where E is a quadratic field extension over a p-adic field F. Denote

$$^{L}G^{op}=\mathrm{GL}_2(\mathbb{C})\rtimes\langle\sigma\rangle,$$

where σ-action on $\mathrm{GL}_2(\mathbb{C})$ is given by

$$\sigma(g)=\omega_0(g^t)^{-1}\omega_0^{-1}=g\cdot\det(g)^{-1},$$

$\omega_0=\begin{pmatrix}&1\\-1&\end{pmatrix}$ and $g\in\mathrm{GL}_2(\mathbb{C})$, g^t denotes its transpose matrix. Given an irreducible representation π of $\mathrm{GL}_2(E)$ with $\phi=\phi_\pi$ irreducible (for simplicity), there is no other pure inner form for GL_2. Then

$$\dim\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi,\omega_{E/F})=|F(\phi)|,$$

where $F(\phi)=\{\tilde{\phi}:WD_F\to{}^{L}G^{op}\mid\tilde{\phi}|_{WD_E}=\phi\}$ and $|F(\phi)|$ denotes its cardinality.

Proposition 6.2.1. *The following statements are equivalent*:

(i) $\dim\mathrm{Hom}_{\mathrm{GL}_2(F)}(\pi,\omega_{E/F})=1$.

(ii) *The Langlands parameter ϕ is conjugate-symplectic.*

(iii) *There is only one extension $\tilde{\phi}\in F(\phi)$.*

Proof. We only prove the direction (ii)⇒(iii) and the rest follows from Flicker's results [1991]. If ϕ is conjugate-symplectic, then

$$\phi^s=\phi^\vee=\phi(\det\phi)^{-1},$$

where $s\in W_F\setminus W_E$ is fixed. There exists $A\in\mathrm{GL}_2(\mathbb{C})$ such that

$$\phi(sts^{-1})=\phi^s(t)=A\cdot\phi(t)\det(\phi(t))^{-1}\cdot A^{-1}$$

for all $t\in WD_E$. Set

$$\tilde{\phi}(s)=A\cdot\sigma$$

and $\tilde{\phi}(t)=\phi(t)$ for $t\in WD_E$. Then

$$\tilde{\phi}(sts^{-1})=\tilde{\phi}(s)\cdot\phi(t)\cdot\tilde{\phi}(s)^{-1}$$

and $\tilde{\phi}(s^2)=\phi(s^2)=(\tilde{\phi}(s))^2$ due to the sign of ϕ. More precisely, assuming that $\langle -,-\rangle$ is the WD_E-equivariant bilinear form associated to $\phi:WD_E\to\mathrm{GSp}(V,\langle -,-\rangle)$, we define

$$B:V\times V\to\mathbb{C}$$

by $B(v_1,v_2)=\langle v_1,A^{-1}v_2\rangle$ for $v_1,v_2\in V$. Then

$$B(\phi(t)v_1,\phi^s(t)v_2)=\langle\phi(t)v_1,\phi^\vee(t)A^{-1}v_2\rangle=B(v_1,v_2)$$

and so B gives a conjugate-self-dual bilinear form on V. By Schur's lemma, B has sign -1, i.e.,

$$B(v_1,\phi(s^2)v_2)=-B(v_2,v_1)$$

for all $v_1, v_2 \in V$. Thus $B(Av_1, \phi(s^2)v_2) = -B(v_2, Av_1)$, i.e.,

$$\langle Av_1, A^{-1}\phi(s^2)v_2\rangle = -\langle v_2, A^{-1}Av_1\rangle = \langle v_1, v_2\rangle$$

for all $v_i \in V$. Then $\det(A)\cdot A^{-2}\phi(s^2) = 1$, i.e., $\phi(s^2) = A\cdot\det(A)^{-1}A = (\tilde{\phi}(s))^2$.

Therefore $\tilde{\phi} \in F(\sigma)$. If there are two extensions $\tilde{\phi}_i$ with $A_i \in \mathrm{GL}_2(\mathbb{C})$ such that $\tilde{\phi}_i|_{WD_E} = \phi$, then $A_1A_2^{-1} \in Z(\phi) \cong \mathbb{C}^\times$ by Schur's lemma, so that $\phi_1 = \phi_2$. □

Remark 6.2.2. This method will appear again when we study the Prasad conjecture for $\boldsymbol{G} = \mathrm{GSp}_4$ in Section 6C1 The key idea is to choose a proper element A such that the lift

$$\tilde{\phi} : WD_F \to {}^LG_0$$

satisfies $\tilde{\phi}(s) = A\cdot\sigma$ and $\tilde{\phi}|_{WD_E} = \phi$.

6C. ***The Prasad conjecture for*** $\mathbf{GSp_4}$**.** The aim of this subsection is to verify the Prasad conjecture for GSp_4. Now we consider the generic representation $\tau = \theta(\Pi\boxtimes\chi)$ of $\mathrm{GSp}_4(E)$, with ϕ_Π conjugate-symplectic and $\chi|_{F^\times} = 1$. Note that the Langlands parameter ϕ_Π is equal to $i\circ\phi_\tau$, where

$$i : \mathrm{GSp}_4(\mathbb{C}) \to \mathrm{GL}_4(\mathbb{C})$$

is the embedding between L-groups. Furthermore, χ is the similitude character of ϕ_τ. If ϕ_Π is conjugate-symplectic (resp. conjugate-orthogonal), we say that ϕ_τ is conjugate-symplectic (resp. conjugate-orthogonal). There are two cases: ϕ_Π is irreducible and ϕ_Π is reducible.

Lemma 6.3.1. *Assume that $\tau = \theta(\Pi\boxtimes\chi)$ is a generic representation of* $\mathrm{GSp}_4(E)$ *and $\omega_\tau|_{F^\times} = \mathbf{1}$. Then τ is* $(\mathrm{GSp}_4(F), \omega_{E/F})$*-distinguished if and only if ϕ_Π is conjugate-symplectic.*

Proof. Due to Theorem 4.4.9, the following are equivalent:

- τ is $\mathrm{GSp}_4(F)$-distinguished.
- Π is $\mathrm{GL}_4(F)$-distinguished.
- ϕ_Π is conjugate-orthogonal.

Fix a character χ_E of $E^\times$ such that $\chi_E|_{F^\times} = \omega_{E/F}$. Then τ is $(\mathrm{GSp}_4(F), \omega_{E/F})$-distinguished if and only if $\tau\otimes\chi_E\circ\lambda_W$ is $\mathrm{GSp}_4(F)$-distinguished, which is equivalent to that $\phi_\Pi\otimes\chi_E$ is conjugate-orthogonal. Note that χ_E^{-1} is conjugate-symplectic. Hence τ is $(\mathrm{GSp}_4(F), \omega_{E/F})$-distinguished if and only if ϕ_Π is conjugate-symplectic. □

Recall that if $\boldsymbol{G} = \mathrm{GSp}_{2n}$, then $\chi_{\boldsymbol{G}} = \omega_{E/F}$ and

$$G^{op}(F) = \{g \in \mathrm{GSp}_{2n}(E) | \sigma(g) = \theta(g)\},$$

where $\theta(g) = \lambda_W(g)^{-1}g$ is the involution. Note that the σ-actions on $\mathrm{GSp}_4(E)$ and $\mathrm{GSp}_4(\mathbb{C})$ are totally different. (We hope that this will not confuse the reader.) Observe that $H^1(\mathrm{Gal}(E/F), Z(\widehat{G}^{op})^{W_E}) = 1$, which corresponds to the fact that the pure inner form of GSp_{2n} is trivial.

According to Theorem 4.4.9, we will divide the proof of Theorem 1.2 into four parts:

- $i\circ\phi_\tau$ is irreducible;
- $i\circ\phi_\tau=\rho\oplus\rho\nu$ with $\nu\neq\mathbf{1}$;
- the endoscopic case $i\circ\phi_\tau=\phi_{\pi_1}\oplus\phi_{\pi_2}$ and τ is generic;
- $i\circ\phi_\tau=\phi_{\pi_1}\oplus\phi_{\pi_2}$ and τ is nongeneric.

See Section 6C1–Section 6C4.

6C1. *The irreducible L-parameter ϕ_τ.* Given a conjugate-symplectic L-parameter $\phi=\phi_\tau$, which is irreducible, we want to extend ϕ to

$$\tilde{\phi}:WD_F\to {}^LG_0=\mathrm{GSp}_4(\mathbb{C})\rtimes\langle\sigma\rangle,$$

where σ acts on $\mathrm{GSp}_4(\mathbb{C})$ by

$$\sigma(g)=g\cdot\mathrm{sim}(g)^{-1}.$$

Let $s\in W_F\setminus W_E$. The parameter ϕ is conjugate-symplectic, so that $\phi^\vee=\phi^s$ and $\phi^\vee=\phi\chi^{-1}$. Hence there exists an element $A\in\mathrm{GSp}_4(\mathbb{C})$ such that

$$\phi(sts^{-1})=\phi^s(t)=A\cdot\phi(t)\chi^{-1}(t)\cdot A^{-1}\tag{6-1}$$

for all $t\in WD_E$. Set

$$\tilde{\phi}(s)=A\cdot\sigma\quad\text{and}\quad\tilde{\phi}(t)=\phi(t)$$

for $t\in WD_E$. Then $\phi(sts^{-1})=A\phi(t)\chi^{-1}(t)A^{-1}=\tilde{\phi}(s)\cdot\phi(t)\cdot\tilde{\phi}(s)^{-1}$. Moreover, we will show that

$$\tilde{\phi}(s^2)=\phi(s^2)=(\tilde{\phi}(s))^2.$$

Then $\tilde{\phi}\in\mathrm{Hom}(WD_F,{}^LG_0)$ and $\tilde{\phi}|_{WD_E}=\phi$.

Assume that $\langle -,-\rangle$ is the WD_E-equivariant bilinear form associated to

$$\phi_\tau:WD_E\to\mathrm{GSp}_4(\mathbb{C})=\mathrm{GSp}(V,\langle -,-\rangle).$$

Set

$$B(v,w)=\langle v,A^{-1}w\rangle$$

for $v,w\in V$. Then (6-1) implies that

$$B(\phi(t)v,\phi(sts^{-1})w)=\langle\phi(t)v,\phi(t)\chi^{-1}(t)A^{-1}w\rangle=\chi(t)\cdot\langle v,\chi^{-1}(t)A^{-1}w\rangle=B(v,w).$$

Thus B is a conjugate-self-dual bilinear form on ϕ and hence it has sign -1 by Schur's lemma, i.e.,

$$-B(w,v)=B(v,\phi(s^2)w).$$

Therefore we have

$$\begin{aligned}\langle v,w\rangle=-\langle w,v\rangle=-B(w,Av)&=B(Av,\phi(s^2)w)\\&=\langle Av,A^{-1}\phi(s^2)w\rangle=\langle v,\mathrm{sim}(A)A^{-2}\phi(s^2)w\rangle\end{aligned}$$

and so $\phi(s^2)=A\cdot\mathrm{sim}(A)^{-1}A=(\tilde{\phi}(s))^2$.

Proposition 6.3.2. *Assume that $\tau=\theta(\Pi\boxtimes\chi)$ with ϕ_Π irreducible. Then there exists at most one extension $\tilde{\phi}: WD_F \to {}^LG_0$ such that $\tilde{\phi}|_{WD_E}=\phi_\tau$.*

Proof. If there are two extensions $\tilde{\phi}_i (i=1,2)$ such that $\tilde{\phi}_i(s)=A_i\cdot\sigma$ with $A_i\in \mathrm{GSp}_4(\mathbb{C})$ and

$$\tilde{\phi}_i(sts^{-1})=\tilde{\phi}_i(s)\cdot\phi_\tau(t)\cdot\tilde{\phi}_i(s)^{-1}$$

for all $t\in WD_E$, then $A_1A_2^{-1}$ commutes with ϕ_τ. So $A_1A_2^{-1}$ is a scalar by Schur's lemma. Thus $\tilde{\phi}_1=\tilde{\phi}_2$. □

Hence, if $\tau=\theta(\Pi\boxtimes\chi)$ with ϕ_Π irreducible and conjugate-symplectic, then there is one extension $\tilde{\phi}\in F(\phi_\tau)$ and

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau,\omega_{E/F})=1.$$

If $\phi=\phi_\tau$ is conjugate-symplectic and reducible, then there are several cases.

6C2. *$\phi_\tau=\rho+\rho\nu$ with $\nu\neq\mathbf{1}$ and ρ irreducible.* If $\phi_\Pi=\rho+\rho\nu$ with ρ irreducible and $\chi=\nu\cdot\det\rho$ conjugate-orthogonal, then $\chi\chi^s=\mathbf{1}$. Thanks to Theorem 4.4.4, there are two subcases:

- ρ and $\rho\nu$ are both conjugate-symplectic or
- $\rho^s=\rho^\vee\nu^{-1}$.

(i) If ρ and $\rho\nu$ are both conjugate-symplectic, then ν is conjugate-orthogonal and there exist

$$\tilde{\rho}_i: WD_F\to \mathrm{GL}_2(\mathbb{C})\rtimes\langle\sigma\rangle$$

such that $\tilde{\rho}_1|_{WD_E}=\rho,\ \tilde{\rho}_2|_{WD_E}=\rho\nu$ and $\tilde{\rho}_i(s)=A_i\cdot\sigma$ for $A_i\in\mathrm{GL}_2(\mathbb{C})$ due to Proposition 6.2.1. Note that ρ is irreducible. Then given $t\in WD_E$,

$$\tilde{\rho}_1^s(t)\nu^s(t)=\tilde{\rho}_2^s(t)=A_2\sigma(\rho(t)\nu(t))(A_2\sigma)^{-1}=A_2\rho^\vee(t)A_2^{-1}\cdot\nu^{-1}(t)$$

and so $A_1\cdot\sigma\cdot\rho(t)\sigma^{-1}A_1^{-1}=A_2\rho^\vee(t)A_2^{-1}$ (since $\nu\nu^s=\mathbf{1}$) which implies $A_1A_2^{-1}\in\mathbb{C}^\times$. Set

$$\tilde{\phi}(s)=\begin{pmatrix} & A_1\\ A_1 & \end{pmatrix}\cdot\sigma\in\mathrm{GSp}_4(\mathbb{C})\rtimes\langle\sigma\rangle \quad\text{and}\quad \tilde{\phi}(t)=\begin{pmatrix}\rho(t) & \\ & \rho(t)\nu(t)\end{pmatrix}$$

for $t\in WD_E$. Then $\tilde{\phi}\in F(\phi)$ is the unique extension of ϕ_τ.

(ii) If $\rho^s\cong\rho^\vee\nu^{-1}$, there exists an $A\in\mathrm{GL}_2(\mathbb{C})$ such that

$$\rho^s(t)\nu(t)=(\det\rho(t))^{-1}\cdot A\rho(t)A^{-1}$$

for $t\in WD_E$. Then

$$\det\rho^s\cdot\det\rho\cdot\nu^2=\mathbf{1},$$

which implies that $\nu=\nu^s$. Observe that

$$\begin{aligned}\rho^s(sts^{-1})\nu(sts^{-1})&=(\det\rho(sts^{-1}))^{-1}A\rho(sts^{-1})A^{-1}\\&=\det\rho^s(t)^{-1}A\cdot\nu(t)^{-1}\det\rho(t)^{-1}A\rho(t)A^{-1}\cdot A^{-1}\\&=\nu(t)^{-1}\det\rho^s(t)^{-1}\det\rho(t)^{-1}A^2\rho(t)A^{-2}.\end{aligned}$$

Then $\rho(s^2)\rho(t)\rho(s^2)^{-1} = A^2\rho(t)A^{-2}$ since the character $\nu\det\rho$ is conjugate-orthogonal. Note that ρ is irreducible. Then $A^{-2}\rho(s^2)$ is a scalar. Choose a proper A such that $A^{-2}\rho(s^2) = 1$. Set

$$\tilde{\phi}(s) = \begin{pmatrix} A & \\ & A\cdot\det(A^{-1}) \end{pmatrix}\cdot\sigma \quad\text{and}\quad \tilde{\phi}(t) = \begin{pmatrix} \rho(t) & \\ & \rho(t)\nu(t) \end{pmatrix}$$

for $t \in WD_E$. Then

$$\begin{aligned} \tilde{\phi}(s)\cdot\phi(t)\cdot\tilde{\phi}(s)^{-1} &= \begin{pmatrix} A & \\ & A\cdot\det(A)^{-1} \end{pmatrix}\cdot\sigma\cdot\begin{pmatrix} \rho(t) & \\ & \rho(v)\nu(t) \end{pmatrix}\cdot\left(\sigma^{-1}\cdot\begin{pmatrix} A^{-1} & \\ & A^{-1}\det(A) \end{pmatrix}\right) \\ &= \begin{pmatrix} A & \\ & A\cdot\det(A)^{-1} \end{pmatrix}\begin{pmatrix} \rho^\vee(t)\nu(t)^{-1} & \\ & \rho^\vee(t) \end{pmatrix}\cdot\sigma\cdot\sigma^{-1}\cdot\begin{pmatrix} A^{-1} & \\ & A^{-1}\det(A) \end{pmatrix} \\ &= \begin{pmatrix} A\rho^\vee(t)\nu(t)^{-1}A^{-1} & \\ & A\rho^\vee(t)A^{-1} \end{pmatrix} = \begin{pmatrix} \rho^s(t) & \\ & \rho^s(t)\nu(t) \end{pmatrix} = \tilde{\phi}^s(t) \end{aligned} \tag{6-2}$$

and $(\tilde{\phi}(s))^2 = \phi(s^2)$. Thus $\tilde{\phi}$ is a homomorphism from WD_F to ${}^LG_0$ and $\tilde{\phi}|_{WD_E} = \phi$.

Remark 6.3.3. The key point here is to find a proper element $\tilde{\phi}(s)$ such that $\tilde{\phi} \in \mathrm{Hom}(WD_F, {}^LG_0)$. Hence we always need to check the following two conditions: $\tilde{\phi}^s(t) = \tilde{\phi}(s)\cdot\phi(t)\cdot\tilde{\phi}(s)^{-1}$ and $(\tilde{\phi}(s))^2 = \phi(t^2)$. Following the definition, the computation like (6-2) is quite straightforward and we may skip it sometimes.

6C3. *Endoscopic case.* If $\phi_\tau = \rho_1 + \rho_2$ is the endoscopic case, then $\det\rho_1 = \det\rho_2$ are both conjugate-orthogonal. There are several subcases. Assume that $\tau = \theta(\pi_1 \boxtimes \pi_2)$ is generic, $\rho_i = \phi_{\pi_i} (i = 1, 2)$ and $\rho_0 = \chi_1 + \chi_2$, with $\chi_1 \neq \chi_2$ and $\chi_1|_{F^\times} = \chi_2|_{F^\times} = \omega_{E/F}$. There are also 2 cases: $\rho_1 \neq \rho_2$ and $\rho_1 = \rho_2$.

Assume that $\rho_1 \neq \rho_2$. Then

(i) If ρ_1 and ρ_2 are both conjugate-symplectic and $\rho_i \neq \rho_0$ $(i = 1, 2)$, so that both π_1 and π_2 are $(D^\times(F), \omega_{E/F})$-distinguished due to Lemma 4.4.5, then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_{E/F}) = 2.$$

Thanks to Proposition 6.2.1, there exist $\tilde{\rho}_1$ and $\tilde{\rho}_2$ of U(2, E/F) such that $\tilde{\rho}_i|_{WD_E} = \rho_i$. (Here we need to choose A_i properly such that $\det A_1 = \det A_2$ if $\tilde{\rho}_i(s) = A_i\cdot\sigma$.)

If ρ_1 and ρ_2 are both irreducible, then every lift of ϕ should be of the form

$$s \mapsto \begin{pmatrix} \lambda_1\tilde{\rho}_1(s) & \\ & \lambda_2\tilde{\rho}_2(s) \end{pmatrix} \in \mathrm{GSp}_4(\mathbb{C}) \rtimes \langle\sigma\rangle$$

with $\lambda_1^2 = \lambda_2^2$. It is known that $\tilde{\phi} = \omega_{E/F}\cdot\tilde{\phi}$ as parameters of ${}^LG_0$ since

$$\omega_{E/F}\cdot\tilde{\phi} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}\tilde{\phi}\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}^{-1}.$$

Thus there are two lifts $\tilde{\phi}_1 = \tilde{\rho}_1 + \tilde{\rho}_2$ and $\tilde{\phi}_2 = \tilde{\rho}_1\omega_{E/F} + \tilde{\rho}_2$ such that $\tilde{\phi}_i|_{WD_E} = \phi$.

If $\rho_1 = \chi^{-1} + \chi^s$, then the centralizer $Z_{\mathrm{GL}_2(\mathbb{C})}(\rho_1)$ is $\mathbb{C}^\times \times \mathbb{C}^\times$ or $\mathrm{GL}_2(\mathbb{C})$. Moreover,

$$\tilde{\rho}_1(s) = \begin{pmatrix} 1 & \\ & \chi(s^2) \end{pmatrix} \cdot \sigma.$$

In this case, $\tilde{\rho}_1 + \tilde{\rho}_2 \neq \tilde{\rho}_1 \omega_{E/F} + \tilde{\rho}_2$, which will be a different story if $\rho_1 = \rho_0$.

(ii) If $\rho_1 = \rho_0$ and ρ_2 is conjugate-symplectic, then $\tilde{\rho}_1(s) = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \cdot \sigma$. Because

$$\begin{pmatrix} \omega_{E/F}\tilde{\rho}_1 & \\ & \tilde{\rho}_2 \end{pmatrix} = \begin{pmatrix} a & \\ & 1 \end{pmatrix} \begin{pmatrix} \tilde{\rho}_1 & \\ & \tilde{\rho}_2 \end{pmatrix} \begin{pmatrix} a^{-1} & \\ & 1 \end{pmatrix},$$

where $a = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$, we have $\tilde{\phi}_1 = \tilde{\phi}_2$.

(iii) If $\rho_1^\vee = \rho_2^s$, then there exists an $A \in \mathrm{SL}_2(\mathbb{C})$ such that

$$A^{-1} \rho_1^\vee(t) A = \rho_2^s(t)$$

for $t \in WD_E$. Set

$$\tilde{\phi}(s) = \begin{pmatrix} & A\rho_2(s^2) \\ A^{-1} & \end{pmatrix} \cdot \sigma \in \mathrm{Sp}_4(\mathbb{C}) \rtimes \sigma.$$

Then $\tilde{\phi}(sts^{-1}) = \tilde{\phi}(s) \cdot \tilde{\phi}(t) \cdot \tilde{\phi}(s^{-1})$ and

$$[\tilde{\phi}(s)]^2 = \begin{pmatrix} & A\rho_2(s^2) \\ A^{-1} & \end{pmatrix}^2 = \begin{pmatrix} A\rho_2(s^2)A^{-1} & \\ & \rho_2(s^2) \end{pmatrix} = \begin{pmatrix} \rho_1^\vee(s^2) & \\ & \rho_2(s^2) \end{pmatrix} = \phi(s^2).$$

The last equality holds because $\det \rho_1$ is conjugate-orthogonal and so $\det \rho_1(s^2) = 1$.

Now we assume $\rho_1 = \rho_2$. According to ρ_1, we still separate it into 3 cases in a similar way.

(i) If ρ_1 is conjugate-symplectic but $\rho_1 \neq \rho_0$, then $\tilde{\phi}_1 = \tilde{\rho}_1 + \tilde{\rho}_1$ and $\tilde{\phi}_2 = \tilde{\rho}_1 + \tilde{\rho}_1 \omega_{E/F}$, where $\tilde{\rho}_1 : WD_F \to \mathrm{GL}_2(\mathbb{C}) \rtimes \langle \sigma \rangle$ satisfies $\tilde{\rho}_1|_{WD_E} = \rho_1$.

(ii) If $\rho_1 = \rho_0$, there is only one lift $\tilde{\phi} = \tilde{\rho}_1 + \tilde{\rho}_1$.

(iii) If ρ_1 is not conjugate-symplectic but conjugate-orthogonal, set

$$\tilde{\phi}(s) = \begin{pmatrix} & -A \\ A & \end{pmatrix} \cdot \sigma \in \mathrm{GSp}_4(\mathbb{C}) \rtimes \langle \sigma \rangle$$

where $A \in \mathrm{GL}_2(\mathbb{C})$ satisfies $A\rho_1^\vee(t)A^{-1} = \rho_1^s(t)$. Let us verify

$$\phi(s^2) = \tilde{\phi}(s^2) = \tilde{\phi}(s)^2,$$

i.e., $-A^2 \det(A)^{-1} = \rho_1(s^2)$.

- Suppose that ρ_1 is irreducible. Let $\langle -,-\rangle$ be the WD_E-equivariant bilinear form associated to $\rho_1 : WD_E \to \mathrm{GSp}(V, \langle -,-\rangle)$. Set

$$B(m,n) = \langle m, A^{-1}n\rangle$$

for $m, n \in V$. We have

$$B(\rho_1(t)m, \rho_1^s(t)n) = \langle \rho_1(t)m, \rho_1^\vee(t)A^{-1}n\rangle = B(m,n).$$

Note that ρ_1 is conjugate-orthogonal. By Schur's lemma, the conjugate-self-dual bilinear form B has sign 1, i.e.,

$$B(m, \rho_1(s^2)n) = B(n,m)$$

for all $m, n \in V$. Replacing m by Am, we have

$$\langle Am, A^{-1}\rho_1(s^2)n\rangle = \langle n, A^{-1}Am\rangle = \langle n, m\rangle = \langle m, -n\rangle.$$

Therefore $\det(A)\cdot A^{-2}\rho_1(s^2) = -1$. In this case,

$$\begin{aligned}\tilde{\phi}(s)\tilde{\phi}(t)\tilde{\phi}(s)^{-1} &= \begin{pmatrix} & -A \\ A & \end{pmatrix}\begin{pmatrix} \rho_1^\vee(t) & \\ & \rho_1^\vee(t)\end{pmatrix}\begin{pmatrix} & -A \\ A & \end{pmatrix}^{-1} \\ &= \begin{pmatrix} A\rho_1^\vee(t)A^{-1} & \\ & A\rho_1^\vee(t)A^{-1}\end{pmatrix} = \tilde{\phi}^s(t)\end{aligned}$$

for all $t \in WD_E$.

- If $\rho_1 = \mu_1 + \mu_2$ with $\mu_1\mu_2^s = \mathbf{1}$, then ρ_1 is conjugate-symplectic, which contradicts the assumption.
- If $\rho_1 = \mu_1 + \mu_2$ with $\mu_1 \neq \mu_2$ and $\mu_1|_{F^\times} = \mu_2|_{F^\times} = \mathbf{1}$, then $A = \begin{pmatrix} 1 & \\ & -1\end{pmatrix}$ and $A^2 = 1 = \rho_1(s^2)$.

6C4. *Nongeneric tempered.* Let τ be an irreducible nongeneric tempered representation of $\mathrm{GSp}_4(E)$ and $\tau = \theta(\pi_1 \boxtimes \pi_2)$, where each π_i is an irreducible representations of $D_E^\times(E)$. If the enhanced L-parameter of τ is (ϕ_τ, λ), where $\phi_\tau = \rho_1 + \rho_2$, $\rho_i = \phi_{\pi_i}$ and λ is the nontrivial character of the component group $\pi_0(Z_{\phi_\tau}/Z_{\mathrm{GSp}_4(\mathbb{C})})$, then

$$\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_{E/F}) = 0.$$

On the Galois side, if $\phi_\pi = \rho_1 + \rho_2$, then for arbitrary parameter $\tilde{\phi}$ satisfying $\tilde{\phi}|_{WD_E} = \phi_\tau$, the restricted representation $\lambda|_{\pi_0(Z(\tilde{\phi}))}$ does not contain the trivial character $\mathbf{1}$, i.e.,

$$m(\lambda, \tilde{\phi}) = 0.$$

Finally we can prove Theorem 1.2.

Proof of Theorem 1.2. It is obvious if τ is a nongeneric tempered representation of $\mathrm{GSp}_4(E)$. (See Section 6C4.) Since the Levi subgroup of a parabolic subgroup in GSp_4 are GL-type, [Prasad 2015, Lemma 10] implies that $\deg \Phi(\tilde{\phi}) = 1$ in our case. By the above discussions, we know that if τ is generic, then the multiplicity $\dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_{E/F})$ equals to the number of inequivalent lifts $|F(\phi_\tau)|$. $\square$

7. Proof of Theorem 1.3

This section focuses on the Prasad conjecture for PGSp_4. Let $\bar{\tau}$ be a representation of $\mathrm{PGSp}_4(E)$, i.e., a representation τ of $\mathrm{GSp}_4(E)$ with trivial central character. If the multiplicity

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}) = \dim \mathrm{Hom}_{\mathrm{GSp}_4(F)}(\tau, \omega_{E/F})$$

is nonzero, then we say $\bar{\tau}$ is $(\mathrm{PGSp}_4(F), \omega_{E/F})$-distinguished. Let $\mathrm{PGSp}_{1,1} = \mathrm{PGU}_2(D)$ be the pure inner form of PGSp_4 defined over F. Similarly,

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = \dim \mathrm{Hom}_{\mathrm{GSp}_{1,1}(F)}(\tau, \omega_{E/F})$$

for a representation τ of $\mathrm{GSp}_4(E)$ with trivial central character.

7A. *Notation.*

- $\bar{\tau}, \pi^{++}, \pi^{--}, \pi^{+}$ and π^{-} are representations of $\mathrm{PGSp}_4(E)$.
- $s \in W_F \setminus W_E$ and $\phi_\tau^s(t) = \phi_\tau(sts^{-1})$ for $t \in WD_E$.
- $S_\phi = \pi_0(Z(\phi))$ is the component group associated to ϕ.
- $\tilde{\phi} : WD_F \to \mathrm{Sp}_4(\mathbb{C})$ and $\tilde{\phi}_i$ are Langlands parameters of $\mathrm{PGSp}_4(F)$.
- $\mathcal{C}_i$ is a coset of $A_G(\tilde{\phi}_i)$ in $H^1(F, \mathrm{PGSp}_4)$ and $1_{\mathcal{C}_i}$ denotes its characteristic function.
- $\mathrm{PGSp}_{1,1}$ (resp. $PD^\times$) is the pure inner form of PGSp_4 (resp. PGL_2) defined over F.

7B. *The Prasad conjecture for* $\mathbf{PGL_2}$.

If $G = \mathrm{PGL}_2$, then $\chi_G = \omega_{E/F}$ and $G^{op} = \mathrm{PGL}_2$.

Theorem 7.2.1. *Let $\bar{\pi}$ be a generic irreducible representation of* $\mathrm{PGL}_2(E)$. *Then the following are equivalent*:

(i) $\dim \mathrm{Hom}_{\mathrm{PGL}_2(F)}(\bar{\pi}, \omega_{E/F}) = 1$.

(ii) *The Langlands parameter $\phi_{\bar{\pi}}$ is conjugate-symplectic.*

(iii) *There exists a parameter $\tilde{\phi} : WD_F \to \mathrm{SL}_2(\mathbb{C})$ such that $\tilde{\phi}|_{WD_E} = \phi_{\bar{\pi}}$.*

(iv) *$\bar{\pi}$ is $(PD^\times(F), \omega_{E/F})$-distinguished or $\bar{\pi} = \pi(\chi_E, \chi_E^{-1})$ with $\chi_E|_{F^\times} = \omega_{E/F}$ and $\chi_E^2 \neq \mathbf{1}$.*

Proof. See [Gan and Raghuram 2013, Theorem 6.2; Lu 2017b, Main Theorem (local)]. □

7C. *The Prasad conjecture for* $\mathbf{PGSp_4}$.

Recall that if $G = \mathrm{PGSp}_4$, then $\widehat{G} = \mathrm{Spin}_5(\mathbb{C}) \cong \mathrm{Sp}_4(\mathbb{C})$, $G^{op} = \mathrm{PGSp}_4$ and $\chi_G = \omega_{E/F}$. Let $\bar{\tau}$ be a representation of $\mathrm{PGSp}_4(E)$ with enhanced L-parameter $(\phi_{\bar{\tau}}, \lambda_{\bar{\tau}})$. Assume that the L-packet $\Pi_{\phi_{\bar{\tau}}}$ is generic. The Prasad conjecture for PGSp_4 implies the following:

P(i) If $\bar{\tau}$ is $(\mathrm{PGSp}_4(F), \omega_{E/F})$-distinguished, then

- $\Pi_{\phi_{\bar{\tau}}^s} = \Pi_{\phi_{\bar{\tau}}^\vee}$, an equality of L-packets and
- $\phi_{\bar{\tau}} = \tilde{\phi}|_{WD_E}$ for some parameter $\tilde{\phi} : WD_F \to \mathrm{Sp}_4(\mathbb{C})$.

P(ii) If $\bar{\tau}$ is generic and there exists $\tilde{\phi}: WD_F \to \mathrm{Sp}_4(\mathbb{C})$ such that $\tilde{\phi}|_{WD_E} = \phi_{\bar{\tau}}$, then we have that $\bar{\tau}$ is $(\mathrm{PGSp}_4(F), \omega_{E/F})$-distinguished.

P(iii) Assume that $\phi_{\bar{\tau}} = \tilde{\phi}|_{WD_E}$ for some parameter $\tilde{\phi}: WD_F \to \mathrm{Sp}_4(\mathbb{C})$. If $\bar{\tau}$ is a discrete series representation, then we set

$$F(\phi_{\bar{\tau}}) = \{\tilde{\phi} : \tilde{\phi}|_{WD_E} = \phi_{\bar{\tau}}\} = \bigsqcup_i \mathcal{O}(\tilde{\phi}_i),$$

where $\mathcal{O}(\tilde{\phi}_i) = \{\tilde{\phi}_i, \omega_{E/F} \cdot \tilde{\phi}_i\}$ which may be a singleton. Given a parameter $\tilde{\phi}_i : W_F \to \mathrm{Sp}_4(\mathbb{C})$ with $\phi_{\bar{\tau}}$ its restriction to WD_E and $\tilde{\phi}_i \cdot \omega_{E/F} = \tilde{\phi}_i$, there exists an element $g_i \in Z(\phi_{\bar{\tau}})$ such that

$$(\tilde{\phi}_i \cdot \omega_{E/F})(x) = g_i \tilde{\phi}_i(x) g_i^{-1}$$

for all $x \in WD_F$ and so g_i normalizes $Z(\tilde{\phi}_i)$. Then $\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}) \neq 0$ if $\lambda_{\bar{\tau}}(g_i) = 1$ and $\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) \neq 0$ if $\lambda_{\bar{\tau}}(g_i) = -1$. In this case, $A_G(\tilde{\phi}_i) \subset H^1(F, \mathrm{PGSp}_4)$ is trivial and

$$\mathcal{C}_i = \begin{cases} \{\mathrm{PGSp}_4\} & \text{if } \lambda_{\bar{\tau}}(g_i) = 1, \\ \{\mathrm{PGSp}_{1,1}\} & \text{if } \lambda_{\bar{\tau}}(g_i) = -1. \end{cases}$$

If $\tilde{\phi}_i \neq \tilde{\phi}_i \cdot \omega_{E/F}$, then $A_G(\tilde{\phi}_i) = H^1(F, \mathrm{PGSp}_4)$ and $\mathcal{C}_i = \{\mathrm{PGSp}_4, \mathrm{PGSp}_{1,1}\}$. Set G_α to be PGSp_4 or $\mathrm{PGSp}_{1,1}$. Then

$$\dim \mathrm{Hom}_{G_\alpha(F)}(\bar{\tau}, \omega_{E/F}) = \sum_i m(\lambda_{\bar{\tau}}, \tilde{\phi}_i) 1_{\mathcal{C}_i}(G_\alpha) / d_0(\tilde{\phi}_i),$$

where $m(\lambda_{\bar{\tau}}, \tilde{\phi}_i)$ is the multiplicity of the trivial representation contained in the restricted representation $\lambda_{\bar{\tau}}|_{\pi_0(Z(\tilde{\phi}_i))}$.

P(iv) If $\Pi_{\phi_{\bar{\tau}}}$ is generic, then we have (1-3), i.e.,

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}) + \dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = \sum_{\varphi \in F(\phi_{\bar{\tau}})} m(\lambda_{\bar{\tau}}, \varphi) \cdot \frac{\deg \Phi(\varphi)}{d_0(\varphi)}.$$

Let us start to verify the Langlands functoriality lift in the Prasad conjecture for PGSp_4, i.e., part **P**(i) and **P**(ii) listed above. Part **P**(iii) is the same with Theorem 1.3. Part **P**(iv) will be studied in detail in the next subsection.

Theorem 7.3.1. *Let $\bar{\tau}$ be a generic representation of* $\mathrm{PGSp}_4(E)$. *It is* $(\mathrm{PGSp}_4(F), \omega_{E/F})$*-distinguished if and only if there exists a parameter* $\tilde{\phi}: WD_F \to \mathrm{Sp}_4(\mathbb{C})$ *such that* $\tilde{\phi}|_{WD_E} = \phi_{\bar{\tau}}$.

Proof. Assume that $\tau = \theta(\Pi \boxtimes \chi)$ with $\chi = \mathbf{1}$, i.e., $\omega_\tau = \mathbf{1}$. Fix $s \in W_F \setminus W_E$.

(i) If $\bar{\tau}$ is $(\mathrm{PGSp}_4(F), \omega_{E/F})$-distinguished, then ϕ_Π is conjugate-symplectic and so $\Pi_{\phi^s_{\bar{\tau}}} = \Pi_{\phi^\vee_{\bar{\tau}}} = \Pi_{\phi_{\bar{\tau}}}$. If ϕ_Π is irreducible, then we can repeat the process in Section 6C1 to obtain that there exists a parameter $\tilde{\phi}: WD_F \to \mathrm{Sp}_4(\mathbb{C})$ such that $\tilde{\phi}|_{WD_E} = \phi_{\bar{\tau}}$. If $\phi_\Pi = \rho_1 \oplus \rho_2$ is reducible and ρ_1 is irreducible, then

$$\rho_1 \oplus \rho_2 = \rho_1^\vee \oplus \rho_2^\vee = \rho_1^s \oplus \rho_2^s$$

and either $\rho_1^s = \rho_2^\vee$ or both ρ_1 and ρ_2 are conjugate-symplectic.

- If $\rho_1^s = \rho_2^\vee$, then there are two subcases. If $\rho_2^\vee = \rho_2$, then $\rho_1^s = \rho_2$. Set $\tilde{\phi} = \mathrm{Ind}_{WD_E}^{WD_F}\rho_1$ if $\rho_1 \neq \rho_2$. If $\rho_1 = \rho_2 = \rho_2^\vee$, then $\rho_1^s = \rho_1$ and so there exists a parameter $\tilde{\rho}_1 : WD_F \to \mathrm{GL}_2(\mathbb{C})$ such that $\tilde{\rho}_1|_{WD_E} = \rho_1$. Set $\tilde{\phi} = \tilde{\rho}_1 \oplus \tilde{\rho}_1^\vee$. If $\rho_2^\vee \neq \rho_2$, then $\rho_2^\vee = \rho_1$. Thus $\rho_1^s = \rho_1$ and $\tilde{\phi} = \tilde{\rho}_1 \oplus \tilde{\rho}_1^\vee$.
- If both ρ_1 and ρ_2 are conjugate-symplectic, then

$$\tilde{\phi} = \begin{cases} \mathrm{Ind}_{WD_E}^{WD_F}\rho_1 & \text{if } \rho_1^s = \rho_2 \neq \rho_1, \\ \tilde{\rho}_1 \oplus \tilde{\rho}_1^\vee & \text{if } \rho_1^s = \rho_1. \end{cases}$$

If neither ρ_1 nor ρ_2 is irreducible, then $\phi_{\tilde{\tau}}$ belongs to the endoscopic case. Thanks to Theorem 4.4.9(ii), either $\rho_1^s = \rho_2^\vee$ or both ρ_1 and ρ_2 are conjugate-symplectic. The argument is similar and we omit it here. Therefore, there exists $\tilde{\phi} : WD_F \to \mathrm{Sp}_4(\mathbb{C})$ such that $\tilde{\phi}|_{WD_E} = \phi_{\tilde{\tau}}$.

(ii) Conversely, if there exists $\tilde{\phi} : WD_F \to \mathrm{Sp}_4(\mathbb{C})$ such that $\tilde{\phi}|_{WD_E} = \phi_{\tilde{\tau}}$, then it suffices to show that ϕ_Π is conjugate-symplectic. (See Lemma 6.3.1.) The nongeneric member in the L-packet $\Pi_{\phi_{\tilde{\tau}}}$ is not $(\mathrm{GSp}_4(F), \omega_{E/F})$-distinguished due to Theorem 4.4.9(i) if $|\Pi_{\phi_{\tilde{\tau}}}| = 2$. Assume that

$$\phi_{\tilde{\tau}} : WD_E \to \mathrm{Sp}(V, \langle -, - \rangle) = \mathrm{Sp}_4(\mathbb{C}) \quad \text{and} \quad \phi_\Pi = i \circ \phi_{\tilde{\tau}} : WD_E \to \mathrm{GL}(V),$$

where $i : \mathrm{Sp}_4(\mathbb{C}) \to \mathrm{GL}(V)$ is the embedding between the L-groups. Then we set

$$B(m, n) = \langle m, \tilde{\phi}(s)^{-1} n \rangle$$

for $m, n \in V$. It is easy to check that $B(\phi_\Pi(t)m, \phi_\Pi^s(t)n) = B(m, n)$ and

$$B(m, \phi_\Pi(s^2)n) = \langle m, \tilde{\phi}(s)n \rangle = -\langle \tilde{\phi}(s)n, m \rangle = -\langle n, \tilde{\phi}(s)^{-1}m \rangle = -B(n, m).$$

Therefore, the bilinear form B on V implies that ϕ_Π is conjugate-symplectic.

We have finished the proof. □

However, in order to verify (1-3), we will need many more results from Theorems 4.4.9 and 5.3.1. We will give the full detail in the next subsection.

7D. ***Proof of Theorem 1.3.*** This subsection focuses on the proof of Theorem 1.3. Before we give the proof of Theorem 1.3, we will use the results in Theorems 4.4.9 and 5.3.1 to study the equality (1-3) in detail. Then Theorem 1.3 will follow automatically. According to the Langlands parameter $\phi_{\tilde{\tau}}$, we divide them into three cases:

- the endoscopic case,
- the discrete series but nonendoscopic case and
- $\phi_{\tilde{\tau}} = \rho + \rho\nu$ with $\nu \neq \mathbf{1}$ and $\nu \det \rho = \mathbf{1}$.

Set $S_\phi = \pi_0(Z(\phi))$ to be the component group. We identify the characters of W_F and the characters of $F^\times$ via the local class field theory.

7D1. *Endoscopic case.* Given $\phi_{\tilde{\tau}} = \phi_1 \oplus \phi_2$, there are two cases: $\phi_1 = \phi_2$ and $\phi_1 \neq \phi_2$.

(A) If $\phi_1 = \phi_2 = \rho$ are irreducible, then the L-packet $\Pi_{\phi_{\tilde{\tau}}}$ equals $\{\pi^+, \pi^-\}$ and $S_{\phi_{\tilde{\tau}}}$ equals $\mathbb{Z}/2\mathbb{Z}$, where π^- (resp. π^+) is a nongeneric (resp. generic) representation of $\mathrm{PGSp}_4(E)$. There are two subcases:

(A1) If ρ is conjugate-orthogonal, then

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^+, \omega_{E/F}) = 0 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^-, \omega_{E/F})$$

and

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^-, \omega_{E/F}) = 1 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^+, \omega_{E/F}).$$

On the Galois side, there is only one extension $\tilde{\phi} = \bar{\rho} \oplus \bar{\rho} \cdot \omega_{E/F}$ with

$$\deg \Phi(\tilde{\phi}) = 2 \quad \text{and} \quad S_{\tilde{\phi}} = \{\mathbf{1}\} \to S_{\phi_{\tilde{\tau}}},$$

where $\bar{\rho} : WD_F \to \mathrm{GL}_2(\mathbb{C}) \times W_F$ with $\det \bar{\rho} = \omega_{E/F}$. Note that $\tilde{\phi} = \tilde{\phi} \cdot \omega_{E/F}$. Then π^+ supports a period on the trivial pure inner form and π^- supports a period on a nontrivial pure inner form.

(A2) If ρ is conjugate-symplectic, then

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^-, \omega_{E/F}) = 0 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^-, \omega_{E/F})$$

and

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^+, \omega_{E/F}) = 1, \quad \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^+, \omega_{E/F}) = 2.$$

In this case, ρ has two extensions $\bar{\rho}$ and $\bar{\rho} \cdot \omega_{E/F}$, where $\bar{\rho} : WD_F \to \mathrm{SL}_2(\mathbb{C})$. There are three choices for the extension $\tilde{\phi} : WD_F \to \mathrm{Sp}_4(\mathbb{C})$ with $\deg \Phi(\tilde{\phi}) = 1$:

- $\tilde{\phi}^{++} = \bar{\rho} \oplus \bar{\rho}$ with $S_{\tilde{\phi}^{++}} = \mathbb{Z}/2\mathbb{Z} \cong S_{\phi_{\tilde{\tau}}}$;
- $\tilde{\phi}^{+-} = \bar{\rho} \oplus \bar{\rho} \cdot \omega_{E/F}$ with $S_{\tilde{\phi}^{+-}} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \to S_{\phi_{\tilde{\tau}}}$(sum map);
- $\tilde{\phi}^{--} = \bar{\rho} \cdot \omega_{E/F} \oplus \bar{\rho} \cdot \omega_{E/F}$ with $S_{\tilde{\phi}^{--}} = \mathbb{Z}/2\mathbb{Z} \cong S_{\phi_{\tilde{\tau}}}$.

The parameters $\tilde{\phi}^{++}$ and ϕ^{--} are in the same orbit under the twisting by $\omega_{E/F}$, which corresponds to both pure inner forms. The parameter $\tilde{\phi}^{+-}$ is fixed under twisting by $\omega_{E/F}$, which supports a period on the trivial pure inner form.

(A3) If ρ is not conjugate-self-dual, then both the Galois side and the automorphic side are 0.

(B) If $\phi_1 \neq \phi_2$ are both irreducible, then the L-packet of PGSp_4 is $\Pi_{\phi_{\tilde{\tau}}} = \{\pi^{++}, \pi^{--}\}$ and

$$S_{\phi_{\tilde{\tau}}} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

(B1) If ϕ_1 and ϕ_2 both extend to L-parameters of $\mathrm{PGL}_2(F)$, i.e., both are conjugate-symplectic, then one has $\phi_1^s \neq \phi_2$,

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^{++}, \omega_{E/F}) = 2 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^{++}, \omega_{E/F})$$

and

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^{--}, \omega_{E/F}) = 0 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^{--}, \omega_{E/F}).$$

On the Galois side, there are also four ways of extending $\phi_{\bar{\tau}}$. For each such extension $\tilde{\phi}$, one has $\deg\Phi(\tilde{\phi})=1$ and the equality of component group

$$S_{\tilde{\phi}}=S_{\phi_{\bar{\tau}}}=\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}.$$

Therefore only the representation π^{++} in the L-packet can support a period. And there are 2 orbits in $F(\phi_{\bar{\tau}})$ under twisting by $\omega_{E/F}$, each of size 2.

(B2) If ϕ_1 and ϕ_2 do not extend to L-parameters of $\mathrm{PGL}_2(F)$, but $\phi_1^s=\phi_2=\phi_2^\vee$, then

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^{++},\omega_{E/F})=0=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^{--},\omega_{E/F})$$

and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\pi^{--},\omega_{E/F})=1=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\pi^{++},\omega_{E/F})$$

There is a unique way of extending $\phi_{\bar{\tau}}=\phi_1\oplus\phi_2$ to $\tilde{\phi}:WD_F\to\mathrm{Sp}_4(\mathbb{C})$. Namely, $\tilde{\phi}=\mathrm{Ind}_{WD_E}^{WD_F}\phi_1$ is an irreducible 4-dimensional symplectic representation, with a component group

$$S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}\hookrightarrow S_{\phi_{\bar{\tau}}}(\text{diagonal embedding}).$$

And $S_{\phi_{\bar{\tau}}}^{\mathrm{Gal}(E/F)}=S_{\tilde{\phi}}$. Thus π^{++} supports a period on the trivial pure inner form and π^{--} supports a period on the nontrivial pure inner form.

(C) If $\phi_1=\chi_1\oplus\chi_1^{-1}$ is reducible, then there is only one element in the L-packet, i.e., $|\Pi_{\phi_{\bar{\tau}}}|=1$. There are two cases: $\phi_1=\phi_2$ and $\phi_1\neq\phi_2$.

(C1) If $\phi_1=\phi_2$, there are three subcases.

(C1.i) If $\chi_1=\chi_1^s=\chi_F|_{W_E}$, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=2=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F}).$$

- If $\chi_F^2\neq\omega_{E/F}$, then there are two ways to extend L-parameters of $\mathrm{PGL}_2(F)$, denoted by $\bar{\rho}$ and $\bar{\rho}\cdot\omega_{E/F}$. Thus there are 3 ways of extending $\phi_{\bar{\tau}}$, which are $\tilde{\phi}^{++}$, $\tilde{\phi}^{--}$ and $\tilde{\phi}^{+-}$. Moreover, $\deg\Phi(\tilde{\phi}^{++})=1=\deg\Phi(\tilde{\phi}^{--})$ and $\deg\Phi(\tilde{\phi}^{+-})=2$.
- If $\chi_F^2=\omega_{E/F}$, then there is only one way to extend $\phi_{\bar{\tau}}$. Denote it by $\tilde{\phi}$. Then

$$\deg\Phi(\tilde{\phi})=4.$$

(C1.ii) If $\chi_1\neq\chi_1^{-1}$ but $\chi_1|_{F^\times}=\omega_{E/F}$, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=0\text{ and }\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F})=1.$$

There is only one way to extend ϕ_1, denoted by

$$\bar{\rho}=\mathrm{Ind}_{WD_E}^{WD_F}\chi_1:WD_F\to\mathrm{SL}_2(\mathbb{C}).$$

Then $\tilde{\phi}=\bar{\rho}\oplus\bar{\rho}$ with $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}$ and $\deg\Phi(\tilde{\phi})=1$. Note that $\tilde{\phi}\cdot\omega_{E/F}=\tilde{\phi}$. Then $\tilde{\phi}$ supports a period on the trivial pure inner form.

(C1.iii) If $\chi_1\neq\chi_1^{-1}$ but $\chi_1|_{F^\times}=\mathbf{1}$, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=0 \text{ and } \dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F})=1.$$

On the Galois side, there is only one choice $\tilde{\phi}=\bar{\rho}\oplus\bar{\rho}$ and $S_{\tilde{\phi}}=\mathbf{1}$, where

$$\bar{\rho}=\mathrm{Ind}_{WD_E}^{WD_F}\chi_1 : WD_F\to\mathrm{GL}_2(\mathbb{C})$$

with $\det\rho=\omega_{E/F}$. Since $\tilde{\phi}=\tilde{\phi}\cdot\omega_{E/F}$, it picks up only the trivial pure inner form.

(C2) If $\phi_1\neq\phi_2$, there are several subcases:

(C2.i) If $\chi_1=\chi_1^s=\chi_F|_{W_E}$ and ϕ_2 is irreducible and conjugate-symplectic, then $S_{\phi_{\bar{\tau}}}=\mathbb{Z}/2\mathbb{Z}$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=2=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F}).$$

- If $\chi_F^2\neq\omega_{E/F}$, then there are four ways of extending $\phi_{\bar{\tau}}$ and for each such extension $\tilde{\phi}$, one has $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}\cong S_{\phi_{\bar{\tau}}}$. There are two orbits under the twisting by $\omega_{E/F}$, each of size 2.
- If $\chi_F^2=\omega_{E/F}$, then there are two ways of extending $\phi_{\bar{\tau}}$. For each such extension $\tilde{\phi}$, one has $\deg\Phi(\tilde{\phi})=2$. There is one orbit under the twisting by $\omega_{E/F}$.

In this case, the identity

$$\dim\mathrm{Hom}_{G_\alpha(F)}(\bar{\tau},\chi_G)=\sum_i m(\lambda,\tilde{\phi}_i)\mathbf{1}_{\mathcal{C}_i}(G_\alpha)\cdot\frac{\deg\Phi(\tilde{\phi}_i)}{d_0(\tilde{\phi}_i)} \tag{7-1}$$

holds for $G_\alpha=\mathrm{PGSp}_4$ and $\mathrm{PGSp}_{1,1}$.

(C2.ii) If $\chi_1=\chi_1^s=\chi_F|_{W_E}$ and $\chi_2=\chi_2^s=\chi'_F|_{W_E}$, where $\phi_2=\chi_2\oplus\chi_2^{-1}$, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=2=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F}).$$

- If neither χ_F^2 nor ${\chi'_F}^2$ equals $\omega_{E/F}$, then there are four ways of extending $\phi_{\bar{\tau}}$. There are two orbits under the twisting by $\omega_{E/F}$, each of size 2.
- If $\chi_F^2=\omega_{E/F}$ and $\chi_F'^2\neq\omega_{E/F}$, then there are two ways to extend $\phi_{\bar{\tau}}$ and for each such extension $\tilde{\phi}$, one has $S_{\tilde{\phi}}=1=S_{\phi_{\bar{\tau}}}$ and $\deg\Phi(\tilde{\phi})=2$. There is one orbit under the twisting by $\omega_{E/F}$, which corresponds to both pure inner forms.
- If $\chi_F^2=\chi_F'^2=\omega_{E/F}$, then there is only one way to extend $\phi_{\bar{\tau}}$. For this extension $\tilde{\phi}$, one has $\deg\Phi(\tilde{\phi})=4$.

(C2.iii) If $\chi_1\neq\chi_1^{-1}$ but χ_1 is conjugate-symplectic, and ϕ_2 is irreducible and conjugate-symplectic, then $S_{\phi_{\bar{\tau}}}=\mathbb{Z}/2\mathbb{Z}$ and

$$\dim\mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau},\omega_{E/F})=1=\dim\mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau},\omega_{E/F}).$$

There are two extensions $\tilde{\phi}=\bar{\rho}_1\oplus\bar{\rho}_2$ or $\bar{\rho}_1\oplus\bar{\rho}_2\omega_{E/F}$ with $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, where $\bar{\rho}_i : WD_F\to \mathrm{SL}_2(\mathbb{C})$ satisfies $\bar{\rho}_i|_{WD_E}=\phi_i$. Here the map $S_{\tilde{\phi}}\to S_{\phi_{\bar{\tau}}}$ is given by

$$(x, y)\mapsto x+y.$$

There is one orbit under the twisting by $\omega_{E/F}$, which corresponds to both pure inner forms.

(C2.iv) If $\chi_1\neq\chi_1^{-1}$ but χ_1 is conjugate-symplectic, and $\chi_2=\chi_2^s=\chi'_F|_{W_E}$ where $\phi_2=\chi_2\oplus\chi_2^{-1}$, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F})=1=\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}).$$

- If $\chi_F'^2\neq\omega_{E/F}$, then there are two ways to extend $\phi_{\bar{\tau}}$. Set $\tilde{\phi}=\bar{\rho}_1\oplus\bar{\rho}_2$ or $\bar{\rho}_1\oplus\bar{\rho}_2\omega_{E/F}$ with $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}$. There is one orbit under the twisting by $\omega_{E/F}$, which corresponds to both pure inner forms.
- If $\chi_F'^2=\omega_{E/F}$, there is one way to extend $\phi_{\bar{\tau}}$. Set $\tilde{\phi}=\bar{\rho}_1\oplus\chi'_F\oplus\chi'_F\omega_{E/F}$, and

$$\deg\Phi(\tilde{\phi})=2.$$

Note that the identity (7-1) fails in this case while the identity (1-3) still holds.

(C2.v) If ϕ_1 and ϕ_2 are reducible and four different characters χ_1, χ_1^{-1}, χ_2 and χ_2^{-1} satisfy

$$\chi_1|_{F^\times}=\omega_{E/F}=\chi_2|_{F^\times},$$

then $S_{\phi_{\bar{\tau}}}$ is trivial,

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F})=0,$$

and $\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F})=1$. There is only one extension $\tilde{\phi}=\bar{\rho}_1\oplus\bar{\rho}_2$ with $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$. Since $\tilde{\phi}=\tilde{\phi}\cdot\omega_{E/F}$, it picks up the trivial pure inner form.

(C2.vi) If $\phi_1^s=\phi_2^\vee=\phi_2$ and ϕ_1 is not conjugate-symplectic, then $S_{\phi_{\bar{\tau}}}=1$ and

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F})=0, \quad \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F})=1.$$

There is only one extension

$$\tilde{\phi}=\mathrm{Ind}_{WD_E}^{WD_F}\phi_1 : WD_F\to \mathrm{Sp}_4(\mathbb{C})$$

with the component group $S_{\tilde{\phi}}=\mathbb{Z}/2\mathbb{Z}$. Since $\tilde{\phi}=\tilde{\phi}\cdot\omega_{E/F}$, it picks up the trivial pure inner form.

It is easy to check that the identity (1-3) holds when $\Pi_{\phi_{\bar{\tau}}}$ is generic, i.e.,

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F})+\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F})=\sum_{\tilde{\phi}\in F(\phi_{\bar{\tau}})} m(\lambda, \tilde{\phi})\cdot\frac{\deg\Phi(\tilde{\phi})}{d_0(\tilde{\phi})}.$$

7D2. *Discrete and nonendoscopic case.* Assume that $\phi_{\bar{\tau}}$ is irreducible and so $\Pi_{\phi_{\bar{\tau}}}$ is a singleton. Given a parameter $\phi_{\bar{\tau}}$, which is nonendoscopic, the theta lift $\Theta_4^+(\tau)$ from $\mathrm{PGSp}_4(E)$ to $\mathrm{PGSO}_{2,2}$(E) is zero.

If $\phi_{\bar{\tau}}$ is conjugate-symplectic, then

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = 1 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}).$$

There are two extensions $\tilde{\phi}$ and $\tilde{\phi} \cdot \omega_{E/F}$ with a component group $S_{\tilde{\phi}} = S_{\phi_{\bar{\tau}}} = \mathbb{Z}/2\mathbb{Z}$. There is one orbit under the twisting by $\omega_{E/F}$, which corresponds to both pure inner forms.

7D3. *Generic but neither discrete nor endoscopic case.* If $\phi_{\bar{\tau}} = \rho \oplus \rho\nu$, $\det \rho = \nu^{-1} \neq \mathbf{1}$, then $S_{\phi_{\bar{\tau}}} = 1$. There are two cases:

- If $\phi_{\bar{\tau}}$ is conjugate-symplectic and $\rho^s = \rho$, then

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = 1 = \dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}).$$

  There are two extensions $\tilde{\phi} = \tilde{\rho} + \tilde{\rho}^\vee$ and $\tilde{\phi} \cdot \omega_{E/F}$ where $\tilde{\rho} : WD_F \to \mathrm{GL}_2(\mathbb{C})$ satisfies $\tilde{\rho}|_{WD_E} = \rho$.

- If $\phi_{\bar{\tau}}$ is conjugate-symplectic and $\rho^s \neq \rho$, then

$$\dim \mathrm{Hom}_{\mathrm{PGSp}_4(F)}(\bar{\tau}, \omega_{E/F}) = 1 \text{ and } \dim \mathrm{Hom}_{\mathrm{PGSp}_{1,1}(F)}(\bar{\tau}, \omega_{E/F}) = 0.$$

  There is only one extension $\tilde{\phi} = \mathrm{Ind}_{WD_E}^{WD_F} \rho$ such that $\tilde{\phi}|_{WD_E} = \phi_{\bar{\tau}}$.

Proof of Theorem 1.3. It follows from the discussions in the endoscopic cases (B)enumz in Section 7D1 and the discrete and nonendoscopic case in Section 7D2. $\square$

7E. ***Further discussion.*** Let E be a quadratic extension over a nonarchimedean local field F. Let $\boldsymbol{G}$ be a quasisplit reductive group defined over F. Let τ be an irreducible representation of $\boldsymbol{G}(E)$ with an enhanced L-parameter (ϕ_τ, λ). Assume that $F(\phi_\tau) = \sqcup_i \mathcal{O}(\tilde{\phi}_i)$ where $\tilde{\phi}_i|_{WD_E} = \phi_\tau$.

If for each orbit $\mathcal{O}(\tilde{\phi}_i)$, the coset $\mathcal{C}_i \subset H^1(W_F, \boldsymbol{G})$ contains all pure inner forms satisfying $G_\alpha(E) = \boldsymbol{G}(E)$, then ϕ_τ is called a "full" L-parameter of $\boldsymbol{G}(E)$, in which case $1_{\mathcal{C}_i}(G_\alpha) \equiv 1$ in (7-1).

Assume that τ belongs to a generic L-packet with Langlands parameter $\phi_\tau : WD_E \to {}^L\boldsymbol{G}$ and that ϕ_τ is "full". Then there is a conjectural identity

$$\dim \mathrm{Hom}_{G_\alpha}(\tau, \chi_{\boldsymbol{G}}) = \sum_i m(\lambda, \tilde{\phi}_i) \cdot \frac{\deg \Phi(\tilde{\phi}_i)}{d_0(\tilde{\phi}_i)} \tag{7-2}$$

for any pure inner form $G_\alpha \in H^1(W_F, \boldsymbol{G})$ satisfying $G_\alpha(E) = \boldsymbol{G}(E)$.

If $H^1(W_F, \boldsymbol{G})$ is trivial, then any L-parameter ϕ_τ is "full". So the conjectural identity (7-2) holds for $\boldsymbol{G} = \mathrm{GL}_2$. In fact, it holds for $\boldsymbol{G} = \mathrm{PGL}_2$ as well.

Acknowledgments

This paper contains part of the author's Ph.D. thesis [Lu 2017a]. He is grateful to Wee Teck Gan for the guidance and numerous discussions when he was studying in Singapore. He would like to thank Dipendra Prasad for proposing this conjecture and fruitful discussions. He wants to thank Dmitry Gourevitch and Lei Zhang for useful comments as well. He is also grateful to the anonymous referee for a careful reading of the manuscript and numerous suggestions which greatly improved the exposition of this paper. This work was partially supported by the ERC, StG grant number 637912.

References

[Adler and Prasad 2006] J. D. Adler and D. Prasad, "On certain multiplicity one theorems", *Israel J. Math.* **153** (2006), 221–245. MR Zbl

[Adler and Prasad 2012] J. D. Adler and D. Prasad, "Extensions of representations of p-adic groups", *Nagoya Math. J.* **208** (2012), 171–199. MR Zbl

[Arthur 2013] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, Amer. Math. Soc. Colloq. Publ. **61**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Atobe and Gan 2017] H. Atobe and W. T. Gan, "Local theta correspondence of tempered representations and Langlands parameters", *Invent. Math.* **210**:2 (2017), 341–415. MR Zbl

[Beuzart-Plessis 2018] R. Beuzart-Plessis, "On distinguished square-integrable representations for Galois pairs and a conjecture of Prasad", *Invent. Math.* **214**:1 (2018), 437–521. MR Zbl

[Casselman and Miličić 1982] W. Casselman and D. Miličić, "Asymptotic behavior of matrix coefficients of admissible representations", *Duke Math. J.* **49**:4 (1982), 869–930. MR Zbl

[Flicker 1991] Y. Z. Flicker, "On distinguished representations", *J. Reine Angew. Math.* **418** (1991), 139–172. MR Zbl

[Flicker and Hakim 1994] Y. Z. Flicker and J. L. Hakim, "Quaternionic distinguished representations", *Amer. J. Math.* **116**:3 (1994), 683–736. MR Zbl

[Gan 2011] W. T. Gan, "The Shimura correspondence à la Waldspurger", lecture notes, 2011, available at https://tinyurl.com/ganshimura.

[Gan 2019] W. T. Gan, "Periods and theta correspondence", pp. 113–132 in *Representations of reductive groups* (Jerusalem/Rehovot, Israel, 2017), edited by A. Aizenbud et al., Proc. Sympos. Pure Math. **101**, Amer. Math. Soc., Providence, RI, 2019. MR

[Gan and Ichino 2011] W. T. Gan and A. Ichino, "On endoscopy and the refined Gross–Prasad conjecture for $(\mathrm{SO}_5, \mathrm{SO}_4)$", *J. Inst. Math. Jussieu* **10**:2 (2011), 235–324. MR Zbl

[Gan and Ichino 2014] W. T. Gan and A. Ichino, "Formal degrees and local theta correspondence", *Invent. Math.* **195**:3 (2014), 509–672. MR Zbl

[Gan and Raghuram 2013] W. T. Gan and A. Raghuram, "Arithmeticity for periods of automorphic forms", pp. 187–229 in *Automorphic representations and L-functions* (Mumbai, 2012), edited by D. Prasad et al., Tata Inst. Fundam. Res. Stud. Math. **22**, Tata Inst. Fund. Res., Mumbai, 2013. MR Zbl

[Gan and Sun 2017] W. T. Gan and B. Sun, "The Howe duality conjecture: quaternionic case", pp. 175–192 in *Representation theory, number theory, and invariant theory*, edited by J. Cogdell et al., Progr. Math. **323**, Birkhäuser, Cham, 2017. MR Zbl

[Gan and Takeda 2011a] W. T. Gan and S. Takeda, "The local Langlands conjecture for GSp(4)", *Ann. of Math.* (2) **173**:3 (2011), 1841–1882. MR Zbl

[Gan and Takeda 2011b] W. T. Gan and S. Takeda, "Theta correspondences for GSp(4)", *Represent. Theory* **15** (2011), 670–718. MR Zbl

[Gan and Takeda 2016a] W. T. Gan and S. Takeda, "On the Howe duality conjecture in classical theta correspondence", pp. 105–117 in *Advances in the theory of automorphic forms and their L-functions*, edited by D. Jiang et al., Contemp. Math. **664**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[Gan and Takeda 2016b] W. T. Gan and S. Takeda, “A proof of the Howe duality conjecture”, *J. Amer. Math. Soc.* **29**:2 (2016), 473–493. MR Zbl

[Gan and Tantono 2014] W. T. Gan and W. Tantono, “The local Langlands conjecture for GSp(4), II: The case of inner forms”, *Amer. J. Math.* **136**:3 (2014), 761–805. MR Zbl

[Gan et al. 2012] W. T. Gan, B. H. Gross, and D. Prasad, “Symplectic local root numbers, central critical L values, and restriction problems in the representation theory of classical groups”, pp. 1–109 in *Sur les conjectures de Gross et Prasad, I*, Astérisque **346**, Soc. Math. France, Paris, 2012. MR Zbl

[Gourevitch et al. 2019] D. Gourevitch, S. Sahi, and E. Sayag, “Analytic continuation of equivariant distributions”, *Int. Math. Res. Not.* **2019**:23 (2019), 7160–7192. MR Zbl

[Gurevich and Szpruch 2015] N. Gurevich and D. Szpruch, “The non-tempered θ_{10} Arthur parameter and Gross–Prasad conjectures”, *J. Number Theory* **153** (2015), 372–426. MR Zbl

[Gurevich et al. 2018] M. Gurevich, J.-J. Ma, and A. Mitra, “On two questions concerning representations distinguished by the Galois involution”, *Forum Math.* **30**:1 (2018), 141–157. MR Zbl

[Heiermann 2016] V. Heiermann, “A note on standard modules and Vogan L-packets”, *Manuscripta Math.* **150**:3-4 (2016), 571–583. MR Zbl

[Heiermann and Opdam 2013] V. Heiermann and E. Opdam, “On the tempered L-functions conjecture”, *Amer. J. Math.* **135**:3 (2013), 777–799. MR Zbl

[Ichino and Prasanna 2016] A. Ichino and K. Prasanna, “Periods of quaternionic Shimura varieties, I”, 2016. To appear in Contemp. Math. arXiv

[Jantzen and Liu 2014] C. Jantzen and B. Liu, “The generic dual of p-adic split SO_{2n} and local Langlands parameters”, *Israel J. Math.* **204**:1 (2014), 199–260. MR Zbl

[Kudla 1994] S. S. Kudla, “Splitting metaplectic covers of dual reductive pairs”, *Israel J. Math.* **87**:1-3 (1994), 361–401. MR Zbl

[Kudla 1996] S. S. Kudla, “Notes on the local theta correspondence”, preprint, 1996, available at https://tinyurl.com/kudnotes.

[Kudla and Rallis 1992] S. S. Kudla and S. Rallis, “Ramified degenerate principal series representations for $Sp(n)$”, *Israel J. Math.* **78**:2-3 (1992), 209–256. MR Zbl

[Kudla and Rallis 2005] S. S. Kudla and S. Rallis, “On first occurrence in the local theta correspondence”, pp. 273–308 in *Automorphic representations, L-functions and applications: progress and prospects*, edited by J. W. Cogdell et al., Ohio State Univ. Math. Res. Inst. Publ. **11**, de Gruyter, Berlin, 2005. MR Zbl

[Liu 2011] B. Liu, “Genericity of representations of p-adic Sp_{2n} and local Langlands parameters”, *Canad. J. Math.* **63**:5 (2011), 1107–1136. MR Zbl

[Lu 2017a] H. Lu, GSp(4)*-period problems over a quadratic field extension*, Ph.D. thesis, National University of Singapore, 2017.

[Lu 2017b] H. Lu, “A new proof to the period problems of GL(2)”, *J. Number Theory* **180** (2017), 1–25. MR Zbl

[Matringe 2011] N. Matringe, “Distinguished generic representations of $GL(n)$ over p-adic fields”, *Int. Math. Res. Not.* **2011**:1 (2011), 74–95. MR Zbl

[Mœglin and Waldspurger 2016a] C. Moeglin and J.-L. Waldspurger, *Stabilisation de la formule des traces tordue, I*, Progr. Math. **316**, Birkhäuser, Cham, 2016. MR Zbl

[Mœglin and Waldspurger 2016b] C. Moeglin and J.-L. Waldspurger, *Stabilisation de la formule des traces tordue, II*, Progr. Math. **317**, Birkhäuser, Cham, 2016. MR Zbl

[Mœglin and Waldspurger 2018] C. Mœglin and J.-L. Waldspurger, *La formule des traces locale tordue*, Mem. Amer. Math. Soc. **1198**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Mœglin et al. 1987] C. Mœglin, M.-F. Vignéras, and J.-L. Waldspurger, *Correspondances de Howe sur un corps p-adique*, Lecture Notes in Math. **1291**, Springer, 1987. MR Zbl

[Ólafsson 1987] G. Ólafsson, “Fourier and Poisson transformation associated to a semisimple symmetric space”, *Invent. Math.* **90**:3 (1987), 605–629. MR Zbl

[Prasad 1992] D. Prasad, "Invariant forms for representations of GL_2 over a local field", *Amer. J. Math.* **114**:6 (1992), 1317–1363. MR Zbl

[Prasad 1993] D. Prasad, "On the local Howe duality correspondence", *Int. Math. Res. Not.* **1993**:11 (1993), 279–287. MR Zbl

[Prasad 1996] D. Prasad, "Some applications of seesaw duality to branching laws", *Math. Ann.* **304**:1 (1996), 1–20. MR Zbl

[Prasad 2015] D. Prasad, "A 'relative' local Langlands correspondence", preprint, 2015. arXiv

[Roberts 1996] B. Roberts, "The theta correspondence for similitudes", *Israel J. Math.* **94** (1996), 285–317. MR Zbl

[Roberts 2001] B. Roberts, "Global L-packets for GSp(2) and theta lifts", *Doc. Math.* **6** (2001), 247–314. MR Zbl

[Sun and Zhu 2015] B. Sun and C.-B. Zhu, "Conservation relations for local theta correspondence", *J. Amer. Math. Soc.* **28**:4 (2015), 939–983. MR Zbl

[Vogan 1993] D. A. Vogan, Jr., "The local Langlands conjecture", pp. 305–379 in *Representation theory of groups and algebras*, edited by J. Adams et al., Contemp. Math. **145**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

[Waldspurger 1990] J.-L. Waldspurger, "Démonstration d'une conjecture de dualité de Howe dans le cas p-adique, $p \neq 2$", pp. 267–324 in *Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, I* (Ramat Aviv, Israel, 1989), edited by R. Howe and P. Sarnak, Israel Math. Conf. Proc. **2**, Weizmann, Jerusalem, 1990. MR Zbl

[Yamana 2011] S. Yamana, "Degenerate principal series representations for quaternionic unitary groups", *Israel J. Math.* **185** (2011), 77–124. MR Zbl

Communicated by Shou-Wu Zhang
Received 2019-05-07 Revised 2020-02-08 Accepted 2020-05-04

hengfei.lu@weizmann.ac.il *Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel*
Current address: *Department of Mathematics, University of Vienna, Vienna, Austria*

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2481

Invertible functions on nonarchimedean symmetric spaces

Ernst-Ulrich Gekeler

Let u be a nowhere vanishing holomorphic function on the Drinfeld space Ω^r of dimension $r-1$, where $r \geq 2$. The logarithm $\log_q|u|$ of its absolute value may be regarded as an affine function on the attached Bruhat–Tits building $\mathcal{BT}^r$. Generalizing a construction of van der Put in case $r = 2$, we relate the group $\mathcal{O}(\Omega^r)^*$ of such u with the group $\mathbf{H}(\mathcal{BT}^r, \mathbb{Z})$ of integer-valued harmonic 1-cochains on $\mathcal{BT}^r$. This also gives rise to a natural $\mathbb{Z}$-structure on the first (ℓ-adic or de Rham) cohomology of Ω^r.

0. Introduction

The nonarchimedean symmetric spaces $\Omega = \Omega^r$ introduced by Drinfeld [1974] have shown great importance in the theories of modular and automorphic forms and of Shimura varieties, in the analytic uniformization of algebraic varieties, in the representation theory of $\mathrm{GL}(r, K)$, in the local Langlands correspondence, and in several other topics of the arithmetic of nonarchimedean local fields K. An incomplete list of a few references is [Manin and Drinfeld 1973; Mustafin 1978; Gerritzen and van der Put 1980; Schneider and Stuhler 1991; Laumon 1996; de Shalit 2001].

For a complete nonarchimedean local field K with finite residue class field $\mathbb{F}$ and completed algebraic closure C, the space Ω is defined as the complement of the K-rational hyperplanes in $\mathbb{P}^{r-1}(C)$. It carries a natural structure as a rigid-analytic space defined over K, and is supplied with an action of the group $\mathrm{PGL}(r, K)$. In contrast with the case of real symmetric spaces, it fails to be simply connected (in the étale topology, see [Fresnel and van der Put 2004, pages 160–161], but has a rich cohomological structure. Its cohomology (for cohomology theories satisfying some natural axioms) has been calculated by Schneider and Stuhler [1991]; see also [de Shalit 2001] and [Iovita and Spiess 2001].

Suppose for the moment that $r = 2$. In this case, $\Omega = \Omega^2$ has dimension 1, and a coarse combinatorial picture is provided by the Bruhat–Tits tree $\mathcal{T}$ of $\mathrm{PGL}(2, K)$, a $(q+1)$-regular tree, where $q = \#(\mathbb{F})$ is the residue class cardinality of K. A map φ from the set $\boldsymbol{A}(\mathcal{T})$ of oriented 1-simplices ("arrows") of $\mathcal{T}$ to $\mathbb{Z}$ that satisfies

(A) $\varphi(e) + \varphi(\bar{e}) = 0$ for each $e \in \boldsymbol{A}(\mathcal{T})$ with inverse $\bar{e}$, and

(B) $\sum \varphi(e) = 0$ for each vertex v of $\mathcal{T}$, where e runs through the arrows emanating from v,

MSC2010: primary 32P05; secondary 11F23, 11F85, 32C30, 32C36.
Keywords: Drinfeld symmetric space, van der Put transform, Bruhat–Tits building.

is called a ($\mathbb{Z}$-valued) *harmonic cochain* on $\mathcal{T}$. The group $\mathbf{H}(\mathcal{T}, \mathbb{Z})$ of all such yields upon tensoring with $\mathbb{Z}_\ell$ (ℓ a prime coprime with q) the first étale cohomology group $H^1_{ét}(\Omega^2, \mathbb{Z}_\ell)$ of Ω^2 [Drinfeld 1974, Proposition 10.2]. Marius van der Put [1983] (see also [Fresnel and van der Put 1981, I.8.9]) established a short exact sequence

$$1 \to C^* \to \mathcal{O}(\Omega^2)^* \xrightarrow{P} \mathbf{H}(\mathcal{T}, \mathbb{Z}) \to 0 \tag{0.1}$$

of $\mathrm{PGL}(2, K)$-modules, where $\mathcal{O}(\Omega^2)$ is the C-algebra of holomorphic functions on Ω^2 with multiplicative group $\mathcal{O}(\Omega^2)^*$. The van der Put transform $P(u)$ of an invertible function u is a substitute for the logarithmic derivative u'/u, and (0.1) provides the starting point for a study of the "Riemann surface" $\Gamma \setminus \Omega^2$, where $\Gamma \subset \mathrm{PGL}(2, K)$ is a discrete subgroup [Gerritzen and van der Put 1980; Gekeler and Reversat 1996].

It is the aim of the present paper to develop a higher-rank (i.e., $r > 2$) analogue of (0.1). In [Gekeler 2019] it was shown that the absolute value $|u|$ of $u \in \mathcal{O}(\Omega^r)^*$ factors over the building map

$$\lambda : \Omega^r \to \mathcal{BT}^r$$

and that its logarithm $\log_q |u|$ defines an affine map on $\mathcal{BT}^r(\mathbb{Q})$. Here $\mathcal{BT}^r$ is the Bruhat–Tits building of $\mathrm{PGL}(r, K)$ (the higher-dimensional analogue of $\mathcal{BT}^2 = \mathcal{T}$) and $\mathcal{BT}^r(\mathbb{Q})$ is the set of $\mathbb{Q}$-points of its realization $\mathcal{BT}^r(\mathbb{R})$. This makes it feasible that $u \mapsto \log_q |u|$ gives rise to a construction of P generalizing van der Put's in the case $r = 2$. The transform $P(u)$ of u will be a $\mathbb{Z}$-valued function on the set of arrows $\boldsymbol{A}(\mathcal{BT}^r)$ of $\mathcal{BT}^r$ subject to (obvious generalizations of) the conditions (A) and (B) above.

Our first result, Proposition 3.1, is that $P(u)$ satisfies one more relation (condition (C) in Corollary 2.12) not visible if $r = 2$. We then define $\mathbf{H}(\mathcal{BT}^r, \mathbb{Z})$ as the group of those $\varphi : \boldsymbol{A}(\mathcal{BT}^r) \to \mathbb{Z}$ which satisfy (A), (B) and (C).

The principal result of the present paper is the fact that the set of these relations is complete:

Theorem 3.11. *The map $P : \mathcal{O}(\Omega^r)^* \to \mathbf{H}(\mathcal{BT}^r, \mathbb{Z})$ is surjective, and the van der Put sequence*

$$1 \to C^* \to \mathcal{O}(\Omega^r)^* \to \mathbf{H}(\mathcal{BT}^r, \mathbb{Z}) \to 0 \tag{0.2}$$

is an exact sequence of $\mathrm{PGL}(r, K)$*-modules.*

The proof requires the construction of certain functions $u = f_{H,H',n}$ whose transforms $P(u)$ have a prescribed behavior on the finite subcomplex $\mathcal{BT}^r(n)$ of $\mathcal{BT}^r$, and a crucial technical result (Proposition 3.10), which solely refers to the geometry of $\mathcal{BT}^r$.

Still, $\mathbf{H}(\mathcal{BT}^r, \mathbb{Z})$ is a torsion-free abelian group of complicated appearance. However, as a further consequence of Proposition 3.10, we are able to describe it in Theorem 4.16

- either as $\mathbf{H}(\mathcal{T}_{v_0}, \mathbb{Z})$, where $\mathcal{T}_{v_0}$ is a subcomplex of dimension 1 of $\mathcal{BT}^r$ (in fact, a tree, which for $r = 2$ agrees with the Bruhat–Tits tree $\mathcal{T} = \mathcal{BT}^2$), and where only conditions (A) and (B) are involved,
- or as the group $\mathbf{D}^0(\mathbb{P}(V), \mathbb{Z})$ of $\mathbb{Z}$-valued distributions of total mass 0 on the projective space $\mathbb{P}(V)$, or by duality, on the compact space $\mathbb{P}(V^\wedge)$ of hyperplanes of the K-vector space $V = K^r$.

As the corresponding group $\mathbf{D}^0(\mathbb{P}(V^\wedge), A)$ with coefficients in some ring A depending on the cohomology theory used (e.g., $A = \mathbb{Z}_\ell$ for étale cohomology, or, in characteristic zero, $A = K$ for de Rham cohomology) has been shown to agree with the first cohomology $H^1(\Omega^r, A)$ [Schneider and Stuhler 1991, Section 3, Theorem 1], we get in some cases a natural integral structure on $H^1(\Omega^r, A)$ along with a concrete arithmetic interpretation.

1. Background

1.1. Throughout, K denotes a nonarchimedean local field with ring O of integers, a fixed uniformizer π, and finite residue class field $O/(\pi) = \mathbb{F} = \mathbb{F}_q$ of cardinality q. Hence K is a finite extension of either a p-adic field $\mathbb{Q}_p$ or of a Laurent series field $\mathbb{F}_p((X))$. We normalize its absolute value $|\cdot|$ by $|\pi| = q^{-1}$, and let $C = \hat{\overline{K}}$ be its completed algebraic closure with respect to the unique extension of $|\cdot|$ to $\overline{K}$. The ring of integers of C and its maximal ideal are denoted by O_C and $\mathfrak{m}_C$. Note that the residue class field $O_C/\mathfrak{m}_C$ is an algebraic closure $\bar{\mathbb{F}}$ of $\mathbb{F}$. Further, $\log\colon C^* \to \mathbb{Q}$ is the map $z \mapsto \log_q |z|$.

1.2. Given a natural number $r \geq 2$, the Drinfeld symmetric space $\Omega = \Omega^r$ of dimension $r-1$ is the complement $\Omega = \mathbb{P}^{r-1} \setminus \bigcup H$ of the K-rational hyperplanes H in projective space $\mathbb{P}^{r-1}$. Hence the set of C-valued points of Ω (for which we briefly write Ω) is

$$\Omega = \{(\omega_1 : \cdots : \omega_r) \in \mathbb{P}^{r-1}(C) \mid \text{the } \omega_i \text{ are } K\text{-linearly independent}\}.$$

If not indicated otherwise, we always suppose that projective coordinates $(\omega_1 : \cdots : \omega_r)$ are *unimodular*, that is $\max_i |\omega_i| = 1$. The set Ω carries a natural structure as a rigid-analytic space defined over K (see [Drinfeld 1974; Deligne and Husemoller 1987; Schneider and Stuhler 1991]); in fact, it is an admissible open subspace of $\mathbb{P}^{r-1}$, and even a Stein domain [Schneider and Stuhler 1991, Section 1, Proposition 14]; see [Kiehl 1967] for the notion of nonarchimedean Stein domain.

1.3. Let G be the group scheme $\mathrm{GL}(r)$ with center Z; hence $G(K) = \mathrm{GL}(r, K)$, $Z(K) \cong K^*$, etc. The Bruhat–Tits building [Bruhat and Tits 1972] $\mathcal{BT} = \mathcal{BT}^r$ of $G(K)/Z(K) = \mathrm{PGL}(r, K)$ is a contractible simplicial complex with set of vertices

$$V(\mathcal{BT}) = \{[L] \mid L \text{ an } O\text{-lattice in } V\}, \tag{1.3.1}$$

where L runs through the set of O-lattices in the K-vector space $V = K^r$ and $[L]$ is the similarity class of L. (An O*-lattice* is a free O-submodule of rank r of V, two such, L and L', are *similar* if there exists $0 \neq c \in K$ such that $L' = cL$.) The classes $[L_0], \ldots, [L_s]$ form an s-simplex if and only if they are represented by lattices L_i such that

$$L_0 \supsetneq L_1 \supsetneq \cdots \supsetneq L_s \supsetneq \pi L_0. \tag{1.3.2}$$

The *combinatorial distance* $d(v, v')$ of two vertices $v, v' \in \boldsymbol{V}(\mathcal{BT})$ is the length of a shortest path connecting them in the 1-skeleton of $\mathcal{BT}$. It is easily verified that

$$d(v, v') = \min\{n \mid \exists \text{ representatives } L, L' \text{ for } v, v' \text{ such that } L \supset L' \supset \pi^n L\}. \tag{1.3.3}$$

The *star* $\mathrm{st}(v)$ of $v \in \boldsymbol{V}(\mathcal{BT})$ will always denote the full subcomplex of $\mathcal{BT}$ with set of vertices

$$\boldsymbol{V}(\mathrm{st}(v)) = \{w \in \boldsymbol{V}(\mathcal{BT}) \mid d(v, w) \leq 1\}. \tag{1.3.4}$$

We regard V as a space of row vectors, on which $G(K)$ acts as a matrix group from the right. Hence $G(K)$ acts also from the right on $\mathcal{BT}$. If the syntax requires a left action, we shift this action to the left by the usual formula $\gamma x := x\gamma^{-1}$.

1.4. The relationship between Ω and $\mathcal{BT}$ is as follows: By the Goldman–Iwahori theorem [Goldman and Iwahori 1963], the realization $\mathcal{BT}(\mathbb{R})$ of $\mathcal{BT}$ is in a natural one-to-one correspondence with the set of similarity classes of real-valued nonarchimedean norms on V, where a vertex $v = [L] \in \boldsymbol{V}(\mathcal{BT}) = \mathcal{BT}(\mathbb{Z})$ corresponds to the class of a norm with unit ball $L \subset V$. Now the *building map*

$$\begin{aligned} \lambda\colon \Omega &\to \mathcal{BT}(\mathbb{R}) \\ \boldsymbol{\omega} = (\omega_1 : \ldots : \omega_r) &\mapsto [\nu_{\boldsymbol{\omega}}] \end{aligned} \tag{1.4.1}$$

is well-defined, where the norm $\nu_{\boldsymbol{\omega}}$ maps $\boldsymbol{x} = (x_1, \ldots, x_r) \in V$ to

$$\nu_{\boldsymbol{\omega}}(\boldsymbol{x}) = \Big|\sum_{1 \leq i \leq r} x_i \omega_i\Big|,$$

and $[\nu_{\boldsymbol{\omega}}]$ is its similarity class. Since the value group is $|C^*| = q^{\mathbb{Q}}$, λ maps to $\mathcal{BT}(\mathbb{Q})$, and is in fact onto $\mathcal{BT}(\mathbb{Q})$, the set of points of $\mathcal{BT}(\mathbb{R})$ with rational barycentric coordinates. $G(K)$ acts from the left on the set of norms via

$$\gamma\nu(\boldsymbol{x}) := \nu(\boldsymbol{x}\gamma) \tag{1.4.2}$$

for $\boldsymbol{x} \in V$, a norm ν, and $\gamma \in G(K)$; the reader may verify that λ is $G(K)$-equivariant, where the action on Ω is the standard one through left matrix multiplication. The preimages under λ of simplices of $\mathcal{BT}$ yield an admissible covering of Ω; see e.g., [de Shalit 2001, (6.2) and (6.3)]. We therefore consider $\mathcal{BT}$ as a combinatorial picture of Ω.

We cite the following results from [Gekeler 2017; 2019].

Theorem 1.5 [Gekeler 2019, Theorem 2.4]. *Let u be an invertible holomorphic function on Ω. Then $|u(\boldsymbol{\omega})|$ depends only on the image $\lambda(\boldsymbol{\omega})$ of $\boldsymbol{\omega} \in \Omega$ in $\mathcal{BT}(\mathbb{Q})$.*

1.5.1. We thus define the *spectral norm* $\|u\|_x$ as the common absolute value $|u(\boldsymbol{\omega})|$ for all $\omega \in \lambda^{-1}(x)$, where $x \in \mathcal{BT}(\mathbb{Q})$.

Theorem 1.6 [Gekeler 2019, Theorem 2.6]. *Let u be an invertible holomorphic function on Ω. Then $\log u = \log_q |u|$ regarded as a function on $\mathcal{BT}(\mathbb{Q})$ is affine, that is, interpolates linearly in simplices.*

1.7. Let $\boldsymbol{A}(\mathcal{BT})$ be the set of *arrows*, i.e., of oriented 1-simplices of $\mathcal{BT}$. For each arrow $e = (v, v') = ([L], [L'])$ we write

$$o(e) = \text{origin of } e := v, \quad t(e) = \text{terminus of } e := v', \quad \text{and} \quad \text{type}(e) := \dim_{\mathbb{F}}(L'/\pi L),$$

where L, L' are representatives with $L \supset L' \supset \pi L$. Then $1 \le \text{type}(e) \le r-1$ and $\text{type}(e) + \text{type}(\bar{e}) = r$, where $\bar{e} = (v', v)$ is e with reverse orientation. We let

$$\boldsymbol{A}_v = \dot{\bigcup_{1 \le t \le r-1}} \boldsymbol{A}_{v,t} \tag{1.7.1}$$

be the arrows e with $o(e) = v$, grouped according to their types t. For an invertible function u on Ω and an arrow $e = (v, w)$, define the *van der Put value* $P(u)(e)$ of u on e as

$$P(u)(e) = \log_q \|u\|_w - \log_q \|u\|_v \tag{1.7.2}$$

with the spectral norm of Section 1.5.1.

Proposition 1.8 [Gekeler 2017, Proposition 2.9]. *The* van der Put transform

$$\begin{aligned} P(u) \colon \boldsymbol{A}(\mathcal{BT}) &\to \mathbb{Q} \\ e &\mapsto P(u)(e) \end{aligned}$$

of u has in fact values in $\mathbb{Z}$ and satisfies

$$\sum_{e \in \boldsymbol{A}_{v,1}} P(u)(e) = 0 \tag{1.8.1}$$

for all $v \in \boldsymbol{V}(\mathcal{BT})$. Here the sum is over the arrows e with $o(e) = v$ and $\text{type}(e) = 1$.

Actually, in [Gekeler 2017] the condition $\sum_{e \in \boldsymbol{A}_{v,r-1}} P(u)(e) = 0$ is given instead of (1.8.1), due to another choice of orientation. We will discuss this in more detail in Section 2.6 and Remarks 3.3(i), which will also show that both conditions are equivalent in our framework.

Remarks 1.9. (i) In the case $r = 2$, Theorems 1.5, 1.6 and Proposition 1.8 have been known for quite some time; see [van der Put 1983] and e.g., [Fresnel and van der Put 1981, I.8.9]. For general r, they are shown in [Gekeler 2017; 2019] in the framework of these papers, where $\text{char}(K) = \text{char}(\mathbb{F}) = p$. However, the proofs make no use of this assumption, and are therefore valid for $\text{char}(K) = 0$, too.

(ii) The three cited results are local in the sense that they do not require u to be a global unit. If, e.g., u is a holomorphic function without zeroes on the affinoid $\lambda^{-1}(x)$ with $x \in \mathcal{BT}(\mathbb{Q})$, then $|u(\boldsymbol{\omega})|$ is constant on $\lambda^{-1}(x)$; if u is invertible on $\lambda^{-1}(\sigma)$ with a closed simplex σ of $\mathcal{BT}$, then $\log u$ is affine there, and if u is invertible on $\lambda^{-1}(\text{st}(v))$, where $\text{st}(v)$ is the star of $v \in \boldsymbol{V}(\mathcal{BT})$ (see (1.3.4)), then $P(u)(e)$ is defined for all $e \in \boldsymbol{A}_v$ and satisfies (1.8.1).

(iii) It is immediate from the definitions that for invertible functions u, u' and arrows e,

$$P(u)(e) + P(u)(\bar{e}) = 0, \tag{1.9.1}$$

and more generally

$$\sum P(u)(e) = 0, \quad \text{if } e \text{ runs through the arrows of a closed path in } \mathcal{BT}, \tag{1.9.2}$$

as well as

$$P(uu') = P(u) + P(u'). \tag{1.9.3}$$

Hence the van der Put transform $P: u \mapsto P(u)$ is a homomorphism from the multiplicative group $\mathcal{O}(\Omega)^*$ of invertible holomorphic functions on Ω to the additive group of maps $\varphi\colon \boldsymbol{A}(\mathcal{BT}) \to \mathbb{Z}$ that satisfy (1.9.1), (1.9.2) and (1.8.1). Moreover, for $\gamma \in G(K)$,

$$P(u)(e\gamma) = P(u \circ \gamma^{-1})(e), \tag{1.9.4}$$

i.e., $\gamma(P(u)) = P(\gamma u) := P(u \circ \gamma^{-1})$ holds; whence P is $G(K)$-equivariant.

In Theorem 3.11 we will find exact conditions that characterize the image of P. This will yield the exact sequence (0.2) of $G(K)$-modules that generalizes (0.1).

2. Evaluation of P on elementary rational functions

2.1. Let U be a subspace of $V = K^r$ of dimension t, where $1 \leq t \leq r-1$. We define the shift toward U on $\boldsymbol{V}(\mathcal{BT})$ by

$$\begin{aligned} \tau_U\colon \boldsymbol{V}(\mathcal{BT}) &\to \boldsymbol{V}(\mathcal{BT}), \\ v = [L] &\mapsto [L'] \end{aligned} \tag{2.1.1}$$

where $L' = (L \cap U) + \pi L$. Obviously, $e = (v, \tau_U(v))$ is a well-defined arrow of type $\mathrm{type}(e) = \dim U = t$. We say that *$e$ points to U*.

2.1.2. For a local ring R (in practice: $R = K$, or O, or a finite quotient $O_n := O/(\pi^n)$) and a free R-module F of finite rank, let $\mathrm{Gr}_{R,t}(F)$ be the Grassmannian of direct summands F' such that $\mathrm{rank}_R\, F' = t$. Fixing $v = [L] \in \boldsymbol{V}(\mathcal{BT})$, there is a natural surjective map

$$\begin{aligned} \mathrm{Gr}_{K,t}(V) &\to \boldsymbol{A}_{v,t} \\ U &\mapsto (v, \tau_U(v)) \end{aligned} \tag{2.1.3}$$

and a canonical bijection

$$\boldsymbol{A}_{v,t} \xrightarrow{\cong} \mathrm{Gr}_{\mathbb{F},t}(L/\pi L) \tag{2.1.4}$$

given by $e = (v, w) = ([L], [M]) \mapsto \overline{M} := M/\pi L$, where $L \supset M \supset \pi L$. We denote the image of e by $\overline{M}_e$ and the preimage of $\overline{M}$ in $\boldsymbol{A}_{v,t}$ by $e_{\overline{M}}$.

2.1.5. For two arrows $e = e_{\overline{M}}$ and $e' = e_{\overline{M}'}$ with the same origin, we write $e \prec e'$ (e' *dominates* e) if and only if $\overline{M} \subset \overline{M}'$.

2.1.6. Fix $n \in \mathbb{N}$, let O_n be the ring $O/(\pi^n)$ and let $t \in \{1, r-1\}$. Then, as a generalization of the above, $U \mapsto (v, \tau_U(v), \ldots, \tau_U^n(v))$ is surjective from $\mathrm{Gr}_{K,t}(V)$ onto the set $\boldsymbol{A}_{v,t,n}$ of paths of length n in $\mathcal{BT}$ which emanate from v, are composed of arrows of type t, and whose endpoints w have distance $d(v, w) = n$ (e.g., $\boldsymbol{A}_{v,t,1} = \boldsymbol{A}_{v,t}$). The set $\boldsymbol{A}_{v,t,n}$ corresponds one-to-one to $\mathrm{Gr}_{O_n,t}(L/\pi^n L)$, where the composite map from $\mathrm{Gr}_{K,t}(V)$ to $\mathrm{Gr}_{O_n,t}(L/\pi^n L)$ is given by $U \mapsto ((L \cap U) + \pi^n L)/\pi^n L$. This yields in the limit the canonical bijections

$$\mathrm{Gr}_{K,t}(V) \xrightarrow{\cong} \varprojlim_n \boldsymbol{A}_{v,t,n} = \varprojlim_n \mathrm{Gr}_{O_n,t}(L/\pi^n L) = \mathrm{Gr}_{O,t}(L), \tag{2.1.7}$$

whose composition is simply $U \mapsto U \cap L$. Let e be an arrow of type t. Then

$$\mathrm{Gr}_{K,t}(e) := \{U \in \mathrm{Gr}_{K,t}(V) \mid e \text{ points to } U\} \tag{2.1.8}$$

is compact and open in the compact space $\mathrm{Gr}_{K,t}(V)$, and it follows from the considerations above that the set of all $\mathrm{Gr}_{K,t}(e)$, where v is fixed and e belongs to $\boldsymbol{A}_{v,t,n}$ for some $n \in \mathbb{N}$, forms a basis for the topology on $\mathrm{Gr}_{K,t}(V)$.

2.2. Given a hyperplane H in V, we let $\ell_H : V \to K$ be a linear form with kernel H. We denote by the same symbol its extension $\ell_H : V \otimes_K C = C^r \to C$. The quotients

$$\ell_{H,H'} := \ell_H/\ell_{H'} \tag{2.2.1}$$

of two such are rational functions on $\mathbb{P}^{r-1}(C)$ without zeroes or poles on $\Omega \hookrightarrow \mathbb{P}^{r-1}(C)$. Note that ℓ_H is determined up to multiplication by a nonzero scalar in K; hence $P(\ell_{H,H'})$ depends only on H and H', but not on the scaling of ℓ_H and $\ell_{H'}$. Our first task will be to describe $P(\ell_{H,H'})$.

2.3. We start with a closer look to the building map λ. Let $v_0 = [L_0]$ be the standard vertex, where L_0 is the standard lattice O^r in V. Let us first recall the easily verified fact (where the unimodularity normalization of $\boldsymbol{\omega} \in \Omega$ is used):

$$\begin{aligned} \Omega_{v_0} := \lambda^{-1}(v_0) &= \{\boldsymbol{\omega} \in \Omega \mid \nu_{\boldsymbol{\omega}} \text{ has unit ball } L_0\} \\ &= \{\boldsymbol{\omega} \in \Omega \mid \text{ the } \omega_i \text{ are orthogonal and } |\omega_i| = 1 \text{ for } 1 \le i \le r\}. \end{aligned} \tag{2.3.1}$$

($z_1, \ldots, z_n \in C$ are *orthogonal* if and only if $|\sum_{1 \le i \le r} a_i z_i| = \max_i |a_i z_i|$ for arbitrary coefficients $a_i \in K$.) Hence the canonical reduction of Ω_{v_0} equals

$$\overline{\Omega}_{v_0} = (\mathbb{P}^{r-1}/\mathbb{F}) \setminus \bigcup \overline{H}, \tag{2.3.2}$$

where $\overline{H}$ runs through the hyperplanes defined over $O/(\pi) = \mathbb{F}$. A similar description holds for $\overline{\lambda^{-1}(v)}$ if v is an arbitrary vertex, but we need some preparations.

2.4. Write $\langle \cdot , \cdot \rangle$ for the standard bilinear form on V given by

$$\langle \boldsymbol{x}', \boldsymbol{x} \rangle = \sum_{1 \le i \le r} x_i' x_i,$$

which we extend to a form $\langle\,\cdot\,,\cdot\,\rangle$ on C^r. It identifies $V = K^r$ with its dual $V^\wedge$. For each K-subspace U of V, let

$$U^\perp := \{\boldsymbol{x} \in V \mid \langle \boldsymbol{x}, U\rangle = 0\} \tag{2.4.1}$$

be its orthogonal with respect to $\langle\,\cdot\,,\cdot\,\rangle$. For an O-lattice L in V,

$$L^\wedge := \{\boldsymbol{x} \in V \mid \langle \boldsymbol{x}, L\rangle \subset O\} \tag{2.4.2}$$

is the dual lattice. We put $\tilde{\Omega}$ for the preimage of Ω in C^r. Then $L^\wedge \otimes_O O_C$ embeds into C^r, and by (1.4.1) we find:

2.4.3. The image of $(L^\wedge \otimes O_C) \cap \tilde{\Omega}$ in Ω equals $\Omega_v := \lambda^{-1}(v)$, where $v = [L]$ is the vertex of $\mathcal{BT}$ corresponding to L.

Similarly,

$$(L^\wedge \otimes O_C) \otimes O_C/\mathfrak{m}_C \xrightarrow{\cong} (L^\wedge/\pi L^\wedge) \otimes_{\mathbb{F}} \bar{\mathbb{F}} = (L/\pi L)^\wedge \otimes_{\mathbb{F}} \bar{\mathbb{F}}. \tag{2.4.4}$$

2.5. Let $\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}$ be the projective space attached to the r-dimensional vector space $L^\wedge/\pi L^\wedge$, regarded as a scheme over $O/(\pi) = \mathbb{F}$. Its $\mathbb{F}$-rational hyperplanes correspond to those of the vector space $(L^\wedge/\pi L^\wedge) \otimes_{\mathbb{F}} \bar{\mathbb{F}}$, or, by duality, to the $\mathbb{F}$-lines (one-dimensional $\mathbb{F}$-subspaces) $\overline{G}$ in $L/\pi L$. Therefore, the canonical reduction $\overline{\Omega}_v$ of Ω_v is

$$\overline{\Omega}_v = (\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}) \setminus \bigcup \overline{H}, \tag{2.5.1}$$

where $\overline{H}$ runs through the hyperplanes defined over $\mathbb{F}$. The set of these is in canonical bijection with the set of $\mathbb{F}$-lines in $L/\pi L$, that is, with $\boldsymbol{A}_{v,1}$. For each $e \in \boldsymbol{A}_{v,1}$ let $\overline{H}_e$ be the corresponding $\mathbb{F}$-hyperplane in (2.5.1).

2.5.2. The object $\overline{H}_e$ (an $\mathbb{F}$-subspace of $L^\wedge/\pi L^\wedge$ or the corresponding hyperplane in $\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}$, described through the same symbol) mustn't be confused with the $\overline{M}_e$ of (2.1.4), which is an $\mathbb{F}$-subspace of $L/\pi L$. The relationship is as follows. The form $\langle\,\cdot\,,\cdot\,\rangle$ induces an $\mathbb{F}$-bilinear form

$$\overline{\langle\,\cdot\,,\cdot\,\rangle}\colon L/\pi L \times L^\wedge/\pi L^\wedge \to \mathbb{F}.$$

For an $\mathbb{F}$-subspace $\overline{M}$ of $L/\pi L$, $\overline{M}^\perp$ denotes its orthogonal with respect to $\overline{\langle\,\cdot\,,\cdot\,\rangle}$ in $L^\wedge/\pi L^\wedge$. Let $\overline{G}_e \subset L/\pi L$ be the line defined by $e \in \boldsymbol{A}_{v,1}$ as in (2.1.4). Then $\overline{H}_e = (\overline{G}_e)^\perp$.

2.6. Let u be an invertible holomorphic function on Ω, scaled such that $\|u\|_v = 1$. Its reduction $\bar{u}$ at v is a rational function on $\overline{\Omega}_v$ without zeroes or poles. For each $e \in \boldsymbol{A}_{v,1}$ let m_e be the vanishing order of $\bar{u}$ along $\overline{H}_e$ (negative, if $\bar{u}$ has a pole along $\overline{H}_e$), and let ℓ_e be a linear form on $\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}$ with vanishing locus $\overline{H}_e$. Up to a multiplicative constant, $\bar{u}$ equals $\prod_{e\in\boldsymbol{A}_{v,1}} \ell_e^{m_e}$, and so

$$\sum_{e\in\boldsymbol{A}_{v,1}} m_e = \text{weight of the form } \bar{u} = 0. \tag{2.6.1}$$

Now the value of the van der Put transform on $e \in \boldsymbol{A}_{v,1}$ is (with notation above)

$$P(u)(e) = -m_e. \tag{2.6.2}$$

To see this, we may assume (by (1.9.4), and since the action of $G(K)$ on arrows of type 1 is transitive) that $e = (v_0, v_1)$ with $v_0 = [L_0]$, $v_1 = [L_1]$, $L_0 = O^r$, $L_1 = (\pi) \times \cdots \times (\pi) \times O$. Then $L_0^\wedge = L_0$, $L_1^\wedge = (\pi^{-1}) \times \cdots \times (\pi^{-1}) \times O$, $\overline{H}_e$ is the hyperplane $\{(* : \cdots : * : 0)\}$ in $\mathbb{P}(L_0^\wedge/\pi L_0^\wedge)/\mathbb{F} = \mathbb{P}^{r-1}/\mathbb{F}$, and we may choose $\ell_e \colon \boldsymbol{x} = (x_1 : \cdots : x_r) \mapsto x_r$, which is the reduction of the global form $\tilde{\ell} \colon \boldsymbol{\omega} = (\omega_1 : \cdots : \omega_r) \mapsto \omega_r$ on Ω. In order to get "functions" instead of "forms", we work with ℓ_e/ℓ_1 (resp. $\tilde{\ell}/\tilde{\ell}_1$), where $\ell_1 \colon \boldsymbol{x} \mapsto x_1$ with lift $\tilde{\ell}_1 \colon \boldsymbol{\omega} \mapsto \omega_1$. If $\bar{u}$ has a zero of order m along $\overline{H}_e$ (i.e., $\bar{u} = \bar{u}_0(\ell_e/\ell_1)^m$, where $\bar{u}_0$ has neither zeroes nor poles on $\overline{\Omega}_{v_0}$ and along $\overline{H}_e$), then u grows like $(\tilde{\ell}/\tilde{\ell}_1)^m$ when moving from Ω_{v_0} to Ω_{v_1}. But the absolute value of $\tilde{\ell}/\tilde{\ell}_1$ on Ω_{v_0} is 1, while it is $|\pi| = q^{-1}$ on Ω_{v_1}, which shows (2.6.2).

Finally, combining the above yields

$$\sum_{e \in \boldsymbol{A}_{v,1}} P(u)(e) = 0, \tag{2.6.3}$$

that is, the assertion of (1.8.1).

2.7. Each hyperplane H of V is given as the kernel of a linear form

$$\ell_H = \ell_{\boldsymbol{y}} \colon \boldsymbol{x} \mapsto \langle \boldsymbol{y}, \boldsymbol{x} \rangle \tag{2.7.1}$$

with some $\boldsymbol{y} \in L_0 \setminus \pi L_0$. Let $G = K\boldsymbol{y} = H^\perp$ be the line spanned by $\boldsymbol{y}$. The arrow $(v_0, \tau_G(v_0)) \in \boldsymbol{A}_{v_0,1}$ equals $e_{\overline{G}}$ with

$$\overline{G} = (O\boldsymbol{y} + \pi L_0)/\pi L_0.$$

Two such vectors $\boldsymbol{y}$, $\boldsymbol{y}'$ give rise to the same $e_{\overline{G}}$ if and only if $\boldsymbol{y}' \equiv c \cdot \boldsymbol{y} \pmod{\pi}$ with some unit $c \in O^*$. More generally, $\boldsymbol{y}$ and $\boldsymbol{y}'$ give rise to the same path $(v_0, \tau_G(v_0), \ldots, \tau_G^n(v_0)) \in \boldsymbol{A}_{v_0,1,n}$ if and only if

$$\boldsymbol{y}' \equiv c \cdot \boldsymbol{y} \pmod{\pi^n} \tag{2.7.2}$$

with $c \in O^*$. In this case we call $\boldsymbol{y}$ and $\boldsymbol{y}'$ *n-equivalent*; the respective equivalence classes are briefly the *n-classes* of $\boldsymbol{y}$, $\boldsymbol{y}'$.

2.8. Let now hyperplanes H, H' of V be given by $\boldsymbol{y}$, $\boldsymbol{y}'$ as above. Put $G = H^\perp = K\boldsymbol{y}$, $G' = K\boldsymbol{y}'$. The function $\ell_{H,H'} = \ell_{\boldsymbol{y}}/\ell_{\boldsymbol{y}'}$ has constant absolute value 1 on Ω_{v_0} and therefore, by reduction, gives a rational function $\bar{\ell}_{H,H'}$ without zeroes or poles on $\overline{\Omega}_{v_0} \hookrightarrow \mathbb{P}^{r-1}/\mathbb{F}$. Put

$$\overline{H} = ((L_0 \cap H) + \pi L_0)/\pi L_0,$$

and ditto $\overline{H}'$. By definition, it is an $\mathbb{F}$-subvector space of $L_0/\pi L_0 \xrightarrow{\cong} \mathbb{F}^r$. As usual, we denote by the same symbol the corresponding $\mathbb{F}$-rational linear subvariety of $\mathbb{P}^{r-1}/\mathbb{F}$ that appears e.g., in (2.3.2). Suppose that $\overline{H}$ differs from $\overline{H}'$. Then $\bar{\ell}_{H,H'}$ has vanishing order 1 along $\overline{H}$, vanishing order -1 along $\overline{H}'$, and vanishing order 0 along the other hyperplanes in the boundary of $\overline{\Omega}_{v_0}$ (see (2.3.2)). If however $\overline{H} = \overline{H}'$, then $\bar{\ell}_{H,H'}$ has neither zeroes nor poles along the boundary (and is therefore constant). According to the recipe discussed in Section 2.6, we find the following description.

Proposition 2.9. *Let e be an arrow in $\boldsymbol{A}_{v_0,1}$. Then*

$$P(\ell_{H,H'})(e) = \begin{cases} -1 & e = (v_0, \tau_G(v_0)) \neq (v_0, \tau_{G'}(v_0)), \\ +1 & e = (v_0, \tau_{G'}(v_0)) \neq (v_0, \tau_G(v_0)), \\ \;\;0 & \textit{otherwise.} \end{cases} \qquad \square$$

Again by the transitivity of the action of $G(K)$, we may transfer Proposition 2.9 to arbitrary arrows of type 1, and thus get:

Corollary 2.10. *Let $e \in \boldsymbol{A}_{v,1}$ be an arrow of type 1 with arbitrary origin $v \in \boldsymbol{V}(\mathcal{BT})$. Write $e_H^\perp$ (resp. $e_{H'}^\perp$) for the arrow $(v, \tau_{H^\perp}(v))$ (resp. $(v, \tau_{H'^\perp}(v))$). Then*

$$P(\ell_{H,H'})(e) = \begin{cases} -1 & e = e_H^\perp \neq e_{H'}^\perp, \\ +1 & e = e_{H'}^\perp \neq e_H^\perp, \\ \;\;0 & \textit{otherwise.} \end{cases} \qquad \square$$

Next, we deal with arrows of arbitrary type.

Proposition 2.11. *Given hyperplanes H, H' of V and an arrow e of $\mathcal{BT}$ with origin $v \in \boldsymbol{V}(\mathcal{BT})$, let $e_H^\perp$ (resp. $e_{H'}^\perp$) be the arrow with origin v pointing to $G = H^\perp$ (resp. to $G' = H'^\perp$). The transform $P(\ell_{H,H'})$ evaluates on e as follows*:

$$P(\ell_{H,H'})(e) = \begin{cases} -1 & e_H^\perp \prec e, e_{H'}^\perp \not\prec e, \\ +1 & e_{H'}^\perp \prec e, e_H^\perp \not\prec e, \\ \;\;0 & \textit{otherwise.} \end{cases}$$

Proof. Let L be a lattice with $[L] = v$ and $e = e_{\overline{M}}$, where $\overline{M}$ is a subspace of $L/\pi L$ of dimension $t = \text{type}(e)$. Since the case $t = 1$ is given by the last corollary, we may assume that $t \geq 2$. Suppose that $e_H^\perp \prec e$, i.e.,

$$\overline{G} = ((L \cap G) + \pi L)/\pi L \subset \overline{M} \subset L/\pi L.$$

Let $\overline{M}_0 = 0 \subsetneq \overline{M}_1 = \overline{G} \subsetneq \cdots \subsetneq \overline{M}_t = \overline{M}$ be a complete flag connecting 0 to $\overline{M}$, where $\dim \overline{M}_i = i$ for $0 \leq i \leq t$. It corresponds to a path $(v_0, v_1, \ldots, v_t)$ in $\mathcal{BT}$, where $v_0 = v = [L]$, $v_t = t(e_{\overline{M}})$, and all the arrows $e_1 = (v_0, v_1), \ldots, e_t = (v_{t-1}, v_t)$ of type 1. As $\{v_0, \ldots, v_t\}$ is a t-simplex, $d(v_0, v_i) = 1$ for $1 \leq i \leq t$, and therefore no e_i different from $e_1 = e_{\overline{G}}$ points to G.

Suppose that moreover $e_{H'}^\perp \not\prec e$, that is,

$$((L \cap G') + \pi L)/\pi L \not\subset \overline{M}.$$

Then none of the e_i $(1 \leq i \leq t)$ points to G', so

$$P(\ell_{H,H'})(e) = \sum_{1 \leq i \leq t} P(\ell_{H,H'})(e_i) = P(\ell_{H,H'})(e_1) = -1$$

by (1.9.2) and Corollary 2.10. If $e_H^{\perp} \neq e_{H'}^{\perp} \prec e$, then we can arrange the flag $\overline{M}_0 \subsetneq \cdots \subsetneq \overline{M}_t$ such that as before e_1 points to G, e_2 points to G', and no e_i $(3 \leq i \leq t)$ points to G or G'. In this case

$$P(\ell_{H,H'})(e) = P(\ell_{H,H'})(e_1) + P(\ell_{H,H'})(e_2) = -1 + 1 = 0.$$

If $e_H^{\perp} = e_{H'}^{\perp} \prec e$, then

$$P(\ell_{H,H'})(e) = P(\ell_{H,H'})(e_1) = 0 \quad \text{by Corollary 2.10.}$$

If neither $e_H^{\perp} \prec e$ nor $e_{H'}^{\perp} \prec e$, neither of the arrows e_i $(1 \leq i \leq t)$ corresponding to a flag $\overline{M}_0 = 0 \subsetneq \cdots \subsetneq \overline{M}_t = \overline{M}$ points to G or to G', and so $P(\ell_{H,H'})(e) = 0$ results. The case $e_{H'}^{\perp} \prec e$, $e_H^{\perp} \not\prec e$ comes out by symmetry. □

Corollary 2.12. *Let $H_1, \dots, H_n$ be finitely many hyperplanes of V with corresponding linear forms $\ell_i = \ell_{H_i}$, $\ker(\ell_i) = H_i$, and multiplicities $m_i \in \mathbb{Z}$ such that $\sum_{1 \leq i \leq n} m_i = 0$. The function*

$$u := \prod_{1 \leq i \leq n} \ell_i^{m_i}$$

is a unit on Ω, whose van der Put transform $P(u)$ satisfies the condition:

(C) *For each arrow $e \in \boldsymbol{A}(\mathcal{BT})$ with $o(e) = v \in \boldsymbol{V}(\mathcal{BT})$,*

$$P(u)(e) = \sum_{\substack{e' \in \boldsymbol{A}_{v,1} \\ e' \prec e}} P(u)(e').$$

Proof. (C) is satisfied for $u = \ell_{H,H'} = \ell_H / \ell_{H'}$ by Proposition 2.11. The general case follows as condition (C) is linear (it holds for $u \cdot u'$ if it holds for u and u') and $\prod \ell_i^{m_i}$ is a product of functions of type $\ell_{H,H'}$. □

3. The van der Put sequence

Proposition 3.1. *Let u be an invertible holomorphic function on Ω. Then its van der Put transform $P(u)$ satisfies condition* (C) *from Corollary 2.12.*

Proof. Again by (1.9.4) we may suppose that the origin $o(e)$ of the arrow in question is equal to $v_0 = [L_0]$. So $e = e_{\overline{M}}$ with some nontrivial $\mathbb{F}$-subspace $\overline{M}$ of $L_0/\pi L_0$. As in Section 2.8 we use the same letter $\overline{M}$ for the corresponding linear subvariety of $\mathbb{P}^{r-1}/\mathbb{F}$ of dimension $t-1$, where $t = \mathrm{type}(e) = \dim \overline{M}$.

Multiplying u by suitable functions of type $\ell_{H,H'}$ (which doesn't alter the (non)validity of (C) for u), we may assume that $P(u)(e') = 0$ for all $e' \in \boldsymbol{A}_{v_0,1}$ dominated by e. Then we must show that $P(u)(e) = 0$ too. Let u be normalized such that $\|u\|_{v_0} = 1$, and let $\bar{u}$ be its reduction as a rational function on $\mathbb{P}^{r-1}/\mathbb{F}$, see (2.3.2).

If $P(u)(e) < 0$ then $|u|$ decays along $e = e_{\overline{M}}$ and $\bar{u}$ vanishes along $\overline{M}$. Correspondingly, if $P(u)(e) > 0$ then $(\bar{u})^{-1} = \overline{(u^{-1})}$ vanishes along $\overline{M}$. Hence it suffices to show that, under our assumptions, $\bar{u}$ restricts to a well-defined rational function on $\overline{M}$, i.e., $\overline{M}$ is neither contained in the vanishing locus $V(\bar{u})$ nor in $V(\bar{u}^{-1})$. But the latter is obvious: With a suitable constant $c \neq 0$ we have

$$\bar{u} = c \cdot \prod \ell_{\overline{H}}^{m(\overline{H})},$$

where $\overline{H}$ runs through the boundary components of $\overline{\Omega}_{v_0}$ as in (2.3.2), $\ell_{\overline{H}}$ is a linear form vanishing on $\overline{H}$, $\sum m(\overline{H}) = 0$, and $m(\overline{H}) = -P(u)(e_{\overline{H}}^{\perp}) = 0$ if $\overline{H}^{\perp} \subset \overline{M}$. Hence neither the rational function $\bar{u}$ nor its reciprocal vanishes identically on $\overline{M}$. □

3.2. The proposition motivates the following definition. Let A be any additively written abelian group. The group of *A-valued harmonic 1-cochains* $\mathbf{H}(\mathcal{BT}, A)$ is the group of maps $\varphi\colon \boldsymbol{A}(\mathcal{BT}) \to A$ that satisfy

(A) $\sum \varphi(e) = 0$, whenever e ranges through the arrows of a closed path in $\mathcal{BT}$;

(B) for each type t, $1 \le t \le r-1$, and each $v \in V(\mathcal{BT})$, the condition

$$\sum_{e \in \boldsymbol{A}_{v,t}} \varphi(e) = 0 \quad \text{holds;} \tag{B$_t$}$$

(C) for each $v \in \boldsymbol{V}(\mathcal{BT})$ and each $e \in \boldsymbol{A}_v$,

$$\sum_{e' \in \boldsymbol{A}_{v,1},\ e' \prec e} \varphi(e') = \varphi(e).$$

Remarks 3.3. (i) In the case where the coefficient group A equals $\mathbb{Z}$, condition (A) is (1.9.2), (B$_1$) is (1.8.1), and (C) is the condition dealt with in Corollary 2.12 and Proposition 3.1. (A) in particular implies that φ is alternating, i.e., $\varphi(\bar{e}) = -\varphi(e)$. Further, (B$_1$) together with (C) implies (B$_t$) for all types t, as

$$\sum_{e \in \boldsymbol{A}_{v,t}} \varphi(e) = \sum_{e' \in \boldsymbol{A}_{v,1}} \varphi(e') \#\{e \in \boldsymbol{A}_{v,t} \mid e' \prec e\},$$

where $\#\{\cdots\}$, the cardinality of some finite Grassmannian, is independent of e'.

(ii) Note that the current $\mathbf{H}(\mathcal{BT}, \mathbb{Z})$ differs from the group defined in [Gekeler 2017], as condition (C) is absent there.

(iii) Proposition 3.1 together with the preceding considerations shows that

$$\begin{aligned} P\colon \mathcal{O}(\Omega)^* &\to \mathbf{H}(\mathcal{BT}, \mathbb{Z}) \\ u &\mapsto P(u) \end{aligned}$$

is well-defined. Its kernel consists of the invertible holomorphic functions f on Ω with constant absolute value, which equals the constants C^*, as will be shown in Proposition 3.4. Hence, by (1.9.4), we have the exact sequence of $G(K)$-modules

$$1 \to C^* \to \mathcal{O}(\Omega)^* \xrightarrow{P} \mathbf{H}(\mathcal{BT}, \mathbb{Z}).$$

In fact, we will show that P is also surjective, which yields our principal result Theorem 3.11.

(iv) Beyond the natural coefficient domains $A = \mathbb{Z}$ or $\mathbb{Q}$ for $\mathbf{H}(\mathcal{BT}, A)$, at least the torsion groups $A = \mathbb{Z}/(N)$ deserve interest. For example, in the case $r = 2$ and $\mathrm{char}(C) = \mathrm{char}(\mathbb{F}) = p$, the invariants $\mathbf{H}(\mathcal{BT}, \mathbb{F}_p)^\Gamma$ under an arithmetic subgroup $\Gamma \subset G(K)$ differ in general from $\mathbf{H}(\mathcal{BT}, \mathbb{Z})^\Gamma \otimes \mathbb{F}_p$; see [Gekeler and Reversat 1996, Section 6]. The coefficient rings $A = \mathbb{Z}_\ell$ (ℓ a prime number) and $A = K$ come into the game by relating $\mathbf{H}(\mathcal{BT}, \mathbb{Z})$ with the first cohomology of Ω; see Section 5.5.

Proposition 3.4. *Each bounded holomorphic function f on Ω is constant. In particular, the kernel of the map P equals the constants C^*.*

Proof. Let $\boldsymbol{\omega} = (\omega_1 : \cdots : \omega_r)$ be an element of Ω. We are going to show that f as a function in ω_i, where $\omega_1, \dots, \omega_{i-1}, \omega_{i+1}, \dots, \omega_r$ are fixed, is constant for each i, which will give the result. Let

$$\alpha_{\boldsymbol{\omega}}^{(i)} : \mathbb{P}^{r-1}(K) \to \mathbb{P}^1(C) \quad \text{be the map} \quad (x_1 : \cdots : x_r) \mapsto \Bigl(\sum_{\substack{1 \le j \le r \\ i \ne j}} x_j \omega_j : x_i \Bigr).$$

It is well-defined (since the ω_j are K-linearly independent) and continuous with respect to the nonarchimedean topologies on both sides, whence its image $\mathrm{im}(\alpha_{\boldsymbol{\omega}}^{(i)})$ is compact. Moreover, the complement $\Omega_{\boldsymbol{\omega}}^{(i)} := \mathbb{P}^1(C) \setminus \mathrm{im}(\alpha_{\boldsymbol{\omega}}^{(i)})$ in $\mathbb{P}^1(C)$ equals the set of those $\omega \in C = \{(\omega : 1)\} \hookrightarrow \mathbb{P}^1(C)$ which are eligible for $(\omega_1 : \cdots : \omega_{i-1} : \omega : \omega_{i+1} : \cdots : \omega_r)$ to lie in Ω. Analytic spaces of this shape are extensively discussed in [Fresnel and van der Put 2004, Chapter II]. Notably, their Proposition 2.7.9 states that bounded functions on $\Omega_{\boldsymbol{\omega}}^{(i)}$ are constant as wanted. □

3.5. The strategy of proof of the surjectivity of P will be to approximate a given $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ by linear combinations of elements $P(u)$, where u is a function of type $\ell_{H,H'}$, or a relative of it.

Given two hyperplanes $H \ne H'$ of V and $n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$, define

$$f_{H,H',n} := 1 + \pi^n \ell_{H,H'}. \tag{3.5.1}$$

Here $\ell_{H,H'} = \ell_H/\ell_{H'} = \ell_{\boldsymbol{y}}/\ell_{\boldsymbol{y}'}$, where $\boldsymbol{y}, \boldsymbol{y}' \in L_0 \setminus \pi L_0$, $H = \ker(\ell_{\boldsymbol{y}})$, $H' = \ker(\ell_{\boldsymbol{y}'})$. Like $\ell_{H,H'}$, $f_{H,H',n}$ is a unit on Ω. We denote by

$$\mathcal{BT}(n) \subset \mathcal{BT} \tag{3.5.2}$$

the full subcomplex with vertices $\boldsymbol{V}(\mathcal{BT}(n)) = \{v \in \boldsymbol{V}(\mathcal{BT}) \mid d(v_0, v) \le n\}$. Hence $\mathcal{BT}(0) = \{v_0\}$, $\mathcal{BT}(1) = \mathrm{st}(v_0)$, etc. Further,

$$\Omega(n) := \lambda^{-1}(\mathcal{BT}(n)). \tag{3.5.3}$$

Then $\Omega(n)$ is an admissible affinoid subspace of Ω and $\Omega = \bigcup_{n \ge 0} \Omega(n)$. (In [Schneider and Stuhler 1991, Section 1, Proposition 4] $\Omega(n)$ is called $\overline{\Omega}_n$, and a system of generators of its affinoid algebra is constructed.)

Lemma 3.6. *For $n \in \mathbb{N}_0$, the following hold on $\Omega(n)$:*

(i) $\log \ell_{H,H'} \le n$.

(ii) $|f_{H,H',n}| = 1$.

Proof. (i) By our normalization, $|\ell_{H,H'}(\boldsymbol{\omega})| = 1$ for $\boldsymbol{\omega} \in \lambda^{-1}(v_0)$. Then by Proposition 2.11, $\|\ell_{H,H'}(\boldsymbol{\omega})\|_v \leq q^n$ for $v \in \boldsymbol{V}(\mathcal{BT})$ whenever $d(v_0, v) \leq n$, which gives the assertion.

(ii) $|f_{H,H',n}(\boldsymbol{\omega})| = |1 + \pi^n \ell_{H,H'}(\boldsymbol{\omega})| \leq 1$ on $\Omega(n)$ by (i), with equality at least if $n = 0$ or $\boldsymbol{\omega}$ doesn't belong to $\lambda^{-1}(v)$, where v is a vertex with $d(v_0, v) = n$, since in this case $\log \ell_{H,H'}(\boldsymbol{\omega}) < n$. But the equality must also hold for $\boldsymbol{\omega}$ with $\lambda(\boldsymbol{\omega}) =$ such a v, due to the linear interpolation property Theorem 1.6 of $\log_q \|f_{H,H',n}\|_x$ for x belonging to an arrow $e = (v', v)$ with $d(v_0, v') = n - 1$. □

Definition 3.7. A vertex $v \in \boldsymbol{V}(\mathcal{BT})$ is called *n-special* ($n \in \mathbb{N}_0$) if there exists a (necessarily uniquely determined) path $(v_0, v_1, \ldots, v_n = v) \in \boldsymbol{A}_{v_0,1,n}$, i.e., the arrows $e_i = (v_{i-1}, v_i)$, $i = 1, 2, \ldots, n$ all have type 1, and $d(v_0, v) = n$. (By definition, v_0 is 0-special.) An arrow $e \in \boldsymbol{A}(\mathcal{BT})$ is *n-special* ($n \in \mathbb{N}$) if $o(e)$ is $(n-1)$-special and $t(e)$ is n-special, that is, if it appears as some e_n as above. Also, the path $(v_0, \ldots, v_n) = (e_1, \ldots, e_n)$ is called *n-special*. An arrow e with $d(v_0, o(e)) = n$ is *inbound* (of level n) if it belongs to $\mathcal{BT}(n)$, and *outbound* otherwise. That is, e is inbound $\Leftrightarrow d(v_0, t(e)) \leq n$.

3.8. Next, we describe the restriction of $P(f_{H,H',n})$ to $(n+1)$-special arrows e. Let $n \in \mathbb{N}_0$, and choose hyperplanes H, H' of V, given as $H = \ker(\ell_{\boldsymbol{y}})$, $H' = \ker(\ell_{\boldsymbol{y}'})$ as in (3.5.1), $G = H^\perp = K\boldsymbol{y}$, $G' = K\boldsymbol{y}'$. Assume that $\boldsymbol{y}$ and $\boldsymbol{y}'$ are not 1-equivalent (2.7.2), that is, $\tau_G(v_0) \neq \tau_{G'}(v_0)$.

(i) According to Corollary 2.10, $\ell_{H,H'} = \ell_{\boldsymbol{y}}/\ell_{\boldsymbol{y}'}$ has the property that $\log \ell_{H,H'}$ grows by 1 in each step of the $(n+1)$-special path

$$(v_0, v_1, \ldots, v_n, v_{n+1}) = (e_1, e_2, \ldots, e_{n+1}) \tag{3.8.1}$$

from v_0 toward G'. Together with Lemma 3.6(ii), this implies that $P(f_{H,H',n})(e_{n+1}) = 1$.

(ii) On the other hand, again by Corollary 2.10, $\log \ell_{H,H'} < n$ on $\lambda^{-1}(v)$ for each n-special v different from v_n. By a variation of the linear interpolation argument in the proof of Lemma 3.6(ii), $P(f_{H,H',n})(e) = 0$ for each $(n+1)$-special arrow e with $o(e) \neq v_n$.

(iii) The function $u := f_{H,H',n} = (\ell_{\boldsymbol{y}'} + \pi^n \ell_{\boldsymbol{y}})/\ell_{\boldsymbol{y}'}$ satisfies $\|u\|_{v_n} = 1$. Its reduction $\bar{u}$ as a rational function on the reduction

$$\overline{\Omega}_{v_n} = (\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}) \setminus \bigcup \overline{H} \quad \text{(see (2.5.1), here } v_n = [L]) \tag{3.8.2}$$

of $\Omega_{v_n} = \lambda^{-1}(v_n)$ has a simple pole along the hyperplane $\overline{H}_{e_{n+1}}$ corresponding to the arrow e_{n+1}, a simple zero along a unique $\overline{H}_e$, where $e = (v_n, w)$, and neither zeroes nor poles along the other hyperplanes that appear in (3.8.2). The hyperplane $\overline{H}_e$ is the vanishing locus in $\mathbb{P}(L^\wedge/\pi L^\wedge)/\mathbb{F}$ of the reduction of the form $\ell_{\boldsymbol{y}'} + \pi^n \ell_{\boldsymbol{y}} = \ell_{\boldsymbol{y}''}$; accordingly, $w = \tau_{G''}(v_n)$, where $G'' = K\boldsymbol{y}''$ and

$$\boldsymbol{y}'' = \boldsymbol{y}' + \pi^n \boldsymbol{y}. \tag{3.8.3}$$

(iv) If $\boldsymbol{y}'$ is fixed and $\boldsymbol{y}$ runs through the elements of $L_0 \setminus \pi L_0$ not 1-equivalent with $\boldsymbol{y}'$, then the corresponding $\boldsymbol{y}''$ are n-equivalent but not $(n+1)$-equivalent with $\boldsymbol{y}'$ (see (2.7.2)). In this way we get all the $(n+1)$-classes with this property, that is, all the $(n+1)$-special paths $(e_1, e_2, \dots, e_n, e)$ which agree with the path $(e_1, \dots, e_n, e_{n+1})$ of (3.8.1) except for the last arrow. We collect what has been shown.

Proposition 3.9. (i) *Let H, H' be two hyperplanes in V, $G = H^{\perp}$, $G' = H'^{\perp}$, with $\tau_G(v_0) \neq \tau_{G'}(v_0)$ and $n \in \mathbb{N}_0$. Put $v_i := (\tau_{G'})^i(v_0)$. If e is an $(n+1)$-special arrow then*

$$P(f_{H,H',n})(e) = \begin{cases} +1 & \text{if } e = (v_n, v_{n+1}), \\ -1 & \text{if } e = (v_n, w), \\ 0 & \text{otherwise.} \end{cases} \tag{3.9.1}$$

Here $w = \tau_{G''}(v_n) \neq v_{n+1}$, where $G'' = K\mathbf{y}''$ with $\mathbf{y}'' = \mathbf{y}' + \pi^n \mathbf{y}$ as described in Section 3.8, notably in (3.8.3).

(ii) *If H' is fixed, each $(n+1)$-special arrow $e \neq (v_n, v_{n+1})$ with $o(e) = v_n$ occurs through a suitable choice of H as the arrow $e = (v, w)$ where $P(f_{H,H',n})$ evaluates to -1.* □

The next result, technical in nature, is crucial for the proof of Theorem 3.11. Its proof is postponed to the next section.

Proposition 3.10. *Let $n \in \mathbb{N}_0$ and $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ be such that $\varphi(e) = 0$ for arrows e that either belong to $\mathcal{BT}(n)$ or are $(n+1)$-special. Then $\varphi(e) = 0$ for all arrows e of $\mathcal{BT}(n+1)$.*

Now we are able to show (modulo Proposition 3.10) the principal result.

Theorem 3.11. *The van der Put map $P\colon \mathcal{O}(\Omega)^* \to \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ is surjective, and so the sequence*

$$1 \to C^* \to \mathcal{O}(\Omega)^* \to \mathbf{H}(\mathcal{BT}, \mathbb{Z}) \to 0 \tag{0.2}$$

is a short exact sequence of $G(K)$-modules.

Proof. (i) Let $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ be given. By successively subtracting $P(u_n)$ from φ, where $(u_n)_{n \in \mathbb{N}}$ is a suitable series of functions in $\mathcal{O}(\Omega)^*$ with $u_n \to 1$ locally uniformly (i.e., uniformly on affinoids) we will achieve that

$$\varphi - P\Big(\prod_{1 \le i \le n} u_i\Big) \equiv 0 \quad \text{on } \mathcal{BT}(n).$$

Then $\varphi = P(u)$, where $u = \lim_{n\to\infty} \prod_{1 \le i \le n} u_i$ is the limit function.

(ii) From condition (B_1) for φ and Proposition 2.9 we find a function u_1, namely a suitable finite product of functions of type $\ell_{H,H'}$, such that $(\varphi - P(u_1))(e) = 0$ for each $e \in \boldsymbol{A}_{v_0,1}$. By condition (C), $\varphi - P(u_1)$ vanishes on all $e \in \boldsymbol{A}_{v_0}$, and thus by (A) on all e that belong to $\mathcal{BT}(1) = \mathrm{st}(v_0)$.

(iii) Suppose that $u_1, \dots, u_n \in \mathcal{O}(\Omega)^*$ are constructed ($n \in \mathbb{N}$) such that for $1 \le i \le n$

(a) $P(u_i) \equiv 0$ on $\mathcal{BT}(i-1)$,

(b) $u_i \equiv 1 \pmod{\pi^{[(i-1)/2]}}$ on $\mathcal{BT}([(i-1)/2])$, here $[\cdot]$ is the Gauss bracket,

(c) $\varphi - P(\prod_{1 \le i \le n} u_i) \equiv 0$ on $\mathcal{BT}(n)$

hold. (Condition (a) is empty for $i = 1$ and therefore trivially fulfilled.) We are going to construct u_{n+1} such that $u_1, \dots, u_{n+1}$ fulfill the conditions on level $n+1$.

(iv) From (c) and (B_1) we have for n-special vertices v and $\psi := \varphi - P(\prod_{1 \le i \le n} u_i) \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$:

$$\sum_{e \in \boldsymbol{A}_{v,1} \text{ outbound}} \psi(e) = \sum_{e \in \boldsymbol{A}_{v,1}} \psi(e) = 0.$$

(v) According to Proposition 3.9, we find u_{n+1}, viz, a suitable product of functions $f_{H,H',n}$, such that

$$(\psi - P(u_{n+1}))(e) = \Big(\varphi - P\Big(\prod_{1 \le i \le n+1} u_i\Big)\Big)(e) = 0$$

on all $(n+1)$-special arrows e. Furthermore, that u_{n+1} (like the functions $f_{H,H',n}$, see Lemma 3.6(ii)) satisfies $P(u_{n+1}) \equiv 0$ on $\mathcal{BT}(n)$, i.e., condition (a), and condition (b): $u_{n+1} \equiv 1 \pmod{\pi^{[n/2]}}$ on $\mathcal{BT}([n/2])$. Hence $\varphi - P(\prod_{1 \le i \le n+1} u_i)$ vanishes on arrows which belong to $\mathcal{BT}(n)$ or are $(n+1)$-special. Using Proposition 3.10, we see that $\varphi - P(\prod_{1 \le i \le n+1} u_i)$ vanishes on $\mathcal{BT}(n+1)$. That is, conditions (a), (b), (c) hold for $u_1, \dots, u_{n+1}$, and we have inductively constructed an infinite series $u_1, u_2, \dots$ with (a), (b) and (c) for all n.

(vi) It follows from (b) that the infinite product

$$u = \prod_{i \in \mathbb{N}} u_i$$

is normally convergent on each $\Omega(n)$ and thus defines a holomorphic invertible function u on Ω. Its van der Put transform $P(u)$ restricted to $\mathcal{BT}(n)$ depends only on $u_1, \dots, u_n$, due to (c), and thus agrees with φ reduced to $\mathcal{BT}(n)$. Therefore, $\varphi = P(u)$, and the result is shown. □

4. The group $\mathbf{H}(\mathcal{BT}, \mathbb{Z})$

4.1.

Proof of Proposition 3.10. (i) The requirements of Proposition 3.10 for $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ on level $n \in \mathbb{N}_0$ will be labeled by R(n).

(ii) Suppose that R(n) holds for φ. Then φ vanishes on all arrows $\boldsymbol{A}_{v,1}$ whenever v is n-special, since such an e is either $(n+1)$-special or belongs to $\mathcal{BT}(n)$. Hence by conditions (C) and (A) of Section 3.2, $\varphi(e) = 0$ whenever e is contiguous with v, i.e., if e belongs to st(v). This shows, in particular, that Proposition 3.10 holds for $n = 0$.

(iii) Let $v \in \boldsymbol{V}(\mathcal{BT})$ have distance $d(v_0, v) = n$, but be not necessarily n-special. For the same reason as in (ii), φ vanishes identically on st(v) if it vanishes on all outbound arrows $e \in \mathrm{A}_{v,1}$. Hence it suffices to show

$$\varphi(e) = 0 \quad \text{for outbound arrows } e \text{ of type 1 and level } n. \tag{O}$$

(iv) For a vertex v with $d(v_0, v) = n$, we let $s(v)$ be the distance to the next $w \in \boldsymbol{V}(\mathcal{BT})$ which is n-special. We are going to show assertion (O) by induction on $s(o(e))$.

(v) By R(n), (O) holds if $s = s(o(e)) = 0$, i.e., if $o(e)$ is n-special. Therefore, suppose that $s > 0$. By the preceding we are reduced to showing:

$$\begin{array}{l}\text{Let } e \text{ be an outbound arrow of type 1, level } n, \text{ and with } s = s(o(e)) > 0.\\ \text{Then } e \text{ belongs to } \mathrm{st}(\tilde{v}), \text{ where } d(v_0, \tilde{v}) = n \text{ and } s(\tilde{v}) < s.\end{array} \tag{P}$$

(vi) We reformulate (P) in lattice terms. Representing $v_0 = [L_0]$ through $L_0 = O^r$, the vertices $v \in \boldsymbol{V}(\mathcal{BT})$ correspond one-to-one to sublattices L of full rank r which satisfy $L \subset L_0$, $L \not\subset \pi L_0$. For such a vertex v or its lattice L, we let $(n_1, n_2, \dots, n_r)$ with $n_1 \geq n_2 \geq \dots \geq n_r = 0$ be the sequence of elementary divisors (*sed*) of L_0/L ($n_r = 0$ as $L \not\subset \pi L_0$). That is,

$$L_0/L \cong O/(\pi^{n_1}) \times \dots \times O/(\pi^{n_r}).$$

Then $n_1 = d(v_0, v)$, and v is n-special if and only if its sed is $(n, \dots, n, 0)$.

(vii) Let $e = (v, v')$ be given as required for (P), $v = [L]$, $v' = [L']$, where $\pi^{n+1} L_0 \subset L' \subset L \subset L_0$. Let $(n_1 = n, n_2, \dots, n_r)$ be the sed of L_0/L. Then, as $\dim_{\mathbb{F}}(L/L') = r - 1$ and $d(v_0, v') = n + 1$, $(n'_1 = n+1, \dots, n'_{r-1} = n_{r-1}+1, n_r)$ is the sed of L_0/L'. This means that L_0 has an ordered O-basis $\{x_1, \dots, x_r\}$ such that $\{\pi^{n+1} x_1, \pi^{n_2+1} x_2, \dots, \pi^{n_{r-1}+1} x_{r-1}, \pi^{n_r} x_r\}$ is a basis of L' and $\{\pi^n x_1, \pi^{n_2} x_2, \dots, \pi^{n_r} x_r\}$ is a basis of L. Assume that k with $1 \leq k \leq r - 1$ is minimal with $n_{r-1} = n_k$. Let M be the sublattice of L_0 with basis $\{\pi^n x_1, \dots, \pi^n x_{r-1}, x_r\}$. Then $w = [M]$ is n-special and $s(v) = d(v, w) = n - n_{r-1}$, which by assumption is positive. Put $\tilde{L}$ for the lattice with basis

$$\{\pi^n x_1, \pi^{n_2} x_2, \dots, \pi^{n_{k-1}} x_{k-1}, \pi^{n_k+1} x_k, \pi^{n_{k+1}+1} x_{k+1}, \dots, \pi^{n_{r-1}+1} x_{r-1}, \pi^{n_r} x_r\}.$$

The vertex $\tilde{v} := [\tilde{L}]$ satisfies

$$d(v_0, \tilde{v}) = n, \quad d(v, \tilde{v}) = 1 = d(v', \tilde{v}) \quad \text{and} \quad s(\tilde{v}) = d(w, \tilde{v}) = n - n_{r-1} - 1 = s(v) - 1. \tag{4.1.1}$$

Hence $e = (v, v')$ belongs to $\mathrm{st}(\tilde{v})$, where $\tilde{v}$ is as wanted for assertion (P).

This finishes the proof of Proposition 3.10. □

Corollary 4.2. *Let $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ be such that $\varphi(e) = 0$ for all i-special arrows e, where $1 \leq i \leq n$. Then $\varphi \equiv 0$ on $\mathcal{BT}(n)$.*

Proof. This follows by induction from Proposition 3.10. □

4.3. Next we give a different description of $\mathbf{H}(\mathcal{BT}, \mathbb{Z})$, see Theorem 4.16. Let v be an n-special vertex ($n \geq 1$), v^* its predecessor on the uniquely determined n-special path $(v_0, v_1, \dots, v_{n-1} = v^*, v)$ from v_0 to v, and e^* the n-special arrow (v^*, v). Its inverse $\overline{e^*} = (v, v^*)$ belongs to $\boldsymbol{A}_{v,r-1}$.

Lemma 4.4. *In the given situation, $e \in \boldsymbol{A}_{v,1}$ is inbound if and only if $e \prec \overline{e^*}$.*

Proof. As the stabilizer $\mathrm{GL}(r, O)$ of $L_0 = O^r$ acts transitively on n-special vertices or arrows, we may suppose that $v = [L_n]$, where L_n is the O-lattice with basis $\{\pi^n x_1, \dots, \pi^n x_{r-1}, x_r\}$, and thus $v^* = [L_{n-1}]$. (Here $\{x_1, \dots, x_r\}$ is the standard basis of L_0.) Under (2.1.4), $\overline{e^*}$ corresponds to the $(r-1)$-dimensional subspace $\pi L_{n-1}/\pi L_n$ of the r-dimensional $\mathbb{F}$-space $L_n/\pi L_n$, which has the $(\overline{\pi^n x_i}) = \pi^n x_i \pmod{\pi L_n}$, $1 \le i < r$, as a basis. Let $\overline{G}$ be a line in $L_n/\pi L_n$ with preimage G in L_n, and let $e_{\overline{G}} = (v, v_{\overline{G}})$ be the arrow of type 1 determined by $\overline{G}$. Then $v_{\overline{G}} = [G]$ and

$$e_{\overline{G}} \prec e^* \Leftrightarrow \overline{G} \subsetneq \pi L_{n-1}/\pi L_n \Leftrightarrow G \subset \pi L_{n-1}.$$

If this is the case, $\pi^n L_0 \subset L_n \subset \pi^{-1} G \subset L_{n-1} \subset L_0$, that is, $d(v_0, [G]) \le n$, and $e_{\overline{G}}$ is inbound. On the other hand, if $G \not\subset \pi L_{n-1}$, then $\pi^{-1} G \not\subset L_0$. Since $\pi^n L_0 \not\subset G$, we then have $d(v_0, [G]) = n+1$, and $e_{\overline{G}}$ is outbound. $\square$

4.5. We may now reformulate condition (B_1) for $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ at the n-special vertex v of level $n \ge 1$ as follows: Splitting

$$\boldsymbol{A}_{v,1} = \boldsymbol{A}_{v,1,\mathrm{in}} \,\dot\cup\, \boldsymbol{A}_{v,1,\mathrm{out}} \tag{4.5.1}$$

into the subsets of inbound / outbound arrows (note that $e \in \boldsymbol{A}_{v,1}$ is outbound if and only if it is $(n+1)$-special), (B_1) reads

$$0 = \sum_{e \in \boldsymbol{A}_{v,1}} \varphi(e) = \sum_{e \in \boldsymbol{A}_{v,1,\mathrm{in}}} \varphi(e) + \sum_{e \in \boldsymbol{A}_{v,1,\mathrm{out}}} \varphi(e) = \varphi(\overline{e^*}) + \sum_{e \in \boldsymbol{A}_{v,1,\mathrm{out}}} \varphi(e)$$

(where we used Lemma 4.4 and condition (C) for $\varphi(\overline{e^*})$), i.e., as the flow condition

$$\varphi(e^*) = \sum_{e \in \boldsymbol{A}_{v,1,\mathrm{out}}} \varphi(e). \tag{4.5.2}$$

The number of terms in the sum is

$$\#\boldsymbol{A}_{v,1,\mathrm{out}} = \#\boldsymbol{A}_{v,1} - \#\boldsymbol{A}_{v,1,\mathrm{in}} = \#\mathbb{P}^{r-1}(\mathbb{F}) - \#\mathbb{P}^{r-2}(\mathbb{F}) = q^{r-1}. \tag{4.5.3}$$

4.6. Let $\mathcal{T}_{v_0}$ be the full subcomplex of $\mathcal{BT}$ composed of the n-special vertices ($n \in \mathbb{N}_0$) along with the 1-simplices connecting them. In other words, $\mathcal{T}_{v_0}$ is the union of the paths $\boldsymbol{A}_{v_0,1,n}$, where $n \in \mathbb{N}$, see Section 2.1.6. It is connected, one-dimensional and cycle-free, hence a tree. The valence (= number of neighbors) of v_0 is $\#\mathbb{P}^{r-1}(\mathbb{F}) = (q^r - 1)/(q-1)$, the valence of each other vertex $v \ne v_0$ is $q^{r-1}+1$, as we read off from (4.5.3). Let further $\mathcal{T}_{v_0}(n) := \mathcal{T}_{v_0} \cap \mathcal{BT}(n)$.

4.6.1. We define $\mathbf{H}(n)$ as the image of $\mathbf{H}(\mathcal{BT}, \mathbb{Z})$ in $\{\varphi \colon \boldsymbol{A}(\mathcal{BT}(n)) \to \mathbb{Z}\}$ obtained by restriction. Hence

$$\mathbf{H}(\mathcal{BT}, \mathbb{Z}) = \varprojlim_{n \in \mathbb{N}} \mathbf{H}(n). \tag{4.6.2}$$

Put further

$$\mathbf{H}'(n) := \{\varphi\colon \boldsymbol{A}(\mathcal{T}_{v_0}(n)) \to \mathbb{Z} \mid \varphi \text{ is subject to (4.6.4) and (4.6.5)}(v) \text{ for each } i\text{-special } v,\ 0 \le i < n\}. \tag{4.6.3}$$

Here $\boldsymbol{A}(\mathcal{S})$ is the set of arrows (oriented 1-simplices) of the simplicial complex $\mathcal{S}$, and the conditions are

$$\varphi(e) + \varphi(\bar{e}) = 0 \quad \text{for each arrow } e \text{ with inverse } \bar{e}; \tag{4.6.4}$$

$$\sum_{\substack{e \in \boldsymbol{A}(\mathcal{T}_{v_0}) \\ o(e)=v}} \varphi(e) = 0. \tag{4.6.5)(v}$$

4.7. Equality (4.5.2) together with the condition (B_1) at v_0 states that the restriction of $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ to $\mathcal{T}_{v_0}(n)$ is an element of $\mathbf{H}'(n)$. Therefore, restriction defines homomorphisms $r_n\colon \mathbf{H}(n) \to \mathbf{H}'(n)$, which make the diagram (with natural maps q_n, q'_n)

$$\begin{tikzcd} \mathbf{H}(n+1) \arrow[r, "r_{n+1}"] \arrow[d, "q_n"] & \mathbf{H}'(n+1) \arrow[d, "q'_n"] \\ \mathbf{H}(n) \arrow[r, "r_n"] & \mathbf{H}'(n) \end{tikzcd} \tag{4.7.1}$$

commutative. Note that both q_n and q'_n are surjective, the first by definition, the second one since $\mathcal{T}_{v_0}$ is a tree. Corollary 4.2 may be rephrased as

Proposition 4.8. *r_n is injective for $n \in \mathbb{N}$.* □

Lemma 4.9. *r_n is also surjective.*

Proof. For $n = 1$, this is implicit in the proof of Theorem 3.11 (i.e., one may arbitrarily prescribe the value of $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$ on $e \in \boldsymbol{A}_{v_0,1}$, subject only to ($\mathrm{B}_1$) at v_0).

For $n \ge 1$, let Q_{n+1} (respectively Q'_{n+1}) be the kernel of q_n (respectively q'_n). Then $r_{n+1}(Q_{n+1}) \subset Q'_{n+1}$, and we have the commutative diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & Q_{n+1} \arrow[r] \arrow[d] & \mathbf{H}(n+1) \arrow[r] \arrow[d, "r_{n+1}"] & \mathbf{H}(n) \arrow[r] \arrow[d, "r_n"] & 0 \\ 0 \arrow[r] & Q'_{n+1} \arrow[r] & \mathbf{H}'(n+1) \arrow[r] & \mathbf{H}'(n) \arrow[r] & 0. \end{tikzcd}$$

By induction hypothesis, r_n is surjective, so the surjectivity of r_{n+1} is implied by

$$r_{n+1}(Q_{n+1}) = Q'_{n+1}. \tag{$*$}$$

But

$$Q_{n+1} = \{\varphi \in \mathbf{H}(n+1) \mid \varphi \equiv 0 \text{ on } \mathcal{BT}(n)\} \quad \text{and} \quad Q'_{n+1} = \{\varphi \in \mathbf{H}'(n+1) \mid \varphi \equiv 0 \text{ on } \mathcal{T}_{v_0}(n)\},$$

so ($*$) follows from the existence of sufficiently many elements of Q_{n+1} (e.g., the classes in $\mathbf{H}(n+1)$ of the $P(f_{H,H',n})$) which have sufficiently independent values on the arrows in $\mathcal{T}_{v_0}(n+1)$ not in $\mathcal{T}_{v_0}(n)$. See also the proof of Theorem 3.11, steps (iv) and (v). □

4.10. Let $\mathbf{H}(\mathcal{T}_{v_0}, \mathbb{Z}) = \varprojlim_{n\in\mathbb{N}} \mathbf{H}'(n)$ be the group of functions $\varphi\colon \boldsymbol{A}(\mathcal{T}_{v_0}) \to \mathbb{Z}$ which satisfy (4.6.4) and (4.6.5)(v) for all vertices v of $\mathcal{T}_{v_0}$. Similarly, we define $\mathbf{H}(\mathcal{T}_{v_0}, A)$ for an arbitrary abelian group A instead of $\mathbb{Z}$. That is, elements of $\mathbf{H}(\mathcal{T}_{v_0}, A)$ are characterized by conditions analogous with (A) and (B) of Section 3.2, while (C) is not applicable. Putting together the considerations of Section 4.5 with Proposition 4.8 and Lemma 4.9, we find

$$\mathbf{H}(\mathcal{BT}, \mathbb{Z}) \xrightarrow{\cong} \mathbf{H}(\mathcal{T}_{v_0}, \mathbb{Z}), \tag{4.11}$$

where the canonical isomorphism is given by restricting $\varphi \in \mathbf{H}(\mathcal{BT}, \mathbb{Z})$, $\varphi\colon \boldsymbol{A}(\mathcal{BT}) \to \mathbb{Z}$ to the subset $\boldsymbol{A}(\mathcal{T}_{v_0})$ of $\boldsymbol{A}(\mathcal{BT})$.

In what follows, A is an arbitrary abelian group. The next result is a consequence of the above.

Proposition 4.12. *Restriction to the arrows of $\mathcal{T}_{v_0}$ yields an isomorphism*

$$\mathbf{H}(\mathcal{BT}, A) \xrightarrow{\cong} \mathbf{H}(\mathcal{T}_{v_0}, A). \tag{4.12.1}$$

Proof. It suffices to observe that the preceding results Proposition 3.10, Corollary 4.2, Proposition 4.8, and Lemma 4.9 remain valid — with identical proofs — for A-valued functions instead of $\mathbb{Z}$-valued functions. □

4.13. Recall that an A-valued distribution on a compact totally disconnected topological space X is a map $\delta\colon U \mapsto \delta(U) \in A$ from the set of compact-open subspaces U of X to A which is additive in finite disjoint unions. We call $\delta(U)$ the *volume* of U *with respect to* δ. The *total mass* (or volume) of δ is $\delta(X)$.

We apply this to the situation (see Section 2.1.6, (2.1.7) and (2.1.8)) where

$$X = \mathrm{Gr}_{K,1}(V) = \{\text{lines } G \text{ of the } K\text{-space } V\} = \mathbb{P}(V). \tag{4.13.1}$$

As we have identified V with its dual $V^\wedge$ through the bilinear form $\langle\cdot,\cdot\rangle$, we also have an identification

$$\mathrm{Gr}_{K,1}(V) = \mathbb{P}(V) \xrightarrow{\cong} \mathbb{P}(V^\wedge) = \mathrm{Gr}_{K,r-1}(V)$$

given by $G \mapsto G^\perp$. Hence we could state the following assertions concerning distributions on $\mathbb{P}(V)$ for distributions on $\mathbb{P}(V^\wedge)$.

4.13.2. Let $\mathbf{D}(\mathbb{P}(V), A)$ be the group of A-valued distributions on $\mathbb{P}(V)$ with subgroup $\mathbf{D}^0(\mathbb{P}(V), A)$ of distributions with total mass 0. By (2.1.8), the sets $\mathbb{P}(V)(e) = \mathrm{Gr}_{K,1}(e)$, where e runs through the outbound arrows of $\boldsymbol{A}_{v_0,1,n}$ ($n \in \mathbb{N}$), i.e., through the set

$$\boldsymbol{A}^+(\mathcal{T}_{v_0}) = \{e \in \boldsymbol{A}(\mathcal{T}_{v_0}) \mid e \text{ oriented away from } v_0\}, \tag{4.13.3}$$

form a basis for the topology on $\mathbb{P}(V)$. Therefore, an element δ of $\mathbf{D}(\mathbb{P}(V), A)$ is an assignment

$$\delta\colon \boldsymbol{A}^+(\mathcal{T}_{v_0}) \to A$$

(where we interpret $\delta(e)$ as the volume of $\mathbb{P}(V)(e)$ with respect to δ) subject to the requirement

$$\delta(e^*) = \sum_{\substack{e \in \boldsymbol{A}^+(\mathcal{T}_{v_0}) \\ o(e)=t(e^*)}} \delta(e) \tag{4.13.4}$$

for each $e^* \in \boldsymbol{A}^+(\mathcal{T}_{v_0})$. The total mass of δ is

$$\delta(\mathbb{P}(V)) = \sum_{\substack{e \in \boldsymbol{A}^+(\mathcal{T}_{v_0}) \\ o(e)=v_0}} \delta(e) = \sum_{e \in \boldsymbol{A}_{v_0,1}} \delta(e). \tag{4.13.5}$$

In view of (4.5.2) and (4.6.5)(v), we find that

$$\mathbf{D}^0(\mathbb{P}(V), A) \xrightarrow{\cong} \mathbf{H}(\mathcal{T}_{v_0}, A), \tag{4.14}$$

where some $\delta\colon \boldsymbol{A}^+(\mathcal{T}_{v_0}) \to A$ in the left hand side is completed to a map on $\boldsymbol{A}(\mathcal{T}_{v_0})$ by (4.6.4), i.e., by $\varphi(\bar{e}) = -\varphi(e)$.

While both isomorphisms in (4.11) (or (4.12.1)) and (4.14) fail to be $G(K)$-equivariant (as $G(K)$ fixes neither v_0 nor $\mathcal{T}_{v_0}$), the resulting isomorphism

$$\begin{aligned} \mathbf{H}(\mathcal{BT}, \boldsymbol{A}) &\xrightarrow{\cong} \mathbf{D}^0(\mathbb{P}(V), A) \\ \varphi &\longmapsto \tilde{\varphi} \end{aligned} \tag{4.15}$$

is. Here the distribution $\tilde{\varphi}$ evaluates on $\mathbb{P}(V)(e)$ as $\varphi(e)$ whenever e is an arrow of $\mathcal{BT}$ of type 1 and $\mathbb{P}(V)(e)$ is the compact-open subset of lines G of V such that e points to G.

We summarize what has been shown.

Theorem 4.16. *Let A be an arbitrary abelian group. Restricting the evaluation of $\varphi \in \mathbf{H}(\mathcal{BT}, A)$ to arrows of $\mathcal{T}_{v_0}$ (resp. arrows of type 1 of $\mathcal{BT}$) yields canonical isomorphisms*

$$\mathbf{H}(\mathcal{BT}, A) \xrightarrow{\cong} \mathbf{H}(\mathcal{T}_{v_0}, A) \quad \textit{resp.} \quad \mathbf{H}(\mathcal{BT}, A) \xrightarrow{\cong} \mathbf{D}^0(\mathbb{P}(V), A).$$

The second of these is equivariant for the natural actions of $G(K) = \mathrm{GL}(r, K)$ on both sides, while the first isomorphism is equivariant for the actions of the stabilizer $G(O)Z(K)$ of $v_0 \in G(K)$.

As a direct consequence of the first isomorphism, i.e., of (4.12.1), we find the following corollary, which is in keeping with the fact that bounded holomorphic functions on Ω are constant (see Proposition 3.4).

Corollary 4.17. *If $\varphi \in \mathbf{H}(\mathcal{BT}, A)$ has finite support, it vanishes identically.*

Proof. Suppose that φ has support in $\mathcal{BT}(n)$ with $n \in \mathbb{N}$. Then its restriction to $\mathcal{T}_{v_0}(n+1)$ satisfies (4.6.4) and (4.6.5)(v) at all vertices v of $\mathcal{T}_{v_0}(n+1)$. As $\mathcal{T}_{v_0}(n+1)$ is a finite tree, this forces φ to vanish identically on $\mathcal{T}_{v_0}(n+1)$, thus on $\mathcal{BT}$. $\square$

5. Concluding remarks

5.1. Ehud de Shalit [2001, Section 3.1] postulated four conditions $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$ for what he calls harmonic k-cochains on $\mathcal{BT}$. These conditions specialized to $k = 1$ are essentially our conditions (A), (B), (C) from Section 3.2. Grosso modo, de Shalit's $\mathfrak{B}$ corresponds to (B), $\mathfrak{C}$ to (C) and $\mathfrak{D}$ to (A), while $\mathfrak{A}$ is a special case of (A).

5.2. In fact, the relationship with de Shalit's work is as follows. Suppose that $\mathrm{char}(K) = 0$, and consider the diagram

$$\begin{array}{ccccc} u & & \mathcal{O}(\Omega)^* & \xrightarrow{\quad P \quad} & \mathbf{H}(\mathcal{BT}, \mathbb{Z}) \\ \downarrow & & \downarrow & & \downarrow \\ d\log u = u^{-1}du & & \{\text{closed 1-forms on } \Omega\} & \xrightarrow{\text{res}} & \mathbf{H}(\mathcal{BT}, K)(= C^1_{\text{har}} \text{ of [de Shalit 2001]}), \end{array} \tag{5.2.1}$$

where "res" is de Shalit's residue mapping. Its commutativity follows for $u = \ell_{H,H'}$ from Corollary 7.6 and Theorem 8.2 of [de Shalit 2001] (along with the explanations given there, and our description of $P(u)$), and may be verified for general u by approximating. Hence the van der Put transform P yields a concrete description of the residue mapping on logarithmic 1-forms. On the other hand, in characteristic p the van der Put transform is finer than "$d\log$", as the latter kills all p-powers.

5.3. Now suppose that $\mathrm{char}(K) = p > 0$, and that moreover $r = 2$. Then $\mathcal{BT}$ is the Bruhat–Tits tree $\mathcal{T}$, and the residue mapping

$$\mathrm{res}\colon \{\text{1-forms on } \Omega = \Omega^2\} \to \mathbf{H}(\mathcal{T}, C)$$

(see [Gekeler and Reversat 1996, 1.8]) is such that the diagram analogous with (5.2.1)

$$\begin{array}{ccccc} u & & \mathcal{O}(\Omega)^* & \xrightarrow{P} & \mathbf{H}(\mathcal{T}, \mathbb{Z}) \\ \downarrow & & \downarrow & & \downarrow \\ d\log u & & \{\text{1-forms on } \Omega\} & \xrightarrow{\text{res}} & \mathbf{H}(\mathcal{T}, C) \end{array} \tag{5.3.1}$$

commutes, with remarkable arithmetic consequences [loc. cit., Sections 6 and 7]. A similar residue map for $r > 2$ unfortunately lacks so far. In any case, we should regard P as a substitute for the logarithmic derivation operator

$$u \mapsto d\log u = u^{-1}du$$

in characteristic 0.

5.4. Peter Schneider and Ulrich Stuhler [1991] described the cohomology $H^*(\Omega, A)$ of $\Omega = \Omega^r$ with respect to an abstract cohomology theory, where $A = H^0(\mathrm{Sp}(K))$. That theory is required to satisfy four natural axioms, [loc. cit., Section 2]. As they explain, these axioms are fulfilled at least

- for the étale ℓ-adic cohomology of rigid-analytic spaces over K, where ℓ is a prime different from $p = \mathrm{char}(\mathbb{F})$, and $A = \mathbb{Z}_\ell$, and
- for the de Rham cohomology (where one must moreover assume that $\mathrm{char}(K) = 0$); here $A = K$.

Their result is stated [loc. cit., Section 3, Theorem 1], which in dimension 1 is (in our notation)

$$H^1(\Omega^r, A) \xrightarrow{\cong} \mathbf{D}^0(\mathbb{P}(V^\wedge), A). \tag{5.4.1}$$

Theorem 8.2 in [de Shalit 2001] gives that (in the case where $\operatorname{char}(K) = 0$ and $H^* = H^*_{\mathrm{dR}}$ is the de Rham cohomology)

$$H^k_{\mathrm{dR}}(\Omega^r) \xrightarrow{\cong} C^k_{\mathrm{har}}, \tag{5.4.2}$$

where C^1_{har} is our $\mathbf{H}(\mathcal{BT}, K)$. Hence our Theorems 3.11 and 4.16 refine the above in the case $k = 1$. In [Alon and de Shalit 2002], the authors relate the approaches of [Schneider and Stuhler 1991; de Shalit 2001; Iovita and Spiess 2001] to the de Rham cohomology of Ω. Specialized to $k = 1$, this gives some more insight into our situation. In particular, it is possible to derive the surjectivity of the map P in Theorem 3.11 also with the methods of [Alon and de Shalit 2002], at least if $\operatorname{char}(K) = 0$.

5.5. Let now Γ be a discrete subgroup of $G(K)$. The most interesting cases are those where the image of Γ in $G(K)/Z(K) = \mathrm{PGL}(r, K)$ has finite covolume with respect to Haar measure, or is even cocompact. Examples are given as Schottky groups in $\mathrm{PGL}(2, K)$ [Gerritzen and van der Put 1980] or as arithmetic subgroups of $G(K)$ of different types, when K is the completion k_∞ of a global field k at a nonarchimedean place ∞ [Drinfeld 1974; Reiner 1975]. Then often the quotient analytic space $\Gamma \setminus \Omega$ is the set of C-points of an algebraic variety [Goldman and Iwahori 1963; Drinfeld 1974; Mustafin 1978], which may be studied via a spectral sequence relating the cohomologies of Ω and Γ with that of $\Gamma \setminus \Omega$ [Schneider and Stuhler 1991, Section 5]. For $r = 2$, this essentially boils down to a study of the Γ-cohomology sequence of (0.2) [Gekeler and Reversat 1996, Section 5]. But also for $r > 2$, (0.2) with its Γ-action will be useful, which is the topic of ongoing work.

Acknowledgement

The author is grateful to the referees for their accurate reading and helpful comments. These resulted in a number of improvements and corrections and enhanced the paper's readability. He also wishes to thank Lennart Gehrmann for pointing out an error in Proposition 4.12 of an earlier version of the paper.

References

[Alon and de Shalit 2002] G. Alon and E. de Shalit, "On the cohomology of Drinfeld's p-adic symmetric domain", *Israel J. Math.* **129** (2002), 1–20. MR Zbl

[Bruhat and Tits 1972] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local", *Inst. Hautes Études Sci. Publ. Math.* **41** (1972), 5–251. MR Zbl

[Deligne and Husemoller 1987] P. Deligne and D. Husemoller, "Survey of Drinfeld modules", pp. 25–91 in *Current trends in arithmetical algebraic geometry*, vol. Arcata, CA, 1985, edited by K. A. Ribet, Contemp. Math. **67**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Drinfeld 1974] V. G. Drinfeld, "Elliptic modules", *Mat. Sb.* (*N.S.*) **94(136)** (1974), 594–627. In Russian; translated in *Math. USSR-Sb.* **23**:4 (1974), 561–592. MR

[Fresnel and van der Put 1981] J. Fresnel and M. van der Put, *Géométrie analytique rigide et applications*, Progr. Math. **18**, Birkhäuser, Boston, 1981. MR Zbl

[Fresnel and van der Put 2004] J. Fresnel and M. van der Put, *Rigid analytic geometry and its applications*, Progr. Math. **218**, Birkhäuser, Boston, 2004. MR Zbl

[Gekeler 2017] E.-U. Gekeler, "On Drinfeld modular forms of higher rank", *J. Théor. Nombres Bordeaux* **29**:3 (2017), 875–902. MR Zbl

[Gekeler 2019] E.-U. Gekeler, "On Drinfeld modular forms of higher rank, II", *J. Number Theory* (online publication January 2019).

[Gekeler and Reversat 1996] E.-U. Gekeler and M. Reversat, "Jacobians of Drinfeld modular curves", *J. Reine Angew. Math.* **476** (1996), 27–93. MR Zbl

[Gerritzen and van der Put 1980] L. Gerritzen and M. van der Put, *Schottky groups and Mumford curves*, Lecture Notes in Math. **817**, Springer, 1980. MR Zbl

[Goldman and Iwahori 1963] O. Goldman and N. Iwahori, "The space of p-adic norms", *Acta Math.* **109** (1963), 137–177. MR Zbl

[Iovita and Spiess 2001] A. Iovita and M. Spiess, "Logarithmic differential forms on p-adic symmetric spaces", *Duke Math. J.* **110**:2 (2001), 253–278. MR Zbl

[Kiehl 1967] R. Kiehl, "Theorem A und Theorem B in der nichtarchimedischen Funktionentheorie", *Invent. Math.* **2** (1967), 256–273. MR Zbl

[Laumon 1996] G. Laumon, *Cohomology of Drinfeld modular varieties, I: Geometry, counting of points and local harmonic analysis*, Cambridge Stud. Adv. Math. **41**, Cambridge Univ. Press, 1996. MR Zbl

[Manin and Drinfeld 1973] Y. Manin and V. Drinfeld, "Periods of p-adic Schottky groups", *J. Reine Angew. Math.* **262/263** (1973), 239–247. MR Zbl

[Mustafin 1978] G. A. Mustafin, "Non-Archimedean uniformization", *Mat. Sb.* (*N.S.*) **105(147)**:2 (1978), 207–237. In Russian; translated in *Math. USSR-Sb.* **34**:2 (1978), 187–214. MR Zbl

[van der Put 1983] M. van der Put, "Les fonctions thêta d'une courbe de Mumford", pp. 1–12 in *Study group on ultrametric analysis, IX*, edited by Y. Amice et al., Inst. Henri Poincaré, Paris, 1983. MR Zbl

[Reiner 1975] I. Reiner, *Maximal orders*, Lond. Math. Soc. Monogr. **5**, Academic Press, London, 1975. MR Zbl

[Schneider and Stuhler 1991] P. Schneider and U. Stuhler, "The cohomology of p-adic symmetric spaces", *Invent. Math.* **105**:1 (1991), 47–122. MR Zbl

[de Shalit 2001] E. de Shalit, "Residues on buildings and de Rham cohomology of p-adic symmetric domains", *Duke Math. J.* **106**:1 (2001), 123–191. MR Zbl

Communicated by Michael Rapoport
Received 2019-09-16 Revised 2020-03-30 Accepted 2020-05-11

gekeler@math.uni-sb.de *Mathematik Universität des Saarlandes, Saarbrücken, Germany*

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2505

On a cohomological generalization of the Shafarevich conjecture for K3 surfaces

Teppei Takamatsu

The Shafarevich conjecture for K3 surfaces asserts the finiteness of isomorphism classes of K3 surfaces over a fixed number field admitting good reduction away from a fixed finite set of finite places. André proved this conjecture for polarized K3 surfaces of fixed degree, and recently She proved it for polarized K3 surfaces of unspecified degree. We prove a certain generalization of their results, which is stated by the unramifiedness of ℓ-adic étale cohomology groups for K3 surfaces over finitely generated fields of characteristic 0. As a corollary, we get the original Shafarevich conjecture for K3 surfaces without assuming the extendability of polarization, which is stronger than the results of André and She. Moreover, as an application, we get the finiteness of twists of K3 surfaces via a finite extension of characteristic 0 fields.

1. Introduction

The Shafarevich conjecture for abelian varieties is a remarkable result which asserts the finiteness of isomorphism classes of abelian varieties of a fixed dimension over a fixed number field admitting good reduction away from a fixed finite set of finite places. This theorem was proved by Faltings [1983] in the polarized case, and Zarhin [1985] in the unpolarized case.

In this paper, we shall prove an analogue of this theorem for K3 surfaces. For any discrete valuation field K and a K3 surface X over K, we say X has good reduction if X admits a smooth proper model over the valuation ring of K, as an algebraic space (see [Liedtke and Matsumoto 2018, Section 1]).[1] Then one can formulate the analogue of the Shafarevich conjecture for K3 surfaces. Previously, this conjecture was studied in [André 1996; She 2017] for polarized K3 surfaces. The goal of this paper is to generalize these results in terms of the unramifiedness of ℓ-adic étale cohomology groups. Our main theorem is the following (for a more generalized form, see Theorem 6.1.1).

Theorem 1.0.1 (compare with Theorem 6.1.1). *Let F be a finitely generated field over $\mathbb{Q}$, and R be a finite type algebra over $\mathbb{Z}$ which is a normal domain with the fraction field F. Then, the set*

$$\mathrm{Shaf}(F, R) := \left\{ X \;\middle|\; \begin{array}{l} X : K3 \textit{ surface over } F, \\ \textit{for any height } 1 \textit{ prime ideal } \mathfrak{p} \in \operatorname{Spec} R, \\ \textit{there exists a prime number } \ell \notin \mathfrak{p} \\ \textit{such that } H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_\ell) \textit{ is unramified at } \mathfrak{p} \end{array} \right\} / F\textit{-isom}$$

is finite.

MSC2010: primary 14J28; secondary 11F80, 11G18, 11G25, 11G35.
Keywords: K3 surfaces, Shafarevich conjecture, good reduction.

As a corollary, we have the original Shafarevich conjecture for K3 surfaces over finitely generated fields of characteristic 0.

Corollary 1.0.2 (Corollary 6.1.4). *Let F be a finitely generated field over $\mathbb{Q}$, and R be a finite type algebra over $\mathbb{Z}$ which is a normal domain with the fraction field F. Then, the set*

$$\bigl\{X \mid X : K3 \textit{ surface over } F,\ X \textit{ has good reduction at any height } 1 \textit{ prime ideal } \mathfrak{p} \in \operatorname{Spec} R\bigr\}/F\text{-}isom$$

is finite.

Note that our results are stronger than results of André and She (see Remark 1.0.4 for details). Moreover, as an application of our cohomological generalization, we get the following corollary, which asserts the finiteness of twists of a K3 surface via a finite extension of characteristic 0 fields.

Corollary 1.0.3 (Corollary 6.2.1). *Let F be a field of characteristic 0, E/F be a finite extension, and X be a K3 surface over F. Then, the set*

$$\mathrm{Tw}_{E/F}(X) := \{Y : K3 \textit{ surface over } F \mid Y_E \simeq_E X_E\}/F\text{-}isom$$

is finite. Here $Y_E \simeq_E X_E$ means the K3 surfaces $Y_E := Y \otimes_F E$ and $X_E := X \otimes_F E$ are isomorphic over E.

We note that our cohomological generalization is necessary for this application, i.e., the original statement of the Shafarevich conjecture (Corollary 1.0.2) is not enough to show Corollary 1.0.3.

Let us give some comments on the statement of Theorem 1.0.1. Theorem 1.0.1 is motivated by the good reduction criterion for K3 surfaces given by Liedtke and Matsumoto [2018]. For K3 surfaces over a Henselian discrete valuation field satisfying some assumptions, they showed the equivalence between the unramifiedness of ℓ-adic étale cohomology groups and admitting good reduction after a finite unramified extension [Liedtke and Matsumoto 2018, Theorem 1.3]. Note that the latter condition cannot be replaced by "admitting good reduction" (see [Liedtke and Matsumoto 2018, Theorem 1.6]), so our cohomological generalization is stronger than the original Shafarevich conjecture. Moreover, we deal with finitely generated fields of characteristic 0 rather than number fields, motivated by the application to Corollary 1.0.3. In fact, André also proved the Shafarevich conjecture for polarized K3 surfaces in this way (see [André 1996, Theorem 9.1.1], and also Remark 1.0.4).

Remark 1.0.4. Our results are stronger than previous results obtained by André and She. To explain this, we briefly recall their results. André [1996, Theorem 9.1.1] proved the Shafarevich conjecture for polarized K3 surfaces, i.e., the finiteness of isomorphism classes of polarized K3 surfaces of fixed degree over a fixed number field which admit good reduction away from a fixed finite set of finite places (actually, as stated above, André dealt with finitely generated fields of characteristic 0). Here, André said that a polarized K3 surface (X, L) admits good reduction if there exists a smooth proper model $\mathcal{X}$ of X as a scheme such that the ample line bundle L extends to an ample line bundle on $\mathcal{X}$. She [2017, Theorem 1.1.5] proved it for

[1] Note that it is natural to admit an integral model being an algebraic space rather than a scheme in the case of K3 surfaces (see [Matsumoto 2015, Section 5.2]).

polarized K3 surfaces of unspecified degree. More correctly, She proved the finiteness of K3 surfaces over a fixed number field which admit good reduction as polarized K3 surfaces (without fixing polarization degree) away from a fixed finite set of finite places. Here, we remark that She's result does not cover K3 surfaces admitting a smooth proper model only as an algebraic space. Moreover, there exists an example of a K3 surface admitting good reduction such that no smooth proper model has a polarization (therefore this K3 surface does not admit good reduction as polarized K3 surfaces) (see [Matsumoto 2015, Section 5.2]). Therefore, Corollary 1.0.2 is also stronger than previous results, even in the number field case.

The strategy of the proof of Theorem 1.0.1 is as follows. We basically take the approach of André and She. We first show the polarized version of Theorem 1.0.1 before dealing with the unpolarized case. To generalize the result obtained by André, we should formulate the Kuga–Satake construction as preserving the finiteness. We achieve this by using the moduli interpretation of the Kuga–Satake construction introduced by Rizov [2010]. In the perspective of the unpolarized case, we use the uniform Kuga–Satake construction introduced by She to study K3 surfaces of all degrees simultaneously. Our proof is slightly different from She's proof, and here we will sketch the differences. In [She 2017], it is crucial to show that K3 surfaces admitting good reduction are sent to abelian varieties admitting good reduction via the uniform Kuga–Satake map. She proves this using integral canonical models of certain Shimura varieties (the argument like "$\mathcal{O}$-valued points go to $\mathcal{O}$-valued points"). However, in our case, we do not assume that each K3 surface admits a smooth proper model, so instead of She's method, we use the Néron–Ogg–Shafarevich criterion for abelian varieties (this approach is already known by [André 1996]; see also [Imai and Mieda 2020]). For this purpose, we study She's uniform Kuga–Satake construction in detail in Section 3. Note that our proof does not require the theory of integral canonical models of Shimura varieties. To simplify the arguments, we need a section of the natural map from the GSpin Shimura variety to the SO Shimura variety (see Section 3.3). To get such a section, we should work with a level structure that has a sufficiently small $\mathbb{Z}_2$-component (see Remark 3.1.7). So we should suppose the unramifiedness of 2-adic representation to overcome that general K3 surfaces may not admit level structure (see Proposition 4.2.3). However, this assumption is not essential since ℓ-independence of the unramifiedness is true in a general situation (see Lemma 5.0.1). Note that Lemma 5.0.1 is essentially known by Madapusi Pera, Matsumoto [2016] and Imai and Mieda [2020] (see Remark 5.0.2).

The outline of this paper is as follows. In Section 2, we will recall the basic results on K3 surfaces, and define the moduli space of K3 surfaces introduced by Rizov and Madapusi Pera. In Section 3, we will define several algebraic groups to introduce the uniform Kuga–Satake abelian varieties, and study their basic properties. In Section 4, we will prove the main theorem in a little weaker form (i.e., only considering 2-adic cohomology) by using the results of Section 3 and the arguments given by André and She. In Section 5, we will see an ℓ-independence of the unramifiedness by using Matsumoto's result on weight filtrations [2016, Theorem 3.3]. We also prove a crystalline analogue of it by using Ochiai's ℓ-independence results [1999] and the Kuga–Satake abelian varieties (as above, this result is essentially proved in [Imai and Mieda 2020]). In Section 6, we will complete the proof of the main theorem combining the results in Sections 4 and 5, and prove the finiteness of twists via a finite extension of characteristic 0 fields.

2. K3 surfaces and their moduli

2.1. ***Basic definitions for K3 surfaces.*** In this subsection, we give definitions and basic notations about K3 surfaces.

Definition 2.1.1. (1) For any field k, a *K3 surface over* k is a smooth proper surface X over k with $\Omega^2_{X/k} \simeq \mathcal{O}_X$ and $H^1(X, \mathcal{O}_X) = 0$.

(2) For any scheme S, a *K3 family over* S is a smooth proper algebraic space $\mathcal{X}$ over S whose geometric fibers are K3 surfaces.

Remark 2.1.2. For any field k, a K3 family over k is automatically a K3 surface over k since smooth proper algebraic spaces of dimension 2 over a field are schemes.

Definition 2.1.3 [Rizov 2006, Definitions 3.2.2, 3.2.3]. (1) A *polarization* on a K3 family $\pi : \mathcal{X} \to S$ is an element $\lambda \in \mathrm{Pic}_{\mathcal{X}/S}(S)$ whose pullback by any geometric point of S is an ample line bundle. Here $\mathrm{Pic}_{\mathcal{X}/S}$ is the relative Picard functor.

(2) A polarization λ is *primitive* if its pullback by any geometric point of S is primitive, i.e., not divisible by an integer greater than 1.

(3) A polarization λ is *of degree* $2d$ if its pullback by any geometric point of S has degree $2d$, i.e., its self intersection number is $2d$.

Remark 2.1.4. Let F be a subfield of $\mathbb{C}$. For a K3 surface X over F, the relative Picard functor $\mathrm{Pic}_{X/F}$ is represented by a scheme, thus $\mathrm{Pic}_{X/F}(F) = \mathrm{Pic}(X_{\bar{F}})^{\mathrm{Gal}(\bar{F}/F)}$ is a primitive sublattice in $H^2(X(\mathbb{C}), \mathbb{Z}(1))$ via the Chern class map as in [She 2017, Lemma 2.2.3]. Hence there exists a primitive polarization for each X (by dividing a polarization by an integer greater than 1 if necessary). Note that the inclusion $\mathrm{Pic}(X) \subset \mathrm{Pic}_{X/F}(F)$ may be proper in general, though it always has a finite cokernel (see [Huybrechts 2016, Chapter 17, Section 2.2]).

Definition 2.1.5. (1) A *K3 lattice* $\mathcal{L}_{K3}$ is a unimodular lattice[2] of signature (19, 3) which is defined as

$$\mathcal{L}_{K3} := \mathbb{E}_8^{\oplus 2} \oplus \mathbb{H}^{\oplus 3},$$

where $\mathbb{E}_8$ is the (positive signature) E_8-lattice as in [Huybrechts 2016, Chapter 14, Example 0.3], and $\mathbb{H}$ is the hyperbolic plane.

(2) Consider the last component $\mathbb{H} \subset \mathcal{L}_{K3}$, and take $e, f \in \mathbb{H} \subset \mathcal{L}_{K3}$ satisfying

$$(e, f) = (f, e) = 1, \quad (e, e) = (f, f) = 0.$$

Let $v_d := e - df$. Then the *degree* $2d$ *primitive part of* $\mathcal{L}_{K3}$ is defined as

$$\mathcal{L}_d := v_d^{\perp} \simeq \mathbb{E}_8^2 \oplus \mathbb{H}^2 \oplus \langle 2d \rangle.$$

The lattice $\mathcal{L}_d$ is a primitive sublattice of $\mathcal{L}_{K3}$, and $\mathrm{disc}(\mathcal{L}_d) = 2d$.

[2] In this paper, a ($\mathbb{Z}$-)lattice means a finite free $\mathbb{Z}$-module with a symmetric bilinear pairing valued in $\mathbb{Z}$, and $\langle c \rangle$ means the $\mathbb{Z}$-lattice of rank 1 given by $(a, b) = cab$.

Remark 2.1.6 [Rizov 2006, Remark 2.3.2]. For a K3 surface X over $\mathbb{C}$ and its primitive polarization L of degree $2d$, there exists an isomorphism

$$(H^2(X(\mathbb{C}), \mathbb{Z}(1)), -\cup) \simeq \mathcal{L}_{K3}$$

which sends $\mathrm{ch}_{\mathbb{Z}}(L)$ to v_d. Here $-\cup$ denotes the minus of the cup product. Therefore, for a primitively polarized K3 surface (X, L) of degree $2d$ over a field F which is contained in $\mathbb{C}$, we sometimes identify $H^2(X(\mathbb{C}), \mathbb{Z}(1))$ with $\mathcal{L}_{K3}$, $H^2_{\text{ét}}(X_{\overline{F}}, \widehat{\mathbb{Z}}(1))$ with $\mathcal{L}_{K3,\widehat{\mathbb{Z}}} := \mathcal{L}_{K3} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$, $P^2((X(\mathbb{C}), L_{\mathbb{C}}), \mathbb{Z}(1))$ with $\mathcal{L}_d$, and $P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \widehat{\mathbb{Z}}(1))$ with $\mathcal{L}_{d,\widehat{\mathbb{Z}}} := \mathcal{L}_d \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$. Here, we denote the primitive parts of the singular cohomology group and the étale cohomology group by

$$P^2((X(\mathbb{C}), L_{\mathbb{C}}), \mathbb{Z}(1)) := \mathrm{ch}(L_{\mathbb{C}})^{\perp} \subset H^2(X(\mathbb{C}), \mathbb{Z}),$$
$$P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \widehat{\mathbb{Z}}(1)) := \mathrm{ch}_{\widehat{\mathbb{Z}}}(L)^{\perp} \subset H^2_{\text{ét}}(X_{\overline{F}}, \widehat{\mathbb{Z}}(1)).$$

To simplify notation, we omit the pairing in the rest of the paper. We denote $(H^2(X(\mathbb{C}), \mathbb{Z}(1)), -\cup)$ by $H^2(X(\mathbb{C}), \mathbb{Z}(1))$, and same with others.

Definition 2.1.7. The *discriminant kernel of* $\mathcal{L}_d$ is

$$D_d := \{g \in \mathrm{SO}(\mathcal{L}_{d,\widehat{\mathbb{Z}}}) \mid g \text{ acts trivially on } \mathcal{L}^{\vee}_{d,\widehat{\mathbb{Z}}}/\mathcal{L}_{d,\widehat{\mathbb{Z}}}\}.$$

Note that D_d is a compact open subgroup of $\mathrm{SO}(\mathcal{L}_{d,\mathbb{A}_f})$. For any prime number ℓ, we denote its $\mathbb{Z}_\ell$-component by $(D_d)_\ell$.

Proposition 2.1.8 [Madapusi Pera 2016, Lemma 2.6]. *There is a natural identification*

$$D_d = \{\tilde{g} \in \mathrm{SO}(\mathcal{L}_{K3,\widehat{\mathbb{Z}}}) \mid \tilde{g}(v_d) = v_d\}.$$

Proof. This is proved in [Madapusi Pera 2016, Lemma 2.6]. We include its proof because we need to recall the identification explicitly. Let ℓ be any prime number, and we will verify this claim for each $\mathbb{Z}_\ell$-component. First, we will define a map from the left-hand side to the right-hand side. For

$$g_\ell \in (D_d)_\ell = \{g \in \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \mid g \text{ acts trivially on } \mathcal{L}^{\vee}_{d,\mathbb{Z}_\ell}/\mathcal{L}_{d,\mathbb{Z}_\ell}\},$$

define $\tilde{g}_\ell$ as the image of g_ℓ via the composition of

$$\mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \hookrightarrow \mathrm{SO}(\mathcal{L}_{d,\mathbb{Q}_\ell}) \hookrightarrow \mathrm{SO}(\mathcal{L}_{K3,\mathbb{Q}_\ell}).$$

Then we have $\tilde{g}_\ell v_d = v_d$. We will show that $\tilde{g}_\ell \mathcal{L}_{K3,\mathbb{Z}_\ell} = \mathcal{L}_{K3,\mathbb{Z}_\ell}$. Consider the morphisms

$$\mathcal{L}_{K3,\mathbb{Z}_\ell} \simeq \mathcal{L}^{\vee}_{K3,\mathbb{Z}_\ell} \hookrightarrow \mathcal{L}^{\vee}_{d,\mathbb{Z}_\ell} \oplus \langle v_d \rangle^{\vee} \twoheadrightarrow \mathcal{L}^{\vee}_{d,\mathbb{Z}_\ell}. \tag{1}$$

For any $v \in \mathcal{L}_{K3,\mathbb{Z}_\ell}$, denote its image in $\mathcal{L}^{\vee}_{d,\mathbb{Z}_\ell} \oplus \langle v_d \rangle^{\vee}$ by $u_1 + u_2$. Then we have

$$\tilde{g}_\ell(u_1 + u_2) = g_\ell(u_1) + u_2 = (g_\ell(u_1) - u_1) + (u_1 + u_2) \in \mathcal{L}_{K3,\mathbb{Z}_\ell},$$

because g_ℓ acts trivially on $\mathcal{L}^{\vee}_{d,\mathbb{Z}_\ell}/\mathcal{L}_{d,\mathbb{Z}_\ell}$. Hence $\tilde{g}_\ell^{\pm 1} \mathcal{L}_{K3,\mathbb{Z}_\ell} \subset \mathcal{L}_{K3,\mathbb{Z}_\ell}$, thus $\tilde{g}_\ell \mathcal{L}_{K3,\mathbb{Z}_\ell} = \mathcal{L}_{K3,\mathbb{Z}_\ell}$, and we can define the desired map.

Next, we will define a map from the right-hand side to the left-hand side. For $\tilde{h}_\ell \in \mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_\ell})$ such that $\tilde{h}_\ell v_d = v_d$, we can associate $h_\ell \in \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell})$ as the restriction of $\tilde{h}_\ell$. We can show that h_ℓ acts trivially on $\mathcal{L}^\vee_{d,\mathbb{Z}_\ell}/\mathcal{L}_{d,\mathbb{Z}_\ell}$. Indeed, because the embedding $\mathcal{L}_d \hookrightarrow \mathcal{L}_{K3}$ is primitive, the composition of (1) is surjective, so for any $u_1 \in \mathcal{L}^\vee_{d,\mathbb{Z}_\ell}$, there exists $u_2 \in \langle v_d\rangle^\vee$ such that $u_1 + u_2 \in \mathcal{L}_{K3,\mathbb{Z}_\ell}$. Thus we have

$$h_\ell(u_1) - u_1 = \tilde{h}_\ell(u_1+u_2) - (u_1+u_2) \in \mathcal{L}_{K3,\mathbb{Z}_\ell} \cap \mathcal{L}_{d,\mathbb{Q}_\ell} = \mathcal{L}_{d,\mathbb{Z}_\ell}.$$

Clearly, the above maps are inverses of each other. This finishes the proof. □

2.2. ***Moduli spaces of K3 surfaces and the Torelli theorem.*** In this subsection, we recall the Torelli theorem in terms of moduli spaces. First, we recall the definition of the moduli space of K3 surfaces with oriented level structures. See [Rizov 2006, Section 6; Madapusi Pera 2015, Section 3; Ito et al. 2018, Section 5] for details.

We define a groupoid-valued moduli functor $M^\circ_{2d,\mathbb{Q}}$ by

$$M^\circ_{2d,\mathbb{Q}}(S) := \bigl\{(\pi : \mathcal{X} \to S, \lambda \in \mathrm{Pic}_{X/S}(S)) \mid \pi : \text{K3 family over } S,\ \lambda : \text{primitive polarization of degree } 2d\bigr\}$$

for any $\mathbb{Q}$-scheme S. Let $\widetilde{M}^\circ_{2d,\mathbb{Q}}$ be the twofold finite étale cover constructed by Madapusi Pera [2015, Section 5] which parametrizes orientations. Then, for any $S \to \widetilde{M}^\circ_{2d,\mathbb{Q}}$ we get (π, λ, ν), where (π, λ) is as above, and ν is an isometry of $\widehat{\mathbb{Z}}$-local systems

$$\nu : \underline{\det \mathcal{L}_{d,\widehat{\mathbb{Z}}}} \simeq \det P^2\pi_*\widehat{\mathbb{Z}}$$

such that for any $s \in S(\mathbb{C})$, the isometry ν restricts to an isometry

$$\nu_s : \det \mathcal{L}_d \simeq \det P^2(\mathcal{X}_s, \mathbb{Z}).$$

Here, we put

$$P^2\pi_*\widehat{\mathbb{Z}}(1) := \mathrm{ch}_{\widehat{\mathbb{Z}}}(\lambda)^\perp \subset R^2\pi_*\widehat{\mathbb{Z}}(1),$$

where $\mathrm{ch}_{\widehat{\mathbb{Z}}}(\lambda)$ is the Chern class of λ [Madapusi Pera 2015, Section 3.10]. Let $\mathbb{K} \subset D_d$ be a compact open subgroup. For any scheme S over $\widetilde{M}^\circ_{2d,\mathbb{Q}}$, one can define the étale sheaf I by

$$I(T) := \bigl\{g : \underline{\mathcal{L}_{K3,\widehat{\mathbb{Z}}}} \to R^2\pi|_{T*}\widehat{\mathbb{Z}}(1) \bigm| g : \text{isometry},\ g(v_d) = \mathrm{ch}_{\widehat{\mathbb{Z}}}(\lambda),\ \det g \text{ induces } \nu|_T\bigr\},$$

for any étale morphism $T \to S$. A $\mathbb{K}$-level structure on $S \to \widetilde{M}^\circ_{2d,\mathbb{Q}}$ is a section $\alpha \in H^0(S, I/\mathbb{K})$, where $\mathbb{K}$ acts on I through $\mathcal{L}_{K3,\widehat{\mathbb{Z}}}$. Then, one can define the moduli functor $M^\circ_{2d,\mathbb{K},\mathbb{Q}}$ over $\widetilde{M}^\circ_{2d,\mathbb{Q}}$ which parametrizes $\mathbb{K}$-level structures. For simplicity, we write an each element of $M^\circ_{2d,\mathbb{K},\mathbb{Q}}(S)$ as $(\mathcal{X}, \lambda, \nu, \alpha)$. Moreover, for any field F of characteristic 0, we denote the base change by $M^\circ_{2d,\mathbb{K},F}$.

Definition 2.2.1. (1) $\mathrm{SO}_{\mathcal{L}_d}$ is an algebraic group over $\mathbb{Q}$ whose R-valued points are given by

$$\mathrm{SO}_{\mathcal{L}_d}(R) := \{g \in \mathrm{SL}(\mathcal{L}_{d,R}) \mid (gv, gw) = (v, w),\ \text{for any } v, w \in \mathcal{L}_{d,R}\}.$$

(2) We put

$$\Omega^{\pm}_{\mathrm{SO}_{\mathcal{L}_d}} := \{\text{oriented negative definite planes in } \mathcal{L}_{d,\mathbb{R}}\}.$$

Then $\Omega^{\pm}_{\mathrm{SO}_{\mathcal{L}_d}}$ is naturally identified with $X_{\mathrm{SO}_{\mathcal{L}_d}}$ which gives the Shimura datum $(\mathrm{SO}_{\mathcal{L}_d}, X_{\mathrm{SO}_{\mathcal{L}_d}})$ with a reflex field $\mathbb{Q}$. Actually, $X_{\mathrm{SO}_{\mathcal{L}_d}}$ is isomorphic to $X_{\mathrm{GSpin}_{\mathcal{L}_d}}$ which is defined as in Definition 3.2.3 via the adjoint representation.

Here, we quickly state the moduli interpretation of the Torelli theorem over $\mathbb{Q}$.

Proposition 2.2.2 (the Torelli theorem [Madapusi Pera 2015, Corollary 5.4, Theorem 5.8]). *Let $\mathbb{K} \subset D_d$ be a compact open subgroup. Moreover, assume that $\mathbb{K}$ is contained in the principal level n congruence subgroup of $\mathrm{SO}(\mathcal{L}_{d,\hat{\mathbb{Z}}})$ with $n \geq 3$. Then $M^{\circ}_{2d,\mathbb{K},\mathbb{Q}}$ is representable by a scheme, and moreover there is the period map which is an étale morphism between $\mathbb{Q}$-schemes*

$$j : M^{\circ}_{2d,\mathbb{K},\mathbb{Q}} \to \mathrm{Sh}_{\mathbb{K}}(\mathrm{SO}_{\mathcal{L}_d}, X_{\mathrm{SO}_{\mathcal{L}_d}}).$$

Here $\mathrm{Sh}_{\mathbb{K}}(\mathrm{SO}_{\mathcal{L}_d}, X_{\mathrm{SO}_{\mathcal{L}_d}})$ is the canonical model of the Shimura variety over $\mathbb{Q}$.

In Proposition 3.3.3, we will use the more detailed properties of the period map j.

3. The uniform Kuga–Satake construction

In this section, we recall the definition and properties of the Kuga–Satake construction. In this section, we use only the uniform Kuga–Satake construction introduced by She. In fact, the classical Kuga–Satake construction is enough for proving the polarized case (Theorem 4.1.3), but we need She's methods to prove the unpolarized case (Theorem 4.1.4). Hence we omit the classical Kuga–Satake construction to avoid some repetitions.

3.1. *Preparation.* In this subsection and the next, we will define several algebraic groups and their adelic subgroups which play an important role in the Kuga–Satake construction. In this subsection, we discuss objects related with the lattice $\mathcal{L}_d$.

For any algebra R and any quadratic space $\mathcal{N}$ over R, we denote the Clifford algebra (resp. even Clifford algebra) of $\mathcal{N}$ by $C(\mathcal{N})$ (resp. $C^{+}(\mathcal{N})$).

Definition 3.1.1. $\mathrm{GSpin}_{\mathcal{L}_d}$ is an algebraic group over $\mathbb{Q}$, whose R-valued points are given by

$$\mathrm{GSpin}_{\mathcal{L}_d}(R) := \{z \in C^{+}(\mathcal{L}_{d,R})^{\times} \mid z\mathcal{L}_{d,R}z^{-1} = \mathcal{L}_{d,R}\}.$$

Remark 3.1.2. (1) There exists the following natural homomorphism of algebraic groups over $\mathbb{Q}$

$$f_d : \mathrm{GSpin}_{\mathcal{L}_d} \to \mathrm{SO}_{\mathcal{L}_d};\, g \mapsto (l \mapsto glg^{-1}).$$

(2) For any $\mathbb{Z}$-algebra R, we put

$$\mathrm{GSpin}(\mathcal{L}_{d,R}) := \{z \in C^{+}(\mathcal{L}_{d,R})^{\times} \mid z\mathcal{L}_{d,R}z^{-1} = \mathcal{L}_{d,R}\}.$$

Then, for any prime number ℓ, we can define

$$f_d : \mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \to \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell})$$

by the conjugation. Moreover, it is easy to confirm the identity

$$\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) = \mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Q}_\ell}) \cap C^+(\mathcal{L}_{d,\mathbb{Z}_\ell})^\times.$$

(3) For any $\mathbb{Z}$-algebra R, we will use the notation $\mathrm{GSpin}(\mathcal{L}_{K3,R})$ in a similar sense to (2). Moreover, for any prime number ℓ, we denote the conjugation map $\mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_\ell}) \to \mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_\ell})$ by f_{K3}. As in (2), it follows that

$$\mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_\ell}) = \mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Q}_\ell}) \cap C^+(\mathcal{L}_{K3,\mathbb{Z}_\ell})^\times.$$

Lemma 3.1.3 [Madapusi Pera 2016, (2.6.1)]. *Let ℓ be any prime number. Through the natural inclusion $C^+(\mathcal{L}_{d,\mathbb{Z}_\ell}) \subset C^+(\mathcal{L}_{K3,\mathbb{Z}_\ell})$, we have*

$$C^+(\mathcal{L}_{d,\mathbb{Z}_\ell}) = \{z \in C^+(\mathcal{L}_{K3,\mathbb{Z}_\ell}) \mid v_d z = z v_d\}.$$

Moreover, the above inclusion induces an embedding

$$\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \subset \mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_\ell}).$$

Proof. The first claim is essentially proved in [Madapusi Pera 2016, (2.6.1)]. For the sake of completeness, we recall the proof. For the first claim, both sides of the desired identity are primitive $\mathbb{Z}_\ell$-modules in $C^+(\mathcal{L}_{K3,\mathbb{Z}_\ell})$. Thus, it is enough to show that

$$C^+(\mathcal{L}_{d,\mathbb{Q}_\ell}) = \{z \in C^+(\mathcal{L}_{K3,\mathbb{Q}_\ell}) \mid v_d z = z v_d\}.$$

It can be easily verified by using a basis of $\mathcal{L}_{K3,\mathbb{Q}_\ell}$ which is given by a basis of $\mathcal{L}_{d,\mathbb{Q}_\ell}$ and v_d. For the second claim, by Remark 3.1.2 (2) and (3), we can reduce the problem to the obvious inclusion $\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Q}_\ell}) \subset \mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Q}_\ell})$. □

Definition 3.1.4 [André 1996, Section 4.4; Rizov 2006, Example 5.1.4]. For any positive integer n, we define a compact open subgroup $\mathbb{K}^{\mathrm{sp}}_{d,n} \subset \mathrm{GSpin}_{\mathcal{L}_d}(\mathbb{A}_f)$ by

$$\mathbb{K}^{\mathrm{sp}}_{d,n} := \{g \in \mathrm{GSpin}(\mathcal{L}_{d,\widehat{\mathbb{Z}}}) \mid g = 1 \text{ in } C^+(\mathcal{L}_{d,\widehat{\mathbb{Z}}/n\widehat{\mathbb{Z}}})\}.$$

Proposition 3.1.5 (cf. [André 1996, Section 4.4; Madapusi Pera 2015, Section 4.4]).

$$D_d(n) := f_d(\mathbb{K}^{\mathrm{sp}}_{d,n}) \subset \mathrm{SO}(\mathcal{L}_{d,\widehat{\mathbb{Z}}})$$

is a compact open subgroup of D_d.

Proof. First, we shall show that $D_d(n)$ is contained in D_d. Lemma 3.1.3 shows that

$$f_{K3}(\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell})) \subset \{g \in \mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_\ell}) \mid g v_d = v_d\},$$

thus the desired inclusion follows from Proposition 2.1.8.

For the openness, it is enough to show that for any ℓ not dividing $2dn$, the $\mathbb{Z}_\ell$-component of $D_d(n)$ is equal to $\mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell})$. This follows from [André 1996, Section 4.4]. □

The following proposition gives more information about $D_d(n)$.

Proposition 3.1.6. *For any odd prime number $\ell \neq 2$, we have*

$$f_d(\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell})) = (D_d)_\ell.$$

If $\ell = 2$, as a subset of $\mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_2})$, we have

$$f_d(\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_2})) = (D_d)_2 \cap f_{K3}(\mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_2})).$$

Proof. If ℓ does not divide $2d$, these results are essentially shown in the proof of Proposition 3.1.5.

First, for any prime number ℓ, we have $f_d(\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell})) \subset (D_d)_\ell$ as in the proof of Proposition 3.1.5. We assume $\ell \neq 2$. For any $g \in (D_d)_\ell \subset \mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_\ell})$, by the same argument as in [André 1996, Section 4.4] (here we use $\ell \neq 2$), there exists $z \in \mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_\ell})$ such that $f_{K3}(z) = g$. Proposition 2.1.8 implies $z v_d z^{-1} = v_d$, and so in fact, $z \in C^+(\mathcal{L}_{d,\mathbb{Z}_\ell})^\times$ by Lemma 3.1.3. By Proposition 2.1.8, z stabilizes $\mathcal{L}_{d,\mathbb{Z}_\ell}$ via conjugation; thus $z \in \mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell})$, finishing the proof of the first claim. If $\ell = 2$, the second claim follows by the same arguments. □

Remark 3.1.7. Unfortunately, if $\ell = 2$, we have $f_d(\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_2})) \neq (D_d)_2$. Indeed, there exists $g_2 \in (D_d)_2$ which is nontrivial in $\mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}/2\mathbb{Z}})$ (for example, permutation of two components $\mathbb{H}_{\mathbb{Z}_2} \subset \mathcal{L}_{d,\mathbb{Z}_2}$), though any element in the image of f_d is trivial there.

Corollary 3.1.8. *Let $(D_d(n))_\ell$ be the $\mathbb{Z}_\ell$-component of $D_d(n)$, and n_ℓ be the ℓ-part of n. Then, for any prime number $\ell \neq 2$, we have*

$$[(D_d)_\ell : (D_d(n))_\ell] \leq n_\ell^{(2^{20})}.$$

Moreover, there exists a positive integer N which is independent of d and n such that

$$[(D_d)_2 : (D_d(n))_2] \leq N \cdot n_2^{(2^{20})}.$$

Proof. Assume $\ell \neq 2$. We have the following commutative diagram:

$$\begin{tikzcd} \mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \arrow[r, two heads] & (D_d)_\ell \\ (\mathbb{K}^{\mathrm{sp}}_{d,n})_\ell \arrow[u, hook] \arrow[r, two heads] & (D_d(n))_\ell \arrow[u, hook] \end{tikzcd}$$

Here we have

$$(\mathbb{K}^{\mathrm{sp}}_{d,n})_\ell = \{g \in \mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \mid g = 1 \text{ in } C^+(\mathcal{L}_{d,\mathbb{Z}_\ell/n_\ell\mathbb{Z}_\ell})\}.$$

Since

$$\#(C^+(\mathcal{L}_{d,\mathbb{Z}_\ell/n_\ell\mathbb{Z}_\ell})^\times) \leq \#(C^+(\mathcal{L}_{d,\mathbb{Z}_\ell/n_\ell\mathbb{Z}_\ell})) = n_\ell^{(2^{20})},$$

the index of $(\mathbb{K}^{\mathrm{sp}}_{d,n})_\ell$ in $\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell})$ is bounded by $n_\ell^{(2^{20})}$. This finishes the proof of the first claim.

For the second claim, we put

$$N := [\mathrm{SO}(\mathcal{L}_{K3,\mathbb{Z}_2}) : f_{K3}(\mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_2}))].$$

Then, by the second claim of Proposition 3.1.6 and the above arguments, we have

$$[(D_d)_2 : D_d(n)_2] \leq [(D_d)_2 : (D_d)_2 \cap f_{K3}(\mathrm{GSpin}(\mathcal{L}_{K3,\mathbb{Z}_2}))] \cdot n_2^{(2^{20})} \leq N \cdot n_2^{(2^{20})}. \qquad \square$$

3.2. ***Preparation II.*** Here, we will introduce an even unimodular lattice $\mathcal{L}$ of signature $(26, 2)$ which contains all $\mathcal{L}_d$. Then we will define related objects as in the previous subsection.

Proposition 3.2.1 (see also [She 2017, Lemma 3.3.1]). *We put*

$$\mathcal{L} := \mathbb{E}_8^3 \oplus \mathbb{H}^2.$$

For any positive integer d, there exists a primitive embedding of lattices

$$i_d : \mathcal{L}_d \hookrightarrow \mathcal{L}.$$

Proof. See [Nikulin 1979, Corollary 1.12.3]. $\square$

Remark 3.2.2. (1) Since $\mathcal{L}$ is unimodular, the group $\mathrm{SO}(\mathcal{L}_{\hat{\mathbb{Z}}})$ is the discriminant kernel of $\mathcal{L}$, which is defined as in Definition 2.1.7.

(2) To get a primitive embedding into a self-dual lattice, the lattice $\mathbb{E}_8^2 \oplus \mathbb{H}^2 \oplus \langle 1 \rangle^5$ is enough (see [She 2017, Lemma 3.3.1]). However, we require that $\mathcal{L}$ comes from a quadratic space for using the usual definition of Clifford algebras.

Next, we will define related algebraic groups and Shimura data for $\mathcal{L}$ as in Definitions 2.2.1 and 3.1.1.

Definition 3.2.3. (1) $\mathrm{GSpin}_{\mathcal{L}}$ is the algebraic group over $\mathbb{Q}$ whose R-valued points are given by

$$\mathrm{GSpin}_{\mathcal{L}}(R) := \{z \in C^+(\mathcal{L}_R)^\times \mid z\mathcal{L}_R z^{-1} = \mathcal{L}_R\}.$$

(2) Take a 2-dimensional negative definite subspace of $\mathcal{L}_{\mathbb{Q}}$, and let e_1, e_2 be its orthogonal basis. Let e_1', e_2' be an orthonormal basis over $\mathbb{R}$ which are given by constant multiples of e_1, e_2, and $J := e_1' e_2' \in C^+(\mathcal{L}_{\mathbb{R}})$. Let ψ be the map

$$\psi : \mathbb{S} \to \mathrm{GSpin}_{\mathcal{L},\mathbb{R}};\ \alpha + \beta i \mapsto \alpha + \beta J,$$

and $X_{\mathrm{GSpin}_{\mathcal{L}}}$ be a $\mathrm{GSpin}_{\mathcal{L}}(\mathbb{R})$-conjugacy class containing ψ.

(3) $\mathrm{SO}_{\mathcal{L}}$ is the algebraic group over $\mathbb{Q}$ whose R-valued points are given by

$$\mathrm{SO}_{\mathcal{L}}(R) := \{g \in \mathrm{SL}(\mathcal{L}_R) \mid (gv, gw) = (v, w) \text{ for any } v, w \in \mathcal{L}_R\}.$$

(4) $X_{\mathrm{SO}_{\mathcal{L}}}$ is the (isomorphic) image of $X_{\mathrm{GSpin}_{\mathcal{L}}}$ via the adjoint representation $\mathrm{GSpin}_{\mathcal{L}} \to \mathrm{SO}_{\mathcal{L}}$.

(5) For $V := C(\mathcal{L})$ and a fixed $a \in V$ which is a constant multiple of $e_1 e_2$, define $\phi_a : V \times V \to \mathbb{Z}$ as $\phi_a(x, y) := \mathrm{tr}_{V/\mathbb{Q}}(xay^*)$. Here $\mathrm{tr}_{V/\mathbb{Q}}(x)$ means the trace of a left multiplication map by x as in [Huybrechts 2016, Chapter 4, Section 2.2], and $*$ denotes the natural anti-automorphism on the Clifford algebra. Then ϕ_a is a nondegenerate alternative form. We denote its degree by r. Let $\mathrm{GSp}_{V,a}$ be the algebraic group over $\mathbb{Q}$ whose R-valued points are given by

$$\mathrm{GSp}_{V,a}(R) := \left\{ g \in \mathrm{GL}(V_R) \mid \text{there exists } c \in R^\times \text{ such that } \phi_a(gx, gy) = c\phi_a(x, y) \text{ for any } x, y \in V_R \right\}.$$

Let $(\mathrm{GSp}_{V,a}, X_{\mathrm{GSp}_{V,a}})$ be the Shimura datum associated with (V, ϕ_a).

Remark 3.2.4. (1) As in Remark 3.1.2(1), we can define a homomorphism

$$f : \mathrm{GSpin}_{\mathcal{L}} \to \mathrm{SO}_{\mathcal{L}};\ g \mapsto (l \mapsto glg^{-1}).$$

Moreover, it induces a morphism of Shimura data

$$(\mathrm{GSpin}_{\mathcal{L}}, X_{\mathrm{GSpin}_{\mathcal{L}}}) \to (\mathrm{SO}_{\mathcal{L}}, X_{\mathrm{SO}_{\mathcal{L}}}).$$

(2) We can define a homomorphism

$$h : \mathrm{GSpin}_{\mathcal{L}} \to \mathrm{GSp}_{V,a};\ g \mapsto (v \mapsto gv).$$

Moreover, it induces an embedding of Shimura data

$$(\mathrm{GSpin}_{\mathcal{L}}, X_{\mathrm{GSpin}_{\mathcal{L}}}) \to (\mathrm{GSpin}_{V,a}, X_{\mathrm{GSpin}_{V,a}})$$

by our definition of a (see [Huybrechts 2016, Chapter 4, Section 2.2]).

(3) We will use a similar notation as in Remark 3.1.2 (2), (3) for $\mathcal{L}$.

Definition 3.2.5. For any positive integer n, we define compact open subgroups $\mathbb{K}_n^{\mathrm{sp}} \subset \mathrm{GSpin}_{\mathcal{L}}(\mathbb{A}_f)$ and $\mathbb{K}_n \subset \mathrm{GSp}_{V,a}(\mathbb{A}_f)$ by

$$\mathbb{K}_n^{\mathrm{sp}} := \{ g \in \mathrm{GSpin}(\mathcal{L}_{\widehat{\mathbb{Z}}}) \mid g = 1 \text{ in } C^+(\mathcal{L}_{\widehat{\mathbb{Z}}/n\widehat{\mathbb{Z}}}) \},$$
$$\mathbb{K}_n := \{ g \in \mathrm{GSp}_{V,a}(\mathbb{A}_f) \mid gV_{\widehat{\mathbb{Z}}} = V_{\widehat{\mathbb{Z}}},\ g \text{ acts trivial on } V_{\widehat{\mathbb{Z}}/n\widehat{\mathbb{Z}}} \}.$$

Remark 3.2.6. (1) One can show that $h(\mathbb{K}_n^{\mathrm{sp}}) \subset \mathbb{K}_n$ and $h^{-1}(\mathbb{K}_n) = \mathbb{K}_n^{\mathrm{sp}}$. Moreover, our definition of $\mathbb{K}_n$ coincides with Λ_n in [Rizov 2010, Section 5.5]. Therefore, as in [Rizov 2010, Section 5.5], we have an embedding

$$\mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a}, X_{\mathrm{GSp}_{V,a}}) \hookrightarrow \mathcal{A}_{g,\sqrt{r},n,\mathbb{Q}}.$$

Here, we put $g := 2^{27}$, and $\mathcal{A}_{g,\sqrt{r},n,\mathbb{Q}}$ is the moduli space of g-dimensional degree r polarized abelian schemes with level n-structure.

(2) The lattice embedding $i_d : \mathcal{L}_d \hookrightarrow \mathcal{L}$ induces a morphism of algebraic groups $i_d : \mathrm{SO}_{\mathcal{L}_d} \to \mathrm{SO}_{\mathcal{L}}$. It induces an embedding of Shimura data

$$(\mathrm{SO}_{\mathcal{L}_d}, X_{\mathrm{SO}_{\mathcal{L}_d}}) \to (\mathrm{SO}_{\mathcal{L}}, X_{\mathrm{SO}_{\mathcal{L}}}).$$

(3) One can show that $D(n) := f(\mathbb{K}_n^{\mathrm{sp}})$ is a compact open subgroup of $\mathrm{SO}(\mathcal{L}_{\hat{\mathbb{Z}}})$ as in Proposition 3.1.5. Moreover, it is clear that $i_d(D_d(n)) \subset D(n)$ because $\mathrm{GSpin}(\mathcal{L}_{d,\mathbb{Z}_\ell}) \subset \mathrm{GSpin}(\mathcal{L}_{\mathbb{Z}_\ell})$ as in Lemma 3.1.3.

3.3. ***The uniform Kuga–Satake construction.*** In this subsection, we assume that a positive integer n is sufficiently large (in our application, n would be a sufficiently large power of 2). The previous two subsections imply that there exists the following diagram of schemes over $\mathbb{Q}$:

$$\begin{tikzcd} & & \mathrm{Sh}_{\mathbb{K}_n^{\mathrm{sp}}}(\mathrm{GSpin}_{\mathcal{L}}, X_{\mathrm{GSpin}_{\mathcal{L}}}) \arrow[d, "f"] \arrow[dr, "h"] & \\ M^{\circ}_{2d, D_d(n), \mathbb{Q}} \arrow[r, "j"] & \mathrm{Sh}_{D_d(n)}(\mathrm{SO}_{\mathcal{L}_d}, X_{\mathrm{SO}_{\mathcal{L}_d}}) \arrow[r, "i_d"] & \mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}}, X_{\mathrm{SO}_{\mathcal{L}}}) & \mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a}, X_{\mathrm{GSp}_{V,a}}) \end{tikzcd}$$

Here $\mathrm{Sh}_{\mathbb{K}}(G, X)$ means the canonical model of a Shimura variety of level $\mathbb{K}$ associated with (G, X) over $\mathbb{Q}$, which is the reflex field of (G, X). Then, by the arguments in [Rizov 2010, Section 5.5], we can find δ which is a section of f over a certain number field E_n. Indeed, as in [Rizov 2010, Section 5.5], our definition of $D(n)$ guarantees that f in the above diagram induces isomorphisms between geometric connected components of the above Shimura varieties. Hence we can find a section of f over a number field on which all geometric connected components are defined.

In the following of this subsection, we fix a field F containing E_n. We consider the base change from $\mathbb{Q}$ to F of the above diagram.

$$\begin{tikzcd} & & \mathrm{Sh}_{\mathbb{K}_n^{\mathrm{sp}}}(\mathrm{GSpin}_{\mathcal{L}}) \arrow[d, shift left, "f"] \arrow[dr, "h"] & \\ M^{\circ}_{2d, D_d(n), F} \arrow[r, "j"] & \mathrm{Sh}_{D_d(n)}(\mathrm{SO}_{\mathcal{L}_d}) \arrow[r, "i_d"] & \mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}}) \arrow[u, shift left, "\delta"] & \mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a}) \end{tikzcd} \tag{$*$}$$

Here, and in the following of this paper, for simplicity, we denote $(\mathrm{Sh}_{\mathbb{K}}(G, X))_F$ by $\mathrm{Sh}_{\mathbb{K}}(G)$. Moreover, we denote the composition $h \circ \delta \circ i_d \circ j$ by Δ_d.

Remark 3.2.6 implies that there exists the universal abelian scheme $\mathcal{A}$ over $\mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a})$ possessing the degree r polarization and the level n-structure. Then, for $(X, L, \nu, \alpha) \in M^{\circ}_{2d, D_d(n), F}(F)$ which corresponds to a morphism $t : \operatorname{Spec} F \to M^{\circ}_{2d, D_d(n), F}$, we can associate an abelian variety $A^{(X,L,\alpha)}$ by pulling buck $\mathcal{A}$ via $\Delta_d \circ t$. We will quickly recall the properties of $A^{(X,L,\alpha)}$.

Definition 3.3.1. Let ℓ be any prime number.

(1) Let S be any (schematic) connected component of $\mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}})$, and $\bar{s} \to S$ be a geometric point. Then, as in [Milne 1990, III, Remark 6.1], we can show that

$$\varprojlim_{\mathbb{K}} \mathrm{Sh}_{\mathbb{K}}(\mathrm{SO}_{\mathcal{L}}) \to \mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}})$$

is a Galois covering with a Galois group $D(n)$, and so we can associate the representation

$$\pi_1(S, \bar{s}) \to (D(n))_\ell \to \mathrm{SO}(\mathcal{L}_{\mathbb{Z}_\ell}).$$

We define $\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell}$ as the corresponding $\mathbb{Z}_\ell$-sheaf on $\mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}})$, which has a symmetric pairing structure.

(2) Similarly, we define $\mathcal{L}^{\mathrm{shf}}_{d,\mathbb{Z}_\ell}$ as the $\mathbb{Z}_\ell$-sheaf on $\mathrm{Sh}_{D_d(n)}(\mathrm{SO}_{\mathcal{L}_d})$ corresponding to the representation $(D_d(n))_\ell \to \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell})$. The sheaf $\mathcal{L}^{\mathrm{shf}}_{d,\mathbb{Z}_\ell}$ has a symmetric pairing structure.

(3) Similarly, we define $V^{\mathrm{shf}}_{\mathbb{Z}_\ell}$ as the $\mathbb{Z}_\ell$-sheaf on $\mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a})$ corresponding to the representation $(\mathbb{K}_n)_\ell \to \mathrm{GSp}(V_{\mathbb{Z}_\ell}, \phi_a)$. The sheaf $V^{\mathrm{shf}}_{\mathbb{Z}_\ell}$ has a symplectic pairing structure.

Lemma 3.3.2. (1) *There exists the natural injection of étale sheaves* $\mathcal{L}^{\mathrm{shf}}_{d,\mathbb{Z}_\ell} \to i_d^*\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell}$ *preserving the pairing. Moreover,* $(\mathcal{L}^{\mathrm{shf}}_{d,\mathbb{Z}_\ell})^\perp$ *is trivial as a* $\mathbb{Z}_\ell$*-sheaf.*

(2) *There exists the natural injection of étale sheaves* $f^*\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell} \to \mathrm{End}(h^*(V^{\mathrm{shf}}_{\mathbb{Z}_\ell}))$*, which induces a "left multiplication" on a stalk.*

Proof. For (1), it is enough to show that $i_d : \mathcal{L}_{d,\mathbb{Z}_\ell} \to \mathcal{L}_{\mathbb{Z}_\ell}$ is $\pi_1(S,\bar{s})$-equivariant, and $(\mathcal{L}_{d,\mathbb{Z}_\ell})^\perp$ is a trivial $\pi_1(S,\bar{s})$-module. Here S is any connected component of $\mathrm{Sh}_{D_d(n)}(\mathrm{SO}_{\mathcal{L}_d})$, $\bar{s}$ is a geometric point of S, and $\pi_1(S,\bar{s})$-module structures on $\mathcal{L}_{d,\mathbb{Z}_\ell}$, $\mathcal{L}_{\mathbb{Z}_\ell}$ correspond to $\mathcal{L}^{\mathrm{shf}}_{d,\mathbb{Z}_\ell}$, $i_d^*\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell}$. In regard to a $\mathbb{Z}_\ell$-sheaf given by a representation of adelic subgroup, a pullback of a $\mathbb{Z}_\ell$-sheaf corresponds to a pullback of a representation. Thus $\pi_1(S,\bar{s})$-module structure on $\mathcal{L}_{\mathbb{Z}_\ell}$ is given by

$$\pi_1(S,\bar{s}) \to (D_d(n))_\ell \hookrightarrow (D(n))_\ell \hookrightarrow \mathrm{SO}(\mathcal{L}_{\mathbb{Z}_\ell}).$$

Hence the desired claim is clear.

For (2), it is enough to show that the morphism

$$\mathcal{L}_{\mathbb{Z}_\ell} \to \mathrm{End}(V_{\mathbb{Z}_\ell});\ v \mapsto (z \mapsto vz)$$

is $\pi_1(S,\bar{s})$-equivariant, where S is any connected component of $\mathrm{Sh}_{\mathbb{K}_n^{\mathrm{sp}}}(\mathrm{GSpin}_{\mathcal{L}})$, and $\pi_1(S,\bar{s})$-module structures on $\mathcal{L}_{\mathbb{Z}_\ell}$, $V_{\mathbb{Z}_\ell}$ correspond to $f^*(\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell})$, $h^*(V^{\mathrm{shf}}_{\mathbb{Z}_\ell})$. By the same reason as (1), these structures are given by

$$\pi_1(S,\bar{s}) \to (\mathbb{K}_n^{\mathrm{sp}})_\ell \xrightarrow{f} (D(n))_\ell \hookrightarrow \mathrm{SO}(\mathcal{L}_{\mathbb{Z}_\ell}),$$
$$\pi_1(S,\bar{s}) \to (\mathbb{K}_n^{\mathrm{sp}})_\ell \xrightarrow{h} (\mathbb{K}_n)_\ell \to \mathrm{GSp}(V_{\mathbb{Z}_\ell}, \phi_a).$$

Hence if we denote the first arrows of the both by σ, these actions are described as

$$\gamma(v) = \sigma(\gamma)v\sigma(\gamma)^{-1}, \quad \gamma(z) = \sigma(\gamma)(z),$$

for $\gamma \in \pi_1(S,\bar{s})$, $v \in \mathcal{L}_{\mathbb{Z}_\ell}$, and $z \in V_{\mathbb{Z}_\ell}$. Thus the desired equivariance is clear. $\square$

Proposition 3.3.3. *Let ℓ be any prime number, $t : \mathrm{Spec}\, F \to M^\circ_{2d,D_d(n),F}$ be the point corresponding to $(X,L,\nu,\alpha) \in M^\circ_{2d,D_d(n),F}(F)$, $A^{(X,L,\alpha)}$ be the abelian variety given by $(\Delta_d \circ t)^*(\mathcal{A})$, and $\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}$ be the* $\mathrm{Gal}(\overline{F}/F)$*-lattice identified with $(i_d \circ j \circ t)^*(\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell})$. Then the following hold.*

(1) *There exists a Galois equivariant lattice embedding*

$$P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_\ell(1)) \subset \mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}$$

such that $\mathrm{Gal}(\overline{F}/F)$ *acts trivially on the orthogonal complement*

$$P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_\ell(1))^\perp \subset \mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}.$$

(2) *The abelian variety* $A^{(X,L,\alpha)}$ *has a level* n*-structure defined over* F*. Thus each* n*-torsion point of* $A^{(X,L,\alpha)}$ *is* F*-rational.*

(3) *The abelian variety* $A^{(X,L,\alpha)}$ *admits a left* $C(\mathcal{L})$*-action over* F*, and moreover there exists an isomorphism of* $\mathbb{Z}_\ell$*-modules*

$$H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_\ell) \simeq C(\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)})$$

which identifies the algebra

$$C(\mathcal{L}_{\mathbb{Z}_\ell})^{\text{op}} \subset \operatorname{End}(H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_\ell))$$

with

$$C(\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)})^{\text{op}} \subset \operatorname{End}(C(\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)})).$$

Here, the former inclusion of algebras is induced by the above $C(\mathcal{L})$*-action, and the latter is induced by the right multiplication.*

(4) *The left multiplication by* $C(\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)})$ *on the right-hand side of the isomorphism in* (3) *induces a Galois equivariant isomorphism*

$$C(\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}) \simeq \operatorname{End}_{C(\mathcal{L}_{\mathbb{Z}_\ell})^{\text{op}}}(H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_\ell)).$$

Here, the (*left*) $C(\mathcal{L}_{\mathbb{Z}_\ell})^{\text{op}}$*-module structure is induced by the left* $C(\mathcal{L})$*-action on* $A^{(X,L,\alpha)}$ *as in* (3).

Proof. These results are essentially proved in [She 2017, Proposition 3.5.8].

Statement (1) follows from Lemma 3.3.2(1) and the fact

$$(j \circ t)^*(\mathcal{L}^{\text{shf}}_{d,\mathbb{Z}_\ell}) \simeq P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_\ell(1))$$

(see [Madapusi Pera 2015, Proposition 5.6 (1)]).

Statement (2) is clear because the universal family $\mathcal{A}$ admits a level n-structure.

Before proving (3) and (4), we note that for the universal abelian scheme $u : \mathcal{A} \to \mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a})$, we have $R^1 u_* \mathbb{Z}_\ell \simeq V^{\text{shf}}_{\mathbb{Z}_\ell}$.

For (3), as in [Madapusi Pera 2016, Section 3.10], $h^*(\mathcal{A})_{\mathrm{Sh}_{\mathbb{K}_n^{\text{sp}}}(\mathrm{GSpin}_{\mathcal{L}})_{\mathbb{C}}}$ admits a $C(\mathcal{L})$-action which corresponds to a right multiplication on the cohomology, since our definition of h guarantees that the right multiplication preserves the Hodge structure. This action descends to F by [Madapusi Pera 2016, Proposition 3.11] and induces a $C(\mathcal{L})$-action on $A^{(X,L,\alpha)}$ with desired properties.

For (4), the statement (3) of this proposition and Lemma 3.3.2(2) imply the well-definedness and the Galois equivariance of our morphism, and it is clearly bijective. □

4. Proof of the Main Theorems

4.1. *Statements.*

Lemma 4.1.1. *Let* F *be a finitely generated field over* $\mathbb{Q}$*,* R *be a smooth algebra over* $\mathbb{Z}$ *which is an integral domain with the fraction field* F*, and* $\bar{s}$ *be a geometric point corresponding to an algebraic* *closure $\overline{F}$ over F. For any $\pi_1(\operatorname{Spec} F, \bar{s})$-module M such that* $\operatorname{Ker}(\pi_1(\operatorname{Spec} F, \bar{s}) \to \operatorname{Aut}(M))$ *is closed, the following are equivalent.*

(1) *The $\pi_1(\operatorname{Spec} F, \bar{s})$-action on M arises from a $\pi_1(\operatorname{Spec} R, \bar{s})$-action on M.*

(2) *For any height 1 prime ideal $\mathfrak{p} \in \operatorname{Spec} R$, the $\pi_1(\operatorname{Spec} F, \bar{s})$-action on M arises from a $\pi_1(\operatorname{Spec} R_{\mathfrak{p}}, \bar{s})$-action on M.*

(3) *For any height 1 prime ideal $\mathfrak{p} \in \operatorname{Spec} R$, M is unramified at $\mathfrak{p}$, i.e., if we take $\bar{v}$ which is an extension of valuation $\mathfrak{p}$ to $\overline{F}$, the inertia group $I_{\bar{v}}$ acts trivially on M.*

When M satisfies the above equivalent conditions, we say M is unramified over $\operatorname{Spec} R$.

Proof. First, we recall that $\pi_1(\operatorname{Spec} F, \bar{s}) \to \pi_1(\operatorname{Spec} R, \bar{s})$ is surjective and its kernel is identified with $\operatorname{Gal}(\overline{F}/F_R^{\mathrm{ur}})$, where F_R^{ur} is the composite of finite extensions E/F which are unramified over $\operatorname{Spec} R$ [Fu 2015, Proposition 3.3.6]. Here, we say E/F is unramified over $\operatorname{Spec} R$ if the normalization of $\operatorname{Spec} R$ in E is unramified over $\operatorname{Spec} R$. The same results hold for $R_{\mathfrak{p}}$.

(1)$\Leftrightarrow$(2) By the assumption on M, it suffices to show that

$$\operatorname{Ker}(\pi_1(\operatorname{Spec} F, \bar{s}) \to \pi_1(\operatorname{Spec} R, \bar{s}))$$

is generated by

$$(\operatorname{Ker}(\pi_1(\operatorname{Spec} F, \bar{s}) \to \pi_1(\operatorname{Spec} R_{\mathfrak{p}}, \bar{s})))_{\mathfrak{p}}$$

as a topological group. By the above remark, it is enough to show that $F_R^{\mathrm{ur}} = \bigcap_{\mathrm{ht}(\mathfrak{p})=1} F_{R_{\mathfrak{p}}}^{\mathrm{ur}}$. The inclusion $F_R^{\mathrm{ur}} \subset \bigcap_{\mathrm{ht}(\mathfrak{p})=1} F_{R_{\mathfrak{p}}}^{\mathrm{ur}}$ is obvious, and the other direction follows from the Zariski–Nagata purity.

(2)$\Leftrightarrow$(3) By the assumption on M, it suffices to show that $\operatorname{Ker}(\pi_1(\operatorname{Spec} F, \bar{s}) \to \pi_1(\operatorname{Spec} R_{\mathfrak{p}}, \bar{s}))$ is generated by $(I_{\bar{v}})_{\bar{v} \text{ over } \mathfrak{p}}$ as a topological group, but it follows from the above remark. $\square$

Remark 4.1.2. The condition "M is unramified at $\mathfrak{p}$" does not depend on a choice of $\bar{v}$. Indeed, for each $\mathfrak{p}$, the inertia group $I_{\bar{v}}$ is determined by $\mathfrak{p}$ up to conjugation in $\operatorname{Gal}(\overline{F}/F)$.

The following are the statements of results of this section (for more generalized statements, see Theorem 6.1.1).

Theorem 4.1.3. *Let F be a finitely generated field over $\mathbb{Q}$, R be a smooth algebra over $\mathbb{Z}$ which is an integral domain with the fraction field F, and d be a positive integer. Then, the set*

$$\operatorname{Shaf}(F, R, d) := \left\{ (X, L) \;\middle|\; \begin{array}{l} X : K3 \textit{ surface over } F, \\ L \in \operatorname{Pic}_{X/F}(F) : \textit{primitive ample}, \\ H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2) : \textit{unramified over } \operatorname{Spec} R, \\ \deg L = 2d \end{array} \right\} / F\textit{-isom.}$$

is finite.

Theorem 4.1.4. *Let F be a finitely generated field over $\mathbb{Q}$, and R be a smooth algebra over $\mathbb{Z}$ which is an integral domain with the fraction field F. Then, the set*

$$\mathrm{Shaf}(F, R) := \big\{X \mid X : K3 \text{ surface over } F,\ H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2) : \text{unramified over } \operatorname{Spec} R\big\}/F\text{-isom}$$

is finite.

Remark 4.1.5. For a nonempty open subscheme $\operatorname{Spec}(R') \subset \operatorname{Spec}(R)$, the finiteness of $\mathrm{Shaf}(F, R', d)$ (resp. $\mathrm{Shaf}(F, R')$) clearly implies the finiteness of $\mathrm{Shaf}(F, R, d)$ (resp. $\mathrm{Shaf}(F, R)$). Thus, to prove Theorems 4.1.3 and 4.1.4, we may assume $\frac{1}{2} \in R$. Note that it is equivalent to say that the residual characteristic at any point of $\operatorname{Spec} R$ is different from 2.

4.2. *Proof of Theorem 4.1.3.* In this subsection, we use the same notation as Theorem 4.1.3, unless otherwise noted. First, for using the Kuga–Satake construction, we will replace F by an appropriate finite extension of it to provide a level structure on $(X, L) \in \mathrm{Shaf}(F, R, d)$. The following lemma is essential for justifying this replacement.

Lemma 4.2.1. *Let E/F be a finite extension, X_0 be a $K3$ surface over F, and $L_0 \in \mathrm{Pic}_{X_0/F}(F)$ be a polarization. Then, the set*

$$\big\{(X, L) \mid X \text{ is a } K3 \text{ surface over } F,\ L \in \mathrm{Pic}_{X/F}(F) : \text{ample},\ (X_E, L_E) \simeq_E (X_{0,E}, L_{0,E})\big\}/F\text{-isom}$$

is finite.

Proof. Taking the Galois closure of E in $\overline{F}$, we may assume that E/F is a Galois extension. Then we can identify this set with the Galois cohomology group $H^1(\mathrm{Gal}(E/F), \mathrm{Aut}_E(X_0, L_0))$. The finiteness of this set follows from [Huybrechts 2016, Chapter 5, Proposition 3.3]. □

Lemma 4.2.2 (cf. [André 1996, Lemma 8.4.1]). *Let X be a $K3$ surface over F, $L \in \mathrm{Pic}_{X/F}(F)$ be a primitive polarization of degree $2d$ on X over F, and n be a positive integer. We put*

$$W_{\widehat{\mathbb{Z}}} := P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \widehat{\mathbb{Z}}(1)).$$

Let

$$\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{O}(W_{\widehat{\mathbb{Z}}})$$

be the natural Galois representation. Fix an isometry

$$i_{(X,L)} : \mathcal{L}_{K3,\widehat{\mathbb{Z}}} \simeq H^2_{\text{ét}}(X_{\overline{F}}, \widehat{\mathbb{Z}}(1)),$$

which restricts to an isometry $\mathcal{L}_{K3} \simeq H^2(X(\mathbb{C}), \mathbb{Z}(1))$, and which sends v_d to $\mathrm{ch}_{\widehat{\mathbb{Z}}}(L)$ (see Remark 2.1.6). Using $i_{(X,L)}$, we identify $D_d(n)$ with a compact open subgroup of $\mathrm{SO}(W_{\widehat{\mathbb{Z}}})$. Then, for any finite extension E/F, we have

$$\rho(\mathrm{Gal}(\overline{F}/E)) \subset D_d(n) \Leftrightarrow \rho_\ell(\mathrm{Gal}(\overline{F}/E)) \subset (D_d(n))_\ell \quad \text{for every } \ell \mid 2n.$$

Proof. In the following, we identify $\mathrm{SO}(\mathcal{L}_{d,\widehat{\mathbb{Z}}})$ with $\mathrm{SO}(W_{\widehat{\mathbb{Z}}})$ vie $i_{(X,L)}$. This lemma is essentially shown in [André 1996, Lemma 8.4.1]. The following claim is shown in the proof of [André 1996, Lemma 8.4.1], using specialization arguments and the Weil conjecture.

Claim. *If there exists a prime number ℓ such that $\rho_\ell(\mathrm{Gal}(\overline{F}/E)) \subset \mathrm{SO}(W_{\mathbb{Z}_\ell})$, then $\rho(\mathrm{Gal}(\overline{F}/E)) \subset \mathrm{SO}(W_{\widehat{\mathbb{Z}}})$.*

André stated that the above claim implies the following result.

$$\rho(\mathrm{Gal}(\overline{F}/E)) \subset D_d(n) \Leftrightarrow \rho_\ell(\mathrm{Gal}(\overline{F}/E)) \subset (D_d(n))_\ell \quad \text{for every } \ell \mid 2dn.$$

Indeed, for $\ell \nmid 2dn$, we have $(D_d(n))_\ell = \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_\ell})$.

More generally, for $\ell \nmid 2n$, we have $(D_d(n))_\ell = (D_d)_\ell$ (see Corollary 3.1.8). Therefore, to generalize André's result to our lemma, it is enough to show that if $\rho_\ell(\mathrm{Gal}(\overline{F}/E)) \subset \mathrm{SO}(W_{\mathbb{Z}_\ell})$, then $\rho_\ell(\mathrm{Gal}(\overline{F}/E)) \subset (D_d)_\ell$. However, since $\mathrm{Gal}(\overline{F}/F)$ stabilizes $\mathrm{ch}_{\mathbb{Z}_\ell}(L)$, it follows from our description of the discriminant kernel

$$(D_d)_\ell = \{\tilde{g}_\ell \in \mathrm{SO}(H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Z}_\ell)) \mid \tilde{g}_\ell(\mathrm{ch}_{\mathbb{Z}_\ell}(L)) = \mathrm{ch}_{\mathbb{Z}_\ell}(L)\},$$

which follows from Proposition 2.1.8. □

In the rest of this section, fix a positive integer n which is a sufficiently large power of 2.

Proposition 4.2.3. *To prove Theorem 4.1.3, it is enough to show that*

$$\mathrm{Shaf}'(F, R, d) := \left\{ (X, L) \;\middle|\; \begin{array}{l} X : \textit{K3 surface over } F, \\ L \in \mathrm{Pic}_{X/F}(F) : \textit{ primitive ample}, \\ H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2) : \textit{unramified over } \mathrm{Spec}\, R, \\ \deg L = 2d, \\ (X, L) \textit{ admits a } D_d(n)\textit{-level structure} \end{array} \right\} \Big/ F\textit{-isom}$$

is a finite set for any F, R, d as in Theorem 4.1.3. Moreover, if we fix a number field F', it suffices to show only in the case where $F \supset F'$ and $\frac{1}{2} \in R$.

Here, "(X, L) admits a $D_d(n)$-level structure" means that there exists an element $(X, L, \nu_{(X,L)}, \alpha_{(X,L)})$ in $M^\circ_{2d, D_d(n), F}(F)$.

Proof. We should prove the finiteness of $\mathrm{Shaf}(F, R, d)$. By Remark 4.1.5, we may assume $\frac{1}{2} \in R$ (so the Tate twist $\otimes \mathbb{Z}_2(1)$ does not affect the unramifiedness over $\mathrm{Spec}\, R$; see Lemma 4.1.1).

First, we will show that there exists a finite extension E/F such that for any $(X, L) \in \mathrm{Shaf}(F, R, d)$, the pair (X_E, L_E) admits a $D_d(n)$-level structure. We fix $(X, L) \in \mathrm{Shaf}(F, R, d)$ and $i_{(X,L)}$, moreover we use the same identification as in Lemma 4.2.2. Let

$$\bar{\rho}_2 := \bar{\rho}_{(X,L),2} : \pi_1(\mathrm{Spec}\, R, \bar{s}) \to \mathrm{O}(P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_2(1)))$$

be the representation induced by

$$\rho := \rho_{(X,L)} : \mathrm{Gal}(\overline{F}/F) \to \mathrm{O}(P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \widehat{\mathbb{Z}}(1))).$$

The inverse image $\bar{\rho}_2^{-1}((D_d(n))_2)$ is a finite index subgroup, so we can associate a pointed finite étale cover $\operatorname{Spec}\widetilde{R} \to \operatorname{Spec} R$. Then we have $\bar{\rho}_2(\pi_1(\operatorname{Spec}\widetilde{R}, \bar{s})) \subset (D_d(n))_2$. The former equals $\rho_2(\operatorname{Gal}(\overline{F}/\operatorname{Frac}(\widetilde{R})))$, and by Lemma 4.2.2, we can get the $D_d(n)$-level structure on $(X_{\operatorname{Frac}(\widetilde{R})}, L_{\operatorname{Frac}(\widetilde{R})})$ by $i_{(X,L)}$.

Here, note that

$$[\pi_1(\operatorname{Spec} R, \bar{s}) : \bar{\rho}_2^{-1}((D_d(n))_2)] \le C_d := [\mathrm{O}(\mathcal{L}_{d,\mathbb{Z}_2}) : (D_d(n))_2],$$

where C_d is independent of (X, L) and $i_{(X,L)}$. By the analogue of the Hermite–Minkowski theorem [Harada and Hiranouchi 2009, Proposition 2.3, Theorem 2.9], the family of subsets

$$\mathcal{C} := \{H \subset \pi_1(\operatorname{Spec} R, \bar{s}) : \text{open subgroup} \mid [\pi_1(\operatorname{Spec} R, \bar{s}) : H] \le C_d\}$$

is finite, therefore

$$H_0 := \bigcap_{H \in \mathcal{C}} H$$

is an open subgroup. Let $\operatorname{Spec}\widetilde{R}_0 \to \operatorname{Spec} R$ be the corresponding pointed finite étale covering, then by the above argument, we can get a $D_d(n)$-level structure on $(X_{\operatorname{Frac}(\widetilde{R}_0)}, L_{\operatorname{Frac}(\widetilde{R}_0)})$. Hence we now get a desired finite extension $E := \operatorname{Frac}(\widetilde{R}_0)$.

Thus, by using the assumption for $\operatorname{Shaf}'(E, \widetilde{R}_0, d)$ and Lemma 4.2.1, we can show the finiteness of $\operatorname{Shaf}(F, R, d)$. Note that the latter statement is clear by Lemma 4.2.1. □

The following proposition is essentially known by [André 1996], and one can prove it as a corollary of the theory of potentially good loci of Shimura variety (see [Imai and Mieda 2020]).

Proposition 4.2.4. *Assume $F \supset E_n$, where E_n is as in Section 3.3. For $(X, L, \nu, \alpha) \in M^{\circ}_{2d, D_d(n), F}(F)$, let $A^{(X,L,\alpha)}$ be the Kuga–Satake abelian variety as in Proposition 3.3.3. Let R be a smooth algebra over $\mathbb{Z}$ which is an integral domain with the fraction field F, and assume $\frac{1}{2} \in R$. Assume that $H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2)$ is unramified over* $\operatorname{Spec} R$ *(its Tate twists are unramified too, because $\frac{1}{2} \in R$). Then, for any height* 1 *prime ideal* $\mathfrak{p} \in \operatorname{Spec} R$, *the abelian variety $A^{(X,L,\alpha)}$ has good reduction at* $\mathfrak{p}$.[3]

Proof. We will follow the proof by André [1996, Lemma 9.3.1].[4] By the Néron–Ogg–Shafarevich criterion for abelian varieties (it is true whether a residue field is perfect or not), it is enough to show that $H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_2)$ is unramified at $\mathfrak{p}$ (here we use $\frac{1}{2} \in R$). Let $\bar{v}$ be an extension on $\overline{F}$ of the valuation $\mathfrak{p}$, and $\varphi : I_{\bar{v}} \to \operatorname{Aut}(H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_2))$ be a restriction of the Galois representation. Since $C(\mathcal{L})$-action on $A^{(X,L,\alpha)}$ is defined over F, for any $\gamma \in I_{\bar{v}}$, we have

$$\varphi(\gamma) \in \operatorname{End}_{C(\mathcal{L}_{\mathbb{Z}_2})^{\text{op}}}(H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_2)) \simeq C(\mathcal{L}_{\mathbb{Z}_2,(X,L,\alpha)})$$

(see Proposition 3.3.3 (3), (4)). Thus we also denote its image by $\varphi(\gamma) \in C(\mathcal{L}_{\mathbb{Z}_2,(X,L,\alpha)})$.

---

[3] If n is a sufficiently large power of ℓ, then the same argument work with ℓ in place of 2.

[4] The referee taught me another quick way of seeing this proposition. By the construction, the Galois representation on $H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Q}_2)$ has to factor through $\operatorname{GSpin}(PH^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2))$ and hence the inertia representation has to factor through the center $\mathbb{Q}_2^{\times}$ and in fact through $\mathbb{Z}_2^{\times}$ by the compactness. Moreover, by the level assumption, it factors through $1 + 4\mathbb{Z}_2$, so it finishes proof since $1 + 4\mathbb{Z}_2$ has no nontrivial quasiunipotent elements.

On the other hand, $I_{\bar{v}}$ acts trivially on $P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_2(1))$ by our assumptions (see Lemma 4.1.1), and moreover acts trivially on $P^2_{\text{ét}}((X_{\overline{F}}, L_{\overline{F}}), \mathbb{Z}_2(1))^{\perp} \subset \mathcal{L}_{\mathbb{Z}_2,(X,L,\alpha)}$ by Proposition 3.3.3(1), thus $\gamma(c) = c$ for any $\gamma \in I_{\bar{v}}$ and $c \in C(\mathcal{L}_{\mathbb{Z}_2,(X,L,\alpha)})$. By Proposition 3.3.3(4), we have $\gamma(z \mapsto cz) = (z \mapsto cz)$ in $\mathrm{End}_{C(\mathcal{L}_{\mathbb{Z}_2})^{\mathrm{op}}}(H^1_{\text{ét}}(A^{(X,L,\alpha)}_{\overline{F}}, \mathbb{Z}_2))$, where the left-hand side is $(z \mapsto \varphi(\gamma)c\varphi(\gamma)^{-1}z)$. This implies $\varphi(\gamma)$ is contained in the center of $C(\mathcal{L}_{\mathbb{Q}_2,(X,L,\alpha)})$, which is a reduced algebra.

Proposition 3.3.3(2) and the Raynaud semiabelian reduction criterion [SGA 7_{I} 1972, Exposé IX, Proposition 4.7] imply that $A^{(X,L,\alpha)}$ has semiabelian reduction at $\mathfrak{p}$ (i.e., $A^{(X,L,\alpha)}$ extends to a semiabelian scheme over Spec $R_{\mathfrak{p}}$). Here, we use that $n \geq 3$ is a power of 2, and the residual characteristic of $\mathfrak{p}$ is not 2. Thus for any $\gamma \in I_{\bar{v}}$, $\varphi(\gamma)$ is a unipotent element of a reduced algebra, it is identity. Hence it finishes the proof. □

We now complete the proof of Theorem 4.1.3. By Proposition 4.2.3, it is enough to show the finiteness of $\mathrm{Shaf}'(F, R, d)$ when $F \supset E_n$ and $1/2 \in R$. Here, we take E_n as in Section 3.3. In the following, we identify $(X, L) \in \mathrm{Shaf}'(F, R, d)$ with $(X, L, \nu, \alpha) \in M^{\circ}_{2d, D_d(n), F}(F)$ by choosing a level structure. Hence for $(X, L, \nu, \alpha) \in \mathrm{Shaf}'(F, R, d)$, we can associate $A^{(X,L,\alpha)}$, and since in the diagram $(*)$ of Section 3.3, each fiber of i_d is finite and h is injective (because they are induced by an embedding of Shimura data), it suffices to show the finiteness of $\Delta_d(\mathrm{Shaf}'(F, R, d))$. The image $\Delta_d(X, L, \nu, \alpha)$ corresponds to $A^{(X,L,\alpha)}$ with their degree r polarization and level n-structure. However, by Proposition 4.2.4, the abelian variety $A^{(X,L,\alpha)}$ has good reduction at any height 1 prime of Spec R, so this set is finite by [Faltings 1983, Satz 6] (for finitely generated fields of characteristic 0, see [Faltings et al. 1984, VI, §1, Theorem 2]).

4.3. ***Proof of Theorem 4.1.4.*** In this subsection, we use the same notation as in Theorem 4.1.4, unless otherwise noted. The strategy is the same as [She 2017], i.e., we use Theorem 4.1.3 for reducing the problem to the finiteness of Picard lattices, and use the uniform Kuga–Satake maps for associating $\mathrm{Shaf}(F, R)$ with a finite set of abelian varieties.

Lemma 4.3.1 (cf. [She 2017, Corollary 4.1.3]). *For any $X_0 \in \mathrm{Shaf}(F, R)$, there exist only finitely many $X \in \mathrm{Shaf}(F, R)$ whose Picard lattice $\mathrm{Pic}_{X/F}(F)$ is isometric to the Picard lattice $\mathrm{Pic}_{X_0/F}(F)$.*

Proof. As in [She 2017, Proposition 4.1.2], a K3 surface X over F admits a primitive polarization whose degree bounded by a constant depending only on the isometry class of $\mathrm{Pic}_{X/F}(F)$. Hence this lemma follows from Theorem 4.1.3. □

Lemma 4.3.2 (cf. [She 2017, Lemma 4.1.4]). *Let E/F be a finite extension. For any $X_0 \in \mathrm{Shaf}(F, R)$, the set*

$$\{X \in \mathrm{Shaf}(F, R) \mid X_E \simeq_E X_{0,E}\}$$

is finite.

Proof. Taking a Galois closure, we may assume E/F is a Galois extension. By Lemma 4.3.1, it suffices to show the finiteness of isometry classes of Picard lattices $\mathrm{Pic}_{X/F}(F)$ associated with the considering set. Note that $\mathrm{Pic}_{X/F}(E)^{\mathrm{Gal}(E/F)} = \mathrm{Pic}_{X/F}(F)$ and $\mathrm{Pic}_{X/F}(E)$ is isometric to $\mathrm{Pic}_{X_0/F}(E)$. Since the set of conjugacy classes of subgroups of $\mathrm{O}(\mathrm{Pic}_{X_0/F}(E))$ with the order $[E : F]$ is finite by [Borel 1963, Section 5, (a)], the desired finiteness follows. □

Proposition 4.3.3. *Recall that we fixed a positive integer n which is a power of 2. To show Theorem 4.1.4, it is enough to show that*

$$\mathrm{Shaf}'(F, R) := \left\{ X \;\middle|\; \begin{array}{l} X : K3 \text{ surface over } F, \\ H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_2) : \text{unramified over } \operatorname{Spec} R, \\ \text{there exists } d_X, L_X, \nu_X, \alpha_X \text{ such that} \\ (X, L_X, \nu_X, \alpha_X) \in M^{\circ}_{2d_X, D_{d_X}(n), \mathbb{Q}}(F) \end{array} \right\} / F\text{-isom}$$

is a finite set for any (F, R) as in Theorem 4.1.4. Moreover, if we fix a number field F', it suffices to show only in the case where $F \supset F'$ and $\frac{1}{2} \in R$.

Proof. The proof is similar to Proposition 4.2.3, but we need more precise evaluation since we should discuss all degrees simultaneously.

As in the proof of Proposition 4.2.3, it is enough to show the finiteness of $\mathrm{Shaf}(F, R)$ with $\frac{1}{2} \in R$. First, note that every K3 surface over F admits some primitive polarization over F. Therefore, for any $X \in \mathrm{Shaf}(F, R)$, we can associate a primitive polarization L_X. Let $2d_X$ be the degree of L_X. We will show that there exists a finite extension E/F such that for any $X \in \mathrm{Shaf}(F, R)$, the pair $(X_E, L_{X,E})$ admits a $D_{d_X}(n)$-level structure. For each $X \in \mathrm{Shaf}(F, R)$, we fix $i_{(X,L_X)}$ as in Lemma 4.2.2, and we use the notation $\bar{\rho}_2 := \bar{\rho}_{(X,L_X),2}$ in the same sense as in Proposition 4.2.3. To get a desired extension, we should replace the bound C_{d_X} in the proof of Proposition 4.2.3 by a bound which is independent of X. For $\Gamma := \bar{\rho}_2(\pi_1(\operatorname{Spec} R, \bar{s}))$, we have

$$[\Gamma : \Gamma \cap (D_{d_X}(n))_2] = [\Gamma : \Gamma \cap \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_2})] \cdot [\Gamma \cap \mathrm{SO}(\mathcal{L}_{d_X,\mathbb{Z}_2}) : \Gamma \cap (D_{d_X}(n))_2] \leq 2N \cdot n^{(2^{20})}.$$

Here, we use $\Gamma \cap \mathrm{SO}(\mathcal{L}_{d,\mathbb{Z}_2}) \subset \Gamma \cap (D_d)_2$ (which follows from Proposition 2.1.8, see the proof of Lemma 4.2.2), and Corollary 3.1.8. We note that this bound is independent of X, L_X, and $i_{(X,L_X)}$. Hence replacing C_d by $2N \cdot n^{(2^{20})}$ in the arguments in the proof of Proposition 4.2.3, we get a pointed finite étale covering $\operatorname{Spec} \widetilde{R}_0 \to \operatorname{Spec} R$, whose fraction field E satisfies the desired property. Thus, by using the assumption for $\mathrm{Shaf}'(E, \widetilde{R}_0)$ and Lemma 4.3.2, we get the finiteness of $\mathrm{Shaf}(F, R)$. The latter statement is clear by Lemma 4.3.2. □

Definition 4.3.4 [She 2017, Definition 4.1.10]. Let F be a subfield of $\mathbb{C}$, X be a K3 surface over F, and ℓ be any prime number. We define (relative) *transcendental lattices* by

$$\begin{aligned} T(X) &:= \mathrm{Pic}_{X/F}(F)^{\perp} \subset H^2(X(\mathbb{C}), \mathbb{Z}(1)), \\ T(X)_{\mathbb{Z}_\ell} &:= \mathrm{Pic}_{X/F}(F)^{\perp} \subset H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Z}_\ell(1)), \\ T(X)_{\widehat{\mathbb{Z}}} &:= \mathrm{Pic}_{X/F}(F)^{\perp} \subset H^2_{\text{ét}}(X_{\overline{F}}, \widehat{\mathbb{Z}}(1)). \end{aligned}$$

Here we omit the Chern class map.

Remark 4.3.5 (cf. [She 2017, Corollary 4.1.13]). Recall that $M := H^2(X(\mathbb{C}), \mathbb{Z}(1)) \simeq \mathcal{L}_{K3}$ is unimodular, and $N := \mathrm{Pic}_{X/F}(F)$ is a primitive sublattice. In this situation, one can verify a canonical isomorphism

$$N^\vee/N \simeq M/(N + N^\perp) \simeq (N^\perp)^\vee/N^\perp.$$

Thus we get $\mathrm{disc}(\mathrm{Pic}_{X/F}(F)) = \mathrm{disc}(T(X))$.

Lemma 4.3.6 (cf. [She 2017, Proposition 4.1.11]). *For $(X, L, \nu, \alpha) \in M^\circ_{2d,D_d(n),F}(F)$ and any prime number ℓ, we have*

$$(\mathcal{L}^{\mathrm{Gal}(\bar{F}/F)}_{\mathbb{Z}_\ell,(X,L,\alpha)})^\perp = T(X)_{\mathbb{Z}_\ell}.$$

Here, the orthogonal complement of the left-hand side is taken in $\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}$, and the above equality is as a sublattice of $P^2_{\text{ét}}((X_{\bar{F}}, L_{\bar{F}}), \mathbb{Z}_\ell)$.

Proof. First, we can show that

$$T(X)_{\mathbb{Z}_\ell} = (P^2_{\text{ét}}((X_{\bar{F}}, L_{\bar{F}}), \mathbb{Z}_\ell(1))^{\mathrm{Gal}(\bar{F}/F)})^\perp$$

(the orthogonal complement of the right-hand side is taken in $P^2_{\text{ét}}((X_{\bar{F}}, L_{\bar{F}}), \mathbb{Z}_\ell(1))$). Indeed, since the both sides of this equality are primitive in $P^2_{\text{ét}}((X_{\bar{F}}, L_{\bar{F}}), \mathbb{Z}_\ell(1))$, it suffices to show this equality after inverting ℓ, which follows directly from the Tate conjecture over F [1994, Theorem 5.6(a)].

Hence we have to show that

$$(P^2_{\text{ét}}((X_{\bar{F}}, L_{\bar{F}}), \mathbb{Z}_\ell(1))^{\mathrm{Gal}(\bar{F}/F)})^\perp = (\mathcal{L}^{\mathrm{Gal}(\bar{F}/F)}_{\mathbb{Z}_\ell,(X,L,\alpha)})^\perp$$

(note that the $\perp$ in both sides have different meaning). However, since the both sides are primitive in $\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}$, we may invert ℓ for showing this equality, so it follows obviously from Proposition 3.3.3(1). $\square$

Let us complete the proof of Theorem 4.1.4. As in the previous subsection, by Theorem 4.1.4, it suffices to show the finiteness of $\mathrm{Shaf}'(F, R)$ when $F \supset E_n$ and $1/2 \in R$. By Lemma 4.3.1 and the fact [Cassels 1982, Chapter 9, Theorem 1.1] which asserts the finiteness of isometry classes of lattices with bounded rank and discriminant, it is enough to show that $\mathrm{disc}(\mathrm{Pic}_{X/F}(F))$ $(X \in \mathrm{Shaf}'(F, R))$ is bounded. Using Remark 4.3.5, we can reduce the problem to the finiteness of $\{T(X)_{\hat{\mathbb{Z}}} \mid X \in \mathrm{Shaf}'(F, R)\}/\text{isometry}$.

For $X \in \mathrm{Shaf}'(F, R)$, we choose an element

$$(X, L_X, \nu_X, \alpha_X) \in M^\circ_{2d_X, D_{d_X}(n),F}(F).$$

Then, by Proposition 4.2.4 and [Zarhin 1985, Theorem 1] (for finitely generated fields of characteristic 0, see [Faltings et al. 1984, VI, §1, Theorem 2]), the subset

$$\{\Delta_{d_X}(X, L_X, \nu_X, \alpha_X) \mid X \in \mathrm{Shaf}'(F, R)\} \subset \mathrm{Sh}_{\mathbb{K}_n}(\mathrm{GSp}_{V,a})(F)$$

is finite. We denote them by $t_1, \ldots t_m$, and we put

$$\mathrm{Shaf}'(F, R)_i := \{X \in \mathrm{Shaf}'(F, R) \mid \Delta_{d_X}(X, L_X, \nu_X, \alpha_X) = t_i\}.$$

Thus, the desired finiteness follows from the following lemma.

Lemma 4.3.7. *The $\widehat{\mathbb{Z}}$-lattices $T(X)_{\widehat{\mathbb{Z}}}$ $(X \in \mathrm{Shaf}'(F, R)_i)$ are isometric to each other.*

Proof. By Lemma 4.3.6, it suffices to show that $\mathcal{L}_{\mathbb{Z}_\ell,(X,L_X,\alpha_X)}$ $(X \in \mathrm{Shaf}'(F, R)_i)$ is unique up to a $\mathrm{Gal}(\overline{F}/F)$-equivariant isometry, for any ℓ. We denote the lift of t_i on $\mathrm{Sh}_{D(n)}(\mathrm{SO}_{\mathcal{L}})$ via $h \circ \delta$ (it exists by the definition of t_i, and it is unique because $h \circ \delta$ is injective) by $\tilde{t}_i$. Recall that we have the étale sheaf $\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell}$, which has a symmetric pairing structure, so we get the $\mathrm{Gal}(\overline{F}/F)$-lattice $\tilde{t}_i^*(\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell})$, which depends only on t_i. By our construction of $\mathcal{L}_{\mathbb{Z}_\ell,(X,L,\alpha)}$ in Proposition 3.3.3, for any $X \in \mathrm{Shaf}'(F, R)_i$, the $\mathrm{Gal}(\overline{F}/F)$-lattice $\mathcal{L}_{\mathbb{Z}_\ell,(X,L_X,\alpha_X)}$ is none other than $\tilde{t}_i^*(\mathcal{L}^{\mathrm{shf}}_{\mathbb{Z}_\ell})$, finishing the proof. □

5. ℓ-independence

In this section, we prove ℓ-independence of the unramifiedness for completing the proof of the main theorem.

Lemma 5.0.1. *Let K be a Henselian discrete valuation field, k be the residue field of K, p be the characteristic of k, and X be a smooth proper surface X over K. Then, the following are equivalent.*

(a) *The $\mathrm{Gal}(\overline{K}/K)$-representation on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$ is unramified for some $\ell \neq p$.*

(b) *The $\mathrm{Gal}(\overline{K}/K)$-representation on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$ is unramified for all $\ell \neq p$.*

Moreover, if K is a complete discrete valuation field of mixed characteristic $(0, p)$ with the perfect residue field k and X is a K3 surface over K, then (a)(⇔(b)) is equivalent to the following.

(c) *The $\mathrm{Gal}(\overline{K}/K)$-representation on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_p)$ is crystalline.*

Remark 5.0.2. (1) In fact, for K3 surfaces, Lemma 5.0.1 is already mentioned by Madapusi Pera in [Matsumoto 2015, Remark 4.3] (using the Kuga–Satake construction, we can reduce the problem to the case of abelian varieties). We note that such arguments also appeared in [Imai and Mieda 2020]. So we will prove only the ℓ versus ℓ' part for general smooth proper surfaces as a corollary of Matsumoto's ℓ-independence result [2016, Theorem 3.3(2)].

(2) If we assume that X admits a Kulikov model after a finite extension of K, then Lemma 5.0.1 is known as a corollary of a good reduction criterion for K3 surfaces (see [Chiarellotto et al. 2019, Theorem 1.1] for example).

(3) For any smooth proper surface X over K, one can easily prove the similar assertion for $H^i_{\text{ét}}$ $(i \neq 2)$. Indeed, the case of $i = 0, 4$ is trivial. Moreover, for $i = 1, 3$, by using the Picard variety, we can reduce the problem to the case of abelian varieties. Therefore, it follows from the Néron–Ogg–Shafarevich criterion for abelian varieties (and its crystalline analogue [Coleman and Iovita 1999, Theorem 1]).

5.1. ***Proof of Lemma 5.0.1.*** In this subsection, we prove the Lemma 5.0.1. As in Remark 5.0.2, it is enough to show the equivalence (a)⇔(b) in Lemma 5.0.1. Let K be a Henselian discrete valuation field, k be the residue field of K, p be the characteristic of k, and X be a smooth proper surface over K.

First, we recall the definition of the monodromy operator.

Definition 5.1.1. Let ℓ be a prime number different from p. Consider the representation

$$\rho_\ell : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)).$$

By Grothendieck's monodromy theorem, there exists an open subgroup of the inertia subgroup $J \subset I_K$ and the nilpotent operator

$$N_\ell : H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)(1) \to H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$$

such that for all $\sigma \in J$, we have $\rho_\ell(\sigma) = \exp(t_\ell(\sigma)N_\ell)$, where $t_\ell : I_K \to \mathbb{Z}_\ell(1)$ is a natural projection. By fixing an isomorphism $\mathbb{Q}_\ell(1) \simeq \mathbb{Q}_\ell$, we regard N_ℓ as a linear endomorphism of $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$, which is called the *monodromy operator*.

Remark 5.1.2. By the definition, N_ℓ does not change if we replace K by a finite extension of it.

The following lemma is an elementary fact about ℓ-adic representations.

Lemma 5.1.3. *The following are equivalent.*

(1) *The ℓ-adic representation ρ_ℓ is unramified.*

(2) $N_\ell = 0$ *and* $\mathrm{tr}(\rho_\ell(\sigma)) = \dim(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell))$ *for any* $\sigma \in I_K$.

Proof. (1) $\Rightarrow$ (2) is trivial. Therefore we prove the opposite direction. By the definition of the monodromy operator, we have $\rho_\ell(g) = 1$ for any $g \in J$, where J is an open subgroup of I_K. Hence for any $\sigma \in I_K$, we get $\rho_\ell(\sigma)$ is of finite order, and the trace condition implies that $\rho_\ell(\sigma) = 1$. □

Definition 5.1.4. (1) There exists a unique increasing filtration $\mathrm{M}_r(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell))$ on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$ such that $\mathrm{M}_r = 0$ for $r \ll 0$, $\mathrm{M}_r = H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$ for $r \gg 0$, $N(\mathrm{M}_r) \subset \mathrm{M}_{r-2}$ and N^r induces an isomorphism $\mathrm{gr}^{\mathrm{M}}_r \simeq \mathrm{gr}^{\mathrm{M}}_{-r}$ for any positive integer r. We call M_r the *monodromy filtration* on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$.

(2) If X admits a strictly semistable model over $\mathcal{O}_K$, we get the *weight filtration* $\mathrm{W}_r(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell))$ on $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$ by the weight spectral sequence (see [Saito 2003, Corollary 2.8]). For general X, one can also define the weight filtration W_r by using de Jong's alteration.

Lemma 5.1.5. (1) *For any integer r, we have*

$$\mathrm{M}_r(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)) = \mathrm{W}_{r+2}(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)).$$

(2) *For any integer r, the dimension of* $\mathrm{gr}^{\mathrm{W}}_r(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell))$ *is independent of ℓ.*

Proof. Part (1) is well-known as the weight monodromy conjecture for surfaces (see [Rapoport and Zink 1982, Satz 2.13; Saito 2003, Lemma 3.9)]. Part (2) follows from [Matsumoto 2016, Theorem 3.3(2)]. □

The proof of (*a*) $\Leftrightarrow$ (*b*) *in Lemma 5.0.1.* Taking a completion, we may assume K is complete. We shall prove $(a) \Rightarrow (b)$. Take prime numbers $\ell, \ell' \neq p$. By [Ochiai 1999, Corollary 2.5] (for imperfect residue fields, see [Vidal 2004, Proposition 4.2]), we have $\mathrm{tr}(\rho_\ell(\sigma)) = \mathrm{tr}(\rho_{\ell'}(\sigma))$ for any $\sigma \in I_K$. By the definition of the monodromy filtration, we have

$$N_\ell = 0 \Leftrightarrow \dim(\mathrm{gr}^M_0(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell))) = \dim(H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)).$$

Therefore, by Lemmas 5.1.3 and 5.1.5, we get the desired implication.

6. Corollaries

6.1. *Some remarks.* First, combining Theorem 4.1.4 with Lemma 5.0.1, we obtain the main theorem in a more generalized form.

Theorem 6.1.1. *Let F be a finitely generated field over $\mathbb{Q}$, R be a finite type algebra over $\mathbb{Z}$ which is a normal domain with the fraction field F, and d be a positive integer. Then, the set*

$$S(F, R) := \{X \mid X : K3 \text{ surface over } F \text{ satisfying the condition } (C)\}/F\text{-isom}$$

is finite. Here, the condition (C) is the following.

(C) *For any height 1 prime $\mathfrak{p} \in \operatorname{Spec} R$, take a discrete valuation field $E_{\mathfrak{p}}$ such that $E_{\mathfrak{p}}$ is an algebraic extension of discrete valuation fields over $F = \operatorname{Frac}(R_{\mathfrak{p}})$, the residue field of $E_{\mathfrak{p}}$ is the perfection of the residue field of $R_{\mathfrak{p}}$, and a uniformizer of $R_{\mathfrak{p}}$ is also a uniformizer of $E_{\mathfrak{p}}$. Then, there exists a prime number ℓ different from the residual characteristic of $\mathfrak{p}$ such that $H^2_{\text{ét}}(X_{\overline{E_{\mathfrak{p}}}}, \mathbb{Q}_\ell)$ is an unramified $\operatorname{Gal}(\overline{F}/E_{\mathfrak{p}})$-representation.*

Remark 6.1.2. (1) The field extension $E_{\mathfrak{p}}$ in the condition (C) always exists by [Matsumura 1989, Theorem 29.1].

(2) By Lemma 5.0.1, the unramifiedness assumption in the condition (C) is independent of ℓ. If the residual characteristic of $\mathfrak{p}$ is positive, replacing $E_{\mathfrak{p}}$ by the completion of it, we can replace this condition in terms of crystalline representations.

Proof. Shrinking $\operatorname{Spec} R$ if necessary, we may assume that R is smooth over $\mathbb{Z}$ since the generic fiber $R \otimes_{\mathbb{Z}} \mathbb{Q}$ is generically smooth over $\mathbb{Q}$. Let M be the order of $\mathrm{GL}_{22}(\mathbb{F}_2)$. Shrinking $\operatorname{Spec} R$ again, we may assume that $1/M \in R$. Consider a height 1 prime $\mathfrak{p} \in \operatorname{Spec} R$, and we denote its residual characteristic by $p \geq 0$. Take an extension of valuation $\mathfrak{p}$ to $\overline{F}$, and we denote it by $\bar{v}$. We denote the inertia subgroups by $I_{\bar{v}} \subset \operatorname{Gal}(\overline{F}/F)$, $I'_{\bar{v}} \subset \operatorname{Gal}(\overline{F}/E_{\mathfrak{p}})$. We denote the $\operatorname{Gal}(\overline{F}/F)$-representation $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Z}_2)$ by ρ. Then, by Remark 6.1.2(2), we get $\rho(I'_{\bar{v}}) = 1$. If $p = 0$, we have $\rho(I_{\bar{v}}) = \rho(I'_{\bar{v}}) = 1$. If $p > 0$, for any finite index open normal subgroup H of $\rho(I_{\bar{v}})$, we get $[\rho(I_{\bar{v}}) : H]$ is a p-group. (Here, we use that for any finite extension of discrete valuation fields of characteristic 0, the extension degree is equal to the product of the ramification index and the inertia degree). Therefore we get

$$\rho(I_{\bar{v}}) \cap (1 + 2 \cdot \mathrm{Mat}_{22}(\mathbb{Z}_2)) = 1$$

since the former is pro-p and the latter is pro-2. Moreover, the image of $\rho(I_{\bar{v}})$ in $\mathrm{GL}_{22}(\mathbb{F}_2)$ via the reduction map is trivial because p does not divide M. Therefore, we get $\rho(I_{\bar{v}}) = 1$ even if $p > 0$. Thus we have $S(F, R) \subset \operatorname{Shaf}(F, R)$, so $S(F, R)$ is a finite set. □

Next, as an immediate consequence of Theorem 6.1.1, we obtain the unpolarized Shafarevich conjecture for K3 surfaces over finitely generated fields of characteristic 0.

Definition 6.1.3. Let $R_{\mathfrak{p}}$ be a discrete valuation ring with maximal ideal $\mathfrak{p}$, and F be the fraction field of $R_{\mathfrak{p}}$. For a K3 surface X over F, we say X has good reduction at $\mathfrak{p}$ if there exists a smooth proper algebraic space over $R_{\mathfrak{p}}$ whose generic fiber is isomorphic to X. Note that such model would be automatically a K3 family over $\operatorname{Spec} R_{\mathfrak{p}}$ (see Definition 2.1.1(2)).

Corollary 6.1.4. *Let F be a finitely generated field over $\mathbb{Q}$, and R be a finite type algebra over $\mathbb{Z}$ which is a normal domain with the fraction field F. Then, the set*

$$\bigl\{X \mid X : \textit{K3 surface over } F,\ X \textit{ has good reduction at any height } 1 \textit{ prime ideal } \mathfrak{p} \in \operatorname{Spec} R\bigr\}/F\textit{-isom}$$

is finite.

6.2. ***The finiteness of twists.*** Here, we give the finiteness result of twists of K3 surfaces via a finite extension of characteristic 0 fields.

Corollary 6.2.1. *Let F be a field of characteristic 0, E/F be a finite extension, and X be a $K3$ surface over F. Then, the set*

$$\mathrm{Tw}_{E/F}(X) := \{Y : \textit{K3 surface over } F \mid Y_E \simeq_E X_E\}/F\textit{-isom}$$

is finite.

Proof. Clearly, we may assume E/F is a finite Galois extension. First, we will reduce the problem to the case of finitely generated fields. Since $\operatorname{Aut}(X_{\overline{F}})$ is a finitely generated group [Sterk 1985, Proposition 2.2], extending E if necessary, we may assume $\operatorname{Aut}(X_E) = \operatorname{Aut}(X_{\overline{F}})$. We can take a finitely generated field $E' \subset E$ on which X and any elements of $\operatorname{Aut}(X_E)$ are defined. Moreover, by extending E' if necessary, we may assume E' is $\operatorname{Gal}(E/F)$-stable and $\operatorname{Gal}(E/F) \to \operatorname{Aut}(E')$ is injective. Let F' be the fixed subfield $E'^{\operatorname{Gal}(E/F)}$. Then, the description of twists

$$\mathrm{Tw}_{E/F}(X) \simeq H^1(\operatorname{Gal}(E/F), \operatorname{Aut}(X_E)) \simeq H^1(\operatorname{Gal}(E'/F'), \operatorname{Aut}(X_{E'})).$$

implies that the desired finiteness is reduced to the case of E'/F'.

Thus, in the following of this proof, we assume F is a finitely generated field and E/F is a finite Galois extension. One can take a smooth proper morphism of schemes $\mathcal{X} \to \operatorname{Spec} R$ whose generic fiber is X, where R is a smooth algebra over $\mathbb{Z}$ which is an integral domain with the fraction field F and $\frac{1}{2} \in R$. Then, via a monodromy action, we get $H^2_{\text{ét}}(X_{\overline{F}}, \mathbb{Z}_2)$ is unramified over $\operatorname{Spec} R$. Let $\widetilde{R}$ be the normalization of R in E. Shrinking $\operatorname{Spec} R$ if necessary, we may assume $\operatorname{Spec} \widetilde{R} \to \operatorname{Spec} R$ is a finite étale covering. Since E is unramified over $\operatorname{Spec} R$, by [Fu 2015, Proposition 3.3.6], we have

$$\operatorname{Ker}(\pi_1(\operatorname{Spec} E, \bar{s}) \to \pi_1(\operatorname{Spec} \widetilde{R}, \bar{s})) = \ker(\pi_1(\operatorname{Spec} F, \bar{s}) \to \pi_1(\operatorname{Spec} R, \bar{s})).$$

For any $Y \in \mathrm{Tw}_{E/F}(X)$, the isomorphism $Y_E \simeq_E X_E$ implies that the $\operatorname{Gal}(\overline{F}/E)$-action on $H^2_{\text{ét}}(Y_{\overline{F}}, \mathbb{Z}_2)$ arises from a $\pi_1(\operatorname{Spec} \widetilde{R}, \bar{s})$-action. Moreover, because of the above equality, the $\operatorname{Gal}(\overline{F}/F)$-action on $H^2_{\text{ét}}(Y_{\overline{F}}, \mathbb{Z}_2)$ also arises from a $\pi_1(\operatorname{Spec} R, \bar{s})$-action. Hence we get a natural inclusion $\mathrm{Tw}_{E/F}(X) \hookrightarrow \operatorname{Shaf}(F, R)$, and thus the desired finiteness follows from Theorem 4.1.4. $\square$

Acknowledgments

The author is deeply grateful to his advisor Naoki Imai for his deep encouragement and helpful advice. He carefully read the draft version of this paper and pointed out a lot of mistakes and typos. The author is greatly indebted to Tetsushi Ito for his warm encouragement, many comments, and invaluable suggestions for the proof. Moreover, the author would like to thank Kazuhiro Ito for helpful comments about the definition of the moduli of K3 surfaces, and Yoichi Mieda for valuable comments about Section 5. The author also thanks the referee for many helpful comments, including suggestions on Remark 3.2.2 and Proposition 4.2.4. The author is supported by the FMSP program at the University of Tokyo.

References

[André 1996] Y. André, "On the Shafarevich and Tate conjectures for hyper-Kähler varieties", *Math. Ann.* **305**:2 (1996), 205–248. MR Zbl

[Borel 1963] A. Borel, "Arithmetic properties of linear algebraic groups", pp. 10–22 in *Proceedings of the International Congress of Mathematicians* (Stockholm, 1962), edited by V. Stenström, Inst. Mittag-Leffler, Djursholm, Sweden, 1963. MR Zbl

[Cassels 1982] J. W. S. Cassels, "Rational quadratic forms", pp. 9–26 in *Proceedings of the International Mathematical Conference* (Singapore, 1981), edited by L. H. Y. Chen et al., North-Holland Math. Stud. **74**, North-Holland, Amsterdam, 1982. MR Zbl

[Chiarellotto et al. 2019] B. Chiarellotto, C. Lazda, and C. Liedtke, "A Néron–Ogg–Shafarevich criterion for K3 surfaces", *Proc. Lond. Math. Soc.* (3) **119**:2 (2019), 469–514. MR Zbl

[Coleman and Iovita 1999] R. Coleman and A. Iovita, "The Frobenius and monodromy operators for curves and abelian varieties", *Duke Math. J.* **97**:1 (1999), 171–215. MR Zbl

[Faltings 1983] G. Faltings, "Endlichkeitssätze für abelsche Varietäten über Zahlkörpern", *Invent. Math.* **73**:3 (1983), 349–366. MR Zbl

[Faltings et al. 1984] G. Faltings, G. Wüstholz, F. Grunewald, N. Schappacher, and U. Stuhler, *Rational points* (Bonn, 1983/1984), Aspects Math. **E6**, Vieweg & Sohn, Braunschweig, Germany, 1984. Zbl

[Fu 2015] L. Fu, *Etale cohomology theory*, 2nd revised ed., Nankai Tracts in Math. **14**, World Sci., Hackensack, NJ, 2015. MR Zbl

[Harada and Hiranouchi 2009] S. Harada and T. Hiranouchi, "Smallness of fundamental groups for arithmetic schemes", *J. Number Theory* **129**:11 (2009), 2702–2712. MR Zbl

[Huybrechts 2016] D. Huybrechts, *Lectures on K3 surfaces*, Cambridge Stud. Adv. Math. **158**, Cambridge Univ. Press, 2016. MR Zbl

[Imai and Mieda 2020] N. Imai and Y. Mieda, "Potentially good reduction loci of Shimura varieties", *Tunis. J. Math.* **2**:2 (2020), 399–454. MR Zbl

[Ito et al. 2018] K. Ito, T. Ito, and T. Koshikawa, "CM liftings of K3 surfaces over finite fields and their applications to the Tate conjecture", preprint, 2018. arXiv

[Liedtke and Matsumoto 2018] C. Liedtke and Y. Matsumoto, "Good reduction of K3 surfaces", *Compos. Math.* **154**:1 (2018), 1–35. MR Zbl

[Madapusi Pera 2015] K. Madapusi Pera, "The Tate conjecture for K3 surfaces in odd characteristic", *Invent. Math.* **201**:2 (2015), 625–668. MR Zbl

[Madapusi Pera 2016] K. Madapusi Pera, "Integral canonical models for spin Shimura varieties", *Compos. Math.* **152**:4 (2016), 769–824. MR Zbl

[Matsumoto 2015] Y. Matsumoto, "Good reduction criterion for K3 surfaces", *Math. Z.* **279**:1-2 (2015), 241–266. MR Zbl

[Matsumoto 2016] Y. Matsumoto, "Degeneration of K3 surfaces with non-symplectic automorphisms", preprint, 2016. arXiv

[Matsumura 1989] H. Matsumura, *Commutative ring theory*, 2nd ed., Cambridge Stud. Adv. Math. **8**, Cambridge Univ. Press, 1989. MR Zbl

[Milne 1990] J. S. Milne, "Canonical models of (mixed) Shimura varieties and automorphic vector bundles", pp. 283–414 in *Automorphic forms, Shimura varieties, and L-functions, I* (Ann Arbor, MI, 1988), edited by L. Clozel and J. S. Milne, Perspect. Math. **10**, Academic Press, Boston, 1990. MR Zbl

[Nikulin 1979] V. V. Nikulin, "Integral symmetric bilinear forms and some of their applications", *Izv. Akad. Nauk SSSR Ser. Mat.* **43**:1 (1979), 111–177. In Russian; translated in *Math. USSR-Izv.* **14**:1 (1980), 103–167. MR Zbl

[Ochiai 1999] T. Ochiai, "l-independence of the trace of monodromy", *Math. Ann.* **315**:2 (1999), 321–340. MR Zbl

[Rapoport and Zink 1982] M. Rapoport and T. Zink, "Über die lokale Zetafunktion von Shimuravarietäten: Monodromiefiltration und verschwindende Zyklen in ungleicher Charakteristik", *Invent. Math.* **68**:1 (1982), 21–101. MR Zbl

[Rizov 2006] J. Rizov, "Moduli stacks of polarized K3 surfaces in mixed characteristic", *Serdica Math. J.* **32**:2-3 (2006), 131–178. MR Zbl

[Rizov 2010] J. Rizov, "Kuga–Satake abelian varieties of K3 surfaces in mixed characteristic", *J. Reine Angew. Math.* **648** (2010), 13–67. MR Zbl

[Saito 2003] T. Saito, "Weight spectral sequences and independence of l", *J. Inst. Math. Jussieu* **2**:4 (2003), 583–634. MR Zbl

[SGA 7_{I} 1972] A. Grothendieck, *Groupes de monodromie en géométrie algébrique, I: Exposés I–II, VI–IX* (Séminaire de Géométrie Algébrique du Bois Marie 1967–1969), Lecture Notes in Math. **288**, Springer, 1972. MR Zbl

[She 2017] Y. She, "The unpolarized Shafarevich conjecture for K3 surfaces", preprint, 2017. arXiv

[Sterk 1985] H. Sterk, "Finiteness results for algebraic K3 surfaces", *Math. Z.* **189**:4 (1985), 507–513. MR Zbl

[Tate 1994] J. Tate, "Conjectures on algebraic cycles in l-adic cohomology", pp. 71–83 in *Motives* (Seattle, 1991), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Vidal 2004] I. Vidal, "Théorie de Brauer et conducteur de Swan", *J. Algebraic Geom.* **13**:2 (2004), 349–391. MR Zbl

[Zarhin 1985] Y. G. Zarhin, "A finiteness theorem for unpolarized abelian varieties over number fields with prescribed places of bad reduction", *Invent. Math.* **79**:2 (1985), 309–321. MR Zbl

Communicated by Hélène Esnault
Received 2019-10-15 Revised 2020-03-11 Accepted 2020-05-11

teppei@ms.u-tokyo.ac.jp *Graduate School of Mathematical Sciences, The University of Tokyo, Tokyo, Japan*

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2533

Iterated local cohomology groups and Lyubeznik numbers for determinantal rings

András C. Lőrincz and Claudiu Raicu

We give an explicit recipe for determining iterated local cohomology groups with support in ideals of minors of a generic matrix in characteristic zero, expressing them as direct sums of indecomposable $\mathcal{D}$-modules. For nonsquare matrices these indecomposables are simple, but this is no longer true for square matrices where the relevant indecomposables arise from the pole order filtration associated with the determinant hypersurface. Specializing our results to a single iteration, we determine the Lyubeznik numbers for all generic determinantal rings, thus answering a question of Hochster.

1. Introduction

We consider positive integers $m \geq n \geq 1$ and let $X = \mathbb{C}^{m\times n}$ denote the affine space of $m \times n$ complex matrices, equipped with the natural action of the group $\mathrm{GL} = \mathrm{GL}_m(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$. We denote the orbits of the GL-action by O_p, $0 \leq p \leq n$, where O_p consists of matrices of rank p, and write $H^\bullet_{\overline{O}_p}(-)$ for the functors of *local cohomology with support in the orbit closures*. If we let $S = \mathbb{C}[x_{ij}]$ denote the coordinate ring of X, and let I_{p+1} be the ideal of $(p+1)\times(p+1)$ minors of the matrix of indeterminates (x_{ij}), then I_{p+1} is the ideal of functions vanishing on the variety $\overline{O}_p$, and the functors $H^\bullet_{\overline{O}_p}(-)$ are often denoted by $H^\bullet_{I_{p+1}}(-)$, and referred to as the functors of *local cohomology with support in the ideal* I_{p+1}. The goal of this work is to give an explicit recipe for computing all the *iterated local cohomology groups*

$$H^\bullet_{\overline{O}_{i_1}}(H^\bullet_{\overline{O}_{i_2}}(\cdots H^\bullet_{\overline{O}_{i_r}}(S)\cdots)). \tag{1-1}$$

Specializing our results to the case $H^\bullet_{\overline{O}_0}(H^\bullet_{\overline{O}_p}(S))$ we determine the *Lyubeznik numbers* of the coordinate ring of each $\overline{O}_p$, and observe a dichotomy between the case of square and nonsquare matrices. This is explained geometrically by the way the conormal varieties to the orbits intersect in the two cases, and algebraically by the fact that an appropriate category of modules is semisimple for nonsquare matrices, and quite interesting for square matrices.

The groups (1-1) are finitely generated modules over the *Weyl algebra* $\mathcal{D}_X$ of differential operators on X, which in addition are *equivariant* for the action of the group GL. We will therefore work in the category $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ of GL-equivariant $\mathcal{D}_X$-modules, which is known by a result of Vilonen [1994, Theorem 4.3] to be equivalent to the category of finitely generated modules over a finite-dimensional algebra, or alternatively, to the category of finite-dimensional representations of a quiver with relations. The explicit

MSC2010: primary 13D45; secondary 13D07, 14M12.
Keywords: determinantal varieties, local cohomology, Lyubeznik numbers, equivariant D-modules.

description of the relevant quiver has been obtained in [Lőrincz and Walther 2019, Theorem 5.4], and it is closely related to that of the quiver attached to a slightly larger category considered in [Braden and Grinberg 1999, Section 4.1]. We identify a suitable finite set of indecomposable objects in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ and express each of the local cohomology groups in (1-1) as a direct sum of these indecomposables. The multiplicities of indecomposables are encoded in terms of Gaussian binomial coefficients (reviewed in Section 2B). Our proofs employ the symmetries coming from the GL-action, the inductive structure of determinantal varieties, and the quiver description of $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, as well as a number of vanishing results for local cohomology that we prove by working on appropriate desingularizations of determinantal varieties, and using Grothendieck duality and the Borel–Weil–Bott theorem.

For nonsquare matrices ($m > n$) the category $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is semisimple by [MacPherson and Vilonen 1986, Theorem 6.7], since the conormal varieties to the orbits (described in [Strickland 1982]) intersect in codimension ≥ 2. This has two important implications:

- The indecomposable modules in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ are simple.
- The module structure of $M \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is determined up to isomorphism by its class $[M]_{\mathcal{D}}$ in the *Grothendieck group* $\Gamma_{\mathcal{D}}$ *of* $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ (see Section 2D).

For this reason we begin by considering the simpler problem of determining the class in $\Gamma_{\mathcal{D}}$ of a local cohomology group. We return to the general case $m \geq n$ and let

$$D_0, D_1, \dots, D_n$$

denote the simple objects in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, where D_p has support equal to $\overline{O}_p$, and is often referred to as the *intersection homology* $\mathcal{D}_X$*-module* corresponding to the orbit O_p. When $p = n$, we have that $\overline{O}_n = X$ and $D_n = S$ is the coordinate ring of X. Our first theorem determines the class in $\Gamma_{\mathcal{D}}$ of the local cohomology groups of each D_p, thus generalizing the main result of [Raicu and Weyman 2014] which addresses the case $p = n$.

Theorem 1.1. *For every* $0 \leq t < p \leq n \leq m$ *we have the following equality in* $\Gamma_{\mathcal{D}}[q]$:

$$\sum_{j\geq 0}[H^j_{\overline{O}_t}(D_p)]_{\mathcal{D}} \cdot q^j = \sum_{s=0}^{t}[D_s]_{\mathcal{D}} \cdot q^{(p-t)^2+(p-s)\cdot(m-n)} \cdot \binom{n-s}{p-s}_{q^2} \cdot \binom{p-1-s}{t-s}_{q^2}. \tag{1-2}$$

The restriction to the case $t < p$ is done in order to avoid trivialities. If M is any S-module whose support is contained in $\overline{O}_t$ (such as $M = D_p$ or $M = H^j_{\overline{O}_p}(N)$ for $p \leq t$, $j \geq 0$, and any S-module N), then

$$H^0_{\overline{O}_t}(M) = M \quad \text{and} \quad H^i_{\overline{O}_t}(M) = 0 \qquad \text{for } i > 0. \tag{1-3}$$

For this reason, there is no harm in assuming for instance that $i_1 < i_2 < \cdots < i_r$ in (1-1).

Example 1.2. Consider the case when $m = 3$ and $n = 2$. For $p = 2$ and $t = 1$ we have $D_2 = S$ and

$$\sum_{j\geq 0}[H^j_{\overline{O}_1}(S)]_{\mathcal{D}} \cdot q^j \overset{(1\text{-}2)}{=} \sum_{s=0}^{1}[D_s]_{\mathcal{D}} \cdot q^{3-s} = [D_1]_{\mathcal{D}} \cdot q^2 + [D_0]_{\mathcal{D}} \cdot q^3,$$

which implies that the only nonzero local cohomology groups are in this case (see also [Walther 1999, Example 6.1])

$$H^2_{\overline{O}_1}(S) = D_1 \quad \text{and} \quad H^3_{\overline{O}_1}(S) = D_0.$$

For $p = 1$ and $t = 0$ we obtain

$$\sum_{j\geq 0}[H^j_{\overline{O}_0}(D_1)]_{\mathcal{D}} \cdot q^j \overset{(1\text{-}2)}{=} [D_0]_{\mathcal{D}} \cdot q^2 \cdot \binom{2}{1}_{q^2} = [D_0]_{\mathcal{D}} \cdot q^2 + [D_0]_{\mathcal{D}} \cdot q^4.$$

Combining this with the observation (1-3) it follows that the only nonzero groups $H^\bullet_{\overline{O}_0}(H^\bullet_{\overline{O}_1}(S))$ are

$$H^2_{\overline{O}_0}(H^2_{\overline{O}_1}(S)) = H^4_{\overline{O}_0}(H^2_{\overline{O}_1}(S)) = H^0_{\overline{O}_0}(H^3_{\overline{O}_1}(S)) = D_0. \tag{1-4}$$

Iterated local cohomology groups have been studied in the seminal work [Lyubeznik 1993], which introduced a new set of numerical invariants attached to any local ring which is a quotient of a regular local ring containing a field [Lyubeznik 1993, Theorem-Definition 4.1]. These invariants are known today under the name of *Lyubeznik numbers*, and have been the subject of extensive investigation (see [Núñez Betancourt et al. 2016], for example). For determinantal rings, the question of describing the Lyubeznik numbers was posed by Mel Hochster as part of his list of "Thirteen Open Questions about Local Cohomology". Part of our work here is dedicated to answering this question. For $p < n$ we have that S/I_{p+1} is the coordinate ring of $\overline{O}_p$, and we let $R^{(p)} = (S/I_{p+1})_{\mathfrak{m}}$ denote its localization at the maximal homogeneous ideal. The *Lyubeznik numbers* $\lambda_{i,j}(R^{(p)})$ are characterized by the equalities

$$H^i_{\overline{O}_0}(H^{m\cdot n-j}_{\overline{O}_p}(S)) = D_0^{\oplus\lambda_{i,j}(R^{(p)})}. \tag{1-5}$$

We encode the Lyubeznik numbers of determinantal rings by a bivariate generating function $L_p(q, w) \in \mathbb{Z}[q, w]$,

$$L_p(q, w) = \sum_{i,j\geq 0} \lambda_{i,j}(R^{(p)}) \cdot q^i \cdot w^j. \tag{1-6}$$

We prefer this encoding since it is more compact than the one given by the *Lyubeznik tables*

$$\Lambda(R^{(p)}) = (\lambda_{i,j}(R^{(p)}))_{0\leq i,j\leq \dim(R^{(p)})}$$

which were first considered in [Walther 2001]. We have for instance from (1-4) that when $m = 3$ and $n = 2$,

$$L_1(q, w) = w^3 + q^2 \cdot w^4 + q^4 \cdot w^4, \qquad \text{or equivalently} \quad \Lambda(R^{(1)}) = \begin{pmatrix} 0&0&0&1&0\\ 0&0&0&0&0\\ 0&0&0&0&1\\ 0&0&0&0&0\\ 0&0&0&0&1 \end{pmatrix}.$$

In this example, $R^{(1)}$ is the local ring at the vertex of the affine cone of the Segre embedding $\mathbb{P}^1 \times \mathbb{P}^2 \to \mathbb{P}^5$. Since $\mathbb{P}^1 \times \mathbb{P}^2$ is smooth, it is known that the Lyubeznik numbers have a topological interpretation, being determined by the Betti numbers of $\mathbb{P}^1 \times \mathbb{P}^2$ [García López and Sabbah 1998; Switala 2015]. By contrast, there are singular examples where the Lyubeznik numbers at the cone point depend on the projective embedding [Reichelt et al. 2018; Wang 2020], so the topology of the projective scheme does not control on its own the Lyubeznik numbers. Nevertheless, based on the work [Reichelt et al. 2018], one can show that in the case of (projective) determinantal varieties most of the Lyubeznik numbers do not depend on the choice of embedding into a projective space (see [Reichelt and Walther $\geq$ 2020]).

For nonsquare matrices our Theorem 1.1, together with the fact that $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is semisimple, gives the following description of Lyubeznik numbers.

Theorem 1.3. *If $m > n > p$ then the Lyubeznik numbers for $R^{(p)}$ are computed by*

$$L_p(q,w)=\sum_{s=0}^{p} q^{s^2+s\cdot(m-n)}\cdot\binom{n}{s}_{q^2}\cdot w^{p^2+2p+s\cdot(m+n-2p-2)}\cdot\binom{n-1-s}{p-s}_{w^2}. \tag{1-7}$$

In fact, using Theorem 1.1 and the semisimplicity of $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ we can determine (1-1), and in particular describe the *generalized Lyubeznik numbers* as defined in [Núñez Betancourt et al. 2016, Section 7]. More generally,

$$H^{\bullet}_{\overline{O}_{i_1}}\big(H^{\bullet}_{\overline{O}_{i_2}}(\cdots H^{\bullet}_{\overline{O}_{i_r}}(D_p)\cdots)\big)$$

can be computed for any D_p. We leave the determination of the precise formulas to the interested reader.

When $m=n$ the situation is more subtle, as can be seen already in the following simple example.

Example 1.4. Suppose that $m=n=2$ and let $p=1$. Applying (1-2) we get

$$[H^1_{\overline{O}_1}(S)]_{\mathcal{D}}=[D_0]_{\mathcal{D}}+[D_1]_{\mathcal{D}}$$

but $H^1_{\overline{O}_1}(S)$ is not the direct sum of D_0 and D_1! If we write det for the 2×2 determinant, then $H^1_{\overline{O}_1}(S)=S_{\det}/S$ contains no nonzero elements annihilated by the maximal homogeneous ideal, so it can't contain D_0 (which is supported at 0) as a submodule. This observation is also reflected in the calculation of Lyubeznik numbers, as follows. Since $\overline{O}_1$ is a hypersurface of (affine) dimension 3 (the cone over $\mathbb{P}^1\times\mathbb{P}^1\subset\mathbb{P}^3$), the only nonzero Lyubeznik number is $\lambda_{3,3}(R^{(1)})=1$, that is the only nonzero group $H^{\bullet}_{\overline{O}_0}(H^{\bullet}_{\overline{O}_1}(S))$ is

$$H^3_{\overline{O}_0}(H^1_{\overline{O}_1}(S))=D_0.$$

The nonzero local cohomology groups $H^{\bullet}_{\overline{O}_0}(D_0)$ and $H^{\bullet}_{\overline{O}_0}(D_1)$ are by (1-2) and (1-3)

$$H^0_{\overline{O}_0}(D_0)=H^1_{\overline{O}_0}(D_1)=H^3_{\overline{O}_0}(D_1)=D_0,$$

so the local cohomology groups of $H^1_{\overline{O}_1}(S)$ are not the direct sums of those of D_0 and D_1. In particular, specializing (1-7) to the case when $m=n$ would give the wrong answer! Instead, we have the following.

Theorem 1.5. *If $m=n$ then $L_{n-1}(q,w)=(q\cdot w)^{n^2-1}$ and for $0\leq p\leq n-2$ we have*

$$L_p(q,w)=\sum_{s=0}^{p} q^{s^2+2s}\cdot\binom{n-1}{s}_{q^2}\cdot w^{p^2+2p+s\cdot(2n-2p-2)}\cdot\binom{n-2-s}{p-s}_{w^2}. \tag{1-8}$$

For instance, in the case of 4×4 matrices of rank at most 2 ($m = n = 4$ and $p = 2$) we obtain

$$L_2(q, w) = w^8 + (q^3 + q^5 + q^7) \cdot w^{10} + (q^8 + q^{10} + q^{12}) \cdot w^{12}. \tag{1-9}$$

Analogues of Theorems 1.3 and 1.5 for ideals of Pfaffians of a generic skew-symmetric matrix have been obtained by Mike Perlman [2020], but the corresponding problem for symmetric matrices remains open.

As we saw in Example 1.4, for square matrices the (iterated) local cohomology groups of S are no longer expressible as direct sums of the simple modules D_p. We proceed instead to construct a different set of indecomposables that play the role of the simples. We let $\det = \det(x_{ij})$ denote the determinant of the generic $n \times n$ matrix, and let $\langle \det^{-p} \rangle_{\mathcal{D}}$ denote the $\mathcal{D}_X$-submodule of $S_{\det}$ generated by $\det^{-p}$. It is shown in [Raicu 2016, Theorem 1.1] that

$$0 \subsetneq S \subsetneq \langle \det^{-1} \rangle_{\mathcal{D}} \subsetneq \cdots \subsetneq \langle \det^{-n} \rangle_{\mathcal{D}} = S_{\det} \tag{1-10}$$

is a $\mathcal{D}_X$-module composition series with composition factors $S \simeq D_n$ and

$$\frac{\langle \det^{-p} \rangle_{\mathcal{D}}}{\langle \det^{-p+1} \rangle_{\mathcal{D}}} \simeq D_{n-p} \qquad \text{for } p = 1, \ldots, n. \tag{1-11}$$

We define $Q_n = S_{\det}$ and for $p = 0, \ldots, n-1$, we let

$$Q_p = \frac{S_{\det}}{\langle \det^{p-n+1} \rangle_{\mathcal{D}}}. \tag{1-12}$$

It follows from (1-10) and (1-11) that Q_p has composition factors $D_0, \ldots, D_p$, hence

$$[Q_p]_{\mathcal{D}} = \sum_{s=0}^{p} [D_s]_{\mathcal{D}} \tag{1-13}$$

and the support of Q_p is $\overline{O}_p$. We denote by $\operatorname{add}(Q)$ the full additive subcategory of $\operatorname{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ consisting of modules that are isomorphic to a direct sum of copies of $Q_0, Q_1, \ldots, Q_n$. It follows from (1-13) that $[Q_0]_{\mathcal{D}}, \ldots, [Q_n]_{\mathcal{D}}$ form a basis of the Grothendieck group $\Gamma_{\mathcal{D}}$, so a module $M \in \operatorname{add}(Q)$ is determined up to isomorphism by $[M]_{\mathcal{D}}$. The following result (when combined with (1-2), (1-3), and (1-13)) allows one to determine (1-1) when $m = n$, or more generally to describe arbitrary iterations

$$H^{\bullet}_{\overline{O}_{i_1}}(H^{\bullet}_{\overline{O}_{i_2}}(\cdots H^{\bullet}_{\overline{O}_{i_r}}(M)\cdots)), \qquad \text{where } M = D_p \text{ or } M = Q_p,\ \ p = 0, \ldots, n.$$

Theorem 1.6. *For every $0 \le t < p \le n = m$ and $j \ge 0$ we have that*

$$H^j_{\overline{O}_t}(D_p) \in \operatorname{add}(Q) \quad \textit{and} \quad H^j_{\overline{O}_t}(Q_p) \in \operatorname{add}(Q).$$

Moreover,

$$\sum_{j \ge 0} [H^j_{\overline{O}_t}(Q_p)]_{\mathcal{D}} \cdot q^j = \sum_{s=0}^{t} [Q_s]_{\mathcal{D}} \cdot q^{(p-t)^2 + 2(p-s)} \cdot \binom{n-s-1}{p-s}_{q^2} \cdot \binom{p-s-1}{p-t-1}_{q^2}. \tag{1-14}$$

This theorem is explained in Section 6. A formula analogous to (1-14) holds for the groups $H^j_{\overline{O}_t}(D_p)$, and can be obtained based on (1-2) from the fact that $H^j_{\overline{O}_t}(D_p) \in \mathrm{add}(Q)$ (see Theorem 6.1). To see how Theorem 1.6 allows for the calculation of Lyubeznik numbers, or more general iterated local cohomology groups, we explain next how to derive (1-9).

Example 1.7. If $m = n = 4$ and $p = 2$ then we have

$$\begin{aligned}\sum_{j\geq 0}[H^j_{\overline{O}_2}(S)]_{\mathcal{D}} \cdot q^j &\overset{(1\text{-}2)}{=} [D_0]_{\mathcal{D}} \cdot (q^4+q^6+q^8) + [D_1]_{\mathcal{D}} \cdot (q^4+q^6) + [D_2]_{\mathcal{D}} \cdot q^4 \\ &\overset{(1\text{-}13)}{=} [Q_2]_{\mathcal{D}} \cdot q^4 + [Q_1]_{\mathcal{D}} \cdot q^6 + [Q_0]_{\mathcal{D}} \cdot q^8.\end{aligned}$$

By Theorem 1.6 we have that $H^j_{\overline{O}_2}(S) \in \mathrm{add}(Q)$ for all j; hence

$$H^4_{\overline{O}_2}(S) = Q_2, \quad H^6_{\overline{O}_2}(S) = Q_1, \quad \text{and} \quad H^8_{\overline{O}_2}(S) = Q_0 = D_0.$$

Using (1-3) we get $H^0_{\overline{O}_0}(H^8_{\overline{O}_2}(S)) = D_0$ and therefore $\lambda_{0,8}(R^{(2)}) = 1$. Using (1-14) we have

$$\sum_{j\geq 0}[H^j_{\overline{O}_0}(Q_1)]_{\mathcal{D}} \cdot q^j = [Q_0]_{\mathcal{D}} \cdot q^3 \cdot \binom{3}{1}_{q^2} = [D_0]_{\mathcal{D}} \cdot (q^3+q^5+q^7)$$

and therefore $\lambda_{3,10}(R^{(2)}) = \lambda_{5,10}(R^{(2)}) = \lambda_{7,10}(R^{(2)}) = 1$. Using (1-14) again we have

$$\sum_{j\geq 0}[H^j_{\overline{O}_0}(Q_2)]_{\mathcal{D}} \cdot q^j = [Q_0]_{\mathcal{D}} \cdot q^8 \cdot \binom{3}{2}_{q^2} = [D_0]_{\mathcal{D}} \cdot (q^8+q^{10}+q^{12})$$

and therefore $\lambda_{8,12}(R^{(2)}) = \lambda_{10,12}(R^{(2)}) = \lambda_{12,12}(R^{(2)}) = 1$. All the remaining Lyubeznik numbers vanish, proving (1-9).

The paper is organized as follows. In Section 2 we recall some basic notions regarding weights and Schur functors, q-binomial coefficients, categories of admissible representations and equivariant $\mathcal{D}$-modules, and Bott's theorem for Grassmannians and flag varieties. We also discuss briefly families of determinantal rings over a general base, and the inductive structure of determinantal rings. In Section 3 we prove Theorems 1.1 and 1.3. Sections 4 and 5 are concerned with a number of technical results proving the vanishing of a range of local cohomology groups. In Section 6 we recall the quiver description of the category $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ and use it in conjunction with the vanishing results of the earlier sections to provide an inductive proof of Theorem 1.6. We also derive Theorem 1.5 as a quick corollary of the previous local cohomology calculations.

2. Preliminaries

2A. *Dominant weights and Schur functors.* We write $\mathbb{Z}^n_{\mathrm{dom}}$ for the set of *dominant weights* in $\mathbb{Z}^n$, i.e., tuples $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. When each $\lambda_i \geq 0$ we identify λ with a *partition* with (at most) n parts, and write $\lambda \in \mathbb{N}^n_{\mathrm{dom}}$. When $\lambda \in \mathbb{Z}^n$ is not dominant, it must contain *inversions*, i.e., pairs (i, j) with $i < j$ and $\lambda_i < \lambda_j$. The *size* of λ is $|\lambda| = \lambda_1 + \cdots + \lambda_n$. We sometimes use Greek letters to denote weights $\lambda \in \mathbb{Z}^n_{\mathrm{dom}}$ and underlined Roman letters to denote partitions $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$. We write $\underline{x}'$ for the *conjugate* partition of $\underline{x}$, where $\underline{x}'_i$ counts the number of parts x_j with $x_j \geq i$. We partially order $\mathbb{Z}^n_{\mathrm{dom}}$ (and $\mathbb{N}^n_{\mathrm{dom}}$) by declaring $\lambda \geq \mu$ if $\lambda_i \geq \mu_i$ for all $i = 1, \dots, n$. If $a \geq 0$ then we write $a \times b$ or (b^a) for the sequence $(b, b, \dots, b)$ where b is repeated a times.

If V is a vector space with $\dim(V) = n$ and $\lambda \in \mathbb{Z}^n_{\mathrm{dom}}$ we write $\mathbb{S}_\lambda V$ for the corresponding irreducible representation of $\mathrm{GL}(V)$ (or *Schur functor*). Our conventions are such that if $\lambda = (d, 0, \dots, 0)$ then $\mathbb{S}_\lambda V = \mathrm{Sym}^d V$, and if $\lambda = (1^n)$ then $\mathbb{S}_\lambda V = \bigwedge^n V$. More generally, one can define $\mathbb{S}_\lambda \mathcal{E}$ for any locally free sheaf $\mathcal{E}$ of rank n on some algebraic variety X. We write $\det(\mathcal{E})$ for $\bigwedge^n \mathcal{E}$ and call it the *determinant* of $\mathcal{E}$. For $m > n$ we will always think of $\mathbb{N}^n_{\mathrm{dom}}$ as a subset of $\mathbb{N}^m_{\mathrm{dom}}$ by identifying $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$ with $(\underline{x}, 0^{m-n})$, and in this way $\mathbb{S}_{\underline{x}} V$ (resp. $S_{\underline{x}}\mathcal{E}$) is defined whenever $\dim(V) \geq n$ (resp. $\mathrm{rank}(\mathcal{E}) \geq n$).

2B. *Gaussian binomial coefficients.* For $a \geq b \geq 0$ we define the *Gaussian (or q-)binomial coefficient* $\binom{a}{b}_q$ to be the polynomial in $\mathbb{Z}[q]$ defined by

$$\binom{a}{b}_q = \frac{(1-q^a)\cdot(1-q^{a-1})\cdots(1-q^{a-b+1})}{(1-q^b)\cdot(1-q^{b-1})\cdots(1-q)}.$$

These polynomials are generalizations of the usual binomial coefficients, satisfying the relations

$$\binom{a}{b}_q = \binom{a}{a-b}_q, \quad \binom{a}{a}_q = \binom{a}{0}_q = 1, \quad \text{and} \quad \binom{a}{b}_1 = \binom{a}{b}. \tag{2-1}$$

One significance of the q-binomial coefficients is that $\binom{a}{b}_{q^2}$ describes the Poincaré polynomial of the Grassmannian of b-dimensional subspaces of $\mathbb{C}^a$. As such, the coefficient of q^j in $\binom{a}{b}_q$ computes the number of Schubert classes of (co)dimension j, or equivalently the number of partitions $\underline{x}$ of size j contained inside the rectangular partition $(a-b) \times b$. We get

$$\binom{a}{b}_q = \sum_{\underline{x} \leq (b^{a-b})} q^{|\underline{x}|}. \tag{2-2}$$

Using the fact that the map $\underline{x} \mapsto \underline{x}^\circ := (b - x_{a-b}, b - x_{a-b-1}, \dots, b - x_2, b - x_1)$ defines an involution on the set of partitions $\underline{x} \leq (b^{a-b})$, satisfying $|\underline{x}^\circ| = b \cdot (a-b) - |\underline{x}|$, we get that

$$\binom{a}{b}_{q^{-1}} = \binom{a}{b}_q \cdot q^{-b\cdot(a-b)}. \tag{2-3}$$

The q-binomial coefficients also satisfy recurrence relations analogous to the Pascal identities for usual binomial coefficients, namely

$$\binom{a}{b}_q = q^b \cdot \binom{a-1}{b}_q + \binom{a-1}{b-1}_q. \tag{2-4}$$

2C. *The ring of polynomial functions on $m \times n$ matrices and its equivariant ideals.* We consider positive integers $m \geq n \geq 1$ and let $X = \mathbb{C}^{m\times n}$ denote the affine space of $m \times n$ complex matrices. We let $\mathrm{GL} = \mathrm{GL}_m(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$ and consider its natural action on X via row and column operations. The orbits of this action are the sets O_p consisting of matrices of rank p, for $p = 0, \dots, n$, and their orbit closures are given by

$$\overline{O}_p = \bigcup_{i=0}^{p} O_i.$$

The coordinate ring S of X can be identified with the polynomial ring $S = \mathbb{C}[x_{ij}]$, where $1 \le i \le m$ and $1 \le j \le n$. If we write I_p for the ideal of $p \times p$ minors of the generic matrix (x_{ij}), then I_p is the defining ideal of the closed subvariety $\overline{O}_{p-1}$ of X. To keep track of the equivariance it is convenient to identify the space of linear forms in S with the tensor product $\mathbb{C}^m \otimes \mathbb{C}^n$, which has a natural GL-action. The polynomial ring S can then be thought of as the *symmetric algebra* $\mathrm{Sym}_{\mathbb{C}}(\mathbb{C}^m \otimes \mathbb{C}^n) = \bigoplus_{d \ge 0} \mathrm{Sym}^d(\mathbb{C}^m \otimes \mathbb{C}^n)$, where the component indexed by d corresponds to homogeneous forms of degree d in the variables x_{ij}. The structure of S as a GL-representation is governed by Cauchy's formula [Weyman 2003, Corollary 2.3.3]

$$S = \bigoplus_{\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}} \mathbb{S}_{\underline{x}}\mathbb{C}^m \otimes \mathbb{S}_{\underline{x}}\mathbb{C}^n. \tag{2-5}$$

We write $I_{\underline{x}} \subset S$ for the ideal generated by the component $\mathbb{S}_{\underline{x}}\mathbb{C}^m \otimes \mathbb{S}_{\underline{x}}\mathbb{C}^n$ in the above decomposition. If $\underline{x} = (1^p)$ then the ideal $I_{\underline{x}}$ coincides with the ideal I_p defined earlier. As a GL-representation we have

$$I_{\underline{x}} = \bigoplus_{\underline{y} \ge \underline{x}} \mathbb{S}_{\underline{y}}\mathbb{C}^m \otimes \mathbb{S}_{\underline{y}}\mathbb{C}^n. \tag{2-6}$$

2D. ***Equivariant $\mathcal{D}$-modules and the Grothendieck group $\Gamma_{\mathcal{D}}$.*** We write $X = \mathbb{C}^{m \times n}$ as in the previous section, let $\mathcal{D}_X$ denote the sheaf of differential operators on X, and let $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ denote the category of GL-*equivariant coherent $\mathcal{D}_X$-modules*. The category $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is a full subcategory of the category of coherent $\mathcal{D}_X$-modules, stable under taking subquotients (for more details on categories of equivariant $\mathcal{D}_X$-modules, see [Lőrincz and Walther 2019, Section 2.1]). The simple objects in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ are $D_0, \dots, D_n$, where D_p denotes the intersection homology $\mathcal{D}$-module corresponding to the orbit O_p. As a GL-representation, D_p decomposes as (see [Raicu and Weyman 2014, Theorem 6.1; 2016, Main Theorem(1); Raicu 2017, Theorem 5.1])

$$D_p = \bigoplus_{\substack{\lambda_p \ge p-n \\ \lambda_{p+1} \le p-m}} \mathbb{S}_{\lambda(p)}\mathbb{C}^m \otimes \mathbb{S}_{\lambda}\mathbb{C}^n, \tag{2-7}$$

where

$$\lambda(p) = \big(\lambda_1, \dots, \lambda_p, (p-n)^{m-n}, \lambda_{p+1} + (m-n), \dots, \lambda_n + (m-n)\big). \tag{2-8}$$

We note that for $p = n$ the formulas in (2-5) and (2-7) coincide, which is a reflection of the fact that $D_n = S$.

We write $\Gamma_{\mathcal{D}}$ for the *Grothendieck group of* $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, and write $[M]_{\mathcal{D}}$ for the class in $\Gamma_{\mathcal{D}}$ of an equivariant $\mathcal{D}_X$-module M. We note that the group $\Gamma_{\mathcal{D}}$ is a free abelian group of rank $(n+1)$, with basis given by $[D_p]_{\mathcal{D}}$, for $p = 0, \dots, n$. An important construction of new objects in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ comes from considering the local cohomology groups $H^j_{\overline{O}_t}(M)$ for $j \ge 0$, $0 \le t \le n$, and $M \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$. A first approximation to the structure of these groups is given by their class in $\Gamma_{\mathcal{D}}$. To keep track of this information it is convenient to write $\Gamma_{\mathcal{D}}[q]$ for the additive group of polynomials in the variable q with coefficients in $\Gamma_{\mathcal{D}}$, and define

$$H_t^{\mathcal{D}}(M;q)=\sum_{j\geq 0}[H^j_{\overline{O}_t}(M)]_{\mathcal{D}}\cdot q^j\in\Gamma_{\mathcal{D}}[q]. \tag{2-9}$$

In the case when $M=S$, the main result of [Raicu and Weyman 2014] (as interpreted in [Raicu and Weyman 2016, Main Theorem(1)]) yields

$$H_t^{\mathcal{D}}(S;q)=\sum_{s=0}^{t}[D_s]_{\mathcal{D}}\cdot q^{(n-t)^2+(n-s)\cdot(m-n)}\cdot\binom{n-1-s}{t-s}_{q^2}. \tag{2-10}$$

We define a pairing $\langle\ ,\ \rangle_{\mathcal{D}}:\Gamma_{\mathcal{D}}[q]\times\Gamma_{\mathcal{D}}[q]\to\mathbb{Z}[q]$ given by

$$\langle\gamma(q),\gamma'(q)\rangle_{\mathcal{D}}=\sum_{s=0}^{n}\gamma_s(q)\cdot\gamma'_s(q),$$

where $\gamma(q)=\sum_{s=0}^{n}[D_s]_{\mathcal{D}}\cdot\gamma_s(q)$ and $\gamma'(q)=\sum_{s=0}^{n}[D_s]_{\mathcal{D}}\cdot\gamma'_s(q)$. The assertion (2-10) is then equivalent to

$$\langle H_t^{\mathcal{D}}(S;q),D_s\rangle_{\mathcal{D}}=q^{(n-t)^2+(n-s)\cdot(m-n)}\cdot\binom{n-1-s}{t-s}_{q^2}\text{ for }0\leq s\leq t,\ \text{ and }\ \langle H_t^{\mathcal{D}}(S;q),D_s\rangle_{\mathcal{D}}=0\text{ for }s>t.$$

Notice that in the formula above we have written D_s instead of $[D_s]_{\mathcal{D}}$, to simplify the notation. We will continue to do so as long as there is no possible source of confusion.

2E. *Admissible representations and the Grothendieck group* Γ_{GL}. We define an *admissible representation* of GL to be a representation M that decomposes as

$$M=\bigoplus_{\substack{\lambda\in\mathbb{Z}^m_{\mathrm{dom}}\\ \mu\in\mathbb{Z}^n_{\mathrm{dom}}}}(\mathbb{S}_\lambda\mathbb{C}^m\otimes\mathbb{S}_\mu\mathbb{C}^n)^{\oplus a_{\lambda,\mu}}$$

for some nonnegative integers $a_{\lambda,\mu}$. Examples of such representations include the polynomial ring in (2-5), the ideals (2-6), and the $\mathcal{D}_X$-modules in (2-7). More generally, if M is a finitely generated GL-equivariant S-module or $\mathcal{D}_X$-module then M is an admissible representation.

We write Γ_{GL} for the *Grothendieck group of admissible* GL-*representations*, and write $[M]_{\mathrm{GL}}$ for the class in Γ_{GL} of a representation M, and often refer to $[M]_{\mathrm{GL}}$ as a *character*. The admissible representations form a semisimple category, which implies that $[M]_{\mathrm{GL}}$ determines M up to isomorphism. We have that Γ_{GL} is isomorphic to the product of copies of $\mathbb{Z}$ indexed by $s_{\lambda,\mu}=[\mathbb{S}_\lambda\mathbb{C}^m\otimes\mathbb{S}_\mu\mathbb{C}^n]_{\mathrm{GL}}$, with $\lambda\in\mathbb{Z}^m_{\mathrm{dom}}$ and $\mu\in\mathbb{Z}^n_{\mathrm{dom}}$. We define $\Gamma_{\mathrm{GL}}[q]$ in analogy with $\Gamma_{\mathcal{D}}[q]$, and express any $\gamma(q)\in\Gamma_{\mathrm{GL}}(q)$ as an infinite sum

$$\gamma(q)=\sum_{\lambda,\mu}a_{\lambda,\mu}(q)\cdot s_{\lambda,\mu},\qquad\text{with }a_{\lambda,\mu}(q)\in\mathbb{Z}.$$

We consider the partially defined pairing $\langle\ ,\ \rangle : \Gamma_{\mathrm{GL}}[q] \times \Gamma_{\mathrm{GL}}[q] \to \mathbb{Z}[q]$

$$\langle \gamma(q), \gamma'(q) \rangle_{\mathrm{GL}} = \sum_{\lambda,\mu} a_{\lambda,\mu}(q) \cdot a'_{\lambda,\mu}(q) \tag{2-11}$$

whenever the sum (2-11) involves only finitely many nonzero terms.

We have a forgetful map that associates to a module $M \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ the underlying admissible representation. This induces a homomorphism $\Gamma_{\mathcal{D}} \to \Gamma_{\mathrm{GL}}$ given by $[M]_{\mathcal{D}} \mapsto [M]_{\mathrm{GL}}$. It will be important to note that this homomorphism is injective, since the characters $[D_p]_{\mathrm{GL}}$ described by (2-7) are linearly independent. In other words, the composition factors of a GL-equivariant $\mathcal{D}$-module (and their multiplicities) are uniquely determined by its character. If we combine (1-13) with the case $m = n$ of (2-7) (so that $\lambda(s) = \lambda$ for all s) then it follows that as a GL-representation Q_p decomposes as

$$Q_p = \bigoplus_{\lambda_{p+1} \le p-n} \mathbb{S}_\lambda \mathbb{C}^n \otimes \mathbb{S}_\lambda \mathbb{C}^n. \tag{2-12}$$

We extend the map $\Gamma_{\mathcal{D}} \to \Gamma_{\mathrm{GL}}$ to an injective homomorphism $\Gamma_{\mathcal{D}}[q] \to \Gamma_{\mathrm{GL}}[q]$, and note that for instance the image of (2-9) via this homomorphism is

$$H_t^{\mathrm{GL}}(M; q) = \sum_{j \ge 0} [H^j_{\overline{O}_t}(M)]_{\mathrm{GL}} \cdot q^j \tag{2-13}$$

Taking $W = \mathbb{S}_{\lambda(p)} \mathbb{C}^m \otimes \mathbb{S}_\lambda \mathbb{C}^n$ to be any representation that appears in (2-7) it follows that

$$\langle H_t^{\mathcal{D}}(M; q), D_p \rangle_{\mathcal{D}} = \langle H_t^{\mathrm{GL}}(M; q), W \rangle_{\mathrm{GL}} \tag{2-14}$$

for any $M \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, which will be particularly useful for our calculations in Section 3. Notice again the abuse of notation where we simply write W instead of $[W]_{\mathrm{GL}}$, since there is no possibility of confusion.

2F. ***Flag varieties, Grassmannians, and Bott's Theorem*** [Weyman 2003, Chapters 3 and 4]**.** Consider nonnegative integers $p \le n$ and a complex vector space V with $\dim(V) = n$. We denote by $\mathrm{Flag}([p,n]; V)$ the variety of partial flags

$$V_\bullet : \quad V = V_n \twoheadrightarrow V_{n-1} \cdots \twoheadrightarrow V_p \twoheadrightarrow 0,$$

where V_q is a q–dimensional quotient of V for each $q = p, p+1, \dots, n$. For $q \in [p,n]$ we write $\mathcal{Q}_q(V)$ for the tautological rank q quotient bundle on $\mathrm{Flag}([p,n]; V)$ whose fiber over a point $V_\bullet \in \mathrm{Flag}([p,n]; V)$ is V_q. We consider the natural projection maps

$$\pi_V^{(p)} : \mathrm{Flag}([p,n]; V) \to \mathrm{Flag}([p+1,n]; V), \tag{2-15}$$

defined by forgetting V_p from the flag $V_\bullet$. For $p \le n-1$, this map identifies $\mathrm{Flag}([p,n]; V)$ with the projective bundle $\mathbb{P}_{\mathrm{Flag}([p+1,n]; V)}(\mathcal{Q}_{p+1}(V))$, which comes with a tautological surjection

$$\mathcal{Q}_{p+1}(V) \twoheadrightarrow \mathcal{Q}_p(V). \tag{2-16}$$

The careful reader may have noticed that we are using the same notation $\mathcal{Q}_q(V)$ for the tautological rank q quotient bundle on each of the spaces $\mathrm{Flag}([p,n];\,V)$ with $p \le q \le n$. This should cause no confusion (but has the advantage of simplifying the notation), as the bundle $\mathcal{Q}_q(V)$ on $\mathrm{Flag}([p,n];\,V)$ is simply the pull-back along $\pi^{(p)}$ of the corresponding bundle on $\mathrm{Flag}([p+1,n];\,V)$ when $p \le q-1$.

The kernel of (2-16) is a line bundle which we denote $\mathcal{L}_{p+1}(V)$ and note that

$$\det(\mathcal{Q}_{p+1}(V)) = \mathcal{L}_{p+1}(V) \otimes \det(\mathcal{Q}_p(V)). \tag{2-17}$$

Just as with $\mathcal{Q}_q(V)$, there is one line bundle $\mathcal{L}_q(V)$ on each of the spaces $\mathrm{Flag}([p,n];\,V)$ with $p \le q-1$. When $p > 0$, the Picard group of $\mathrm{Flag}([p,n];\,V)$ is free of rank $(n-p)$, with $\mu \in \mathbb{Z}^{n-p}$ corresponding to the line bundle

$$\mathcal{L}^{\mu}(V) = \bigotimes_{i=1}^{n-p} \mathcal{L}_{p+i}(V)^{\otimes \mu_i}. \tag{2-18}$$

Note that (2-17) can be used to prove inductively that

$$\det(V) \otimes \mathcal{O}_{\mathrm{Flag}([p,n];\,V)} = \mathcal{L}^{(1^{n-p})}(V) \otimes \det(\mathcal{Q}_p(V)). \tag{2-19}$$

In particular for $p = 0$ (when $\mathrm{Flag}([p,n];\,V)$ is the full flag variety) we get that $\mathcal{L}^{(1^n)}$ is (nonequivariantly) isomorphic to the trivial line bundle, and the Picard group has rank $(n-1)$.

If we let $\mathbb{G}(p,\,V)$ denote the Grassmannian of p–dimensional quotients of V then we have a natural map

$$\psi_V^{(p)} : \mathrm{Flag}([p,n];\,V) \to \mathbb{G}(p,\,V), \text{ given by } \psi_V^{(p)}(V_\bullet) = V_p. \tag{2-20}$$

We abuse notation once more and write $\mathcal{Q}_p(V)$ for the tautological rank p quotient bundle on $\mathbb{G}(p,\,V)$, and let $\mathcal{R}_{n-p}(V)$ denote the tautological rank $(n-p)$ subbundle, whose fiber over the point corresponding to V_p is the kernel of the quotient map $V \twoheadrightarrow V_p$. The following formulation of Bott's theorem will be useful for us throughout Section 4 (see [Weyman 2003, Theorem 4.1.8]). For $m > 0$ and $\gamma \in \mathbb{Z}^m$ we let

$$\delta^{(m)} = (m-1, m-2, \dots, 0) \quad \text{and} \quad \tilde{\gamma} = \mathrm{sort}(\gamma + \delta^{(m)}) - \delta^{(m)}, \tag{2-21}$$

where $\mathrm{sort}(\gamma + \delta^{(m)}) \in \mathbb{Z}^m$ is obtained by arranging the entries of $\gamma + \delta^{(m)}$ in nonincreasing order.

Theorem 2.1. *Let $\lambda \in \mathbb{Z}^p_{\mathrm{dom}}$, $\mu \in \mathbb{Z}^{n-p}$, and let $\gamma = (\lambda|\mu) \in \mathbb{Z}^n$ be the concatenation of λ and μ. We write $\mathbb{F} = \mathrm{Flag}([p,n];\,V)$, $\psi = \psi_V^{(p)}$, $\pi = \pi_V^{(p)}$, and let $R^t\psi_*$ (resp. $R^t\pi_*$) denote the right derived functors of ψ_* (resp. π_*). Using* (2-21) *we have*:

(a) *If $\mu + \delta^{(n-p)}$ has repeated entries then $R^t\psi_*(\mathbb{S}_\lambda \mathcal{Q}_p(V) \otimes \mathcal{L}^{\mu}(V)) = 0$ for all t. Otherwise, there exists a unique $l \ge 0$ (equal to the number of inversions in $\mu + \delta^{(n-p)}$) so that*

$$R^t\psi_*(\mathbb{S}_\lambda \mathcal{Q}_p(V) \otimes \mathcal{L}^{\mu}(V)) = \begin{cases} \mathbb{S}_\lambda \mathcal{Q}_p(V) \otimes \mathbb{S}_{\tilde{\mu}} \mathcal{R}_{n-p}(V) & \text{if } t = l, \\ 0 & \text{otherwise.} \end{cases}$$

(b) *If* $\gamma+\delta^{(n)}$ *has repeated entries then* $H^t(\mathbb{F},\mathbb{S}_\lambda\mathcal{Q}_p(V)\otimes\mathcal{L}^\mu(V))=0$ *for all* t. *Otherwise, there exists a unique* $l\geq 0$ (*equal to the number of inversions in* $\gamma+\delta^{(n)}$) *so that*

$$H^t(\mathbb{F},\mathbb{S}_\lambda\mathcal{Q}_p(V)\otimes\mathcal{L}^\mu(V))=\begin{cases}\mathbb{S}_{\tilde{\gamma}}V & \textit{if } t=l,\\ 0 & \textit{otherwise}.\end{cases}$$

(c) *If* $\lambda_p\geq\mu_1$ *and if we let* $\lambda^+=(\lambda_1,\dots,\lambda_p,\mu_1)\in\mathbb{Z}^{p+1}_{\text{dom}}$ *and* $\mu^-=(\mu_2,\dots,\mu_{n-p})\in\mathbb{Z}^{n-p-1}$ *then*

$$R^t\pi_*(\mathbb{S}_\lambda\mathcal{Q}_p(V)\otimes\mathcal{L}^\mu(V))=\begin{cases}\mathbb{S}_{\lambda^+}\mathcal{Q}_{p+1}(V)\otimes\mathcal{L}^{\mu^-}(V) & \textit{if } t=0,\\ 0 & \textit{otherwise}.\end{cases}$$

2G. ***The relative setting.*** It will sometimes be convenient to work with spaces of matrices relative to some base as follows. We let B denote an algebraic variety over $\operatorname{Spec}(\mathbb{C})$ and let $\mathcal{F},\mathcal{G}$ be locally free sheaves on B of ranks m and n respectively. We can form

$$\mathcal{S}=\operatorname{Sym}_{\mathcal{O}_B}(\mathcal{F}\otimes_{\mathcal{O}_B}\mathcal{G})$$

and define $\mathfrak{X}=\underline{\operatorname{Spec}}_B(\mathcal{S})$. We identify freely quasicoherent $\mathcal{O}_{\mathfrak{X}}$-modules $\mathcal{M}$ with quasicoherent sheaves of $\mathcal{S}$-modules on B. We simply refer to such an $\mathcal{M}$ as an $\mathcal{S}$-module, and when $\mathcal{M}\subseteq\mathcal{O}_{\mathfrak{X}}$ is an ideal sheaf, we call $\mathcal{M}$ an ideal in $\mathcal{S}$. An example of such ideal is the one defining locally matrices of rank less than p: we denote by $\mathcal{I}_p\subset\mathcal{S}$ the ideal generated by the subsheaf $\bigwedge^p\mathcal{F}\otimes\bigwedge^p\mathcal{G}\subset\operatorname{Sym}^p(\mathcal{F}\otimes\mathcal{G})\subset\mathcal{S}$. If we let $Z_p\subset\mathfrak{X}$ denote the subvariety cut out by $\mathcal{I}_{p+1}$ then we obtain a decomposition of the local cohomology groups as $\mathcal{O}_B$-modules of the form

$$\mathcal{H}^j_{Z_p}(\mathfrak{X},\mathcal{O}_{\mathfrak{X}})=\bigoplus_{\lambda,\mu}(\mathbb{S}_\lambda\mathcal{F}\otimes\mathbb{S}_\mu\mathcal{G})^{\oplus a_{\lambda,\mu}},$$

where the multiplicities $a_{\lambda,\mu}$ are the same as in the case when $B=\operatorname{Spec}(\mathbb{C})$, $\mathfrak{X}=X$, and $Z_p=\overline{O}_p$.

2H. ***The inductive structure.*** This section builds on a standard localization trick that is often used to study determinantal varieties inductively (see [Bruns and Vetter 1988, Proposition 2.4] or [Lyubeznik et al. 2016]). We let $X=\mathbb{C}^{m\times n}$ and consider the basic open affine $X_1\subset X$ consisting of matrices with $x_{11}\neq 0$, whose coordinate ring is the localization $S_{x_{11}}$. We let $X'=\mathbb{C}^{(m-1)\times(n-1)}$, and identify its coordinate ring with $S'=\mathbb{C}[x'_{ij}]$, with $2\leq i,j\leq n$. We have an isomorphism (given by performing row and column operations in order to eliminate entries on the first row and first column of the generic matrix)

$$X_1\simeq X'\times\mathbb{C}^{m-1}\times\mathbb{C}^{n-1}\times\mathbb{C}^*,$$

where the coordinate functions on $\mathbb{C}^{m-1}$ are x_{i1}, $2\leq i\leq m$, those on $\mathbb{C}^{n-1}$ are x_{1j}, $2\leq j\leq n$, the coordinate function on $\mathbb{C}^*$ is x_{11}, and

$$x'_{ij}=x_{ij}-\frac{x_{i1}\cdot x_{1j}}{x_{11}}.$$

If we let $\pi:X_1\to X'$ denote the projection map, and let O'_p denote the orbit of rank p matrices in X' then

$$\pi^{-1}(O'_p)=O_{p+1}\cap X_1\quad\text{ for all } p=0,\dots,n-1.$$

It follows that if we let D'_p denote the intersection homology $\mathcal{D}_{X'}$-module associated with O'_p then

$$\pi^*(D'_p) = (D_{p+1})_{|X_1} = (D_{p+1})_{x_{11}} \quad \text{for all } p = 0, \dots, n-1.$$

If $m = n$ and if we let $\det' = \det(x'_{ij})$ then $\det = x_{11} \cdot \det'$, so $\pi^*(S'_{\det'}) = (S_{\det})_{|X_1} = S_{\det \cdot x_{11}}$. More generally, if we define the $\mathcal{D}_{X'}$-modules Q'_p in analogy with (1-12) then we obtain

$$\pi^*(Q'_p) = (Q_{p+1})_{|X_1} = (Q_{p+1})_{x_{11}} \quad \text{for all } p = 0, \dots, n-1. \tag{2-22}$$

For every S'-module (resp. $\mathcal{D}_{X'}$-module) M' and every closed subset $Z' \subset X'$, if we let $Z = \pi^{-1}(Z')$ and $M = \pi^*(M')$ then we have isomorphisms of $S_{x_{11}}$-modules (resp. of $\mathcal{D}_{X_1}$-modules)

$$\pi^*(H^j_{Z'}(M')) = H^j_Z(M) \quad \text{for all } j \geq 0.$$

In particular, we obtain

$$\pi^*(H^j_{\overline{O'_p}}(S')) = H^j_{\overline{O}_{p+1} \cap X_1}(S_{x_{11}}) = (H^j_{\overline{O}_{p+1}}(S))_{|X_1} \quad \text{for all } p = 0, \dots, n-1, \text{ and } j \geq 0. \tag{2-23}$$

3. Grothendieck group calculation of the local cohomology of simple $\mathcal{D}$-modules

Recall that $\Gamma_{\mathcal{D}}$ denotes the Grothendieck group of $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, and that if $M \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ then $[M]_{\mathcal{D}}$ denotes its class in $\Gamma_{\mathcal{D}}$. The main result of this section describes the class in $\Gamma_{\mathcal{D}}$ of the local cohomology groups with determinantal support for the modules D_p.

Theorem 3.1. *For every $0 \leq t < p \leq n \leq m$ we have the following equality in $\Gamma_{\mathcal{D}}[q]$:*

$$H^{\mathcal{D}}_t(D_p; q) = \sum_{s=0}^{t} [D_s]_{\mathcal{D}} \cdot q^{(p-t)^2 + (p-s)\cdot(m-n)} \cdot \binom{n-s}{p-s}_{q^2} \cdot \binom{p-1-s}{t-s}_{q^2}$$

We record here a special case of Theorem 3.1, which will be used in Section 6C. If $m = n = p$ and $c_t = (n-t)^2$ is the codimension of the orbit O_t inside $\mathbb{C}^{n \times n}$ then

$$[H^{c_t}_{\overline{O}_t}(S)]_{\mathcal{D}} = [D_0]_{\mathcal{D}} + [D_1]_{\mathcal{D}} + \cdots + [D_t]_{\mathcal{D}}. \tag{3-1}$$

3A. ***A relation between rectangular ideals and simple equivariant $\mathcal{D}$-modules.*** We use the notational conventions from Section 2E, and recall from Section 2A that α' denotes the conjugate of a partition α. For positive integers a, d and partitions $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_a)$ and $\beta = (\beta_1 \geq \beta_2 \geq \cdots \geq \beta_{m-a})$ we let

$$\lambda(a, d; \alpha, \beta) = (d + \alpha_1, d + \alpha_2, \dots, d + \alpha_a, \beta_1, \beta_2, \dots, \beta_{m-a})$$

and consider the polynomial $h_{a \times d}(q) \in \Gamma_{\mathrm{GL}}[q]$ given by

$$h_{a \times d}(q) = \sum_{\alpha, \beta} [\mathbb{S}_{\lambda(a,d;\alpha,\beta)}\mathbb{C}^m \otimes \mathbb{S}_{\lambda(a,d;\beta',\alpha')}\mathbb{C}^n]_{\mathrm{GL}} \cdot q^{|\alpha|+|\beta|},$$

where the sum is over partitions α, β satisfying

$$\alpha_1 \leq n - a, \quad \alpha'_1, \beta_1 \leq \min(a, d) \quad \text{and} \quad \beta'_1 \leq m - a. \tag{3-2}$$

The significance of the polynomials $h_{a\times d}(q)$ is that they describe the GL-equivariant Hilbert series of certain simple modules over the general linear Lie superalgebra $\mathfrak{gl}(m|n)$. As such, they provide the building blocks of the minimal free resolution over the polynomial ring S of the ideals $I_{a\times d}$ (see [Raicu and Weyman 2017, Theorem 3.1] or [Raicu 2017, Theorem 6.1]), namely we have

$$\sum_{j\geq 0}[\mathrm{Tor}_j^S(I_{a\times d},\mathbb{C})]_{\mathrm{GL}}\cdot q^j=\sum_{r=0}^{n-a}h_{(a+r)\times(d+r)}(q)\cdot q^{r^2+2r}\cdot\binom{r+\min(a,d)-1}{r}_{q^2} \tag{3-3}$$

which will be used in Section 3D. For now, we prove the following.

Lemma 3.2. *If we let $V=\mathbb{S}_{(n^m)}\mathbb{C}^m\otimes\mathbb{S}_{(m^n)}\mathbb{C}^n=\det(\mathbb{C}^m\otimes\mathbb{C}^n)$ and let $d\gg 0$ then*

$$\langle V\otimes D_p, h_{a\times d}(q)\rangle_{\mathrm{GL}}=0 \textit{ for } a\neq p \quad \textit{and} \quad \langle V\otimes D_p, h_{p\times d}(q)\rangle_{\mathrm{GL}}=q^{p\cdot(m-n)}\cdot\binom{n}{p}_{q^2}.$$

Proof. To compute $\langle V\otimes D_p, h_{a\times d}(q)\rangle_{\mathrm{GL}}$, we need to characterize the partitions α,β satisfying (3-2) and for which $\mathbb{S}_{\lambda(a,d;\alpha,\beta)}\mathbb{C}^m\otimes\mathbb{S}_{\lambda(a,d;\beta',\alpha')}\mathbb{C}^n$ appears as a subrepresentation of $V\otimes D_p$, i.e., those for which there exists a dominant weight $\mu\in\mathbb{Z}^n$ with $\mu_p\geq p-n$, $\mu_{p+1}\leq p-m$ (see (2-7)), and such that

$$\mu(p)+(n^m)=\lambda(a,d;\alpha,\beta) \quad\text{ and }\quad \mu+(m^n)=\lambda(a,d;\beta',\alpha'). \tag{3-4}$$

If $p<a$ then it follows from (2-8) that

$$p=(p-n)+n\geq\mu(p)_{p+1}+n=\lambda(a,d;\alpha,\beta)_{p+1}=d+\alpha_{p+1}$$

which is in contradiction with the fact that $d\gg 0$. If $p>a$ then

$$a\geq\beta_1=\lambda(a,d;\alpha,\beta)_{a+1}=\mu(p)_{a+1}+n=\mu_{a+1}+n\geq\mu_p+n\geq(p-n)+n=p$$

which is again a contradiction. It follows that $\langle V\otimes D_p, h_{a\times d}(q)\rangle=0$ for $a\neq p$, and it remains to analyze the case $p=a$. The conditions (3-4) imply that

$$\mu_i=d+\alpha_i-n \quad\text{ and }\quad \alpha_i+(m-n)=\beta_i' \qquad \text{for all } i=1,\dots,p.$$

Since $\beta_1\leq\min(p,d)=p$ it follows from the above that β is completely determined by α via the relation $\beta'=\alpha+((m-n)^p)$, which in turn implies that $\beta=(p^{m-n}|\alpha')$ and in particular

$$\beta_1=\cdots=\beta_{m-n}=p.$$

Suppose now that α is any partition with at most p parts (i.e., $\alpha_1'\leq p$) and that $\alpha_1\leq n-p$. If we define $\beta=(p^{m-n}|\alpha')$ then $\beta_1\leq p$ and $\beta_1'=\alpha_1+m-n\leq m-p$, so the conditions (3-2) hold for $a=p$, since $d\gg 0$. We next let

$$\mu_i=d+\alpha_i-n \quad \text{for } i=1,\dots,p, \quad\text{ and }\quad \mu_j=\alpha'_{j-p}-m \quad \text{for } j=p+1,\dots,n,$$

and observe that $\mu_p\geq p-n$ since $d\gg 0$, and that $\mu_{p+1}=\alpha_1'-m\leq p-m$, so $\mathbb{S}_{\mu(p)}\mathbb{C}^m\otimes\mathbb{S}_\mu\mathbb{C}^n$ appears as a subrepresentation of D_p. Once we verify (3-4) it follows that the pair of partitions (α,β) contributes the term $q^{|\alpha|+|\beta|}=q^{2\cdot|\alpha|+p\cdot(m-n)}$ to $\langle V\otimes D_p, h_{p\times d}(q)\rangle$; hence

$$\langle V\otimes D_p, h_{p\times d}(q)\rangle=\sum_{\alpha} q^{2\cdot|\alpha|+p\cdot(m-n)} \overset{(2\text{-}2)}{=} q^{p\cdot(m-n)}\cdot\binom{n}{p}_{q^2},$$

as desired. For $1\le i\le p$ we have that

$$\mu(p)_i+n=d+\alpha_i=\lambda(p,d;\alpha,\beta)_i,\quad\text{and}\quad \mu_i+m=d+\alpha_i+m-n=d+\beta'_i=\lambda(p,d;\beta',\alpha')_i.$$

We have moreover that for $1\le j\le m-n$,

$$\mu(p)_{p+j}+n=(p-n)+n=p=\beta_j=\lambda(p,d;\alpha,\beta)_{p+j},$$

and that for $p+1\le j\le n$,

$$\mu_j+m=\alpha'_{j-p}=\lambda(p,d;\beta',\alpha')_j,$$
$$\mu(p)_{m-n+j}+n=\mu_j+m=\alpha'_{j-p}=\beta_{m-n+j-p}=\lambda(p,d;\alpha,\beta)_{m-n+j},$$

showing that (3-4) holds for $a=p$ and concluding our proof. □

3B. ***A recursive formula for Euler characteristics.*** We use the notational conventions from Sections 2D and 2E, and define the *Euler characteristic maps*

$$\chi:\Gamma_{\mathcal{D}}[q]\to\Gamma_{\mathcal{D}}\quad\text{and}\quad\chi_s:\Gamma_{\mathcal{D}}[q]\to\mathbb{Z}\qquad\text{for } s=0,\dots,n,$$

as follows: if $\gamma(q)\in\Gamma_{\mathcal{D}}[q]$ is expressed as $\gamma(q)=\sum_{s=0}^{n}[D_s]_{\mathcal{D}}\cdot\gamma_s(q)$ with $\gamma_s(q)\in\mathbb{Z}[q]$ then we let

$$\chi(\gamma(q))=\gamma(-1)\quad\text{and}\quad\chi_s(\gamma(q))=\gamma_s(-1).\tag{3-5}$$

We recall the notation (2-9) where the subscript t indicates that we are considering local cohomology with support in the orbit closure $\overline{O}_t$. Using (2-10) and (2-1) we get that

$$\chi_s(H_t^{\mathcal{D}}(S;q))=\begin{cases}(-1)^{(n-t)+(n-s)\cdot(m-n)}\cdot\binom{n-1-s}{t-s} & \text{for } s=0,\dots,t,\\ 0 & \text{for } s>t.\end{cases}\tag{3-6}$$

Lemma 3.3. *For $t<p$ the Euler characteristics $\chi_0(H_t^{\mathcal{D}}(D_p;q))$ satisfy the recurrence relation*

$$\sum_{s=t+1}^{p}\chi_0(H_t^{\mathcal{D}}(D_s;q))\cdot(-1)^{s\cdot(m-n)}\cdot\binom{n-1-s}{p-s}=(-1)^{p-t}\cdot\binom{n-1}{t}-\binom{n-1}{p}.\tag{3-7}$$

Proof. The existence of a spectral sequence

$$E_2^{i,j}=H^i_{\overline{O}_t}(H^j_{\overline{O}_p}(S))\Rightarrow H^{i+j}_{\overline{O}_t}(S)$$

and the fact that Euler characteristic is invariant under taking homology, imply the equality

$$\sum_{s=0}^{p}\chi_0(H_t^{\mathcal{D}}(D_s;q))\cdot\chi_s(H_p^{\mathcal{D}}(S;q))=\chi_0(H_t^{\mathcal{D}}(S;q))$$

which in view of (3-6) can be reformulated as

$$\sum_{s=0}^{p}\chi_0(H_t^{\mathcal{D}}(D_s;q))\cdot(-1)^{(n-p)+(n-s)\cdot(m-n)}\cdot\binom{n-1-s}{p-s}=(-1)^{(n-t)+n\cdot(m-n)}\cdot\binom{n-1}{t}.$$

Dividing both sides by $(-1)^{(n-p)+n\cdot(m-n)}$ and moving the term $s=0$ to the right-hand side yields

$$\sum_{s=1}^{p}\chi_0(H_t^{\mathcal{D}}(D_s;q))\cdot(-1)^{s\cdot(m-n)}\cdot\binom{n-1-s}{p-s}=(-1)^{p-t}\cdot\binom{n-1}{t}-\chi_0(H_t^{\mathcal{D}}(D_0;q))\cdot\binom{n-1}{p}. \quad (3\text{-}8)$$

Note that for $s\le t$ we have that the support of D_s is contained in $\overline{O}_t$ and in particular $H^0_{\overline{O}_t}(D_s)=D_s$ and $H^j_{\overline{O}_t}(D_s)=0$ for $j>0$. It follows that $\chi_0(H_t^{\mathcal{D}}(D_0;q))=1$ and $\chi_0(H_t^{\mathcal{D}}(D_s;q))=0$ for $0<s\le t$, so (3-8) is equivalent to the desired relation (3-7). □

3C. ***A binomial identity.*** The goal of this section is to use the recurrence relation from Lemma 3.3 in order to deduce a closed formula for the Euler characteristic $\chi_0(H_t^{\mathcal{D}}(D_p;q))$. We prove the following.

Proposition 3.4. *For* $0\le t<p\le n$ *we have that*

$$\chi_0(H_t^{\mathcal{D}}(D_p;q))=(-1)^{(p-t)+p\cdot(m-n)}\cdot\binom{n}{p}\cdot\binom{p-1}{t}.$$

Proof. It suffices to check that the right-hand side of the above equality satisfies the recursion in Lemma 3.3, that is (after canceling some signs)

$$\sum_{s=t+1}^{p}(-1)^{(s-t)}\cdot\binom{n}{s}\cdot\binom{s-1}{t}\cdot\binom{n-1-s}{p-s}=(-1)^{p-t}\cdot\binom{n-1}{t}-\binom{n-1}{p}. \quad (3\text{-}9)$$

It suffices to prove that the (bivariate) generating functions of the two sides coincide, so we multiply each side by $x^t\cdot y^p$ and sum over all pairs $0\le t<p$ of nonnegative integers. We have

$$\begin{aligned}
&\sum_{0\le t<p}\left(\sum_{s=t+1}^{p}(-1)^{(s-t)}\cdot\binom{n}{s}\cdot\binom{s-1}{t}\cdot\binom{n-1-s}{p-s}\right)\cdot x^t\cdot y^p\\
&=\sum_{s\ge1}\binom{n}{s}\cdot(-y)^s\cdot\left(\sum_{t=0}^{s-1}\binom{s-1}{t}\cdot(-x)^t\right)\cdot\left(\sum_{p\ge s}\binom{n-1-s}{p-s}\cdot y^{p-s}\right)\\
&=\sum_{s\ge1}\binom{n}{s}\cdot(-y)^s\cdot(1-x)^{s-1}\cdot(1+y)^{n-1-s}=\frac{(1+y)^{n-1}}{1-x}\cdot\left[\sum_{s\ge1}\binom{n}{s}\cdot\left(\frac{-y\cdot(1-x)}{1+y}\right)^s\right]\\
&=\frac{(1+y)^{n-1}}{1-x}\cdot\left[\left(1-\frac{y\cdot(1-x)}{1+y}\right)^n-1\right]=\frac{(1+xy)^n}{(1-x)\cdot(1+y)}-\frac{(1+y)^{n-1}}{1-x}.
\end{aligned} \quad (3\text{-}10)$$

We split the generating function of the right-hand side of (3-9) into two parts, as follows.

$$\sum_{0\le t<p}(-1)^{p-t}\cdot\binom{n-1}{t}\cdot x^t\cdot y^p=\sum_{t\ge 0}\binom{n-1}{t}\cdot(xy)^t\cdot\left(\sum_{p>t}(-y)^{p-t}\right)=(1+xy)^{n-1}\cdot\left(\frac{-y}{1+y}\right),\tag{3-11}$$

$$\sum_{0\le t<p}\binom{n-1}{p}\cdot x^t\cdot y^p=\sum_{p\ge 0}\binom{n-1}{p}\cdot\frac{1-x^p}{1-x}\cdot y^p=\frac{1}{1-x}\cdot\big((1+y)^{n-1}-(1+xy)^{n-1}\big).\tag{3-12}$$

Taking the difference between (3-11) and (3-12) we obtain

$$(1+xy)^{n-1}\cdot\left(\frac{1}{1-x}-\frac{y}{1+y}\right)-\frac{(1+y)^{n-1}}{1-x}=\frac{(1+xy)^n}{(1-x)\cdot(1+y)}-\frac{(1+y)^{n-1}}{1-x}$$

which is the same as (3-10), proving the identity (3-9). ☐

3D. ***The proof of Theorem 3.1.*** The conclusion of Theorem 3.1 can be rephrased using (2-9) as

$$\langle H_t^{\mathcal{D}}(D_p;q),D_s\rangle_{\mathcal{D}}=q^{(p-t)^2+(p-s)\cdot(m-n)}\cdot\binom{n-s}{p-s}_{q^2}\cdot\binom{p-1-s}{t-s}_{q^2}\quad\text{for } s=0,\dots,t.\tag{3-13}$$

The fact that $\langle H_t^{\mathcal{D}}(D_p;q),D_s\rangle_{\mathcal{D}}=0$ for $s>t$ follows since we are considering local cohomology groups with support in $\overline{O}_t$, and the modules D_s with $s>t$ have strictly larger support.

We note that the polynomial on the right-hand side of the above formula is invariant under subtracting one from each of m,n,p,t and s. If we restrict the local cohomology groups to the basic open affine $X_1=(x_{11}\neq 0)$ and use the inductive structure as explained in Section 2H then it follows that for $s>0$

$$\langle H_t^{\mathcal{D}}(D_p;q),[D_s]\rangle_{\mathcal{D}}=\left\langle\sum_{j\ge 0}[H^j_{\overline{O}'_{t-1}}(D'_{p-1})]\cdot q^j,[D'_{s-1}]\right\rangle_{\mathcal{D}}$$

so the desired conclusion follows by induction. We are left with considering the case $s=0$, where we need to verify that

$$\langle H_t^{\mathcal{D}}(D_p;q),[D_0]\rangle_{\mathcal{D}}=q^{(p-t)^2+p\cdot(m-n)}\cdot\binom{n}{p}_{q^2}\cdot\binom{p-1}{t}_{q^2}.$$

We consider a *witness representation* for the module D_0 (as in (2-14)) defined by

$$W=\mathbb{S}_{(-n^m)}\mathbb{C}^m\otimes\mathbb{S}_{(-m^n)}\mathbb{C}^n=\det(\mathbb{C}^m\otimes\mathbb{C}^n)^\vee.$$

As seen in (2-14), the multiplicity of D_0 as a composition factor in some GL-equivariant $\mathcal{D}$-module M is the same as the multiplicity of W as a subrepresentation of M, so W *witnesses* the occurrences of D_0 as a composition factor of M. It therefore suffices to verify that

$$\langle H_t^{\mathrm{GL}}(D_p;q),W\rangle_{\mathrm{GL}}=q^{(p-t)^2+p\cdot(m-n)}\cdot\binom{n}{p}_{q^2}\cdot\binom{p-1}{t}_{q^2}.$$

We prove this equality in two steps:

(1) We show the inequality $\le$, where $\sum a_i\cdot q^i\le\sum b_i\cdot q^i$ if and only if $a_i\le b_i$ for all i.

(2) We show that after plugging in $q=-1$ we obtain an equality.

For the inequality in (1) we begin by recalling that $\overline{O}_t$ is defined by the ideal I_{t+1} of $(t+1)\times(t+1)$ minors of the generic matrix, and that the sequence of ideals $I_{(t+1)\times d}$ is cofinal with the sequence of powers of I_{t+1}. It follows from [Eisenbud 2005, Exercise A1D.1] that

$$H^j_{\overline{O}_t}(D_p) = \varinjlim_d \operatorname{Ext}^j_S(S/I_{(t+1)\times d}, D_p). \tag{3-14}$$

We compute the Ext modules in the above limit from the minimal resolution of $S/I_{(t+1)\times d}$ described in [Raicu and Weyman 2017]. We have that $\operatorname{Ext}^j_S(S/I_{(t+1)\times d}, D_p)$ is the j-th cohomology group of a complex $F^\bullet$, where

$$F^j = \operatorname{Tor}^S_j(S/I_{(t+1)\times d}, \mathbb{C})^\vee \otimes_{\mathbb{C}} D_p.$$

Notice that $\operatorname{Tor}^S_0(S/I_{(t+1)\times d}, \mathbb{C}) = \mathbb{C}$ so that $F^0 = D_p$ and $\langle F^0, W\rangle_{\mathrm{GL}} = 0$ since $p > 0$. Notice also that

$$\operatorname{Tor}^S_j(S/I_{(t+1)\times d}, \mathbb{C}) = \operatorname{Tor}^S_{j-1}(I_{(t+1)\times d}, \mathbb{C}) \quad \text{for } j \geq 1,$$

so taking $d \gg 0$ (in particular $d \geq t+1$) we have that

$$\begin{aligned} \Big\langle \sum_{j\geq 0} [F^j]_{\mathrm{GL}} \cdot q^j, W \Big\rangle_{\mathrm{GL}} &= \Big\langle W^\vee \otimes D_p, \sum_{j\geq 0} [\operatorname{Tor}^S_j(S/I_{(t+1)\times d}, \mathbb{C})]_{\mathrm{GL}} \cdot q^j \Big\rangle_{\mathrm{GL}} \\ &= q \cdot \Big\langle W^\vee \otimes D_p, \sum_{j\geq 0} [\operatorname{Tor}^S_j(I_{(t+1)\times d}, \mathbb{C})]_{\mathrm{GL}} \cdot q^j \Big\rangle_{\mathrm{GL}} \\ &= \sum_{r=0}^{n-1-t} \big\langle W^\vee \otimes D_p, h_{(t+1+r)\times(d+r)}(q) \big\rangle_{\mathrm{GL}} \cdot q^{r^2+2r+1} \cdot \binom{r+t}{t}_{q^2}, \end{aligned} \tag{3-15}$$

where the last equality follows from (3-3) by taking $a = t+1$, using the fact that $\min(t+1, d) = t+1$, and noting that $\binom{r+t}{r}_{q^2} = \binom{r+t}{t}_{q^2}$. Letting $a = t+1+r$ and $V = W^\vee$ in Lemma 3.2 it follows that the only term that survives in (3-15) is the one corresponding to $r = p-t-1$, which yields

$$\Big\langle \sum_{j\geq 0} [F^j]_{\mathrm{GL}} \cdot q^j, W \Big\rangle_{\mathrm{GL}} = q^{p\cdot(m-n)} \cdot \binom{n}{p}_{q^2} \cdot q^{(p-t)^2} \cdot \binom{p-1}{t}_{q^2}.$$

This shows that W can only occur as a subrepresentation in F^j only if $j \equiv p\cdot(m-n) + (p-t)^2 \pmod 2$, and in particular W does not occur in any two consecutive terms of $F^\bullet$. Since $\operatorname{Ext}^j_S(S/I_{(t+1)\times d}, D_p)$ is obtained as the j-th cohomology group of $F^\bullet$, it follows that $\big\langle \operatorname{Ext}^j_S(S/I_{(t+1)\times d}, D_p), W \big\rangle_{\mathrm{GL}} = \langle F^j, W\rangle_{\mathrm{GL}}$ for all j, and using (3-14) we conclude that

$$\langle H^{\mathcal{D}}_t(D_p; q), D_0\rangle_{\mathcal{D}} = \langle H^{\mathrm{GL}}_t(D_p; q), W\rangle_{\mathrm{GL}} \leq q^{(p-t)^2 + p\cdot(m-n)} \cdot \binom{n}{p}_{q^2} \cdot \binom{p-1}{t}_{q^2}. \tag{3-16}$$

Since the exponents of q appearing in (3-16) with nonzero coefficient have the same parity, it follows that in order to prove the equality and conclude Step (2) of our argument, it suffices to check that equality holds in (3-16) after plugging in $q = -1$. In this case the left-hand side becomes $\chi_0(H^{\mathcal{D}}_t(D_p; q))$, while the right-hand side becomes $(-1)^{(p-t)+p\cdot(m-n)} \cdot \binom{n}{p} \cdot \binom{p-1}{t}$, so the conclusion follows from Proposition 3.4. □

One consequence of (3-16) is a vanishing result for the local cohomology groups $H^j_{\overline{O}_t}(D_p)$, based solely on the parity of j. Similar vanishing results, proved using more refined techniques in Sections 4 and 5, will play an important role in analyzing square matrices.

Corollary 3.5. *If $j \not\equiv (p-t) + p\cdot(m-n) \pmod 2$ then $H^j_{\overline{O}_t}(D_p) = 0$. In particular, when $m = n$ we may have $H^j_{\overline{O}_t}(D_p) \neq 0$ only when $j \equiv (p-t) \pmod 2$.*

Proof. By (3-16), q^j may appear with nonzero coefficient only when $j \equiv p\cdot(m-n) + (p-t)^2 \pmod 2$. Since $(p-t)^2$ and $(p-t)$ have the same parity, the conclusion follows. □

3E. ***The proof of Theorem 1.3.*** We have

$$\begin{aligned} L_p(q,w) &= \sum_{i,j\geq 0} \langle H^i_{O_0}(H^{mn-j}_{\overline{O}_p}(S)), D_0\rangle_{\mathcal{D}} \cdot q^i \cdot w^j \\ &= \sum_{i\geq 0}\Bigg(\sum_{s=0}^{p} \langle H^i_{O_0}(D_s), D_0\rangle_{\mathcal{D}} \cdot q^i \cdot \bigg(\sum_{j\geq 0} \langle H^{mn-j}_{\overline{O}_p}(S), D_s\rangle_{\mathcal{D}} \cdot w^j \bigg)\Bigg), \end{aligned}$$

where the first equality follows from (1-5) and the second from the fact that $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is semisimple, and the fact that local cohomology commutes with direct sums. We obtain by reversing the summation order that

$$\begin{aligned} L_p(q,w) &= \sum_{s=0}^{p} \langle H_0^{\mathcal{D}}(D_s;q), D_0\rangle_{\mathcal{D}} \cdot \langle H_p^{\mathcal{D}}(S;w^{-1}), D_s\rangle_{\mathcal{D}} \cdot w^{mn} \\ &\overset{(3\text{-}13),(2\text{-}10)}{=} \sum_{s=0}^{p} q^{s^2+s\cdot(m-n)} \cdot \binom{n}{s}_{q^2} \cdot w^{-(n-p)^2-(n-s)\cdot(m-n)} \cdot \binom{n-1-s}{p-s}_{w^{-2}} \cdot w^{mn} \end{aligned}$$

Using (2-3), it follows that in order to prove (1-7) it suffices to verify the identity

$$p^2 + 2p + s\cdot(m+n-2p-2) = -(n-p)^2 - (n-s)\cdot(m-n) - 2\cdot(p-s)\cdot(n-1-p) + mn$$

which follows by inspection after expanding the products.

4. Vanishing of local cohomology for the subquotients $J_{\underline{x},p}$

Throughout this section we let $m = n$, and in order to keep track of the two distinct copies of $\mathbb{C}^n$ we will denote them by F and G respectively. We will then let $X = (F\otimes G)^\vee$ and $S = \mathrm{Sym}_{\mathbb{C}}(F\otimes G)$ be the coordinate ring of X. Finally, we write $\mathrm{GL} = \mathrm{GL}(F)\times\mathrm{GL}(G)$. The goal of this section is to revisit the construction of a class of GL-equivariant S-modules which have played a prominent role in describing the graded components of Ext and local cohomology modules for determinantal ideals and their thickenings [Raicu and Weyman 2014; Raicu 2018], and to prove vanishing results for some of their local cohomology groups. These modules are indexed by pairs $(\underline{x}, p)$ with $\underline{x}$ a partition and p a nonnegative integer, and are denoted $J_{\underline{x},p}$ (see Section 4A for their construction). We write $\mathfrak{m}$ for the maximal homogeneous ideal of the polynomial ring S, so that $H^j_{\mathfrak{m}}(-) = H^j_{O_0}(-)$. Our key vanishing result below will be proved in Section 4B.

Theorem 4.1. *Suppose that $0 \le p \le n$ and that $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$ with $x_1 = \cdots = x_p$.*

(a) $H^1_{\mathfrak{m}}(\mathrm{Ext}^j_S(J_{\underline{x},p}, S)) = 0$ *for all* $j \ge 0$.

(b) *If* $0 \le t \le p$ *then* $H^k_{\overline{O}_t}(J_{\underline{x},p}) = 0$ *for* $k \not\equiv p - t \pmod 2$.

4A. ***The $J_{\underline{x},p}$-modules and their relative versions.*** Recall the notation from Section 2F. For $0 \le p \le n$ we define

$$X^{(p)} = \mathrm{Flag}([p,n]; F) \times \mathrm{Flag}([p,n]; G),$$

noting that $X^{(n)} = \mathrm{Spec}(\mathbb{C})$. On $X^{(p)}$ we have a natural sheaf of algebras given by

$$\mathcal{S}^{(p)} = \mathrm{Sym}_{\mathcal{O}_{X^{(p)}}}(\mathcal{Q}_p(F) \otimes \mathcal{Q}_p(G)) = \bigoplus_{\underline{x} \in \mathbb{N}^p_{\mathrm{dom}}} \mathbb{S}_{\underline{x}}\mathcal{Q}_p(F) \otimes \mathbb{S}_{\underline{x}}\mathcal{Q}_p(G),$$

where the last equality comes from Cauchy's formula just like (2-5). Note that when $p = n$ we get $\mathcal{S}^{(n)} = S$. We define $Y^{(p)} = \underline{\mathrm{Spec}}_{X^{(p)}}\mathcal{S}^{(p)}$, which is a vector bundle over $X^{(p)}$ whose fiber can be identified locally with the space of $p \times p$ matrices (see Section 2G). For $\underline{x} \in \mathbb{N}^p_{\mathrm{dom}}$ we let $\mathcal{I}^{(p)}_{\underline{x}}$ denote the ideal in $\mathcal{S}^{(p)}$ (see also (2-6)) generated by $\mathbb{S}_{\underline{x}}\mathcal{Q}_p(F) \otimes \mathbb{S}_{\underline{x}}\mathcal{Q}_p(G)$, and define

$$\mathcal{I}^{(p)}_{\mathcal{X}} = \sum_{\underline{x} \in \mathcal{X}} \mathcal{I}^{(p)}_{\underline{x}} \qquad \text{for any subset } \mathcal{X} \subset \mathbb{N}^p_{\mathrm{dom}}.$$

We define for $l < p$ and $\underline{z} \in \mathbb{N}^p_{\mathrm{dom}}$ the subset of $\mathbb{N}^p_{\mathrm{dom}}$

$$\mathfrak{succ}(\underline{z}, l; p) = \{\underline{y} \in \mathbb{N}^p_{\mathrm{dom}} : \underline{y} \ge \underline{z} \text{ and } y_i > z_i \text{ for some } i > l\},$$

and consider the $\mathcal{S}^{(p)}$-modules defined by

$$\mathcal{J}^{(p)}_{\underline{z},l} = \mathcal{I}^{(p)}_{\underline{z}} / \mathcal{I}^{(p)}_{\mathfrak{succ}(\underline{z},l;p)},$$

with the convention that $\mathfrak{succ}(\underline{z}, p; p) = \varnothing$ and $\mathcal{J}^{(p)}_{\underline{z},p} = \mathcal{I}^{(p)}_{\underline{z}}$. When $p = n$ and $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$ we have $I_{\underline{x}} = \mathcal{I}^{(n)}_{\underline{x}}$ as in (2-6), and we write $J_{\underline{x},l} = \mathcal{J}^{(n)}_{\underline{x},l}$. The ideals $I_{\underline{x}}$ and the S-modules $J_{\underline{x},l}$ have been studied in [Raicu and Weyman 2014, Section 2B; Raicu 2018, Section 2.1]. As noted in [Raicu and Weyman 2014, Lemma 3.1(a)], if we consider the line bundle

$$\det{}^{(p)} = \det \mathcal{Q}_p(F) \otimes \det \mathcal{Q}_p(G) \tag{4-1}$$

then we have $\mathcal{J}^{(p)}_{\underline{x},l} \otimes \det^{(p)} = \mathcal{J}^{(p)}_{\underline{x}+(1^p),l}$. This allows us to define $\mathcal{J}^{(p)}_{\lambda,l}$ for any $\lambda \in \mathbb{Z}^p_{\mathrm{dom}}$: if $\lambda = \underline{x} - (d^p)$ for some $d \in \mathbb{Z}_{\ge 0}$ and $\underline{x} \in \mathbb{N}^p_{\mathrm{dom}}$, we let

$$\mathcal{J}^{(p)}_{\lambda,l} = \mathcal{J}^{(p)}_{\underline{x},l} \otimes (\det{}^{(p)})^{\otimes(-d)}. \tag{4-2}$$

For $p + 1 \le q \le n$, we consider the line bundle on $X^{(p)}$ given by (see the notation in (2-17))

$$\mathcal{L}_q = \mathcal{L}_q(F) \otimes \mathcal{L}_q(G)$$

and for $\mu \in \mathbb{Z}^{n-p}$ we define in analogy with (2-18)

$$\mathcal{L}^{\mu} = \bigotimes_{i=1}^{n-p} \mathcal{L}^{\otimes \mu_i}_{p+i}.$$

For $\lambda \in \mathbb{Z}^p_{\mathrm{dom}}$, $l \le p$ and $\mu \in \mathbb{Z}^{n-p}$ we define the $\mathcal{S}^{(p)}$-module (with $\mathcal{S}^{(p)}$-action inherited from $\mathcal{J}^{(p)}_{\lambda,l}$)

$$\mathcal{M}^{(p)}_{\lambda,l;\mu} = \mathcal{J}^{(p)}_{\lambda,l} \otimes \mathcal{L}^{\mu}.$$

As an $\mathcal{O}_{X^{(p)}}$-module, we have a direct sum decomposition

$$\mathcal{M}^{(p)}_{\lambda,l;\mu} = \bigoplus_{\substack{\delta \ge \lambda \\ \delta_i = \lambda_i \text{ for } i > l}} \mathbb{S}_\delta \mathcal{Q}_p(F) \otimes \mathbb{S}_\delta \mathcal{Q}_p(G) \otimes \mathcal{L}^{\mu}. \tag{4-3}$$

We note that if $\underline{y} \in \mathbb{N}^{n-p}_{\mathrm{dom}}$ and $d \ge y_1$, and if we define $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$ by letting

$$x_1 = \cdots = x_p = d \quad \text{and} \quad x_{p+i} = y_i \quad \text{for } i = 1, \dots, n-p, \tag{4-4}$$

then the module $\mathcal{M}^{(p)}_{(d^p),p;\underline{y}}$ coincides with the one denoted by $\mathcal{M}_{\underline{x},p}$ in [Raicu and Weyman 2014, (3-8)]. It follows from [Raicu and Weyman 2014, Lemma 3.2] that if we define $\underline{x}$ as in (4-4) then

$$H^k\big(X^{(p)}, \mathcal{M}^{(p)}_{(d^p),p;\underline{y}}\big) = \begin{cases} J_{\underline{x},p} & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4-5}$$

We will be interested more generally in the cohomology groups of $\mathcal{M}^{(p)}_{\lambda,l;\mu}$ for $l \le p$, which are naturally S-modules. It will be useful to note that (2-19) yields $\det^{(n)} = \det^{(p)} \otimes \mathcal{L}^{(1^{n-p})}$ and therefore

$$\mathcal{M}^{(p)}_{\lambda+(1^p),l;\mu+(1^{n-p})} = \mathcal{M}^{(p)}_{\lambda,l;\mu} \otimes \det{}^{(n)}. \tag{4-6}$$

Theorem 4.2. *Let $0 \le q \le p$ and $k \ge 0$, suppose that $\lambda \in \mathbb{Z}^p_{\mathrm{dom}}$ with $\lambda_1 = \cdots = \lambda_q$, and that $\mu \in \mathbb{Z}^{n-p}$. The cohomology group $H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu})$ admits an S-module composition series with composition factors isomorphic to $J_{\nu,l}$ for $l \le q$ and $\nu \in \mathbb{Z}^n_{\mathrm{dom}}$. Moreover, if $\lambda_p \le \mu_j$ for some j then the composition series can be chosen in such a way that each $J_{\nu,l}$ appearing as a composition factor satisfies $\nu_1 = \cdots = \nu_{l+1}$.*

Proof. Using (4-6) and the fact that $\det^{(n)}$ is a trivial bundle with fiber $\det(F) \otimes \det(G)$ we obtain

$$H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda+(1^p),q;\mu+(1^{n-p})}) = H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu}) \otimes (\det(F) \otimes \det(G)).$$

Since we also have that $J_{\nu+(1^n),l} = J_{\nu,l} \otimes (\det(F) \otimes \det(G))$, it follows that we may assume without loss of generality that $\lambda \in \mathbb{N}^p_{\mathrm{dom}}$ and $\mu \in \mathbb{N}^{n-p}$. We next reduce ourselves to the case when μ is dominant. Consider

$$\mathbb{G}^{(p)} = \mathbb{G}(p, F) \times \mathbb{G}(p, G)$$

and the natural map $\psi^{(p)} = \psi^{(p)}_F \times \psi^{(p)}_G : X^{(p)} \to \mathbb{G}^{(p)}$ (see (2-20)). Using Theorem 2.1(a) we get that

$$R^i \psi^{(p)}_*(\mathcal{M}^{(p)}_{\lambda,q;\mu}) = R^{i-2l} \psi^{(p)}_*(\mathcal{M}^{(p)}_{\lambda,q;\tilde{\mu}}) \quad \text{for all } i \in \mathbb{Z},$$

where l is the number of inversions in $\mu + \delta^{(n-p)}$. We know moreover that $R^i \psi^{(p)}_*(\mathcal{M}^{(p)}_{\lambda,q;\mu})$ is nonzero for at most one value of i, so the Leray spectral sequence degenerates and yields

$$\begin{aligned} H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu}) &= H^{k-i}\big(\mathbb{G}^{(p)}, R^i \psi^{(p)}_*(\mathcal{M}^{(p)}_{\lambda,q;\mu})\big) \\ &= H^{k-i}\big(\mathbb{G}^{(p)}, R^{i-2l} \psi^{(p)}_*(\mathcal{M}^{(p)}_{\lambda,q;\tilde{\mu}})\big) = H^{k-2l}(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\tilde{\mu}}). \end{aligned}$$

Notice that if $\lambda_p \le \mu_j$ for some j, then (2-21) forces $\lambda_p \le \tilde{\mu}_1$. With these reductions, we prove our theorem by induction on p and q. When $p = q = 0$ we have $\mathcal{M}^{(p)}_{\lambda,q;\mu} = \mathcal{L}^{\mu}$. Since μ is dominant, it follows that its higher cohomology groups vanish and

$$H^0(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu}) = \mathbb{S}_\mu F \otimes \mathbb{S}_\mu G = J_{\mu,0},$$

proving the base case. Suppose next that $0 \le q < p$, consider the natural map (see (2-15))

$$\pi^{(p-1)} = \pi_F^{(p-1)} \times \pi_G^{(p-1)} : X^{(p-1)} \to X^{(p)}$$

and define

$$\lambda^- = (\lambda_1, \dots, \lambda_{p-1}) \quad \text{and} \quad \mu^+ = (\lambda_p, \mu_1, \dots, \mu_{n-p}).$$

We have using Theorem 2.1(c) that

$$R^i \pi_*^{(p-1)}(\mathcal{M}^{(p-1)}_{\lambda^-,q;\mu^+}) = \begin{cases} \mathcal{M}^{(p)}_{\lambda,q;\mu} & \text{if } i = 0, \\ 0 & \text{otherwise.} \end{cases}$$

The Leray spectral sequence degenerates again, showing that

$$H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu}) = H^k(X^{(p-1)}, \mathcal{M}^{(p-1)}_{\lambda^-,q;\mu^+})$$

and allowing us to obtain the desired conclusion by induction on p.

Finally, the most interesting situation is when $p = q > 0$, in which case $\lambda = (d^p)$ for some $d \ge 0$. If $d > \mu_1$ then it follows from (4-5) that $\mathcal{M}^{(p)}_{\lambda,q;\mu}$ has no higher cohomology and

$$H^0(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu}) = J_{\nu,p}, \qquad \text{where } \nu = (d^p, \mu_1, \dots, \mu_{n-p}).$$

Note that in this case $\nu_p \ne \nu_{p+1}$! If $d \le \mu_1$ then we obtain a filtration of $\mathcal{M} = \mathcal{M}^{(p)}_{\lambda,p;\mu}$ given by

$$\mathcal{M} = \mathcal{M}_0 \supset \mathcal{M}_1 \supset \cdots \supset \mathcal{M}_{\mu_1 - d}, \qquad \text{where } \mathcal{M}_i = \mathcal{M}^{(p)}_{\lambda+(i^p),p;\mu} \text{ for } i = 0, \dots, \mu_1 - d,$$

where the inclusions are the natural ones, compatible with the decomposition (4-3). In particular, each $\mathcal{M}_{i+1}$ is a direct summand in $\mathcal{M}_i$ (as an $\mathcal{O}_{X^{(p)}}$-module, but not as an $\mathcal{S}^{(p)}$-module!), and we obtain a filtration

$$H^k(X^{(p)}, \mathcal{M}) \supseteq H^k(X^{(p)}, \mathcal{M}_1) \supseteq \cdots \supseteq H^k(X^{(p)}, \mathcal{M}_{\mu_1 - d}). \tag{4-7}$$

It follows from (4-5) that $H^k(X^{(p)}, \mathcal{M}_{\mu_1-d}) = 0$ for $k > 0$ and

$$H^0(X^{(p)}, \mathcal{M}_{\mu_1-d}) = J_{\nu,p}, \qquad \text{where } \nu_1 = \cdots = \nu_{p+1} = \mu_1 \text{ and } \nu_{p+i} = \mu_i \text{ for } i = 2, \dots, n-p.$$

Moreover, since

$$\mathcal{M}_i / \mathcal{M}_{i+1} = \mathcal{M}^{(p)}_{\lambda^i,p-1;\mu^i}, \qquad \text{where } \lambda^i = (d+i)^{p-1} \text{ and } \mu^i = (d+i, \mu_1, \dots, \mu_{n-p}),$$

it follows that the intermediate quotients in the filtration (4-7) have the form

$$H^k(X^{(p)}, \mathcal{M}_i/\mathcal{M}_{i+1}) = H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda^i,p-1;\mu^i})$$

which by induction (on q) have an S-module filtration with composition factors as in the statement of the Theorem. Therefore (4-7) can be further refined to obtain the desired filtration for $H^k(X^{(p)}, \mathcal{M})$. $\square$

We will use Theorem 4.2 in conjunction with the following vanishing result. Recall that $\mathfrak{m}$ is the maximal homogeneous ideal of the polynomial ring S.

Lemma 4.3. *Suppose that* $0 \le l \le n$ *and that* $\nu \in \mathbb{Z}^n_{\mathrm{dom}}$ *is such that* $\nu_1 = \cdots = \nu_l$. *If* $l \neq 1$ *or if* $l = 1$ *and* $\nu_1 = \nu_2$ *then*

$$H^1_{\mathfrak{m}}(J_{\nu,l}) = 0.$$

Proof. Using graded local duality, the desired vanishing is equivalent to

$$\mathrm{Ext}^{n^2-1}_S(J_{\nu,l}, S) = 0.$$

Based on (4-2), we may assume without loss of generality that $\nu \in \mathbb{N}^n_{\mathrm{dom}}$ so we can apply [Raicu and Weyman 2014, Theorem 3.3] which completely describes the graded components of all the modules $\mathrm{Ext}^j_S(J_{\nu,l}, S)$. Based on the said theorem, the vanishing of $\mathrm{Ext}^{n^2-1}_S(J_{\nu,l}, S)$ amounts to proving that it is impossible to find integers $0 \le s \le t_1 \le \cdots \le t_{n-l} \le l$ and dominant weights $\alpha \in \mathbb{Z}^n_{\mathrm{dom}}$ simultaneously satisfying the following conditions:

$$\begin{cases} l^2 + 2\sum_{j=1}^{n-l} t_j = 1, & \\ \alpha_n \ge l - \nu_l - n, & \\ \alpha_{t_j+j} = t_j - \nu_{n+1-j} - n & \text{for } j = 1, \dots, n-l, \\ \alpha_s \ge s - n \text{ and } \alpha_{s+1} \le s - n, & \end{cases}$$

where by convention $\alpha_0 = \infty$. The first condition already forces $l = 1$ and $t_1 = \cdots = t_{n-1} = 0$. Applying the third condition for $j = n-1$ we obtain $\alpha_{n-1} = -\nu_2 - n$. Since α is dominant we must then have

$$-\nu_2 - n = \alpha_{n-1} \ge \alpha_n \ge 1 - \nu_1 - n,$$

which in turn implies $\nu_1 - 1 \ge \nu_2$ and in particular $\nu_1 \neq \nu_2$. It follows that if $l \neq 1$ or if $l = 1$ and $\nu_1 = \nu_2$ the above conditions cannot be satisfied and $\mathrm{Ext}^{n^2-1}_S(J_{\nu,l}, S) = 0$, concluding our proof. $\square$

Remark 4.4. If $l = 1$ and $\nu_1 > \nu_2$ then $H^1_{\mathfrak{m}}(J_{\nu,l}) \neq 0$. As explained in the proof above we may assume that ν is a partition. We can then take $s = t_1 = \cdots = t_{n-1} = 0$ and define $\alpha \in \mathbb{Z}^n_{\mathrm{dom}}$ by letting

$$\alpha_j = -\nu_{n+1-j} - n \text{ for } j = 1, \dots, n-1, \quad \text{and} \quad \alpha_n = 1 - \nu_1 - n.$$

It follows that $\mathbb{S}_\alpha F \otimes \mathbb{S}_\alpha G$ appears as a subrepresentation of $\mathrm{Ext}^{n^2-1}_S(J_{\nu,l}, S)$, proving that $H^1_{\mathfrak{m}}(J_{\nu,l}) \neq 0$.

Corollary 4.5. *Suppose that* p, q, λ, μ *are as in the statement of Theorem 4.2 . If* $\lambda_p \le \mu_j$ *for some* j *then*

$$H^1_{\mathfrak{m}}\big(H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu})\big) = 0 \quad \textit{for all } k.$$

Proof. We know by Theorem 4.2 that each of the groups $H^k(X^{(p)}, \mathcal{M}^{(p)}_{\lambda,q;\mu})$ has an S-module filtration with composition factors isomorphic to $J_{\nu,l}$ where $\nu_1 = \cdots = \nu_{l+1}$, so it suffices to prove that $H^1_{\mathfrak{m}}(J_{\nu,l}) = 0$ for each such factor. Since no factor has $l = 1$ and $\nu_1 \neq \nu_2$, the desired vanishing follows from Lemma 4.3. $\square$

We record for later use one more vanishing result which is a direct consequence of Bott's theorem.

Lemma 4.6. *Suppose that $\mathcal{M}$ decomposes as an $\mathcal{O}_{X^{(p)}}$-module into a direct sum of sheaves of the form*

$$\mathcal{B} = \mathbb{S}_\nu \mathcal{Q}_q(F) \otimes \mathcal{L}^\mu(F) \otimes \mathbb{S}_\nu \mathcal{Q}_q(G) \otimes \mathcal{L}^\mu(G),$$

where $\nu \in \mathbb{Z}^p_{\mathrm{dom}}$ and $\mu \in \mathbb{Z}^{n-p}$. We have that

$$H^k(X^{(p)}, \mathcal{M}) = 0 \quad \textit{for } k \textit{ odd.}$$

Proof. Combining the Künneth theorem with Theorem 2.1(b) we see that $\mathcal{B}$ has nonvanishing cohomology if and only if $(\nu|\mu)+\delta^{(n)}$ has no repeated entries, in which case its only nonvanishing cohomology group is

$$H^{2l}(X^{(p)}, \mathcal{B}) = H^l\big(\mathrm{Flag}([p, n]; F), \mathbb{S}_\nu \mathcal{Q}_q(F) \otimes \mathcal{L}^\mu(F)\big) \otimes H^l\big(\mathrm{Flag}([p, n]; G), \mathbb{S}_\nu \mathcal{Q}_q(G) \otimes \mathcal{L}^\mu(G)\big),$$

where l is the number of inversions in $(\nu|\mu) + \delta^{(n)}$. In particular $H^k(X^{(p)}, \mathcal{B}) = 0$ for k odd, so the same is true for $\mathcal{M}$, concluding the proof. □

Remark 4.7. The above vanishing applies when $\mathcal{M} = \mathcal{M}^{(p)}_{\lambda,q;\mu}$, where $0 \le q \le p$, $\lambda \in \mathbb{Z}^p_{\mathrm{dom}}$ is such that $\lambda_1 = \cdots = \lambda_q$, and $\mu \in \mathbb{Z}^{n-p}$.

4B. ***Proof of Theorem 4.1.*** We fix $0 \le p \le n$ and $\underline{x} \in \mathbb{N}^n_{\mathrm{dom}}$ with $x_1 = \cdots = x_p$. We write $\mathcal{X} = X^{(p)}$, $\mathcal{Y} = Y^{(p)}$, and consider the commutative diagram

$$\begin{tikzcd}
\mathcal{Y} \arrow[r, hook, "\iota"] \arrow[dr, "\phi"'] & T = \mathrm{Spec}\, S \times \mathcal{X} \arrow[r, "\pi_T"] \arrow[d, "\pi"] & \mathcal{X} \\
& \mathrm{Spec}\, S &
\end{tikzcd}$$

We can identify T with the total space of the trivial bundle $(F \otimes G)^\vee$ over $\mathcal{X}$, and $\mathcal{Y}$ with a subbundle of T via the inclusion ι. We write $\pi_{\mathcal{Y}} = \pi_T \circ \iota$ for the projection map $\mathcal{Y} \to \mathcal{X}$.

We define $\underline{y} \in \mathbb{N}^{n-p}_{\mathrm{dom}}$ by letting $y_i = x_{p+i}$ for $i = 1, \dots, n-p$, set $d = x_1$ and let $\mathcal{M} = \mathcal{M}^{(p)}_{(d^p),p;\underline{y}}$. We write $\mathcal{S} = \mathcal{S}^{(p)}$, $\mathcal{D} = \det^{(p)}$, and thinking of $\mathcal{M}$ as an $\mathcal{S}$-module on $\mathcal{X}$ we have that

$$\mathcal{M} = \mathcal{S} \otimes_{\mathcal{O}_\mathcal{X}} \mathcal{V}, \qquad \text{where } \mathcal{V} = \mathcal{D}^{\otimes d} \otimes_{\mathcal{O}_\mathcal{X}} \mathcal{L}^{\underline{y}} \text{ is a line bundle on } \mathcal{X}. \tag{4-8}$$

We can then think of $\mathcal{M}$ as being locally (on the base $\mathcal{X}$) isomorphic to $\mathcal{S}$, or as the invertible sheaf $\pi^*_{\mathcal{Y}}\mathcal{V}$ on $\mathcal{Y}$. The relationship between $\mathcal{M}$ and $J_{\underline{x},p}$ is given by (4-5), which can be interpreted as the equality

$$R\phi_*(\mathcal{M}) = J_{\underline{x},p} \tag{4-9}$$

in the derived category, where $J_{\underline{x},p}$ is considered as a complex concentrated in cohomological degree 0.

Proof of Theorem 4.1(a). Observe that $\mathrm{Ext}^j_S(J_{\underline{x},p}, S) = R^j\,\mathrm{Hom}_S(J_{\underline{x},p}, S)$. Using (4-9) and Grothendieck duality [Hartshorne 1966, Theorem 11.1] we obtain

$$R\,\mathrm{Hom}_S(J_{\underline{x},p}, S) = R\,\mathrm{Hom}_S(R\phi_*(\mathcal{M}), S) = R\phi_*(R\mathcal{H}om_\mathcal{Y}(\mathcal{M}, \phi^! S)) = R\phi_*(\mathcal{M}^\vee \otimes_{\mathcal{O}_\mathcal{Y}} \phi^! S)), \tag{4-10}$$

where the last equality follows from the fact that $\mathcal{M}$ is locally free. By functoriality we have $\phi^! S = \iota^!(\pi^! S)$

and $\pi^! S = \pi_T^* \omega_{\mathcal{X}}[\dim \mathcal{X}]$, where $[-]$ indicates the shift in cohomological degree and $\omega_{\mathcal{X}}$ is the canonical bundle on $\mathcal{X}$ (see [Hartshorne 1966, Section III.2]). Using [Hartshorne 1966, Section III.6], we also have

$$\begin{aligned}\phi^! S = \iota^!(\pi_T^* \omega_{\mathcal{X}}[\dim \mathcal{X}]) &= \iota^* R\mathcal{H}om_T(\iota_* \mathcal{O}_{\mathcal{Y}}, \pi_T^* \omega_{\mathcal{X}}[\dim \mathcal{X}]) \\ &= \det(\mathcal{N}_{\mathcal{Y}|T})[\dim(\mathcal{Y}) - \dim(T)] \otimes_{\mathcal{O}_{\mathcal{Y}}} \pi_{\mathcal{Y}}^* \omega_{\mathcal{X}}[\dim \mathcal{X}], \end{aligned} \tag{4-11}$$

where $\mathcal{N}_{\mathcal{Y}|T}$ is the normal bundle of $\mathcal{Y}$ in $\mathcal{X}$. We have $\mathcal{N}_{\mathcal{Y}|T} = \pi_{\mathcal{Y}}^* \xi^\vee$, where

$$\xi = \ker\big((F \otimes G) \otimes \mathcal{O}_{\mathcal{X}} \to \mathcal{Q}_p(F) \otimes \mathcal{Q}_p(G)\big),$$

so in order to compute $\det(\mathcal{N}_{\mathcal{Y}|T})$ it suffices to compute $\det(\xi)$. We have

$$\begin{aligned}\det(\xi) &= \det(F \otimes G) \otimes \det(\mathcal{Q}_p(F) \otimes \mathcal{Q}_p(G))^\vee \\ &= (\det{}^{(n)})^{\otimes n} \otimes_{\mathcal{O}_{\mathcal{X}}} (\det{}^{(p)})^{\otimes(-p)} \\ &= \mathcal{D}^{\otimes(n-p)} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{L}^{(n^{n-p})},\end{aligned}$$

where the last equality follows from the fact that $\det^{(n)} = \det^{(p)} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{L}^{(1^{n-p})}$. We have moreover that

$$\dim(\mathcal{Y}) - \dim(T) + \dim(\mathcal{X}) = -n + p,$$

and the canonical bundle on $\mathcal{X}$ is given by (see for instance [Weyman 2003, Exercise 13, Chapter 3])

$$\omega_{\mathcal{X}} = \mathcal{D}^{\otimes(p-n)} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{L}^{(2p+1-n, 2p+3-n, \dots, n-1)}.$$

We can therefore rewrite (4-11) as

$$\phi^! S = \pi_{\mathcal{Y}}^* \big(\mathcal{D}^{\otimes(2p-2n)} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{L}^{(2p+1-2n, 2p+3-2n, \dots, -1)}\big)[-n+p]$$

Tensoring this with $\mathcal{M}^\vee = \pi_{\mathcal{Y}}^* \big(\mathcal{D}^{\otimes(-d)} \otimes_{\mathcal{O}_{\mathcal{X}}} \mathcal{L}^{-\underline{y}}\big)$ we obtain

$$\mathcal{M}^\vee \otimes_{\mathcal{O}_{\mathcal{Y}}} \phi^! S = \mathcal{M}^{(p)}_{\lambda, p; \mu}[-n+p],$$

where

$$\lambda = ((2p - 2n - d)^p) \quad \text{and} \quad \mu_i = 2p + 2i - 1 - 2n - y_i \quad \text{for } 1 \le i \le n - p.$$

It follows from (4-10) that

$$\operatorname{Ext}^j_S(J_{\underline{x}, p}, S) = R^j \phi_* (\mathcal{M}^{(p)}_{\lambda, p; \mu}[-n+p]) = H^{j-n+p}(X^{(p)}, \mathcal{M}^{(p)}_{\lambda, p; \mu}).$$

Since $d \ge y_1$ it follows that $\lambda_p = 2p - 2n - d \le \mu_1 = 2p + 1 - 2n - y_1$, so we can apply Corollary 4.5 to conclude that $H^1_{\mathfrak{m}}(\operatorname{Ext}^j_S(J_{\underline{x}, p}, S)) = 0$ for all j. □

Proof of Theorem 4.1(b). We let $Z_t = \phi^{-1}(\overline{O}_t)$ and note that working relative to the base $\mathcal{X}$, Z_t is locally the variety of $p \times p$ matrices of rank at most t. It is cut out inside $\mathcal{Y}$ by the sheaf of ideals $\mathcal{I}_{(t+1)\times 1} \subset \mathcal{S}$.

It follows from the discussion in Section 2G that if we set $m = n = p$ in (2-10) then

$$\mathcal{H}^j_{Z_t}(\mathcal{Y}, \mathcal{O}_{\mathcal{Y}}) = \mathcal{H}^j_{\mathcal{I}_{(t+1)\times 1}}(\mathcal{X}, \mathcal{S}) = \begin{cases} 0 & \text{if } j \not\equiv (p-t) \pmod 2, \\ \bigoplus \mathbb{S}_\lambda \mathcal{Q}_p(F) \otimes \mathbb{S}_\lambda \mathcal{Q}_p(G) & \text{if } j \equiv (p-t) \pmod 2, \end{cases}$$

where the direct sum is over some collection of weights $\lambda \in \mathbb{Z}^p_{\text{dom}}$ (with repetitions allowed), whose precise description follows from (2-7), but is not relevant for the rest of the argument. It follows from (4-8) that

$$\mathcal{H}^j_{Z_t}(\mathcal{Y}, \mathcal{M}) = \begin{cases} 0 & \text{if } j \not\equiv (p-t) \pmod 2, \\ \mathcal{V} \otimes_{\mathcal{O}_{\mathcal{X}}} \left(\bigoplus \mathbb{S}_\lambda \mathcal{Q}_p(F) \otimes \mathbb{S}_\lambda \mathcal{Q}_p(G)\right) & \text{if } j \equiv (p-t) \pmod 2. \end{cases} \tag{4-12}$$

Writing Γ_Z for the functor of sections with support in Z, we get a natural isomorphism

$$\Gamma_{\overline{O}_t} \circ \phi_* = \phi_* \circ \Gamma_{Z_t},$$

which yields in the derived category

$$R\Gamma_{\overline{O}_t}(J_{\underline{x},p}) = R\Gamma_{\overline{O}_t}(R\phi_*\mathcal{M}) = R\phi_*(R\Gamma_{Z_t}(\mathcal{M})).$$

This means that we have a spectral sequence

$$E_2^{i,j} = H^i(\mathcal{Y}, \mathcal{H}^j_{Z_t}(\mathcal{Y}, \mathcal{M})) \Rightarrow H^{i+j}_{\overline{O}_t}(J_{\underline{x},p}).$$

We have noted in (4-12) that $\mathcal{H}^j_{Z_t}(\mathcal{Y}, \mathcal{M}) = 0$ when $j \not\equiv (p-t) \pmod 2$, and it follows from Lemma 4.6 and (4-12) that $H^i(\mathcal{Y}, \mathcal{H}^j_{Z_t}(\mathcal{Y}, \mathcal{M})) = 0$ when i is odd. It follows that

$$E_2^{i,j} = 0 \quad \text{when } i+j \not\equiv (p-t) \pmod 2,$$

proving that $H^k_{\overline{O}_t}(J_{\underline{x},p}) = 0$ for $k \not\equiv (p-t) \pmod 2$, as desired. □

5. More vanishing of local cohomology

The goal of this short section is to prove two vanishing results, which are based on Theorem 4.1 and will constitute important ingredients in describing the module structure of local cohomology groups for square matrices. We continue to assume as in Section 4 that $m = n$.

Theorem 5.1. *For all $p < n$ and all $j \geq 0$ we have that*

$$H^1_{\mathfrak{m}}(H^j_{\overline{O}_p}(S)) = 0.$$

Proof. As in (3-14) we can write

$$H^j_{\overline{O}_p}(S) = \varinjlim_d \operatorname{Ext}^j_S(S/I_{(p+1)\times d}, S).$$

Since local cohomology commutes with direct limits, it is sufficient to prove that

$$H^1_{\mathfrak{m}}(\operatorname{Ext}^j_S(S/I_{(p+1)\times d}, S)) = 0.$$

Using [Raicu and Weyman 2014, Lemma 2.2] (with the notation there, we choose $\underline{x}$ to be the zero partition and $\underline{y} = (d^{p+1})$), we see that the modules $S/I_{(p+1)\times d}$ admit a finite filtration by S-submodules whose successive quotients are of the form $J_{\underline{z},p}$, with $z_1 = \cdots = z_p (= z_{p+1})$. By [Raicu and Weyman 2014, Corollary 3.5] (see also [Raicu 2018, Theorem 3.2]), this induces a filtration on $\operatorname{Ext}^j_S(S/I_{(p+1)\times d}, S)$ with successive quotients $\operatorname{Ext}^j_S(J_{\underline{z},p}, S)$. The conclusion follows now from Theorem 4.1(a). □

Recall the definition of Q_p from (1-12). The following should be seen as an analogue of Corollary 3.5.

Theorem 5.2. *If $t \le p$ then we have that for all $k \not\equiv p - t \pmod 2$*

$$H^k_{\overline{O}_t}(Q_p) = 0.$$

Proof. Note that since $S_{\det} = Q_n$, we have by (2-12) a decomposition

$$S_{\det} = \bigoplus_{\lambda \in \mathbb{Z}^n_{\text{dom}}} \mathbb{S}_\lambda \mathbb{C}^n \otimes \mathbb{S}_\lambda \mathbb{C}^n,$$

analogous to (2-5), with the only difference that λ is allowed to be any dominant weight, as opposed to just a partition. In analogy with $I_{\underline{x}}$, we can then define the *fractional ideals* I_λ to be the S-submodules of $S_{\det}$ generated by $\mathbb{S}_\lambda \mathbb{C}^n \otimes \mathbb{S}_\lambda \mathbb{C}^n$. We have $I_\lambda = \det^{-1} \cdot I_{\lambda+(1^n)}$, and it follows from (2-6) that

$$I_\lambda = \bigoplus_{\mu \ge \lambda} \mathbb{S}_\mu \mathbb{C}^n \otimes \mathbb{S}_\mu \mathbb{C}^n. \tag{5-1}$$

We can write

$$S_{\det} = \varinjlim_d (\det{}^{-d} \cdot S) = \varinjlim_d I_{(-d^n)}.$$

Using (1-12) and (2-12) it follows that

$$\langle \det{}^{p-n+1} \rangle_{\mathcal{D}} = \bigoplus_{\lambda_{p+1} \ge p+1-n} \mathbb{S}_\lambda \mathbb{C}^n \otimes \mathbb{S}_\lambda \mathbb{C}^n$$

and in particular using (5-1) we get

$$I_{(-d^n)} \cap \langle \det{}^{p-n+1} \rangle_{\mathcal{D}} = I_{((p+1-n)^{p+1}, (-d)^{n-p-1})}$$

for $d \gg 0$. We can then rewrite (1-12) as

$$Q_p = \varinjlim_d \frac{I_{(-d^n)}}{I_{((p+1-n)^{p+1}, (-d)^{n-p-1})}} = \varinjlim_d \frac{\det^d \cdot I_{(-d^n)}}{\det^d \cdot I_{((p+1-n)^{p+1}, (-d)^{n-p-1})}} = \varinjlim_d \frac{S}{I_{(p+1)\times(d+p+1-n)}}.$$

Since local cohomology commutes with direct limits, it is enough to show that

$$H^k_{\overline{O}_t}(S/I_{(p+1)\times(p-n+d+1)}) = 0 \quad \text{for } k \not\equiv p-t \pmod 2 \text{ and } d \gg 0.$$

As seen in the proof of Theorem 5.1, the modules $S/I_{(p+1)\times(p-n+d+1)}$ admit a finite composition series by S-submodules, with composition factors of the form $J_{\underline{z},p}$, with $z_1 = \cdots = z_p$. The desired vanishing now follows from Theorem 4.1(b). □

6. Module structure of local cohomology groups

The goal of this section is to describe for $X = \mathbb{C}^{n\times n}$ the decomposition into a sum of indecomposable objects in $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ of the local cohomology groups $H^\bullet_{\overline{O}_t}(D_p)$ and $H^\bullet_{\overline{O}_t}(Q_p)$. In the case of nonsquare matrices ($m > n$) we have noted that $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ is semisimple, so the indecomposable objects are the simple modules $D_0, \dots, D_n$, and the decomposition of the local cohomology groups into a sum of simple modules is already encoded by their class in the Grothendieck group described in Theorem 3.1. We will therefore only be concerned with the case when $m = n$ for the rest of the section.

To state the main results of the section, we begin by considering the full additive subcategory $\mathrm{add}(Q)$ of $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ formed by the $\mathcal{D}_X$-modules that are isomorphic to a direct sum of copies of $Q_0, Q_1, \dots, Q_n$. We let Ψ denote the semigroup of isomorphism classes of objects in $\mathrm{add}(Q)$, where the semigroup operation is given by direct sum. We write $[M]$ for the class in Ψ of a module $M \in \mathrm{add}(Q)$. We have a natural inclusion of Ψ as a subsemigroup of $\Gamma_{\mathcal{D}}$, given by $[M] \mapsto [M]_{\mathcal{D}}$. Our first theorem describes the local cohomology groups $H^\bullet_{\overline{O}_t}(D_p)$ as elements of $\mathrm{add}(Q)$ as follows.

Theorem 6.1. *For every t, p, j with $0 \le t < p \le n$ and $j \ge 0$ we have that $H^j_{\overline{O}_t}(D_p) \in \mathrm{add}(Q)$. Moreover,*

$$\sum_{j\ge 0} [H^j_{\overline{O}_t}(D_p)] \cdot q^j = \sum_{s=0}^{t} [Q_s] \cdot q^{(p-t)^2} \cdot m_s(q^2)$$

holds in $\Psi[q]$, where $m_s(q) \in \mathbb{Z}[q]$ is computed by $m_t(q) = \binom{n-t}{p-t}_q$ and

$$m_s(q) = \binom{n-s}{p-s}_q \cdot \binom{p-1-s}{t-s}_q - \binom{n-s-1}{p-s-1}_q \cdot \binom{p-2-s}{t-1-s}_q \quad \textit{for } s = 0, \dots, t-1.$$

Proof. The main content of the theorem is the assertion that $H^j_{\overline{O}_t}(D_p) \in \mathrm{add}(Q)$, which will be proved in Proposition 6.11. Since Ψ embeds into $\Gamma_{\mathcal{D}}$, we can determine the polynomials $m_s(q)$ by expressing $[H^j_{\overline{O}_t}(D_p)]_{\mathcal{D}}$ in terms of $[Q_s]_{\mathcal{D}}$. Using the fact that $[D_s]_{\mathcal{D}} = [Q_s]_{\mathcal{D}} - [Q_{s-1}]_{\mathcal{D}}$ for $s \ge 1$ and $[D_0]_{\mathcal{D}} = [Q_0]_{\mathcal{D}}$, the desired formula for $m_s(q)$ follows from the case $m = n$ of Theorem 3.1. □

In order to be able to compute iterated local cohomology groups, we need to be able to describe the local cohomology groups of the modules Q_p.

Theorem 6.2. *For every t, p, j with $0 \le t < p \le n$ and $j \ge 0$ we have that $H^j_{\overline{O}_t}(Q_p) \in \mathrm{add}(Q)$. Moreover,*

$$\sum_{j\ge 0} [H^j_{\overline{O}_t}(Q_p)] \cdot q^j = \sum_{s=0}^{t} [Q_s] \cdot q^{(p-t)^2+2(p-s)} \cdot \binom{n-s-1}{p-s}_{q^2} \cdot \binom{p-s-1}{p-t-1}_{q^2} \quad \textit{holds in } \Psi[q]. \tag{6-1}$$

6A. ***The quiver description of*** $\mathbf{mod}_{\mathbf{GL}}(\mathcal{D}_X)$**.** We recall from [Lőrincz and Walther 2019] the quiver-theoretical description of the category $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, referring the reader to [Lőrincz et al. 2019, Section 2.4] for a quick summary of the notation and properties of quiver representations that we will use. In particular, for a quiver representation $\mathfrak{W}$ we write $\mathfrak{W}_x$ for the vector space associated to a vertex x, and write $\mathfrak{W}(\alpha)$ for the linear transformation attached to an arrow α. We consider the quiver with relations pictured as

$$\widehat{AA}_n : \quad (0) \underset{\beta_1}{\overset{\alpha_1}{\rightleftarrows}} (1) \underset{\beta_2}{\overset{\alpha_2}{\rightleftarrows}} \cdots \underset{\beta_{n-1}}{\overset{\alpha_{n-1}}{\rightleftarrows}} (n-1) \underset{\beta_n}{\overset{\alpha_n}{\rightleftarrows}} (n), \tag{6-2}$$

where the relations are given by the condition that all 2-cycles are zero (i.e., $\alpha_i\beta_i = 0 = \beta_i\alpha_i$ for all $i = 1, \dots, n$). By [Lőrincz and Walther 2019, Theorem 5.4] we have an equivalence of categories

$$\operatorname{mod}_{\mathrm{GL}}(\mathcal{D}_X) \simeq \operatorname{rep}(\widehat{AA}_n) \tag{6-3}$$

between $\operatorname{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ and the category of finite-dimensional representations of $\widehat{AA}_n$. For instance, under this equivalence the simple $\mathcal{D}_X$-module D_p corresponds to the irreducible representation

$$\mathfrak{D}^{(p)} : \quad 0 \underset{0}{\overset{0}{\rightleftarrows}} \cdots \underset{0}{\overset{0}{\rightleftarrows}} 0 \underset{0}{\overset{0}{\rightleftarrows}} \mathbb{C} \underset{0}{\overset{0}{\rightleftarrows}} 0 \underset{0}{\overset{0}{\rightleftarrows}} \cdots \underset{0}{\overset{0}{\rightleftarrows}} 0,$$

where a one-dimensional vector space $\mathbb{C}$ is placed at vertex (p), and 0 is placed at all the other vertices. It will be important to identify the quiver representations corresponding to the modules Q_p in (1-12).

Lemma 6.3. *For each $p = 0, \dots, n$, we consider the representation $\mathfrak{Q}^{(p)} \in \operatorname{rep}(\widehat{AA}_n)$ obtained by letting $\mathfrak{Q}^{(p)}_{(i)} = \mathbb{C}$ for $0 \le i \le p$, and $\mathfrak{Q}^{(p)}_{(i)} = 0$ for $i > p$, and with maps*

$$\mathfrak{Q}^{(p)} : \quad \mathbb{C} \underset{0}{\overset{1}{\rightleftarrows}} \mathbb{C} \underset{0}{\overset{1}{\rightleftarrows}} \cdots \underset{0}{\overset{1}{\rightleftarrows}} \mathbb{C} \underset{0}{\overset{0}{\rightleftarrows}} 0 \underset{0}{\overset{0}{\rightleftarrows}} \cdots \underset{0}{\overset{0}{\rightleftarrows}} 0. \tag{6-4}$$

We have that $\mathfrak{Q}^{(p)}$ contains $\mathfrak{D}^{(p)}$ as its unique irreducible subrepresentation, and that $\mathfrak{Q}^{(p)}/\mathfrak{D}^{(p)} \simeq \mathfrak{Q}^{(p-1)}$. Moreover, the $\mathcal{D}_X$-module Q_p corresponds via (6-3) to the representation $\mathfrak{Q}^{(p)}$ for all $0 \le p \le n$.

Proof. The fact that $\mathfrak{D}^{(p)}$ is a subrepresentation of $\mathfrak{Q}^{(p)}$ and the identification $\mathfrak{Q}^{(p)}/\mathfrak{D}^{(p)} \simeq \mathfrak{Q}^{(p-1)}$ follow from the definition of $\mathfrak{Q}^{(p)}$. If $\mathfrak{W} \subseteq \mathfrak{Q}^{(p)}$ is a subrepresentation with $\mathfrak{W}_{(i)} = \mathbb{C}$ for some $i < p$, then $\mathfrak{W}_{(i+1)}$ contains the image under $\mathfrak{Q}^{(p)}(\alpha_{i+1})$ of $\mathfrak{W}_{(i)}$, that is $\mathfrak{W}_{(i+1)} = \mathbb{C}$. It follows that $\mathfrak{W}_{(j)} = \mathbb{C}$ for all $j = i, \dots, p$, and in particular $\mathfrak{W}$ contains $\mathfrak{D}_p$ as a subrepresentation.

To prove that Q_p corresponds to $\mathfrak{Q}^{(p)}$ via (6-3) we argue by descending induction on p. Using (1-10)–(1-12) we get that $Q_{p-1} \simeq Q_p/D_p$, proving the inductive step. It remains to address the base case $p=n$, when $Q_n = S_{\det}$. If we apply [Lőrincz et al. 2019, Lemma 2.4] with $G = \mathrm{GL}$, $Y = \mathbb{C}^{n\times n}$, $U = O = O_n$ the dense orbit of rank n matrices, and $j : U \to Y$ the natural inclusion, it follows that $S_{\det} = j_* j^* S$ is the injective envelope of $S = D_n$ in $\operatorname{mod}_{\mathrm{GL}}(\mathcal{D}_X)$, so $S_{\det}$ corresponds via (6-3) to the injective envelope of $\mathfrak{D}^{(n)}$. Using the quiver description of the injective envelope of a simple representation from [Lőrincz et al. 2019, (2.15)], it follows that $\mathfrak{Q}^{(n)}$ is the injective envelope of $\mathfrak{D}^{(n)}$, concluding the proof. □

For each $p = 0, \dots, n$ we consider the full subcategory

$$\operatorname{mod}^{\overline{O}_p}_{\mathrm{GL}}(\mathcal{D}_X)$$

of $\operatorname{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ consisting of modules with support contained in $\overline{O}_p$. This subcategory is closed under extensions and taking subquotients, and it corresponds via (6-3) to the subcategory $\operatorname{rep}(\widehat{AA}_p)$ of $\operatorname{rep}(\widehat{AA}_n)$, obtained by forgetting the vertices $(p+1), \dots, (n)$ of the quiver $\widehat{AA}_n$. We have the following important observation, which follows from [Lőrincz et al. 2019, (2.15)] and the equivalence with $\mathrm{rep}(\widehat{AA}_p)$.

Lemma 6.4. *Inside the category* $\mathrm{mod}_{\mathrm{GL}}^{\overline{O}_p}(\mathcal{D}_X)$, *the module* Q_p *is the injective envelope of* D_p *and the projective cover of* D_0. *In particular*, Q_p *is indecomposable.*

To describe local cohomology groups we will work mainly in the additive subcategory $\mathrm{add}(Q)$ of $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$. One property important to us is that $\mathrm{add}(Q)$ is closed under taking extensions and quotients.

Lemma 6.5. *For every* $0 \le i, j \le n$ *we have that* $\mathrm{Ext}^1_{\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)}(Q_i, Q_j) = 0$. *In particular, every short exact sequence in* $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$

$$0 \to M_1 \to N \to M_2 \to 0, \tag{6-5}$$

with $M_1, M_2 \in \mathrm{add}(Q)$ *splits, and hence* $N \in \mathrm{add}(Q)$. *More generally, if* $N \in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ *has a composition series with composition factors in* $\mathrm{add}(Q)$, *then* $N \in \mathrm{add}(Q)$.

Proof. For the first assertion, we let $p = \max(i, j)$ and note that since $\mathrm{mod}_{\mathrm{GL}}^{\overline{O}_p}(\mathcal{D}_X)$ is closed under taking extensions it suffices to prove that

$$\mathrm{Ext}^1_{\mathrm{mod}_{\mathrm{GL}}^{\overline{O}_p}(\mathcal{D}_X)}(Q_i, Q_j) = 0.$$

By Lemma 6.4, if $p = i$ then Q_i is projective in $\mathrm{mod}_{\mathrm{GL}}^{\overline{O}_p}(\mathcal{D}_X)$, while if $p = j$ then Q_j is injective, so the above vanishing follows. Since Ext^1 commutes with finite direct sums, it follows that $\mathrm{Ext}^1_{\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)}(M_2, M_1) = 0$ for $M_1, M_2 \in \mathrm{add}(Q)$, and therefore (6-5) splits. To prove the last assertion, we argue by induction on the length of the composition series. We write N as an extension (6-5), where $M_2 \in \mathrm{add}(Q)$ and M_1 has a shorter composition series with composition factors in $\mathrm{add}(Q)$. By induction we have that $M_1 \in \mathrm{add}(Q)$, hence (6-5) splits and N is also in $\mathrm{add}(Q)$. □

Lemma 6.6. *Any quotient of* Q_p *in* $\mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_X)$ *is isomorphic to* Q_q *for some* $0 \le q \le p$. *More generally, if* $M \in \mathrm{add}(Q)$ *then any quotient of* M *is also in* $\mathrm{add}(Q)$.

Proof. We prove the first assertion by induction on p. By Lemma 6.3 and (6-3), D_p is the unique simple submodule of Q_p, and therefore every proper quotient of Q_p factors through $Q_p/D_p = Q_{p-1}$. By induction, every quotient of Q_{p-1} is isomorphic to Q_q for some $0 \le q \le p-1$, so the same must be true about every proper quotient of Q_p.

For the last assertion we argue by induction on the length of M. We consider a quotient $\pi : M \twoheadrightarrow P$ and write $M = Q_p \oplus N$ with $N \in \mathrm{add}(Q)$, and let $P' = \pi(Q_p)$. Using the previous paragraph, $P' \simeq Q_q$ for some $0 \le q \le p$. The map π induces a map of short exact sequences,

$$\begin{tikzcd} 0 \arrow[r] & Q_i \arrow[r] \arrow[d] & M \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & P' \arrow[r] & P \arrow[r] & P'' \arrow[r] & 0 \end{tikzcd}$$

where the vertical maps are surjective and $P'' = P/P'$. Since N has smaller length than M, it follows that $P'' \in \mathrm{add}(Q)$, hence $P \in \mathrm{add}(Q)$ by Lemma 6.5. □

6B. ***Local cohomology of the polynomial ring S.*** The goal of this section is to prove that the local cohomology groups of S are in $\mathrm{add}(Q)$, thus proving the case $p=n$ of Theorem 6.1. Our argument will be inductive, starting with the observations in Section 2H. We let $X_1\subset X$ denote the basic open affine where $x_{11}\neq 0$, let $U=X\setminus\{0\}$, and let $j_1:X_1\to U$ denote the open immersion.

Lemma 6.7. *If $M, N\in \mathrm{mod}_{\mathrm{GL}}(\mathcal{D}_U)$ are such that there exists a $\mathcal{D}_{X_1}$-module isomorphism $j_1^*M\simeq j_1^*N$ then $M\simeq N$.*

Proof. We let $Z=U\setminus X_1$ and consider the exact sequences

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{H}^0_Z(M) \arrow[r] & M \arrow[r, "\alpha"] & j_{1*}j_1^*M \arrow[r, "\alpha'"] \arrow[d, "\phi"', "\simeq"] & \mathcal{H}^1_Z(M) \arrow[r] & 0 \\ 0 \arrow[r] & \mathcal{H}^0_Z(N) \arrow[r] & N \arrow[r, "\beta"] & j_{1*}j_1^*N \arrow[r, "\beta'"] & \mathcal{H}^1_Z(N) \arrow[r] & 0 \end{tikzcd}$$

where ϕ exists by assumption. Since Z contains no invariant closed subset of U, it follows that no nonzero subquotient of M and N can have support in Z. Hence, we have $\mathcal{H}^0_Z(M)=\mathcal{H}^0_Z(N)=0$ and therefore α, β are injective. Moreover, we have $\beta'\circ\phi\circ\alpha=0$ and $\alpha'\circ\phi^{-1}\circ\beta=0$, so that $\phi\circ\alpha$ (resp. $\phi^{-1}\circ\beta$) lifts to an injective $\mathcal{D}_U$-module homomorphism ϕ_1 (resp. ϕ_2). Since M and N have finite length, it follows that their lengths coincide, and ϕ_1 and ϕ_2 must be isomorphisms. ☐

Proposition 6.8. *For all $t<n$ and $i\geq 0$ we have $H^i_{\overline{O}_t}(S)\in\mathrm{add}(Q)$.*

Proof. We proceed by induction on n: if $t=0$ then each $H^i_{\overline{O}_t}(S)$ is a direct sum of copies of $D_0=Q_0$, so it is in $\mathrm{add}(Q)$. We may assume then that $n>t\geq 1$ and let $j:U\to X$ denote the inclusion. For any $\mathcal{D}_X$-module M, we have the exact sequence

$$0\to H^0_{\mathfrak{m}}(M)\to M\to j_*j^*M\to H^1_{\mathfrak{m}}(M)\to 0. \tag{6-6}$$

We let $Q_p^0=j^*Q_p$ for $1\leq p\leq n$ and prove that $j_*Q_p^0=Q_p$. From Lemma 6.3 we see that Q_p has no submodules supported at O_0, so $H^0_{\mathfrak{m}}(Q_p)=0$. Choosing $M=Q_p$ in (6-6) gives then the exact sequence

$$0\to Q_p\to j_*Q_p^0\to H^1_{\mathfrak{m}}(Q_p)\to 0.$$

Since $H^1_{\mathfrak{m}}(Q_p)$ is a direct sum of copies of $D_0=Q_0$, it follows from Lemma 6.5 that the above sequence splits. If we set $M=j_*Q_p^0$ in (6-6) and note that the map $j_*Q_p^0\to j_*j^*j_*Q_p^0$ is an isomorphism, we get $H^0_{\mathfrak{m}}(j_*Q_p^0)=0$. Since $H^1_{\mathfrak{m}}(Q_p)$ is a summand of $j_*Q_p^0$ supported at O_0, this shows that $H^1_{\mathfrak{m}}(Q_p)=0$ and $j_*Q_p^0=Q_p$.

We now claim that each $j^*H^i_{\overline{O}_t}(S)$ is a direct sum of copies of the $\mathcal{D}_U$-modules $Q_1^0,\dots,Q_n^0$. To prove this, it suffices by Lemma 6.7 to show that an isomorphism exists after restricting to X_1. For that we have

$$j_1^*j^*H^i_{\overline{O}_t}(S)=(H^i_{\overline{O}_t}(S))_{|X_1}\overset{\text{(2-23)}}{=}\pi^*(H^i_{\overline{O'}_{t-1}}(S'))\overset{\text{(2-22)}}{=}\bigoplus_{1\leq s\leq t}(Q_s^{\oplus a_s})_{|X_1}=\bigoplus_{1\leq s\leq t}j_1^*(Q_s^0)^{\oplus a_s}$$

where the equality labeled (2-22) uses also the induction hypothesis, and where the numbers a_s are in $\mathbb{Z}_{\geq 0}$.

Since $j^* H^i_{\overline{O}_t}(S)$ is a direct sum of copies of $Q^0_1, \dots, Q^0_n$, and $j_* Q^0_p = Q_p$ for $1 \le p \le n$, it follows now that $j_* j^* H^i_{\overline{O}_t}(S) \in \operatorname{add}(Q)$. Setting $M = H^i_{\overline{O}_t}(S)$ in (6-6) we obtain using Theorem 5.1 the exact sequence

$$0 \to H^0_{\mathfrak{m}}(H^i_{\overline{O}_t}(S)) \to H^i_{\overline{O}_t}(S) \to j_* j^* H^i_{\overline{O}_t}(S) \to 0.$$

Since $H^0_{\mathfrak{m}}(H^i_{\overline{O}_t}(S)) \in \operatorname{add}(Q)$, it follows from Lemma 6.5 that $H^i_{\overline{O}_t}(S) \in \operatorname{add}(Q)$, concluding the proof. $\square$

6C. ***The structure of the modules $H^\bullet_{\overline{O}_{p-1}}(Q_p)$.*** In this section we prove the case $t = p-1$ of Theorem 6.2.

Lemma 6.9. *For all $j \ge 0$ and $t < n$ we have $H^j_{\overline{O}_t}(S_{\det}) = 0$.*

Proof. Multiplication by the polynomial det induces an S-module isomorphism $S_{\det} \xrightarrow{\cdot \det} S_{\det}$, which in turn gives rise to an isomorphism $H^j_{\overline{O}_t}(S_{\det}) \xrightarrow{\cdot \det} H^j_{\overline{O}_t}(S_{\det})$ for each $j \ge 0$. Since the polynomial det vanishes on $\overline{O}_t$ it follows that every element $m \in H^j_{\overline{O}_t}(S_{\det})$ is annihilated by $\det^k$ for some k. Since multiplication by $\det^k$ is an isomorphism, we conclude that $m = 0$ and, since m was arbitrary, that $H^j_{\overline{O}_t}(S_{\det}) = 0$. $\square$

Lemma 6.10. *For all $p \le n$ and $j \ge 0$, we have $H^j_{\overline{O}_{p-1}}(Q_p) \in \operatorname{add}(Q)$ and*

$$H^0_{\overline{O}_{p-1}}(Q_p) = H^1_{\overline{O}_{p-1}}(Q_p) = 0. \tag{6-7}$$

Proof. The case $p = n$ follows from Lemma 6.9, so we may assume that $p < n$. We consider the spectral sequence

$$E_2^{i,j} = H^i_{\overline{O}_{p-1}}(H^j_{\overline{O}_p}(S)) \Rightarrow H^{i+j}_{\overline{O}_{p-1}}(S).$$

If we let $c_p = (n-p)^2$ denote the codimension of O_p in X, then we know that $H^j_{\overline{O}_p}(S)$ has support contained in $\overline{O}_{p-1}$ if $j \ne c_p$, and therefore $E_2^{i,j} = 0$ if $i \ne 0$ and $j \ne c_p$. Moreover, combining Proposition 6.8 with (1-13) and (3-1) we see that $H^{c_p}_{\overline{O}_p}(S) \cong Q_p$, so we have

$$E_2^{i,c_p} = H^i_{\overline{O}_{p-1}}(Q_p) \ \text{ for } i \ge 0 \quad \text{and} \quad E_2^{0,j} = H^j_{\overline{O}_p}(S) \ \text{ for } j \ne c_p. \tag{6-8}$$

It follows that the potentially nonzero groups $E_2^{i,j}$ are arranged along a hook shape centered around the point $(i,j) = (0, c_p)$, and that the only potentially nonzero maps in the spectral sequence are the homomorphisms

$$E_2^{0,c_p+r-1} = E_r^{0,c_p+r-1} \xrightarrow{d_r} E_r^{r,c_p} = E_2^{r,c_p} \quad \text{for } r \ge 2. \tag{6-9}$$

It follows that

$$E_\infty^{0,c_p+r-1} = \ker(d_r) \quad \text{and} \quad E_\infty^{r,c_p} = \operatorname{coker}(d_r) \quad \text{for } r \ge 2.$$

Since $H^k_{\overline{O}_{p-1}}(S) = 0$ for $k \equiv c_p \pmod 2$ by the case $p = m = n$ and $t = p-1$ of Corollary 3.5 it follows that $E_\infty^{0,j} = 0$ when $j \equiv c_p \pmod 2$. Since $E_2^{0,j} = 0$ for $j \not\equiv c_p \pmod 2$ by (6-8) and Corollary 3.5, we conclude that $E_\infty^{0,j} = 0$ for all $j \ge 0$, and in particular that all the maps d_r in (6-9) are injective. The vanishing of $E_\infty^{0,j}$ and the shape of the spectral sequence show that

$$E_\infty^{i,c_p} = H^{i+c_p}_{\overline{O}_{p-1}}(S) \quad \text{for all } i \ge 0, \tag{6-10}$$

and therefore we obtain short exact sequences

$$0 \to E_2^{0,c_p+r-1} \xrightarrow{d_r} E_2^{r,c_p} \to H^{r+c_p}_{\overline{O}_{p-1}}(S) \to 0.$$

Since the modules E_2^{0,c_p+r-1} and $H^{r+c_p}_{\overline{O}_{p-1}}(S)$ are in $\operatorname{add}(Q)$ by Proposition 6.8 and (6-8), it follows from Lemma 6.5 that the same is true for E_2^{r,c_p}, i.e., $H^r_{\overline{O}_{p-1}}(Q_p) \in \operatorname{add}(Q)$ for all $r \geq 2$.

Since the maps (6-9) do not involve any of the modules E_r^{i,c_p} for $i = 0, 1$, it follows that

$$H^i_{\overline{O}_{p-1}}(Q_p) \overset{(6\text{-}8)}{=} E_2^{i,c_p} = E_\infty^{i,c_p} \overset{(6\text{-}10)}{=} H^{i+c_p}_{\overline{O}_{p-1}}(S) = 0 \quad \text{for } i = 0, 1,$$

where the vanishing of $H^{i+c_p}_{\overline{O}_{p-1}}(S)$ follows from the fact that

$$c_{p-1} = (n-p+1)^2 > i + c_p = i + (n-p)^2 \quad \text{for } i = 0, 1 \text{ and } p < n,$$

proving (6-7) and concluding our proof. □

6D. ***Local cohomology of the simples D_p.*** We are now ready to finalize the proof of Theorem 6.1.

Proposition 6.11. *For every $j \geq 0$ and t, p with $0 \leq t < p \leq n$ we have $H^j_{\overline{O}_t}(D_p) \in \operatorname{add}(Q)$.*

Proof. We prove the result by descending induction on the pair $t < p$. We begin with the case when $t = p-1$ and consider the short exact sequence

$$0 \to D_p \to Q_p \to Q_{p-1} \to 0.$$

Since $H^0_{\overline{O}_{p-1}}(Q_{p-1}) = Q_{p-1}$, $H^j_{\overline{O}_{p-1}}(Q_{p-1}) = 0$ for $j > 0$ by (1-3), and $H^j_{\overline{O}_{p-1}}(Q_p) = 0$ for $j = 0, 1$ by (6-7), we obtain by the long exact sequence in cohomology that

$$H^0_{\overline{O}_{p-1}}(D_p) = 0, \quad H^1_{\overline{O}_{p-1}}(D_p) = Q_{p-1}, \quad \text{and} \quad H^j_{\overline{O}_{p-1}}(D_p) = H^j_{\overline{O}_{p-1}}(Q_p) \ \text{ for } j \geq 2. \tag{6-11}$$

It follows from Lemma 6.10 that $H^j_{\overline{O}_{p-1}}(D_p) \in \operatorname{add}(Q)$ for all $j \geq 0$. For the inductive step we consider $1 \leq t < p$ and the spectral sequence

$$E_2^{i,j} = H^i_{\overline{O}_{t-1}}(H^j_{\overline{O}_t}(D_p)) \Rightarrow H^{i+j}_{\overline{O}_{t-1}}(D_p).$$

By induction, the modules $H^j_{\overline{O}_t}(D_p)$ belong to $\operatorname{add}(Q)$, and their summands are among $Q_0, \ldots, Q_t$, since they have support contained in $\overline{O}_t$. Using the fact that for $s \leq t-1$ we have $H^0_{\overline{O}_{t-1}}(Q_s) = Q_s$ and $H^i_{\overline{O}_{t-1}}(Q_s) = 0$, together with the fact that $H^i_{\overline{O}_{t-1}}(Q_t) \in \operatorname{add}(Q)$ proved in Lemma 6.10, we conclude that each $E_2^{i,j}$ belongs to $\operatorname{add}(Q)$. Our final goal is to prove that $E_\infty^{i,j} \in \operatorname{add}(Q)$, since the modules $E_\infty^{i,j}$ constitute the composition factors of $H^{i+j}_{\overline{O}_{t-1}}(D_p)$ with respect to the filtration induced by the spectral sequence. By Lemma 6.5, this will imply that $H^k_{\overline{O}_{t-1}}(D_p) \in \operatorname{add}(Q)$ for all $k \geq 0$, concluding the inductive step.

Using Theorem 3.1 we have that $H^k_{\overline{O}_{t-1}}(D_p) = 0$ for $k \equiv p - t \pmod 2$, so we only need to consider the modules $E_\infty^{i,j}$ when $i + j \not\equiv p - t \pmod 2$. We will prove by induction on $r \geq 2$ that $E_r^{i,j}$ is a quotient of $E_2^{i,j}$ when $i+j \not\equiv p-t \pmod 2$. Since $E_{r+1} = \ker(d_r)/\mathrm{Im}(d_r)$, it suffices to check that the differentials $d_r^{i,j} : E_r^{i,j} \to E_r^{i+r,j-r+1}$ are identically 0 for $i+j \not\equiv p-t \pmod 2$.

Since $i+r \geq 2$ this is in turn is implied by the vanishing

$$E_2^{i,j} = 0 \quad \text{for } i \geq 2 \text{ and } i+j \equiv p-t \pmod 2, \tag{6-12}$$

which we explain next. Theorem 3.1 implies that $H^j_{\overline{O}_t}(D_p) = 0$ for $j \not\equiv p-t \pmod 2$, so we only need to prove (6-12) when $i \geq 2$ is even and $j \equiv p-t \pmod 2$. Since $H^i_{\overline{O}_{t-1}}(Q_s) = 0$ for $i > 0$ and $s \leq t-1$, and since $H^j_{\overline{O}_t}(D_p)$ is a direct sum of copies of $Q_0, \dots, Q_t$, it suffices to check that

$$H^i_{\overline{O}_{t-1}}(Q_t) = 0 \quad \text{for } i \text{ even},$$

which follows from Theorem 5.2, and concludes our proof. □

6E. ***Local cohomology of the indecomposables Q_p.*** The goal of this section is to prove Theorem 6.2.

Proof of Theorem 6.2. If $p = n$ then it follows from Lemma 6.9 that $H^j_{\overline{O}_t}(Q_n) = 0$ for all $0 \leq t < n$, which coincides with the formula (6-1) since $\binom{n-s-1}{n-s}_{q^2} = 0$ for all s. We may therefore assume that $t \leq n-2$, and proceed by induction on p, starting with the case $p = t+1$. Combining (6-7) with (6-11) and Theorem 6.1, we get that $H^j_{\overline{O}_{p-1}}(Q_p) \in \mathrm{add}(Q)$ for all $j \geq 0$ and moreover

$$[Q_{p-1}] \cdot q + \sum_{j \geq 0} [H^j_{\overline{O}_{p-1}}(Q_p)] \cdot q^j = \sum_{j \geq 0} [H^j_{\overline{O}_{p-1}}(D_p)] \cdot q^j = \sum_{s=0}^{p-1} [Q_s] \cdot q \cdot m_s(q^2),$$

where $m_{p-1}(q) = \binom{n-p+1}{1}_q = 1 + q + q^2 + \cdots + q^{n-p}$ and

$$m_s(q) = \binom{n-s}{p-s}_q - \binom{n-s-1}{p-s-1}_q \overset{(2\text{-}4)}{=} q^{p-s} \cdot \binom{n-s-1}{p-s}_q.$$

Using the fact that $m_{p-1}(q) - 1 = q \cdot \binom{n-p}{1}_q$, we obtain

$$\sum_{j \geq 0} [H^j_{\overline{O}_{p-1}}(Q_p)] \cdot q^j = \sum_{s=0}^{p-1} [Q_s] \cdot q^{1+2\cdot(p-s)} \cdot \binom{n-s-1}{p-s}_{q^2},$$

which agrees with (6-1) in the case when $t = p-1$.

For the induction step, we assume that $p \geq t+2$ and consider the short exact sequence

$$0 \to D_p \to Q_p \to Q_{p-1} \to 0. \tag{6-13}$$

Combining Theorem 5.2 with Theorem 6.1 we obtain

$$H^{j-1}_{\overline{O}_t}(Q_{p-1}) = H^j_{\overline{O}_t}(Q_p) = H^j_{\overline{O}_t}(D_p) = 0 \quad \text{for } j \not\equiv p-t \pmod 2.$$

Therefore, the long exact sequence in cohomology associated with (6-13) splits into short exact sequences

$$0 \to H^{j-1}_{\overline{O}_t}(Q_{p-1}) \to H^j_{\overline{O}_t}(D_p) \to H^j_{\overline{O}_t}(Q_p) \to 0. \tag{6-14}$$

Since the module $H^j_{\overline{O}_t}(Q_p)$ is a quotient of $H^j_{\overline{O}_t}(D_p)$, and the latter belongs to $\mathrm{add}(Q)$ by Proposition 6.11, it follows from Lemma 6.6 that the former also belongs to $\mathrm{add}(Q)$. It is then sufficient to verify that (6-1) holds in $\Gamma_{\mathcal{D}}[q]$. Using (6-14), Theorem 6.1, and the induction hypothesis we get

$$\begin{aligned}\sum_{j\geq 0}[H^j_{\overline{O}_t}(Q_p)]_{\mathcal{D}}\cdot q^j &= \sum_{j\geq 0}[H^j_{\overline{O}_t}(D_p)]_{\mathcal{D}}\cdot q^j - q\cdot\sum_{j\geq 0}[H^j_{\overline{O}_t}(Q_{p-1})]_{\mathcal{D}}\cdot q^j \\ &= \sum_{s=0}^{t}[Q_s]_{\mathcal{D}}\cdot\left[q^{(p-t)^2}\cdot m_s(q^2) - q\cdot q^{(p-1-t)^2+2\cdot(p-1-s)}\cdot\binom{n-s-1}{p-s-1}_{q^2}\cdot\binom{p-s-2}{p-t-2}_{q^2}\right]\end{aligned}$$

Since $1+(p-1-t)^2+2\cdot(p-1-s) = (p-t)^2+2\cdot(t-s)$, in order to prove (6-1) it suffices to check that

$$m_s(q) - q^{t-s}\cdot\binom{n-s-1}{p-s-1}_q\cdot\binom{p-s-2}{p-t-2}_q = q^{p-s}\cdot\binom{n-s-1}{p-s}_q\cdot\binom{p-s-1}{p-t-1}_q. \tag{6-15}$$

When $s=t$, we have

$$\binom{p-s-2}{p-t-2}_q = \binom{p-s-1}{p-t-1}_q = 1$$

by (2-1), so (6-15) amounts to the equality

$$\binom{n-t}{p-t}_q - \binom{n-t-1}{p-t-1}_q = q^{p-t}\cdot\binom{n-t-1}{p-t}_q$$

which follows from (2-4). When $s<t$ we get

$$\binom{p-s-2}{p-t-2}_q = \binom{p-s-2}{t-s}_q \quad\text{and}\quad \binom{p-s-1}{p-t-1}_q = \binom{p-s-1}{t-s}_q$$

using (2-1), so we can rewrite (6-15) as

$$\binom{p-s-1}{t-s}_q\cdot\left[\binom{n-s}{p-s}_q - q^{p-s}\cdot\binom{n-s-1}{p-s}_q\right] = \left[q^{t-s}\cdot\binom{p-s-2}{t-s}_q + \binom{p-s-2}{t-s-1}_q\right]\cdot\binom{n-s-1}{p-s-1}_q$$

which follows by applying (2-4) to both sides of the equation. □

6F. ***The proof of Theorem 1.5.*** If $p=n-1$ then $\overline{O}_{n-1}$ is a hypersurface so its only nonzero Lyubeznik number is $\lambda_{n^2-1,n^2-1}(R^{(n-1)})=1$. We assume that $p\leq n-2$ and get as in Section 3E that

$$\begin{aligned}L_p(q,w) &= \sum_{i,j\geq 0}\langle H^i_{O_0}(H^{n^2-j}_{\overline{O}_p}(S)), D_0\rangle_{\mathcal{D}}\cdot q^i\cdot w^j \\ &= \sum_{i\geq 0}\left[\sum_{s=0}^{p}\langle H^i_{O_0}(Q_s), D_0\rangle\cdot q^i\cdot\left(\sum_{j\geq 0}\langle H^{n^2-j}_{\overline{O}_p}(S), D_s - D_{s+1}\rangle_{\mathcal{D}}\cdot w^j\right)\right],\end{aligned}$$

where we used the fact that the groups $H^{n^2-j}_{\overline{O}_p}(S)$ belong to $\mathrm{add}(Q)$, and that the multiplicity of Q_s as a summand in $M \in \mathrm{add}(Q)$ can be computed using (1-13) by the formula $\langle M, D_s - D_{s+1}\rangle_{\mathcal{D}}$. We obtain that

$$L_p(q,w) = \sum_{s=0}^{p} \langle H_0^{\mathcal{D}}(Q_s;q), D_0\rangle_{\mathcal{D}} \cdot \langle H_p^{\mathcal{D}}(S;w^{-1}), D_s - D_{s+1}\rangle_{\mathcal{D}} \cdot w^{n^2}$$

$$\overset{\text{(6-1),(2-10)}}{=} \sum_{s=0}^{p} q^{s^2+2s} \cdot \binom{n-1}{s}_{q^2} \cdot w^{-(n-p)^2} \cdot \left[\binom{n-1-s}{p-s}_{w^{-2}} - \binom{n-2-s}{p-s-1}_{w^{-2}} \right] \cdot w^{n^2}.$$

Using (2-4) we have

$$\binom{n-1-s}{p-s}_{w^{-2}} - \binom{n-2-s}{p-s-1}_{w^{-2}} = w^{-2\cdot(p-s)} \cdot \binom{n-2-s}{p-s},$$

and combining this with (2-3) it follows that in order to prove (1-8) it suffices to verify the identity

$$p^2 + 2p + s\cdot(2n-2p-2) = -(n-p)^2 - 2\cdot(p-s) - 2(p-s)\cdot(n-2-p) + n^2,$$

which follows again by inspection.

Acknowledgements

We are grateful to the organizers of the conference "Local Cohomology in Commutative Algebra and Algebraic Geometry" at University of Minnesota, where this work was initiated. We would like to thank Mike Perlman for correcting some of the formulas in Section 4B, and Uli Walther for fruitful conversations regarding Lyubeznik numbers. We would like to thank the anonymous referees for many helpful suggestions for improving the manuscript. Raicu acknowledges the support of the Alfred P. Sloan Foundation, and of the National Science Foundation Grant No. 1901886.

References

[Núñez Betancourt et al. 2016] L. Núñez Betancourt, E. E. Witt, and W. Zhang, "A survey on the Lyubeznik numbers", pp. 137–163 in *Mexican mathematicians abroad: recent contributions* (Guanajuato, Mexico, 2012), edited by N. Bárcenas et al., Contemp. Math. **657**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[Braden and Grinberg 1999] T. Braden and M. Grinberg, "Perverse sheaves on rank stratifications", *Duke Math. J.* **96**:2 (1999), 317–362. MR Zbl

[Bruns and Vetter 1988] W. Bruns and U. Vetter, *Determinantal rings*, Lecture Notes in Math. **1327**, Springer, 1988. MR Zbl

[Eisenbud 2005] D. Eisenbud, *The geometry of syzygies*, Grad. Texts in Math. **229**, Springer, 2005. MR Zbl

[García López and Sabbah 1998] R. García López and C. Sabbah, "Topological computation of local cohomology multiplicities", *Collect. Math.* **49**:2-3 (1998), 317–324. MR Zbl

[Hartshorne 1966] R. Hartshorne, *Residues and duality*, Lecture Notes in Math. **20**, Springer, 1966. MR Zbl

[Lőrincz and Walther 2019] A. C. Lőrincz and U. Walther, "On categories of equivariant $\mathscr{D}$-modules", *Adv. Math.* **351** (2019), 429–478. MR Zbl

[Lőrincz et al. 2019] A. C. Lőrincz, C. Raicu, and J. Weyman, "Equivariant $\mathscr{D}$-modules on binary cubic forms", *Comm. Algebra* **47**:6 (2019), 2457–2487. MR Zbl

[Lyubeznik 1993] G. Lyubeznik, "Finiteness properties of local cohomology modules (an application of D-modules to commutative algebra)", *Invent. Math.* **113**:1 (1993), 41–55. MR Zbl

[Lyubeznik et al. 2016] G. Lyubeznik, A. K. Singh, and U. Walther, "Local cohomology modules supported at determinantal ideals", *J. Eur. Math. Soc.* **18**:11 (2016), 2545–2578. MR Zbl

[MacPherson and Vilonen 1986] R. MacPherson and K. Vilonen, "Elementary construction of perverse sheaves", *Invent. Math.* **84**:2 (1986), 403–435. MR Zbl

[Perlman 2020] M. Perlman, "Lyubeznik numbers for Pfaffian rings", *J. Pure Appl. Algebra* **224**:5 (2020), art. id. 106247. MR Zbl

[Raicu 2016] C. Raicu, "Characters of equivariant $\mathscr{D}$-modules on spaces of matrices", *Compos. Math.* **152**:9 (2016), 1935–1965. MR Zbl

[Raicu 2017] C. Raicu, "Homological invariants of determinantal thickenings", *Bull. Math. Soc. Sci. Math. Roumanie* (*N.S.*) **60(108)**:4 (2017), 425–446. MR Zbl

[Raicu 2018] C. Raicu, "Regularity and cohomology of determinantal thickenings", *Proc. Lond. Math. Soc.* (3) **116**:2 (2018), 248–280. MR Zbl

[Raicu and Weyman 2014] C. Raicu and J. Weyman, "Local cohomology with support in generic determinantal ideals", *Algebra Number Theory* **8**:5 (2014), 1231–1257. MR Zbl

[Raicu and Weyman 2016] C. Raicu and J. Weyman, "Local cohomology with support in ideals of symmetric minors and Pfaffians", *J. Lond. Math. Soc.* (2) **94**:3 (2016), 709–725. MR Zbl

[Raicu and Weyman 2017] C. Raicu and J. Weyman, "The syzygies of some thickenings of determinantal varieties", *Proc. Amer. Math. Soc.* **145**:1 (2017), 49–59. MR Zbl

[Reichelt and Walther $\geq$ 2020] T. Reichelt and U. Walther, "Projective varieties with intrinsic Lyubeznik numbers", in preparation.

[Reichelt et al. 2018] T. Reichelt, M. Saito, and U. Walther, "Dependence of Lyubeznik numbers of cones of projective schemes on projective embeddings", preprint, 2018. arXiv

[Strickland 1982] E. Strickland, "On the conormal bundle of the determinantal variety", *J. Algebra* **75**:2 (1982), 523–537. MR Zbl

[Switala 2015] N. Switala, "Lyubeznik numbers for nonsingular projective varieties", *Bull. Lond. Math. Soc.* **47**:1 (2015), 1–6. MR Zbl

[Vilonen 1994] K. Vilonen, "Perverse sheaves and finite-dimensional algebras", *Trans. Amer. Math. Soc.* **341**:2 (1994), 665–676. MR Zbl

[Walther 1999] U. Walther, "Algorithmic computation of local cohomology modules and the local cohomological dimension of algebraic varieties", *J. Pure Appl. Algebra* **139**:1-3 (1999), 303–321. MR Zbl

[Walther 2001] U. Walther, "On the Lyubeznik numbers of a local ring", *Proc. Amer. Math. Soc.* **129**:6 (2001), 1631–1634. MR Zbl

[Wang 2020] B. Wang, "Lyubeznik numbers of irreducible projective varieties depend on the embedding", *Proc. Amer. Math. Soc.* **148**:5 (2020), 2091–2096. MR Zbl

[Weyman 2003] J. Weyman, *Cohomology of vector bundles and syzygies*, Cambridge Tracts in Math. **149**, Cambridge Univ. Press, 2003. MR Zbl

Communicated by Bhargav Bhatt
Received 2019-11-20 Revised 2020-03-27 Accepted 2020-04-29

lorincz@mis.mpg.de *Department of Mathematics, Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany*

craicu@nd.edu *Department of Mathematics, University of Notre Dame, Notre Dame, IN, United States*

Institute of Mathematics "Simion Stoilow" of the Romanian Academy, Bucharest, Romania

ALGEBRA AND NUMBER THEORY 14:9 (2020)
https://doi.org/10.2140/ant.2020.14.2571

On asymptotic Fermat over $\mathbb{Z}_p$-extensions of $\mathbb{Q}$

Nuno Freitas, Alain Kraus and Samir Siksek

Let p be a prime and let $\mathbb{Q}_{n,p}$ denote the n-th layer of the cyclotomic $\mathbb{Z}_p$-extension of $\mathbb{Q}$. We prove the effective asymptotic FLT over $\mathbb{Q}_{n,p}$ for all $n \geq 1$ and all primes $p \geq 5$ that are non-Wieferich, i.e., $2^{p-1} \not\equiv 1 \pmod{p^2}$. The effectivity in our result builds on recent work of Thorne proving modularity of elliptic curves over $\mathbb{Q}_{n,p}$.

1. Introduction

Let F be a totally real number field. The *asymptotic Fermat's last theorem over* F is the statement that there exists a constant B_F, depending only on F, such that, for all primes $\ell > B_F$, the only solutions to the equation $x^\ell + y^\ell + z^\ell = 0$, with $x, y, z \in F$ are the trivial ones satisfying $xyz = 0$. If B_F is effectively computable, we refer to this as the *effective asymptotic Fermat's last theorem over* F. Let p be a prime, n a positive integer and write $\mathbb{Q}_{n,p}$ for the n-th layer of the cyclotomic $\mathbb{Z}_p$-extension. In [Freitas et al. 2020], the authors established the following theorem.

Theorem 1. *The effective asymptotic Fermat's last theorem holds over each layer* $\mathbb{Q}_{n,2}$ *of the cyclotomic* $\mathbb{Z}_2$*-extension.*

The proof of Theorem 1 relies heavily on class field theory and the theory of 2-extensions, and the method depends crucially on the fact that 2 is totally ramified in $\mathbb{Q}_{n,2}$. We establish the following.

Theorem 2. *Let* $p \geq 5$ *be a prime. Suppose* p *is non-Wieferich, i.e.,* $2^{p-1} \not\equiv 1 \pmod{p^2}$. *The effective asymptotic Fermat's last theorem holds over each layer* $\mathbb{Q}_{n,p}$ *of the cyclotomic* $\mathbb{Z}_p$*-extension of* $\mathbb{Q}$.

We remark that the only Wieferich primes currently known are 1093 and 3511. It is fascinating to observe that these primes originally arose in connection with historical attempts at proving Fermat's last theorem. Indeed Wieferich [1909] showed that if $2^{p-1} \not\equiv 1 \pmod{p^2}$ then the first case of Fermat's last theorem holds for exponent p.

In contrast to Theorem 1, the proof of Theorem 2 makes use of a criterion (Theorem 3 below) established in [Freitas and Siksek 2015] for asymptotic FLT in terms of solutions to a certain S-unit equation. The proof of that criterion builds on many deep results including modularity lifting theorems due to Breuil, Diamond, Gee, Kisin, and others, and Merel's uniform boundedness theorem, and exploits the strategy of Frey, Serre, Ribet, Wiles and Taylor, utilized in Wiles' proof of Fermat's last theorem [1995]. We use elementary

Freitas is supported by a Ramón y Cajal fellowship (RYC-2017-22262). Siksek is supported by EPSRC grant "Moduli of Elliptic curves and Classical Diophantine Problems" (EP/S031537/1).
MSC2020: primary 11D41; secondary 11R23.
Keywords: Fermat, unit equation, $\mathbb{Z}_p$-extensions.

arguments to study these S-unit equations in $\mathbb{Q}_{n,p}$ and this study, together with the S-unit criterion, quickly yields Theorem 2. The effectivity in Theorem 2 builds on the following great theorem due to Thorne [2019].

Theorem (Thorne). *Elliptic curves over $\mathbb{Q}_{n,p}$ are modular.*

2. An S-unit criterion for asymptotic FLT

The following criterion for asymptotic FLT is a special case of [Freitas and Siksek 2015, Theorem 3].

Theorem 3. *Let F be a totally real number field. Suppose the Eichler–Shimura conjecture over F holds. Assume that 2 is inert in F and write $\mathfrak{q} = 2\mathcal{O}_F$ for the prime ideal above 2. Let $S = \{\mathfrak{q}\}$ and write $\mathcal{O}_S^\times$ for the group of S-units in F. Suppose every solution (λ, μ) to the S-unit equation*

$$\lambda + \mu = 1, \qquad \lambda, \mu \in \mathcal{O}_S^\times \tag{2-1}$$

satisfies both of the following conditions:

$$\max\{|\mathrm{ord}_{\mathfrak{q}}(\lambda)|, |\mathrm{ord}_{\mathfrak{q}}(\mu)|\} \leq 4, \qquad \mathrm{ord}_{\mathfrak{q}}(\lambda\mu) \equiv 1 \pmod 3. \tag{2-2}$$

Then the asymptotic Fermat's last theorem holds over F. Moreover, *if all elliptic curves over F with full 2-torsion are modular*, *then the effective asymptotic Fermat's last theorem holds over F.*

For a discussion of the Eichler–Shimura conjecture see [Freitas and Siksek 2015, Section 2.4], but for the purpose of this paper we note that the conjecture is known to hold for all totally real fields of odd degree. In particular, it holds for $\mathbb{Q}_{n,p}$ for all odd p.

To apply Theorem 3 to $F = \mathbb{Q}_{n,p}$ we need to know for which p is 2 inert in F. The answer is given by the following lemma, which for $n = 1$ is [Washington 1997, Exercise 2.4].

Lemma 2.1. *Let $p \geq 3$, q be distinct primes. Then q is inert in $\mathbb{Q}_{n,p}$ if and only if $q^{p-1} \not\equiv 1 \pmod{p^2}$.*

Proof. Let $L = \mathbb{Q}(\zeta_{p^{n+1}})$ and $F = \mathbb{Q}_{n,p}$. Write σ_q and τ_q for the Frobenius elements corresponding to q in $\mathrm{Gal}(L/\mathbb{Q})$ and $\mathrm{Gal}(F/\mathbb{Q})$. The prime q is inert in F precisely when τ_q has order p^n. The natural surjection $\mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Gal}(F/\mathbb{Q})$ sends σ_q to τ_q and its kernel has order $p-1$. Thus q is inert in F if and only if the order of σ_q is divisible by p^n, which is equivalent to σ_q^{p-1} having order p^n. There is a canonical isomorphism $\mathrm{Gal}(L/\mathbb{Q}) \to (\mathbb{Z}/p^{n+1}\mathbb{Z})^\times$ sending σ_q to $q + p^{n+1}\mathbb{Z}$. Thus q is inert in F if and only if $q^{p-1} + p^{n+1}\mathbb{Z}$ has order p^n. This is equivalent to $q^{p-1} \not\equiv 1 \pmod{p^2}$. □

3. Proof of Theorem 2

Lemma 3.1. *Let $\mathfrak{p}$ be the unique prime above p in $F = \mathbb{Q}_{n,p}$. Let $\lambda \in \mathcal{O}_F$. Then $\lambda \equiv \mathrm{Norm}_{F/\mathbb{Q}}(\lambda) \pmod{\mathfrak{p}}$.*

Proof. As p is totally ramified in F, we know that the residue field $\mathcal{O}_F/\mathfrak{p}$ is $\mathbb{F}_p$. Thus there is some $a \in \mathbb{Z}$ such that $\lambda \equiv a \pmod{\mathfrak{p}}$. Let $\sigma \in G = \mathrm{Gal}(F/\mathbb{Q})$. Since $\mathfrak{p}^\sigma = \mathfrak{p}$, we have $\lambda^\sigma \equiv a \pmod{\mathfrak{p}}$. Hence

$$\mathrm{Norm}_{F/\mathbb{Q}}(\lambda) = \prod_{\sigma \in G} \lambda^\sigma \equiv a^{\#G} \pmod{\mathfrak{p}}.$$

However $\#G = p^n$ so $\mathrm{Norm}_{F/\mathbb{Q}}(\lambda) \equiv a \equiv \lambda \pmod{\mathfrak{p}}$. □

Lemma 3.2. *Let $p \neq 3$ be a rational prime. Let $F = \mathbb{Q}_{n,p}$. Then the unit equation*

$$\lambda + \mu = 1, \quad \lambda, \mu \in \mathcal{O}_F^\times \tag{3-1}$$

has no solutions.

Proof. Let (λ, μ) be a solution to (3-1). By Lemma 3.1, $\lambda \equiv \pm 1 \pmod{\mathfrak{p}}$ and $\mu \equiv \pm 1 \pmod{\mathfrak{p}}$. Thus $\pm 1 \pm 1 \equiv \lambda + \mu = 1 \pmod{\mathfrak{p}}$. This is impossible as $p \neq 3$. □

Remark 3.3. Lemma 3.2 is false for $p = 3$. Indeed, let $p = 3$ and $n = 1$. Then $F = \mathbb{Q}_{1,3} = \mathbb{Q}(\theta)$, where θ satisfies $\theta^3 - 6\theta^2 + 9\theta - 3 = 0$. The unit equation has solution $\lambda = 2 - \theta$ and $\mu = -1 + \theta$. In fact, the unit equation solver of the computer algebra system `Magma` [Bosma et al. 1997] gives a total of 18 solutions.

Lemma 3.4. *Let $p \geq 5$ be a rational prime. Let $F = \mathbb{Q}_{n,p}$. Suppose 2 is inert in F and write $\mathfrak{q} = 2\mathcal{O}_F$ for the unique prime above 2. Let $S = \{\mathfrak{q}\}$ and write $\mathcal{O}_S^\times$ for the group of S-units. Then every solution to the S-unit equation* (2-1) *satisfies one of the following*:

(i) $\mathrm{ord}_{\mathfrak{q}}(\lambda) = 1$, $\mathrm{ord}_{\mathfrak{q}}(\mu) = 0$;

(ii) $\mathrm{ord}_{\mathfrak{q}}(\lambda) = 0$, $\mathrm{ord}_{\mathfrak{q}}(\mu) = 1$;

(iii) $\mathrm{ord}_{\mathfrak{q}}(\lambda) = \mathrm{ord}_{\mathfrak{q}}(\mu) = -1$.

Proof. Write $n_\lambda = \mathrm{ord}_{\mathfrak{q}}(\lambda)$ and $n_\mu = \mathrm{ord}_{\mathfrak{q}}(\mu)$. Suppose first $n_\lambda \geq 2$. Then $n_\mu = 0$ and so $\mu \in \mathcal{O}_F^\times$. Moreover, as $4 \mid \lambda$, we have $\mu \equiv 1 \pmod 4$ and so $\mu^\sigma \equiv 1 \pmod 4$ for all $\sigma \in G = \mathrm{Gal}(F/\mathbb{Q})$. Hence $\mathrm{Norm}_{F/\mathbb{Q}}(\mu) = \prod \mu^\sigma \equiv 1 \pmod 4$. But $\mathrm{Norm}_{F/\mathbb{Q}}(\mu) = \pm 1$, thus $\mathrm{Norm}_{F/\mathbb{Q}}(\mu) = 1$. As before, denote the unique prime above p by $\mathfrak{p}$. By Lemma 3.1 we have $\mu \equiv 1 \pmod{\mathfrak{p}}$. Hence $\mathfrak{p}$ divides $1 - \mu = \lambda$ giving a contradiction.

Thus $n_\lambda \leq 1$. Next suppose $n_\lambda \leq -2$. Then $n_\lambda = n_\mu$. Let $\lambda' = 1/\lambda$ and $\mu' = -\mu/\lambda$. Then (λ', μ') is a solution to the S-unit equation satisfying $n_{\lambda'} \geq 2$, giving a contradiction by the previous case. Hence $-1 \leq n_\lambda \leq 1$ and by symmetry $-1 \leq n_\mu \leq 1$. From Lemma 3.2 either $n_\lambda \neq 0$ or $n_\mu \neq 0$. Thus one of (i), (ii), (iii) must hold. □

Remark 3.5. Possibilities (i), (ii), (iii) cannot be eliminated because of the solutions $(2, -1)$, $(-1, 2)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$ to the S-unit equation.

Proof of Theorem 2. We suppose $p \geq 5$ and non-Wieferich. It follows from Lemma 2.1 that 2 is inert in $F = \mathbb{Q}_{n,p}$. Write $\mathfrak{q} = 2\mathcal{O}_F$. By Lemma 3.4 all solutions (λ, μ) to the S-unit equation (2-1) satisfy (2-2). We now apply Theorem 3. As elliptic curves over $\mathbb{Q}_{n,p}$ are modular thanks to Thorne's theorem, we conclude that the effective Fermat's last theorem holds over $\mathbb{Q}_{n,p}$. □

Remark 3.6. The proof of Theorem 2 for $p = 3$ and for the Wieferich primes seems out of reach at present. There are solutions to the unit equation in $\mathbb{Q}_{1,3}$ (as indicated in Remark 3.3), and therefore in $\mathbb{Q}_{n,3}$ for all n, and these solutions violate the criterion of Theorem 3. For p a Wieferich prime, 2 splits in $\mathbb{Q}_{n,p}$ into at least p prime ideals and we would need to consider the S-unit equation (2-1) with S the set of primes above 2. It appears difficult to treat the S-unit equation in infinite families of number fields where $\#S \geq 2$ (see [Freitas et al. 2020, Theorem 7] and its proof).

4. A generalization

In fact, the proof of Theorem 2 establishes the following more general theorem.

Theorem 4. *Let F be a totally real number field and $p \geq 5$ be a rational prime. Suppose that the following conditions are satisfied.*

(a) *F is a p-extension of $\mathbb{Q}$ (i.e., $F/\mathbb{Q}$ is a Galois extension of degree p^n for some $n \geq 1$).*

(b) *p is totally ramified in F.*

(c) *2 is inert in F.*

Then the asymptotic Fermat's last theorem holds for F.

Example 4.1. A quick search on the L-Functions and Modular Forms Database [LMFDB 2020] yields 153 fields of degree 5 satisfying conditions of the theorem with $p = 5$. The one with smallest discriminant is $\mathbb{Q}_{1,5}$. The one with the next smallest discriminant is $F = \mathbb{Q}(\theta)$ where $\theta^5 - 110\theta^3 - 605\theta^2 - 990\theta - 451 = 0$. The discriminant of F is $5^8 \cdot 11^4$. It is therefore not contained in any $\mathbb{Z}_p$-extension of $\mathbb{Q}$.

References

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. See also Magma computational algebra system. MR Zbl

[Freitas and Siksek 2015] N. Freitas and S. Siksek, "The asymptotic Fermat's last theorem for five-sixths of real quadratic fields", *Compos. Math.* **151**:8 (2015), 1395–1415. MR Zbl

[Freitas et al. 2020] N. Freitas, A. Kraus, and S. Siksek, "Class field theory, Diophantine analysis and the asymptotic Fermat's last theorem", *Adv. Math.* **363** (2020), art. id. 106964. MR Zbl

[LMFDB 2020] The LMFDB Collaboration, "The L-functions and modular forms database", 2020, http://www.lmfdb.org.

[Thorne 2019] J. A. Thorne, "Elliptic curves over $\mathbb{Q}_\infty$ are modular", *J. Eur. Math. Soc.* **21**:7 (2019), 1943–1948. MR Zbl

[Washington 1997] L. C. Washington, *Introduction to cyclotomic fields*, 2nd ed., Grad. Texts in Math. **83**, Springer, 1997. MR Zbl

[Wieferich 1909] A. Wieferich, "Zum letzten Fermatschen Theorem", *J. Reine Angew. Math.* **136** (1909), 293–302. MR Zbl

[Wiles 1995] A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141**:3 (1995), 443–551. MR Zbl

Communicated by Andrew Granville
Received 2020-04-02 Accepted 2020-05-11

nunobfreitas@gmail.com — *Department de Matemàtiques i Informàtica, Universitat de Barcelona, Barcelona, Spain*

alain.kraus@imj-prg.fr — *Institut de Mathématiques de Jussieu - Paris Rive Gauche, Sorbonne Université, UMR 7586 CNRS - Paris Diderot, Paris, France*

s.siksek@warwick.ac.uk — *Mathematics Institute, University of Warwick, Coventry, United Kingdom*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 14 No. 9 2020

1937-0652(2020)14:9;1-T