Algebra & Number Theory

Volume 20
2026
No. 6

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2026 is US $590/year for the electronic version, and $865/year (+$75, 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, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published continuously online.

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 20:6 (2026)
https://doi.org/10.2140/ant.2026.20.1073

Effective multiplicative independence of three singular moduli

Yuri Bilu, Sanoli Gun and Emanuele Tron

Pila and Tsimerman proved in 2017 that for every k there exist at most finitely many k-tuples $(x_1, \dots, x_k)$ of nonzero singular moduli such that $x_1, \dots, x_k$ are multiplicatively dependent, but any proper subset of them is multiplicatively independent. The proof was noneffective, using Siegel's lower bound for the class numbers. In 2019 Riffaut obtained an effective version of this result for $k = 2$. Moreover, he determined all the instances of $x^m y^n \in \mathbb{Q}^\times$, where x, y are distinct singular moduli and m, n are nonzero integers. In this article we obtain a similar result for $k = 3$. We show that $x^m y^n z^r \in \mathbb{Q}^\times$ (where x, y, z are distinct singular moduli and m, n, r nonzero integers) implies that the discriminants of x, y, z do not exceed 10^{10}.

1. Introduction

A *singular modulus* is the j-invariant of an elliptic curve with complex multiplication. Given a singular modulus x we denote by Δ_x the discriminant of the associated imaginary quadratic order. We denote by $h(\Delta)$ the class number of the imaginary quadratic order of discriminant Δ.

Recall that two singular moduli x and y are conjugate over $\mathbb{Q}$ if and only if $\Delta_x = \Delta_y$, and that there are $h(\Delta)$ singular moduli of a given discriminant Δ. In particular, $[\mathbb{Q}(x) : \mathbb{Q}] = h(\Delta_x)$. For details see, for instance, [10, §7 and §11].

There has been much work on diophantine properties of singular moduli in recent years. In particular, studying algebraic equations where the unknowns are singular moduli [2; 6; 7] is interesting by virtue of its connection with the André–Oort property for affine space [5; 16; 17].

Bilu and Gun were supported by SPARC Project P445 (India). Bilu and Tron were supported by ANR project JINVARIANT.
MSC2020: primary 11G15; secondary 11G18.
Keywords: singular moduli.

Pila and Tsimerman [20] proved that for every k there exists at most finitely many k-tuples $(x_1, \dots, x_k)$ of nonzero singular moduli such that $x_1, \dots, x_k$ are multiplicatively dependent, but any proper subset of them is multiplicatively independent. Their argument is fundamentally noneffective.

Riffaut [21, Theorem 1.7] gave an effective (and totally explicit) version of the theorem of Pila and Tsimerman in the case $k = 2$. He in fact classified all cases when $x^m y^n \in \mathbb{Q}^\times$, where x, y are singular moduli and m, n nonzero integers.

Here we obtain an effective result for $k = 3$. As Riffaut did, we prove a stronger statement: we bound explicitly discriminants of singular moduli x, y, z such that $x^m y^n z^r \in \mathbb{Q}^\times$ for some nonzero integers m, n, r. Our bound is as follows.

Theorem 1.1. *Let x, y, z be distinct nonzero singular moduli and m, n, r nonzero integers. Assume that $x^m y^n z^r \in \mathbb{Q}^\times$. Then*

$$\max\{|\Delta_x|, |\Delta_y|, |\Delta_z|\} < 10^{10}.$$

The special case $m = n = r$ has been recently settled by Fowler [13; 14].

There do exist triples of distinct singular moduli x, y, z such that $x^m y^n z^r \in \mathbb{Q}^\times$ for some nonzero $m, n, r \in \mathbb{Z}$. There are three types of currently known examples.

Rational type: Take distinct x, y, z such that

$$h(\Delta_x) = h(\Delta_y) = h(\Delta_z) = 1, \quad \Delta_x, \Delta_y, \Delta_z \neq -3.$$

Then $x, y, z \in \mathbb{Q}^\times$ and $x^m y^n z^r \in \mathbb{Q}^\times$ for any choice of m, n, r. Pila and Tsimerman [20, Example 6.2] even found an example of $x^m y^n z^r = 1$:

$$(2^6 3^3)^{10} (-2^{15})^6 (-2^{15} 3^3)^{-10} = 1,$$

the corresponding discriminants being -4, -11 and -19.

Quadratic type: Take distinct x, y, z such that

$$h(\Delta_x) = 1, \quad \Delta_x \neq -3, \quad \Delta_y = \Delta_z, \quad h(\Delta_y) = h(\Delta_z) = 2.$$

Then $x \in \mathbb{Q}^\times$ and y, z are of degree 2, conjugate over $\mathbb{Q}$. Hence $x^m y^n z^n \in \mathbb{Q}^\times$ for any choice of m, n.

Cubic type: Take distinct x, y, z such that

$$\Delta_x = \Delta_y = \Delta_z, \quad h(\Delta_x) = h(\Delta_y) = h(\Delta_z) = 3.$$

Then x, y, z are of degree 3, forming a full Galois orbit over $\mathbb{Q}$. Hence $xyz \in \mathbb{Q}^\times$.

We believe that, up to permuting x, y, z, there are no other examples, but to justify it, one needs to improve on the numerical bound 10^{10} in Theorem 1.1.

The proof of Theorem 1.1 relies on the following result, which is a partial common generalization (for big discriminants) of [21, Theorem 1.7] and [12, Theorem 1.3].

Theorem 1.2. *Let x, y be distinct nonzero singular moduli and m, n nonzero integers. Assume that*

$$\max\{|\Delta_x|, |\Delta_y|\} \geq 10^8. \tag{1-1}$$

Then $[\mathbb{Q}(x, y) : \mathbb{Q}(x^m y^n)] \leq 2$. More precisely, we have either

$$\mathbb{Q}(x^m y^n) = \mathbb{Q}(x, y) \tag{1-2}$$

or

$$\Delta_x = \Delta_y, \quad m = n, \quad [\mathbb{Q}(x, y) : \mathbb{Q}(x^m y^m)] = 2. \tag{1-3}$$

Moreover, in the latter case x and y are conjugate over the field $\mathbb{Q}(x^m y^m)$.

If $\{\Delta_x, \Delta_y\}$ is not of the form $\{\Delta, 4\Delta\}$ for some $\Delta \equiv 1 \bmod 8$, then condition (1-1) *can be relaxed to*

$$\max\{|\Delta_x|, |\Delta_y|\} \geq 10^6. \tag{1-4}$$

Plan of the article. In Section 2 we collect basic fact about singular moduli to be used throughout the article. In Section 3 we establish our principal tool: a linear relation between the exponents $m_1, \dots, m_k$ stemming from the multiplicative relation $x_1^{m_1} \cdots x_k^{m_k} = 1$. Theorems 1.2 and 1.1 are proved in Sections 4 and 5, respectively.

Notation and conventions. We denote by $\mathbb{H}$ the Poincaré half-plane, and by $\mathcal{F}$ the standard fundamental domain for the action of the modular group; that is, the open hyperbolic triangle with vertices

$$\zeta_6 = \frac{1+\sqrt{-3}}{2}, \quad \zeta_3 = \frac{-1+\sqrt{-3}}{2}, \quad i\infty,$$

together with the hyperbolic geodesics $[i, \zeta_6]$ and $[\zeta_6, i\infty]$.

We denote by log the principal branch of the complex logarithm:

$$-\pi < \arg \log z \leq \pi \quad (z \in \mathbb{C}^\times).$$

We use $O_1(\cdot)$ as a quantitative version of the $O(\cdot)$ notation: $A = O_1(B)$ means that $|A| \leq B$.

We write the Galois action exponentially: $x \mapsto x^\sigma$. In particular, it is a right action: $x^{(\sigma_1\sigma_2)} = (x^{\sigma_1})^{\sigma_2}$. Most of the Galois groups occurring in this article are abelian, so this is not relevant, but in the few cases where the group is not abelian one must be vigilant.

Let R be a commutative ring, and $a \in R$. When this does not lead to confusion, we write R/a instead of R/aR.

We denote by $\mathcal{C}_m$ the cyclic group of order m.

In cross-references, item Y of Proposition X is quoted as Proposition X:Y.

2. Class numbers, denominators, isogenies

Unless the contrary is stated explicitly, the letter Δ stands for an *imaginary quadratic discriminant*, that is, $\Delta < 0$ and $\Delta \equiv 0, 1 \bmod 4$.

We denote by $\mathcal{O}_\Delta$ the imaginary quadratic order of discriminant Δ, that is,

$$\mathcal{O}_\Delta = \mathbb{Z}[(\Delta + \sqrt{\Delta})/2].$$

Then $\Delta = Df^2$, where D is the discriminant of the number field $K = \mathbb{Q}(\sqrt{\Delta})$, called the *fundamental discriminant* of Δ, and $f = [\mathcal{O}_D : \mathcal{O}_\Delta]$ is the *conductor* of Δ.

We denote by $h(\Delta)$ the class number of $\mathcal{O}_\Delta$.

Given a singular modulus x, we denote by Δ_x the discriminant of the associated CM order, and we write $\Delta_x = D_x f_x^2$ with D_x the fundamental discriminant and f_x the conductor. We denote by K_x the associated imaginary quadratic field

$$K_x = \mathbb{Q}(\sqrt{D_x}) = \mathbb{Q}(\sqrt{\Delta_x}).$$

We will call K_x the *CM field* of the singular modulus x.

It is known (see, for instance, §11 in [10]) that a singular modulus x is an algebraic integer of degree $h(\Delta_x)$, and that there are exactly $h(\Delta)$ singular moduli of given discriminant Δ, which form a full Galois orbit over $\mathbb{Q}$.

2.1. ***Class numbers and class groups.*** For a discriminant Δ and a positive integer ℓ set

$$\Psi(\ell, \Delta) = \ell \prod_{p \mid \ell} \left(1 - \frac{(\Delta/p)}{p}\right), \tag{2-1}$$

where (Δ/p) denotes the Kronecker symbol. It is useful to note that

$$\Psi(\ell, \Delta) \geq \varphi(\ell), \tag{2-2}$$

where $\varphi(\,\cdot\,)$ is Euler's totient function. Note also the multiplicativity relation

$$\Psi(\ell_1\ell_2, \Delta) = \Psi(\ell_2, \Delta\ell_1^2)\Psi(\ell_1, \Delta). \tag{2-3}$$

Recall the *class number formula*

$$h(\Delta\ell^2) = \frac{1}{[\mathcal{O}_\Delta^\times : \mathcal{O}_{\Delta\ell^2}^\times]} h(\Delta)\Psi(\ell, \Delta). \tag{2-4}$$

In [10, Theorem 7.24] this formula is proved in the case when $\Delta = D$ is a fundamental discriminant; the general case easily follows using the multiplicativity relation (2-3). Note also that

$$[\mathcal{O}_\Delta^\times : \mathcal{O}_{\Delta\ell^2}^\times] = \begin{cases} 3 & \text{if } \Delta = -3,\, \ell > 1, \\ 2 & \text{if } \Delta = -4,\, \ell > 1, \\ 1 & \text{if } \Delta \neq -3, -4. \end{cases} \tag{2-5}$$

2.1.1. *Discriminants with small class number.* Watkins [24] classified fundamental discriminants D with $h(D) \leq 100$ and found that such discriminants do not exceed 2383747 in absolute value. It turns out that the same upper bound holds true for all discriminants, not only fundamental ones.

n	1	2	3	4	5	6	7	8	10	12	16	25	50	100
max f	420	210	120	90	66	60	42	42	30	30	18	12	6	2
$D_{\max}(n)$	163	427	907	1555	2683	3763	5923	6307	13843	17803	34483	111763	462883	2383747

Table 1. Values of $D_{\max}(n)$ for arguments occurring in (2-7). The first row contains all positive integers n of the form $\lfloor 100/\varphi(f)\rfloor$ for some positive integer f. In the second row, for each n we give the biggest f with the property $100/\varphi(f)\geq n$. Finally, the third row displays $D_{\max}(n)$.

Proposition 2.1. *Let Δ be a negative discriminant with $h(\Delta)\leq 100$. Then we have $|\Delta|\leq 2383747$. If $h(\Delta)\leq 64$, then $|\Delta|\leq 991027$.*

Remark 2.2. As Guy Fowler informed us, Janis Klaise obtained the same result, with a similar proof, in his 2012 Master's thesis [15]. Apparently, this work has not been published.

Proof. Given a positive integer n, denote by

$$D_{\max}(n) := \max\{|D| : D \text{ fundamental}, h(D)\leq n\}$$

the biggest absolute value of a *fundamental* discriminant D with $h(D)\leq n$; the values of $D_{\max}$ for arguments up to 100 can be found in Watkins [24, Table 4]. For the reader's convenience, we give in Table 1 the $D_{\max}$ of the arguments occurring in equation (2-7) below.

Now let $\Delta = Df^2$ be such that $h(\Delta)\leq 100$. Using the class number formula (2-4) (applied with D as Δ and with f as ℓ), and the bound (2-2) we get

$$h(D)\varphi(f)\leq 100[\mathcal{O}_D^\times : \mathcal{O}_\Delta^\times]. \tag{2-6}$$

If $D=-3$ or -4 then this implies $\varphi(f)\leq 300$: the largest such f is $f=1260$, so that in this case $|\Delta|\leq 6350400$.

If $D\neq -3,-4$ then we find from (2-6) that $h(D)\leq 100/\varphi(f)$, and hence

$$|\Delta| = f^2|D|\leq f^2 D_{\max}(\lfloor 100/\varphi(f)\rfloor). \tag{2-7}$$

Plugging in the values from Table 1, we find that the maximum of the right-hand side is attained for $f=420$ and is equal to 28753200. This proves that $|\Delta|\leq 28753200$.

To complete the proof, we ran a PARI script computing the class numbers of all Δ with $|\Delta|\leq 28753200$. It confirms that the biggest Δ with $h(\Delta)\leq 100$ is -2383747, and the biggest Δ with $h(\Delta)\leq 64$ is -991027. ☐

2.1.2. *The* 2*-rank.* Given a finite abelian group G and a prime number p, the *p-rank* of G, denoted by $\rho_p(G)$, is the dimension of the $\mathbb{F}_p$-vector space G/G^p. If Δ is a discriminant, we denote by $\rho_p(\Delta)$ the p-rank of its class group.

The 2-rank of a discriminant was determined by Gauss; see [10, Proposition 3.11 and Theorem 3.15]. As usual, we denote by $\omega(n)$ the number of distinct prime divisors of a nonzero integer n.

Proposition 2.3. *Let Δ be a discriminant. Then*

$$\rho_2(\Delta) = \begin{cases} \omega(\Delta)-1 & \text{if } \Delta \equiv 1 \bmod 4,\\ \omega(\Delta)-2 & \text{if } \Delta \equiv 4 \bmod 16,\\ \omega(\Delta)-1 & \text{if } \Delta \equiv 8,\, 12 \bmod 16,\\ \omega(\Delta)-1 & \text{if } \Delta \equiv 16 \bmod 32,\\ \omega(\Delta) & \text{if } \Delta \equiv 0 \bmod 32. \end{cases}$$

In particular, $\rho_2(\Delta) \in \{\omega(\Delta), \omega(\Delta)-1, \omega(\Delta)-2\}$. If D is a fundamental discriminant, then $\rho_2(D) = \omega(D)-1$.

2.2. *Ring class fields.* If x is a singular modulus with discriminant $\Delta = Df^2$ and $K = \mathbb{Q}(\sqrt{D})$ is its CM field, then $K(x)$ is an abelian extension of K such that $\mathrm{Gal}(K(x)/K)$ is isomorphic to to the class group of Δ; in particular, $[K(x):K] = h(\Delta)$, and the singular moduli of discriminant Δ form a full Galois orbit over K as well.

This leads to the useful notion of *ring class field*. Given an imaginary quadratic field K of discriminant D and a positive integer f, the *ring class field* of K of conductor f, denoted $K[f]$, is, by definition, $K(x)$, where x is some singular modulus of discriminant Df^2. It does not depend on the particular choice of x and is an abelian extension of K.

Proofs of the statements above can be found, for instance, in [10, §§9–11].

The following properties will be systematically used.

Proposition 2.4. *Let K be an imaginary quadratic field and L a ring class field of K. Denote*

$$G = \mathrm{Gal}(L/\mathbb{Q}), \qquad H = \mathrm{Gal}(L/K). \tag{2-8}$$

(As we have just seen, H is an abelian group.) *Then*:

- *Every element of $G \smallsetminus H$ is of order* 2.
- *If $\gamma \in G \smallsetminus H$ and $\eta \in H$ then $\gamma\eta\gamma = \eta^{-1}$.*

Hence, no element of $G \smallsetminus H$ commutes with any element of G of order bigger than 2.

For the proof see [10, Lemma 9.3], for instance.

Proposition 2.5. *Let K be an imaginary quadratic field of discriminant D, and ℓ, m positive integers.*

1. *Assume that either $D \ne -3, -4$ or $\gcd(\ell, m) > 1$. Then the compositum $K[\ell]K[m]$ is equal to $K[\mathrm{lcm}(\ell, m)]$.*
2. *Assume that $D = -3$ and $\gcd(\ell, m) = 1$. Then $K[\ell]K[m]$ is either equal to $K[\mathrm{lcm}(\ell, m)]$ or is a subfield of $K[\mathrm{lcm}(\ell, m)]$ of degree* 3.
3. *Assume that $D = -4$ and $\gcd(\ell, m) = 1$. Then $K[\ell]K[m]$ is either equal to $K[\mathrm{lcm}(\ell, m)]$ or is a subfield of $K[\mathrm{lcm}(\ell, m)]$ of degree* 2.

For the proof see, for instance, [2, Proposition 3.1].

2.2.1. *Two-elementary subfields of ring class fields.* We call a group 2-*elementary* if all its elements are of order dividing 2. A finite 2-elementary group is a product of cyclic groups of order 2. Let K be a field and L a finite extension of K; we say that L is 2-elementary over K if L is Galois over K, with 2-elementary Galois group. We call a number field 2-elementary if it is 2-elementary over $\mathbb{Q}$.

The following is well-known, but we include the proof for the reader's convenience.

Proposition 2.6. *Let F be a number field abelian over $\mathbb{Q}$ and contained in some ring class field. Then F is* 2*-elementary.*

Proof. This is an easy consequence of Proposition 2.4. Let K be an imaginary quadratic field such that its ring class field, denoted L, contains F. We use the notation of (2-8).

For $\gamma \in G$ let $\tilde{\gamma} \in \mathrm{Gal}(F/\mathbb{Q})$ denote the restriction of γ to F. Each element of $\mathrm{Gal}(F/\mathbb{Q})$ is a restriction of either some $\gamma \in G \setminus H$ or some $\eta \in H$. In the former case $\tilde{\gamma}^2 = 1$ because $\gamma^2 = 1$. Now consider $\tilde{\eta}$ for some $\eta \in H$. Pick $\gamma \in G \setminus H$. Then $\tilde{\gamma}\tilde{\eta}\tilde{\gamma} = \tilde{\eta}^{-1}$. But $\mathrm{Gal}(F/\mathbb{Q})$ is abelian, which implies that $\tilde{\gamma}\tilde{\eta}\tilde{\gamma} = \tilde{\gamma}^2\tilde{\eta} = \tilde{\eta}$. Hence $\tilde{\eta}^2 = 1$ as well. Thus, every element of $\mathrm{Gal}(F/\mathbb{Q})$ is of order dividing 2, as wanted. □

The only positive integers m such that the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^\times$ is 2-elementary are the divisors of 24. Hence we have the following corollary.

Corollary 2.7. *The group of roots of unity in a ring class field is of order dividing* 24.

Another important case of 2-elementary fields is the intersection $\mathbb{Q}(x) \cap \mathbb{Q}(y)$, where x and y are singular moduli with distinct fundamental discriminants. This has been known for a long time (see, for instance, the articles of André [4] or Edixhoven [11]), but we again include a proof for the reader's convenience.

Proposition 2.8. *Let x and y be singular moduli with distinct fundamental discriminants*: $D_x \neq D_y$. *Then the field $\mathbb{Q}(x) \cap \mathbb{Q}(y)$ is* 2*-elementary. In particular*, *if $\mathbb{Q}(x) \subset \mathbb{Q}(y)$ then $\mathbb{Q}(x)$ is* 2*-elementary.*

Proof. It suffices to prove that the field $\mathbb{Q}(x) \cap \mathbb{Q}(y)$ is abelian: Proposition 2.6 will then complete the job.

Recall that we let $K_x = \mathbb{Q}(\sqrt{D_x})$ denote the CM field for x. We will denote K_{xy} the compositum of K_x and K_y, that is, the field $\mathbb{Q}(\sqrt{D_x}, \sqrt{D_y})$. Furthermore, we define

$$M = K_{xy}(x, y), \quad L = K_{xy}(x) \cap K_{xy}(y).$$

It suffices to prove that L is abelian, because $L \supset \mathbb{Q}(x) \cap \mathbb{Q}(y)$. We first prove that L is 2-elementary over the field K_{xy}.

Since $K_x \neq K_y$, there exists $\iota \in \mathrm{Gal}(M/\mathbb{Q})$ such that

$$\iota|_{K_x} = \mathrm{id} \quad \text{and} \quad \iota|_{K_y} \neq \mathrm{id}.$$

Proposition 2.4 implies that for $\eta \in \mathrm{Gal}(M/K_{xy})$ we have

$$\iota^{-1}\eta\iota|_{K_x(x)} = \eta|_{K_x(x)} \quad \text{and} \quad \iota^{-1}\eta\iota|_{K_y(y)} = \eta^{-1}|_{K_y(y)}.$$

We also have $\eta|_{K_{xy}} = \mathrm{id}$ by the choice of η. Hence $\eta|_L = \eta^{-1}|_L$. Since every element of $\mathrm{Gal}(L/K_{xy})$ is a restriction to L of some $\eta \in \mathrm{Gal}(M/K_{xy})$, this proves that the Galois group $\mathrm{Gal}(L/K_{xy})$ is 2-elementary.

To complete the proof of the proposition, we must show that L is abelian over $\mathbb{Q}$. Clearly, L is Galois over $\mathbb{Q}$, being the intersection of two Galois extensions. We have to show that $\mathrm{Gal}(K_{xy}/\mathbb{Q})$ acts trivially on $\mathrm{Gal}(L/K_{xy})$. This means proving the following: for every $\eta, \gamma \in \mathrm{Gal}(M/\mathbb{Q})$ such that $\eta|_{K_{xy}} = \mathrm{id}$ we have $\gamma^{-1}\eta\gamma|_L = \eta$.

We denote $\eta^\gamma = \gamma^{-1}\eta\gamma$. Proposition 2.4 implies that

$$\eta^\gamma|_{K_x(x)} \in \left\{\eta|_{K_x(x)}, \eta^{-1}|_{K_x(x)}\right\}.$$

We also have $\eta^\gamma|_{K_y} = \mathrm{id}\,|_{K_y} = \eta|_{K_y} = \eta^{-1}|_{K_y}$. It follows that

$$\eta^\gamma|_{K_{xy}(x)} \in \left\{\eta|_{K_{xy}(x)}, \eta^{-1}|_{K_{xy}(x)}\right\}.$$

In particular, $\eta^\gamma|_L \in \left\{\eta|_L, \eta^{-1}|_L\right\}$. Since $\eta|_{K_{xy}} = \mathrm{id}$ and L/K_{xy} is 2-elementary, we have $\eta|_L = \eta^{-1}|_L$. Hence $\eta^\gamma|_L = \eta|_L$. The proposition is proved. □

2.2.2. (*Almost*) 2-*elementary discriminants.* We will need a slight generalization of the notion of a 2-elementary group. A finite abelian group G will be called *almost* 2-*elementary* if it has a 2-elementary subgroup of index 2. This means that either G is 2-elementary, or it is $\mathcal{C}_4$ times a 2-elementary group. (Recall that $\mathcal{C}_m$ denotes the cyclic group of order m.)

We call a discriminant 2-elementary or almost 2-elementary if its class group has the same property. Such discriminants can be conveniently characterized in terms of the 2-rank (see Section 2.1.2):

$$\Delta \text{ is 2-elementary} \iff h(\Delta) = 2^{\rho_2(\Delta)}; \tag{2-9}$$

$$\Delta \text{ is almost 2-elementary} \iff h(\Delta) \in \{2^{\rho_2(\Delta)}, 2^{\rho_2(\Delta)+1}\}. \tag{2-10}$$

Proposition 2.9. *Let $D \neq -3, -4$ be a fundamental discriminant, and let f be such that Df^2 is an almost 2-elementary discriminant. Then*

$$f \mid 2^4 \cdot 3 \cdot 5 \cdot 7 \cdot 17. \tag{2-11}$$

Proof. As before, we write $\Delta = Df^2$. Since $D \neq -3, -4$, the class number formula (2-4), together with equations (2-1) and (2-5), implies that $h(\Delta) = h(D)\Psi$, where

$$\Psi = \Psi(f, D) = f \prod_{p|f} \left(1 - \frac{(D/p)}{p}\right).$$

When Δ is almost 2-elementary, Ψ is a power of 2, by (2-10). More precisely, we have

$$\Psi \mid 2^{\omega(f)+2}. \tag{2-12}$$

Indeed, if Δ is almost 2-elementary, then so is D, and we have both

$$h(D) \in \{2^{\rho_2(D)}, 2^{\rho_2(D)+1}\} \quad \text{and} \quad h(\Delta) \in \{2^{\rho_2(\Delta)}, 2^{\rho_2(\Delta)+1}\}.$$

Hence,

$$\nu_2(\Psi) \le \rho_2(\Delta) - \rho_2(D) + 1. \tag{2-13}$$

Proposition 2.3 implies that

$$\rho_2(\Delta) - \rho_2(D) \le \omega(\Delta) - \omega(D) + 1 \le \omega(f) + 1, \tag{2-14}$$

which, together with (2-13), proves (2-12).

The following easily shown implications will be systematically used throughout the proof:

$$p \mid f \Longrightarrow p - (D/p) \mid \Psi,$$
$$p^2 \mid f \Longrightarrow p \mid \Psi.$$

This implies very strong constraints on the prime divisors of f.

First of all, D and f cannot have a common prime divisor other than 2. Indeed, if p divides both D and f then $(D/p) = 0$ and $p - (D/p) = p \mid \Psi$. Since Ψ is a power of 2, we must have $p = 2$.

Next, if $p \mid f$ then either $p+1$ or $p-1$ is a power of 2. Indeed, if p is odd then $p \nmid D$ (as we have just seen) which implies that $p - (D/p) \in \{p-1, p+1\}$.

Yet another observation: if $p^2 \mid f$ then $p = 2$. Indeed, in this case we again have $p \mid \Psi$.

Thus,

$$f = 2^k p_1 \cdots p_m, \tag{2-15}$$

where k and m are nonnegative integers and $p_1, \dots, p_m$ are distinct odd primes not dividing D. We claim that

$$k = \nu_2(f) \le 4. \tag{2-16}$$

Indeed, if $k \ge 1$ then $\omega(f) = m + 1$, and

$$2^{\omega(f)+2} = 2^{m+3} \ge \Psi \ge 2^{k-1}(p_1 - 1) \cdots (p_m - 1) \ge 2^{m+k-1},$$

which proves (2-16).

To complete the proof, we have to show that the only possible prime divisors of f are 2, 3, 5, 7, 17. If $w(f) = 1$ then $\Psi \mid 8$, and $f = 2^k$ or $f = p$, an odd prime. Since $p-1$ or $p+1$ divides Ψ, this implies that $p \le 7$.

Now assume that $\omega(f) \ge 2$, and that f has a prime divisor $p \ne 2, 3, 5, 7, 17$. Then $p \ge 31$, because one of $p \pm 1$ must be a power of 2. Writing $f = 2^k p_1 \cdots p_m$ with $2 < p_1 < \cdots < p_m$ and $p_m \ge 31$, we have

$$2^{m+3} \ge 2^{\omega(f)+2} \ge \Psi \ge (p_1 - (D/p_1)) \cdots (p_m - (D/p_m)).$$

Since $p_m \ge 31$ and $p_m - (D/p_m)$ is a power of 2, we have $p_m - (D/p_m) \ge 32$.

If $m \ge 2$ then

$$(p_1 - (D/p_1)) \cdots (p_m - (D/p_m)) \ge (3-1)(5-1)^{m-2} \cdot 32 = 2^{2m+2}.$$

We obtain that $m+3 \ge 2m+2$, which is impossible for $m \ge 2$. It follows that $m = 1$, in which case $\omega(f) = 2$ and $f = 2^k p$, where $k \ge 1$ and $p \ge 31$. We obtain $16 \ge \Psi \ge 2^{k-1} \cdot 32$, a contradiction. The proof is complete. □

One can do some case-by-case analysis and show that f satisfies a stronger (but also more complicated) condition than (2-11). However, (2-11) is sufficient for our purposes.

We also need a similar result for the fundamental discriminants -3 and -4.

Proposition 2.10. *If $\Delta = -4f^2$ is an almost 2-elementary discriminant, then*

$$f \le 8 \quad or \quad f \in \{10, 12, 15, 20\}. \tag{2-17}$$

If $\Delta = -3f^2$ is an almost 2-elementary discriminant, then

$$f \le 5 \quad or \quad f \in \{7, 8, 11, 13, 16\}. \tag{2-18}$$

In both cases, it follows that $h(\Delta) \mid 8$.

Proof. Assume first that $\Delta = -4f^2$ is almost 2-elementary. The class number formula (2-4), together with equations (2-1) and (2-5), implies that $h(\Delta) = \Psi/2$, where

$$\Psi = \Psi(f, -4) = f \prod_{p \mid f} \left(1 - \frac{(-4/p)}{p}\right).$$

As in the proof of Proposition 2.9, this Ψ must be a power of 2; more precisely,

$$\Psi \mid 2^{\omega(f)+2}. \tag{2-19}$$

Indeed, if $2 \nmid f$, then $\Delta \equiv 12 \bmod 16$ and $\rho_2(\Delta) \le \omega(\Delta) - 1 = \omega(f)$, while when $2 \mid f$, we have $\rho_2(\Delta) \le \omega(\Delta) = \omega(f)$. In both cases we obtain

$$\Psi/2 = h(\Delta) \mid 2^{\omega(f)+1},$$

which is (2-19).

Let p be an odd prime divisor of f. As in the proof of Proposition 2.9, we have $p^2 \nmid f$ and $p - (-4/p) \mid \Psi$; thus $p - (-4/p)$ is a power of 2. We again write the prime factorization of f as $f = 2^k p_1 \cdots p_m$, where $2 < p_1 < \cdots < p_m$. We claim that

$$k + m \le 3 \quad \text{and} \quad m \le 2. \tag{2-20}$$

Indeed, if $k \ge 1$ then $\omega(f) = m+1$, and (2-19) implies that

$$2^{m+3} \ge \Psi \ge 2^{k-1}(2 - (-4/2))(3 - (-4/3))^m = 2^{k+2m},$$

which proves (2-20) in the case $k \ge 1$. Similarly, if $k = 0$ then $2^{m+2} \ge \Psi \ge 2^{2m}$, proving (2-20) in this case as well.

In a similar fashion one proves that

$$p_m \le 7. \tag{2-21}$$

Indeed, if $p_m > 7$ then $p_m \geq 17$, because one of $p_m \pm 1$ must be a power of 2. If $k \geq 1$ then

$$2^{m+3} \geq \Psi \geq 2^k \cdot 4^{m-1} \cdot 16 = 2^{k+2m+2},$$

which is impossible; if $k = 0$ then

$$2^{m+2} \geq \Psi \geq 4^{m-1} \cdot 16 = 2^{2m+2},$$

again impossible. This proves (2-21).

It follows from (2-20) and (2-21) that there are finitely many possible f. Checking them all using a PARI script, we obtain (2-17).

Now assume that $\Delta = -3f^2$ is almost 2-elementary. In this case $h(\Delta) = \Psi/3$, where

$$\Psi = \Psi(f, -3) = f \prod_{p \mid f} \left(1 - \frac{(-3/p)}{p}\right).$$

This time Ψ must be 3 times a power of 2. It follows again that for an odd prime p we have $p^2 \nmid f$. This is clear when $p \neq 3$, and if $9 \mid f$ then

$$3(3 - (-3/3)) = 9 \mid \Psi,$$

a contradiction.

For every $p \mid f$ the difference $p - (-3/p)$ must be either a power of 2, or 3 times a power of 2. We claim that$p - (-3/p)$ cannot be a power of 3. This is clear for $p = 2$ and for $p = 3$. Now assume that $p \neq 2, 3$ and $p - (-3/p) = 2^n$. If n is odd then $3 \mid 2^n + 1$, which means that $p = 2^n - 1$. But in this case $p \equiv 1 \bmod 3$ and $p \equiv -1 \bmod 4$, which implies that $(-3/p) = 1$. It follows that $p - (-3/p) = 2^n - 2$, a contradiction. Similarly, when n is even, we have $p = 2^n + 1$ and $(-3/p) = -1$, again a contradiction.

Thus, for every $p \mid f$ we have $p - (-3/p) = 3 \cdot 2^n$ for some n. This implies that f cannot have two distinct prime divisors: if it did, then Ψ would be divisible by 9, a contradiction.

Thus, either $f = 2^k$ for some k, or $f = p$, an odd prime. This implies that $\rho_2(\Delta) \leq \omega(\Delta) \leq 2$, and

$$\Psi/3 = h(\Delta) \mid 2^{\rho_2(\Delta)+1} \mid 8.$$

It follows that either $f \mid 16$, or $f \in \{3, 5, 7, 11, 13, 23\}$. Checking all possible f using a PARI script, we obtain (2-18). □

Proposition 2.11. *There exists a fundamental discriminant D^* such that $h(D^*) \geq 128$ and the following holds. Let $\Delta = Df^2$ be either 2-elementary or almost 2-elementary. Then either $D = D^*$ or*

$$h(\Delta) \leq \begin{cases} 16 & \text{if } \Delta \text{ is 2-elementary,} \\ 64 & \text{if } \Delta \text{ is almost 2-elementary.} \end{cases}$$

This is proved in [1, Corollary 2.5 and Remark 2.6]. It was not included in [2], the published version of the same work, so we reproduce the proof here (adding some details missing in [1]). The proof broadly follows the strategy of Weinberger [25], which rests on a classical bound of Tatuzawa, stated below.

Given a fundamental discriminant D, the *L-function attached to D* is $L(s, \chi)$, where χ is the quadratic character defined by the Kronecker symbol: $\chi(n) = (D/n)$.

Lemma 2.12 (Tatuzawa [23, Theorem 2]). *Let $0 < \varepsilon < \frac{1}{2}$. There exists a fundamental discriminant D^* such that the following holds. Let D be a fundamental discriminant and $L(s, \chi)$ the attached L-function. Then $L(1, \chi) \geq 0.655\varepsilon|D|^{-\varepsilon}$ when $|D| \leq \max\{e^{1/\varepsilon}, 73130\}$ and $D \neq D^*$.*

Proof of Proposition 2.11. If $D = -3$ or -4 then the result follows from Proposition 2.10. Hence we may assume that $D \neq -3, -4$. In this case the analytic class number formula states that $h(D) = \pi^{-1}|D|^{1/2}L(1, \chi)$. If Δ is almost 2-elementary then so is D. By (2-10) and Proposition 2.3 we have $h(D) \leq 2^{\rho_2(D)+1} \leq 2^{\omega(D)}$.

We pick $\varepsilon = 0.048$ throughout and we use the corresponding D^* from Lemma 2.12. Assuming that $D \neq D^*$, Lemma 2.12 implies that

$$2^{\omega(D)} \geq h(D) = \pi^{-1}|D|^{1/2}L(1, \chi) \geq 0.655\pi^{-1}\varepsilon|D|^{1/2-\varepsilon}$$

as long as $|D| \geq 1.2 \cdot 10^9$. This implies that

$$|D| \leq \left((0.655\varepsilon)^{-1}\pi 2^{\omega(D)}\right)^{1/(1/2-\varepsilon)} \leq 26549 \cdot 4.635^{\omega(D)};$$

we conclude that

$$|D| \leq \max\{26549 \cdot 4.635^{\omega(D)}, 1.2 \cdot 10^9\}. \tag{2-22}$$

Moreover, since D is fundamental and

$$1.2 \cdot 10^9 < 4 \cdot (3 \cdot 5 \cdot 7 \cdot 11 \cdots 37)$$

(4 times the product of the first 11 odd primes), we must have $\omega(D) \leq 11$ whenever $|D| \leq 1.2 \cdot 10^9$. More generally, $|D|$ is at least 4 times the product of the first $\omega(D) - 1$ odd primes. Hence, when $\omega(D) \geq 12$, we have

$$|D| \geq 4 \cdot (3 \cdot 5 \cdot 7 \cdot 11 \cdots 37) \cdot 41^{\omega(D)-12}.$$

Combining this observation with the upper bound (2-22), we conclude that, when $\omega(D) \geq 12$, we have

$$4 \cdot (3 \cdot 5 \cdot 7 \cdot 11 \cdots 37) \cdot 41^{\omega(D)-12} \leq |D| \leq 26549 \cdot 4.635^{\omega(D)}.$$

This is easily seen to be a contradiction for $\omega(D) \geq 12$. We conclude that

$$\omega(D) \leq 11 \tag{2-23}$$

for any almost 2-elementary fundamental $D \neq D^*$.

Thus, we are now left with the task of examining discriminants $\Delta = Df^2$ such that the corresponding fundamental D satisfies conditions (2-22) and (2-23). We want to show that

- if such Δ is 2-elementary then $h(\Delta) \leq 16$, and
- if such Δ is almost 2-elementary then $h(\Delta) \leq 64$.

Proving this is a numerical check using PARI. We distinguish two cases: $\omega(D) \le 6$ and $7 \le \omega(D) \le 11$.

When $\omega(D) \le 6$ and D is almost 2-elementary then $h(D) | 64$. Table 4 of Watkins [24] implies that in this case $|D| \le 693067$. Using Proposition 2.9 and a PARI script, we computed all 2-elementary and all almost 2-elementary discriminants $\Delta = Df^2$ such that $|D| \le 693067$. Our script found 101 discriminants that are 2-elementary, the largest being $-7392 = -1848 \cdot 2^2$. The class numbers of all these discriminants do not exceed 16. Similarly, the script found 425 almost 2-elementary discriminants, $-87360 = -5460 \cdot 4^2$ being the largest, and their class numbers do not exceed 64. This completes the proof in the case $\omega(D) \le 6$.

When $7 \le \omega(D) \le 11$, we can no longer use [24]. To complete the proof in this case, for every $n = 7, \dots, 11$ we determine all fundamental discriminants D satisfying

$$\omega(D) = n, \quad |D| \le 26549 \cdot 4.635^n \tag{2-24}$$

(note that $26549 \cdot 4.635^n > 1.2 \cdot 10^9$ for $n \ge 7$), and for each of them we check whether it is almost 2-elementary. Our script found no almost 2-elementary fundamental discriminants satisfying (2-24) with $7 \le n \le 11$. This completes the proof of Proposition 2.11. □

Corollary 2.13. *Let x and y be singular moduli with distinct fundamental discriminants, $D_x \ne D_y$.*

(1) *Assume that $\mathbb{Q}(x) = \mathbb{Q}(y)$. Then $h(\Delta_x) = h(\Delta_y) \le 16$.*

(2) *Assume that $\mathbb{Q}(x) \subset \mathbb{Q}(y)$ and $[\mathbb{Q}(y) : \mathbb{Q}(x)] = 2$. Then $h(\Delta_x) \le 16$ and $h(\Delta_y) \le 32$.*

Proof. If $\mathbb{Q}(x) = \mathbb{Q}(y)$ then both Δ_x and Δ_y are 2-elementary by Proposition 2.8. Since $D_x \ne D_y$, one of the two is distinct from D^*; say, $D_x \ne D^*$. Then $h(\Delta_x) \le 16$. Hence $h(\Delta_y) = h(\Delta_x) \le 16$ as well. This proves item (1).

Now assume that we are in the situation of item (2). Then $\mathrm{Gal}(\mathbb{Q}(x)/\mathbb{Q})$ is 2-elementary by Proposition 2.8. Hence so is $\mathrm{Gal}(K_y(x)/K_y)$. Since

$$[K_y(y) : K_y(x)] \le [\mathbb{Q}(y) : \mathbb{Q}(x)] = 2,$$

the group $\mathrm{Gal}(K_y(y)/K_y)$ is almost 2-elementary. If $D_x \ne D^*$ then $h(\Delta_x) \le 16$ and $h(\Delta_y) \le 32$, so we are done. If $D_y \ne D^*$ then $h(\Delta_y) \le 64$. It follows that $h(\Delta_x) \le 32$ and we must have $D_x \ne D^*$, so we are done again. □

2.3. ***Gauss reduction theory, denominators.*** Denote by T_Δ the set of triples $(a, b, c) \in \mathbb{Z}^3$ with $\Delta = b^2 - 4ac$ satisfying

$$\gcd(a, b, c) = 1 \quad \text{and} \quad (\text{either } -a < b \le a < c \text{ or } 0 \le b \le a = c). \tag{2-25}$$

Condition (2-25) is equivalent to

$$\frac{b + \sqrt{\Delta}}{2a} \in \mathcal{F}.$$

For every singular modulus x of discriminant Δ there exists a unique triple $(a_x, b_x, c_x) \in T_\Delta$ such that, denoting

$$\tau_x = \frac{b_x + \sqrt{\Delta}}{2a_x},$$

we have $x = j(\tau_x)$. This is, essentially, due to Gauss; see [6, Section 2.2] for details.

We will call a_x the *denominator* of the singular modulus x.

Note that, alternatively, τ_x can be defined as the unique $\tau \in \mathcal{F}$ such that $j(\tau) = x$.

We will say that a positive integer a is a denominator for Δ if it is a denominator of some singular modulus of discriminant Δ; equivalently, there exist $b, c \in \mathbb{Z}$ such that $(a, b, c) \in T_\Delta$.

It will often be more convenient to use the notation $a(x)$, $b(x)$, $\tau(x)$ etc. instead of a_x, b_x, τ_x, etc.

Remark 2.14. It is useful to note that b_x and Δ_x are of the same parity: $b_x \equiv \Delta_x \bmod 2$. This is because $\Delta_x = b_x^2 - 4a_xc_x \equiv b_x^2 \bmod 4$.

For every Δ there exists exactly one singular modulus of discriminant Δ and of denominator 1, which will be called the *dominant* singular modulus of discriminant Δ. Singular moduli with denominator 2 will be called *subdominant*.

Proposition 2.15. *Let Δ be a discriminant. Then for every $a \in \{2, 3, 4, 5\}$ there exist at most 2 singular moduli x with $\Delta_x = \Delta$ and $a_x = a$. For every $A \in \{13, 18, 30\}$ there exists at most $S(A)$ singular moduli x with $\Delta_x = \Delta$ and $a_x < A$, where $S(A)$ is given in the following table*:

A	13	18	30
$S(A)$	32	48	99

Proof. Let a be a positive integer. For a residue class $r \bmod 4a$ denote $B(r)$ the number of $b \in \mathbb{Z}$ satisfying $-a < b \le a$ and $b^2 \equiv r \bmod 4a$. Denote $s(a)$ the biggest of all $B(r)$:

$$s(a) = \max\{B(r) : r \bmod 4a\}.$$

The number of triples $(a, b, c) \in T_\Delta$ with given a does not exceed $B(\Delta)$; hence it does not exceed $s(a)$ either. A quick calculation shows that $s(a) = 2$ for $a \in \{2, 3, 4, 5\}$, and

$$\sum_{a<A} s(a) = S(A)$$

for $A \in \{13, 18, 30\}$. The proposition is proved. □

We will also need miscellaneous facts about the (non)existence of singular moduli of some specific shapes. The following proposition will be used in this article only for $p = 3$. We, however, state it for general p, for the sake of further applications.

Proposition 2.16. *Let Δ be a discriminant and p an odd prime number.*

1. *Assume that $(\Delta/p)=1$. If $|\Delta|\geq 4p^2-1$ then Δ admits exactly 2 singular moduli with denominator p. More generally, if $|\Delta|\geq 4p^k-1$ then Δ admits exactly 2 singular moduli with denominator p^k.*
2. *Assume that $p^2\,|\,\Delta$, and let a be a denominator for Δ. Then either $p\nmid a$ or $p^2\,|\,a$. In particular, p is not a denominator for Δ.*

Proof. By Hensel's lemma, the assumption $(\Delta/p)=1$ implies that the congruence $b^2\equiv\Delta \bmod p^k$ has exactly two solutions b satisfying $0<b<p^k$, and exactly one of these solutions satisfies $b^2\equiv\Delta \bmod 4p^k$. If b is this solution and $|\Delta|\geq 4p^k-1$ then the two triples $(p^k, \pm b, (b^2-\Delta)/4p^k)$ belong to T_Δ. This proves item 1.

If $p^2\,|\,\Delta$ and $p\,|\,a$ then $p\,|\,b$ and $p\nmid c$. Hence $p^2\,|4ac=b^2-\Delta$, which implies that $p^2\,|\,a$. This proves item 2. □

Here is an analogue of Proposition 2.16 for $p=2$.

Proposition 2.17. *Let Δ be a discriminant.*

1. *Assume that $\Delta\equiv 1 \bmod 8$. If $|\Delta|>15$ then Δ admits exactly 2 subdominant singular moduli, which are $j\big((\pm1+\sqrt{\Delta})/4\big)$. More generally, if $|\Delta|\geq 4^{k+1}-1$ then Δ admits exactly 2 singular moduli with denominator 2^k.*
2. *If $\Delta\not\equiv 1 \bmod 8$ then it admits at most one subdominant singular modulus.*
3. *Let Δ satisfy $\Delta\equiv 4 \bmod 32$ and $|\Delta|\geq 252$. Then it admits exactly 2 singular moduli of denominator 8. These are*
$$j\left(\frac{\pm b'+\sqrt{\Delta/4}}{8}\right),\qquad b'=\begin{cases}1 & \text{if } \Delta\equiv 36 \bmod 64,\\ 3 & \text{if } \Delta\equiv 4 \bmod 64.\end{cases}\tag{2-26}$$
More generally, if $k\geq 3$ and $|\Delta|\geq 4^{k+1}-4$ then Δ admits exactly 2 singular moduli with denominator 2^k.
4. *Let Δ satisfy $\Delta\equiv 4 \bmod 32$ and let a be a denominator for Δ. Then either a is odd, or $8\,|\,a$. In particular, 2, 4 and 6 are not denominators for Δ.*
5. *Let Δ be divisible by 16 and let a be a denominator for Δ. Then either a is odd, or $4\,|\,a$. In particular, 2 is not a denominator for Δ.*

   Furthermore, Δ admits at most one singular modulus with denominator 4.
6. *Assume that Δ is even, but $\Delta\not\equiv 4 \bmod 32$, and that $|\Delta|>76$. Then 2 or 4 is a denominator for Δ.*

Proof. Items 1 and 3 are proved using Hensel's lemma exactly like item 1 of Proposition 2.16; we omit the details.

Item 2 follows from [6, Proposition 2.6], and item 6 is [9, Proposition 3.1.4]. Note that in [9] denominators are called *suitable integers*.

We are left with items 4 and 5. If $\Delta\equiv 4 \bmod 32$ and $(a, b, c)\in T_\Delta$ with $2\,|\,a$ then $2\,\|\,b$ and c is odd. Hence $b^2\equiv 4 \bmod 32$, which implies that $4ac\equiv 0 \bmod 32$. This shows that $8\,|\,a$, which proves item 4.

Finally, if $16\,|\,\Delta$ and $2\|a$ then $2\,|\,b$ and $2\nmid c$, which implies that

$$b^2 = \Delta + 4ac \equiv 8 \bmod 16,$$

a contradiction. This proves the first statement in item 5. Similarly, if $a = 4$ then $4\,|\,b$ and $2\nmid c$; in particular, $4ac \equiv 16 \bmod 32$. Hence

$$b = \begin{cases} 0 & \text{if } \Delta \equiv 16 \bmod 32, \\ 4 & \text{if } \Delta \equiv 0 \bmod 32. \end{cases}$$

Thus, in any case there is only one choice for b, which proves the second statement in item 5. □

It is useful to note that the dominant singular modulus is real; thus there exists at least one real singular modulus of every discriminant. This has the following consequence.

Proposition 2.18. *Let x be a singular modulus, and let K be a subfield of $\mathbb{Q}(x)$. Assume that K is Galois over $\mathbb{Q}$. Then K is a totally real field.*

Since $\mathbb{Q}(x)$ is Galois over $\mathbb{Q}$ when Δ_x is 2-elementary, this implies that singular moduli of 2-elementary discriminants are all real.

2.4. *Isogenies.* Let Λ and M be lattices in $\mathbb{C}$. We say that Λ and M are isogenous if Λ is isomorphic to a sublattice of M. Specifically, given a positive integer n, Λ and M are *n-isogenous* if M has a sublattice Λ', isomorphic to Λ, such that the quotient group M/Λ' is cyclic of order n. This relation is symmetric.

Two lattices $\langle z, 1\rangle$ and $\langle w, 1\rangle$, for z, w in the Poincaré plane $\mathbb{H}$, are n-isogenous if and only if there exists $\gamma \in \mathrm{M}_2(\mathbb{Z})$ with coprime entries and determinant n such that $w = \gamma(z)$. Imposing upon z and w certain reasonable conditions, one may assume that the matrix γ is upper triangular.

Proposition 2.19. *Let $z, w \in \mathbb{H}$ and let n be a positive integer. Assume that*

$$w \in \mathcal{F} \quad \textit{and} \quad \operatorname{Im} z \geq n. \tag{2-27}$$

Then the following two conditions are equivalent.

(1) *The lattices $\langle z, 1\rangle$ and $\langle w, 1\rangle$ are n-isogenous.*

(2) *We have*

$$w = \frac{pz+q}{s},$$

where $p, q, s \in \mathbb{Z}$ satisfy

$$p, s > 0, \quad ps = n, \quad \gcd(p, q, s) = 1.$$

Proof. The implication (2) ⇒ (1) is trivial and does not require (2-27).

Now assume that (1) holds, and let $\gamma \in M_2(\mathbb{Z})$ be a matrix with coprime entries and determinant n such that $w = \gamma(z)$. There exists $\delta \in \mathrm{SL}_2(\mathbb{Z})$ such that $\delta\gamma$ is an upper triangular matrix. Replacing δ by $\left(\begin{smallmatrix} 1 & \nu \\ 0 & 1 \end{smallmatrix}\right)\delta$ with a suitable $\nu \in \mathbb{Z}$, we may assume that $w' := \delta\gamma(z)$ satisfies

$$-\tfrac{1}{2} < \operatorname{Re} w' \leq \tfrac{1}{2}. \tag{2-28}$$

Write $\delta\gamma = \left(\begin{smallmatrix} p & q \\ 0 & s \end{smallmatrix}\right)$. Replacing δ by $-\delta$ if necessary, we may assume that $p, s > 0$. Since $ps = n$ and

$$w' = \frac{pz+q}{s},$$

we only have to prove that $w' = w$. We have $\operatorname{Im} w' = (\operatorname{Im} z)/s \geq 1$, by (2-27). Together with (2-28) this implies that $w' \in \mathcal{F}$. But w belongs to $\mathcal{F}$ as well, again by (2-27). Since $w' = \delta w$ and each $\mathrm{SL}_2(\mathbb{Z})$-orbit has exactly one point in $\mathcal{F}$, we must have $w' = w$. □

We say that two singular moduli are n-isogenous if, writing $x = j(\tau)$ and $y = j(\upsilon)$, the lattices $\langle \tau, 1\rangle$ and $\langle \upsilon, 1\rangle$ are n-isogenous.

Singular moduli x and y are n-isogenous if and only if $\Phi_n(x, y) = 0$, where $\Phi_n(X, Y)$ denotes the modular polynomial of level n. Since $\Phi_n(X, Y) \in \mathbb{Q}[X, Y]$, being n-isogenous is preserved by Galois conjugation: for any $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ the singular moduli x^σ and y^σ are n-isogenous as long as x and y are.

For a positive integer n define

$$\mathcal{Q}(n) = \left\{ \frac{r}{s} : r, s \in \mathbb{Z}, rs = n \right\}.$$

For example,

$$\mathcal{Q}(12) = \left\{ \tfrac{1}{12}, \tfrac{1}{3}, \tfrac{3}{4}, \tfrac{4}{3}, 3, 12 \right\}.$$

The following property is an immediate consequence of Proposition 2.19.

Corollary 2.20. *Let x and y be n-isogenous singular moduli. Assume that $|\Delta_x|^{1/2} \geq 2na_x$. Then $(a_y/f_y)/(a_x/f_x) \in \mathcal{Q}(n)$. In particular,*

$$\frac{1}{n} \leq \frac{a_y/f_y}{a_x/f_x} \leq n.$$

When $n = p$ is a prime number, we have $(a_y/f_y)/(a_x/f_x) \in \{p, 1/p\}$.

The following simple facts will be repeatedly used, often without special reference.

Proposition 2.21. *Let x and y be singular moduli.*

1. *Assume that $\Delta_x = \Delta_y$ and $\gcd(a_x, a_y) = 1$. Then x and y are $a_x a_y$-isogenous.*
2. *Assume that $a_x = a_y = 1$ and $\Delta_x/e_x^2 = \Delta_y/e_y^2$, where e_x and e_y are coprime positive integers. Then x and y are $e_x e_y$-isogenous.*
3. *Two subdominant singular moduli of the same discriminant are either equal or 4-isogenous.*

Proof. To prove item 1, note that

$$\tau_y = \frac{a_x\tau_x + (b_y - b_x)/2}{a_y}.$$

Since $b_x \equiv b_y \bmod 2$ (see Remark 2.14), this proves that x and y are $a_x a_y$-isogenous.

For item 2 we have

$$\tau_x = \frac{b_x + e_x\sqrt{\Delta}}{2} \quad \text{and} \quad \tau_y = \frac{b_y + e_y\sqrt{\Delta}}{2},$$

where $\Delta = \Delta_x/e_x^2 = \Delta_y/e_y^2$. Hence

$$\tau_y = \frac{e_y\tau_x + (b_ye_x - b_xe_y)/2}{e_x}.$$

Remark 2.14 now implies that $b_xe_y \equiv b_ye_x \bmod 2$, and we conclude that x and y are e_xe_y-isogenous.

To prove item 3, note that distinct subdominant singular moduli of the same discriminant Δ must be of the form $j(\tau)$ and $j(\tau')$, where

$$\tau = \frac{-1+\sqrt{\Delta}}{4} \quad \text{and} \quad \tau' = \frac{1+\sqrt{\Delta}}{4};$$

see Proposition 2.17:1. We have $\tau' = (2\tau+1)/2$, which implies that $j(\tau)$ and $j(\tau')$ are 4-isogenous. □

2.5. *Galois-theoretic lemmas.* In this subsection we collect some lemmas with Galois-theoretic flavor that will be repeatedly used in the proofs of Theorems 1.1 and 1.2.

Lemma 2.22. *Let m be a positive integer and x a singular modulus. Then $\mathbb{Q}(x) = \mathbb{Q}(x^m)$. In other words: if x and y are distinct singular moduli of the same discriminant then $x^m \ne y^m$.*

Proof. See [21, Lemma 2.6]. □

Lemma 2.23. *Let x and y be distinct singular moduli of the same discriminant, $K = K_x = K_y$ their common CM field, $L = K(x) = K(y)$ the ring class field and $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ a Galois morphism. Assume that σ permutes x and y:*

$$x^\sigma = y \quad \text{and} \quad y^\sigma = x.$$

Then σ is of order 2.

Proof. If $\sigma \notin \mathrm{Gal}(L/K)$ then it is of order 2 because every element of $\mathrm{Gal}(L/\mathbb{Q})$ not belonging to $\mathrm{Gal}(L/K)$ is of order 2. And if $\sigma \in \mathrm{Gal}(L/K)$ then $\sigma^2 = 1$ because $x^{\sigma^2} = x$ and $L = K(x)$. □

Lemma 2.24. *Let x, y be distinct singular moduli of the same discriminant and let K and L be as in Lemma 2.23. Let F be a proper subfield of $\mathbb{Q}(x, y)$; we write $G = \mathrm{Gal}(L/F)$. Then one of the following conditions is satisfied.*

(1) *There exists $\sigma \in G$ such that, up to switching x, y, we have $x^\sigma = x$ but $y^\sigma \ne y$.*
(2) *We have $[\mathbb{Q}(x, y) : F] = 2$ and the nontrivial automorphism of $\mathbb{Q}(x, y)/F$ permutes x and y.*
(3) *We have $L = \mathbb{Q}(x, y)$ and $[L : F] = 3$. Moreover, there exists a singular modulus z and $\sigma \in G$ such that*

$$x^\sigma = y, \quad y^\sigma = z, \quad z^\sigma = x.$$

(4) *There exists $\sigma \in G$ such that*

$$x^\sigma, y^\sigma \notin \{x, y\} \tag{2-29}$$

(Versions of this lemma were used, albeit implicitly, in [12] and elsewhere, but it does not seem to have appeared in the literature in this form.)

Proof. We may assume that every element in G which fixes x or y fixes both of them; otherwise we have item (1). We may also assume that x and y are conjugate over F; otherwise, any $\sigma \in G$ not belonging to $\mathrm{Gal}(L/\mathbb{Q}(x, y))$ satisfies (2-29).

Assume first that $L = \mathbb{Q}(x, y)$. Then the only element of G that fixes x or y is the identity. Since x and y are conjugate over F, there is exactly one $\sigma \in G$ with the property $x^\sigma = y$ and exactly one $\sigma' \in G$ with the property $y^{\sigma'} = x$. Hence (4) holds if $[L : F] \geq 4$. And if $[L : F] \leq 3$ then we have one of conditions (2) or (3) is satisfied. This completes the proof in the case $L = \mathbb{Q}(x, y)$.

Now assume that $L \neq \mathbb{Q}(x, y)$. Since $L = K(x)$, the field $\mathbb{Q}(x)$ is a subfield of L of degree 2, and so is $\mathbb{Q}(y)$. If $Q(x) \neq \mathbb{Q}(y)$ then the compositum of these fields must be L, which contradicts the assumption $L \neq \mathbb{Q}(x, y)$. Hence $\mathbb{Q}(x) = \mathbb{Q}(y)$ is a subfield of L of degree 2. (4) holds if $[\mathbb{Q}(x) : F] \geq 4$, and (2) does if $[\mathbb{Q}(x) : F] = 2$.

We are left with the case $[\mathbb{Q}(x) : F] = 3$. In this case the Galois orbit of x over F consists of 3 elements: x, y and a certain z. The group G must be either cyclic $\mathcal{C}_6$ or symmetric S_3. In the latter case G acts by permutations on the set $\{x, y, z\}$. But in this case G has an element fixing x and permuting y, z, which is impossible because $y \in \mathbb{Q}(x)$.

Thus, $G = \mathcal{C}_6$. The group $\mathrm{Gal}(L/\mathbb{Q}(x))$ is a subgroup of G; let γ be the nontrivial element of $\mathrm{Gal}(L/\mathbb{Q}(x))$. Then $\gamma \notin \mathrm{Gal}(L/K)$; otherwise, from $L = K(x)$ and $x^\gamma = x$ we would obtain that γ is the identity. It follows (see Proposition 2.4) that γ does not commute with the elements of G of order 3, which is impossible, because G is an abelian group. The lemma is proved. □

Lemma 2.25. *Let x and y be singular moduli with the same fundamental discriminant D, and let $K = \mathbb{Q}(\sqrt{D})$ be their common CM field. Assume that $K(x) = K(y)$, and that x and y do not both belong to $\mathbb{Q}$. Then*

$$\Delta_x/\Delta_y \in \{4, 1, 1/4\}.$$

Moreover, if (say) $\Delta_x = 4\Delta_y$, then $\Delta_y \equiv 1 \bmod 8$.

Proof. See [2, Proposition 4.3] and [6, Subsection 3.2.2] (where the congruence $\Delta_y \equiv 1 \bmod 8$ is proved). Note that in [2] a formally stronger hypothesis $\mathbb{Q}(x) = \mathbb{Q}(y)$ is imposed, but in the proof it is only used that $K(x) = K(y)$. □

Lemma 2.26. *Let x, x', y, y' be singular moduli. Assume that*

$$\Delta_x = \Delta_{x'}, \quad \Delta_y = \Delta_{y'} \quad \text{and} \quad \mathbb{Q}(x, x') = \mathbb{Q}(y, y').$$

Then:

1. *If $D_x \neq D_y$ then $\mathbb{Q}(x) = \mathbb{Q}(y)$.*
2. *If $D_x = D_y$ then $K(x) = K(y)$, where $K = K_x = K_y$ is the common CM field for x and y.*

Proof. See [8, Lemma 7.1]. □

3. The linear relation

Let $x_1, \dots, x_k$ be nonzero singular moduli of the same fundamental discriminant D, and let $m_1, \dots, m_k \in \mathbb{Z}$. We want to show that, under some reasonable assumptions, the multiplicative relation

$$x_1^{m_1} \cdots x_k^{m_k} = 1 \tag{3-1}$$

implies the linear relation

$$\sum_{i=1}^{k} \frac{f(x_i)}{a(x_i)} m_i = 0. \tag{3-2}$$

To state those assumptions, set

$$X = \max\{|\Delta(x_i)| : 1 \le i \le k\} \quad \text{and} \quad Y = \min\{|\Delta(x_i)| : 1 \le i \le k\}, \tag{3-3}$$

and, as in Section 2.3, let f_x or $f(x)$ mean the conductor of a singular modulus x, and a_x or $a(x)$ its denominator.

Proposition 3.1. *Let A be a positive number such that*

$$a(x_i) \le A \quad (1 \le i \le k). \tag{3-4}$$

Assume that

$$Y^{1/2} > \tfrac{1}{3} Ak(\log X + \log A + \log k + 20). \tag{3-5}$$

Then (3-1) *implies* (3-2).

It often happens that we control only a part of the denominators of $x_1, \dots, x_k$. In this case we cannot expect an identity like (3-2), but we may have good bounds for the part of the sum corresponding to the terms with small denominators.

We need some extra notation. Set $f = \gcd(f_{x_1}, \dots, f_{x_k})$ and $\Delta = Df^2$. We also define

$$e_{x_i} = e(x_i) = f(x_i)/f \quad \text{and} \quad m_i' = e(x_i) m_i \quad (i = 1, \dots, k).$$

Then we have $\Delta(x_i) = e(x_i)^2 \Delta$, and (3-2) can be rewritten as

$$\sum_{i=1}^{k} \frac{m_i'}{a(x_i)} = 0. \tag{3-6}$$

As indicated above, we want to obtain a less precise result, in the form of an inequality, which however holds true without the assumption that all the denominators are small. It will be practical to estimate separately the sums with positive and with negative exponents m_i.

Proposition 3.2. *Let A, ε be real numbers satisfying $A \ge 1$ and $0 < \varepsilon \le 0.5$. Assume that*

$$|\Delta|^{1/2} \ge \max\left\{k\varepsilon^{-1} \log X,\, \tfrac{1}{3} A\big(\log(k\varepsilon^{-1}) + 4\big)\right\}. \tag{3-7}$$

Then

$$\sum_{\substack{a(x_i)<A\\ m_i>0}} \frac{m_i'}{a(x_i)} \le \sum_{m_i<0} \frac{|m_i'|}{\min\{a(x_i),\, A\}} + \varepsilon\|\boldsymbol{m}'\|, \tag{3-8}$$

$$\sum_{\substack{a(x_i)<A\\ m_i<0}} \frac{|m_i'|}{a(x_i)} \le \sum_{m_i>0} \frac{m_i'}{\min\{a(x_i),\, A\}} + \varepsilon\|\boldsymbol{m}'\|. \tag{3-9}$$

Here we denote by $\|\boldsymbol{m}'\|$ *the sup-norm* $\max\{|m_1'|, \dots, |m_k'|\}$.

Propositions 3.1 and 3.2 will be our principal tools in the proofs of Theorems 1.1 and 1.2. They will be proved in Section 3.3, after some preparatory work in Sections 3.1 and 3.2.

3.1. ***Estimates for singular moduli.*** For $\tau \in \mathbb{H}$, set $q = q_\tau := e^{2\pi i\tau}$. Recall that the j-invariant function has the Fourier expansion

$$j(\tau) = \sum_{k=-1}^{\infty} c_k q^k,$$

where the c_k are positive integers, starting with $c_{-1} = 1,\ c_0 = 744$, and $c_1 = 196884$. (That they are positive integers follows from the formula

$$j(\tau) = q^{-1}\left(1 + 240\sum_{k=1}^{\infty}\sigma_3(k)q^k\right)^3 \prod_{n=1}^{\infty}(1-q^n)^{-24}, \quad \text{where} \quad \sigma_3(k) = \sum_{d\,|\,k} d^3;$$

see, for instance, [22, Chapter 1, Proposition 7.4 and Remark 7.4.1]. Clearly, each of the series

$$\left(1 + 240\sum_{k=1}^{\infty}\sigma_3(k)q^k\right)^3 \quad \text{and} \quad (1-q^n)^{-24} \quad (n = 1, 2, 3, \dots)$$

has positive integer coefficients; hence so does the Fourier expansion of j.)

Proposition 3.3. *Let* $\tau \in \mathbb{H}$, *and set* $v = \operatorname{Im}\tau$.

1. *Assume that* $v \ge 5$. *Then*

$$j(\tau) = q^{-1} + 744 + O_1(2\cdot 10^5|q|), \tag{3-10}$$

$$j(\tau) = q^{-1} + 744 + 196884q + O_1(3\cdot 10^7|q|^2), \tag{3-11}$$

$$\log|j(\tau)| = 2\pi v + O_1(800|q|), \tag{3-12}$$

$$\log(qj(\tau)) = 744q + O_1(5\cdot 10^5|q|^2), \tag{3-13}$$

$$\log(qj(\tau)) = 744q - 79884q^2 + O_1(2\cdot 10^8|q|^3). \tag{3-14}$$

2. *Assume that* $\tau \in \mathcal{F}$ *and* $v \le V$, *where* V *is a real number satisfying* $V \ge 5$. *Then*

$$\log|j(\tau)| \le 2\pi V + 3000e^{-2\pi V}. \tag{3-15}$$

We will use the following trivial lemma.

Lemma 3.4. *Let u be a complex number satisfying $|u| < 1$. Then*

$$|\log(1+u)| \le \frac{|u|}{1-|u|}.$$

Furthermore, for $n = 1, 2, \dots$ we have

$$\log(1+u) = \sum_{k=1}^{n} \frac{(-1)^{k-1}u^k}{k} + O_1\left(\frac{1}{n+1}\frac{|u|^{n+1}}{1-|u|}\right).$$

Proof of Proposition 3.3. Write $\tau = u + vi$. Then

$$q = e^{2\pi u i}e^{-2\pi v} \quad \text{and} \quad |q| = e^{-2\pi v}.$$

Let us prove item 1. We have $u \ge 5$, which implies that

$$|q| \le e^{-10\pi}.$$

For $n \ge 0$ write

$$j_n(\tau) = \sum_{k=n+1}^{\infty} c_k q^k$$

(so for example $j_0(\tau) = j(\tau) - q^{-1} - 744$ and $j_1(\tau) = j(\tau) - q^{-1} - 744 - 196884q$). Positivity of the coefficients c_k implies that

$$|j_n(\tau)q^{-n-1}| \le \sum_{k=n+1}^{\infty} c_k|q|^{k-n-1} \le \sum_{k=n+1}^{\infty} c_k e^{-10\pi(k-n-1)} = e^{10\pi(n+1)} j_n(5i).$$

In particular,

$$|j_0(\tau)q^{-1}| \le e^{10\pi} j_0(5i) < 2\cdot 10^5 \quad \text{and} \quad |j_1(\tau)q^{-2}| \le e^{20\pi} j_1(5i) < 3\cdot 10^7,$$

which proves expansions (3-10) and (3-11). Using Lemma 3.4, we deduce from them expansions (3-13) and (3-14), and (3-12) easily follows from (3-13).

Now let us prove item 2. We no longer have $v \ge 5$. However, since $\tau \in \mathcal{F}$, we have $v \ge \sqrt{3}/2$. Hence

$$|j(\tau)| \le |q|^{-1} + 744 + j_0(\sqrt{3}/2) \le e^{2\pi v} + 2079 \le e^{2\pi V}(1 + 2079e^{-2\pi V}).$$

Using Lemma 3.4, we obtain (3-15). □

We want to apply Proposition 3.3 to singular moduli. If x is a singular modulus, then there exists a unique $\tau_x \in \mathcal{F}$ such that $x = j(\tau_x)$. We have

$$\tau_x = \frac{b+\sqrt{\Delta}}{2a},$$

where $\Delta = \Delta_x$ is the discriminant, $a = a_x$ is the denominator of the singular modulus x, and $b \in \mathbb{Z}$; see Section 2.3 for details.

Corollary 3.5. *Let x be a singular modulus of discriminant Δ and denominator a.*

1. *Assume that $a \le 0.1|\Delta|^{1/2}$. Then*
$$\log|x| = \pi\frac{|\Delta|^{1/2}}{a} + O_1(e^{-2|\Delta|^{1/2}/a}).$$
2. *Let $A \ge 1$ be such that $a \ge A$ and $A \le 0.1|\Delta|^{1/2}$. Then*
$$\log|x| \le \pi\frac{|\Delta|^{1/2}}{A} + e^{-2|\Delta|^{1/2}/A}. \tag{3-16}$$
In particular, if $|\Delta| \ge 10^4$ then
$$\log|x| \le \pi|\Delta|^{1/2} + e^{-2|\Delta|^{1/2}} \le 4|\Delta|^{1/2}. \tag{3-17}$$
3. *Define τ_x as in Section 2.3. Assume that $a \le 0.1|\Delta|^{1/2}$. Then*
$$\log(xq) = 744q + O_1(5\cdot 10^5|q|^2), \tag{3-18}$$
$$\log(xq) = 744q - 79884q^2 + O_1(2\cdot 10^8|q|^3). \tag{3-19}$$
where $q = e^{2\pi i\tau_x}$.

Proof. The hypothesis that $a \le 0.1|\Delta|^{1/2}$ implies that
$$\operatorname{Im}\tau_x = \frac{|\Delta|^{1/2}}{2a} \ge 5.$$
Applying (3-12) with $\tau = \tau_x$, we obtain
$$\log|x| = \pi\frac{|\Delta|^{1/2}}{a} + O_1(800e^{-\pi|\Delta|^{1/2}/a}).$$
Since $|\Delta|^{1/2}/a \ge 10$, we have $800e^{-\pi|\Delta|^{1/2}/a} \le e^{-2|\Delta|^{1/2}/a}$. This proves item 1.

Similarly, item 2 follows by applying (3-15) with $V = |\Delta|^{1/2}/2A$. Note that (3-17) is a special case of (3-16) corresponding to $A = 1$.

Finally, (3-18) and (3-19) are obtained by setting $\tau = \tau_x$ in (3-13) and (3-14), respectively. □

We also need a lower bound. The following is a weaker version of [8, Theorem 6.1], applied with $y = 0$.

Proposition 3.6. *Let x be a singular modulus with discriminant $\Delta_x \ne -3$. Then $|x| \ge |\Delta_x|^{-3}$.*

3.2. ***Bounding the exponents.*** Let $\alpha_1, \dots, \alpha_k$ be nonzero algebraic numbers. The set of vectors $\boldsymbol{m} = (m_1, \dots, m_k) \in \mathbb{Z}^k$ such that
$$\alpha_1^{m_1}\cdots\alpha_k^{m_k} = 1$$
is a subgroup in $\mathbb{Z}^k$, denoted here by $\Gamma(\alpha_1, \dots, \alpha_k)$ (or simply Γ if this does not cause confusion).

Masser [18] showed that Γ admits a small $\mathbb{Z}$-basis. To state his result, let us introduce some notation. Let L be a number field. We denote by $\omega = \omega(L)$ the order of the group of roots of unity belonging to L, and by $\eta = \eta(L)$ the smallest positive height of the elements of L:

$$\eta = \min\{\mathrm{h}(\alpha) : \alpha \in L,\ \mathrm{h}(\alpha) > 0\}.^{1}$$

Proposition 3.7 (Masser). *Let $\alpha_1, \dots, \alpha_k$ be elements in $L^\times$. Then $\Gamma(\alpha_1, \dots, \alpha_k)$ has a $\mathbb{Z}$-basis consisting of vectors with norm bounded by $\omega(k\mathrm{h}/\eta)^{k-1}$, where*

$$\mathrm{h} = \max\{\mathrm{h}(\alpha_1), \dots, \mathrm{h}(\alpha_k), \eta\}.$$

We want to adapt this result to the case when our algebraic numbers are singular moduli.

Proposition 3.8. *Let $x_1, \dots, x_k$ be nonzero singular moduli. Set*

$$X = \max\{|\Delta_{x_1}|, \dots, |\Delta_{x_k}|\},$$

and assume that, among $D_{x_1}, \dots, D_{x_k}$, there are ℓ distinct fundamental discriminants; in symbols:

$$\ell = \#\{D_{x_1}, \dots, D_{x_k}\}.$$

Then the group $\Gamma(x_1, \dots, x_k)$ has a $\mathbb{Z}$-basis consisting of vectors with norm bounded by $24(c(\ell)kX^{1/2})^{k-1}$, where $c(\ell) = 3^{4^\ell + 2^{\ell+1} + 8}$. In particular, $c(1) = 3^{16}$.

The proof uses the following result due to Amoroso and Zannier [3, Theorem 1.2].

Lemma 3.9. *Let K be a number field of degree d, and let α be an algebraic number such $K(\alpha)$ is an abelian extension of K. Then either $\mathrm{h}(\alpha) = 0$ or $\mathrm{h}(\alpha) \geq 3^{-d^2-2d-6}$.*

Proof of Proposition 3.8. Set

$$K = \mathbb{Q}(\sqrt{D_{x_1}}, \dots, \sqrt{D_{x_k}}), \qquad L = K(x_1, \dots, x_k).$$

To apply Proposition 3.7, we have to estimate the quantities h, η and ω.

Since x_i is an algebraic integer, and every one of its conjugates x_i^σ satisfies $\log|x_i^\sigma| \leq 4|\Delta|^{1/2}$ (see Corollary 3.5:2), we have $\mathrm{h}(x_i) \leq 4X^{1/2}$. It follows that $\mathrm{h} \leq 4X^{1/2}$.

Since $[K : \mathbb{Q}] \leq 2^\ell$ and L is an abelian extension of K, Lemma 3.9 implies that $\eta \geq 3^{-4^\ell - 2^{\ell+1} - 6}$. Finally, we have $\omega \leq 24$ from Corollary 2.7.

Putting all of this together, the result follows. □

In this article we will often work with relations of the form

$$x_1^{m_1} \cdots x_k^{m_k} = (x_1')^{m_1} \cdots (x_k')^{m_k}.$$

It is useful to have a bounded basis for the group of these relations as well.

[1] Here $\mathrm{h}(\cdot)$ is the usual absolute logarithmic height; there is no risk of confusing it with the class number $h(\cdot)$, not only because the latter is in italics, but because class numbers do not occur in this section, and heights do not occur outside this section.

Proposition 3.10. *Let $x_1, \dots, x_k, x'_1, \dots, x'_k$ be singular moduli of discriminants not exceeding X, and let ℓ be the number of distinct fundamental discriminants among $D_{x_1}, \dots, D_{x'_k}$. Then the group $\Gamma(x_1/x'_1, \dots, x_k/x'_k)$ has a $\mathbb{Z}$-basis consisting of vectors with norm bounded by $24(c(\ell)kX^{1/2})^{k-1}$, where $c(\ell) = 3^{4^\ell+2^{\ell+1}+8}$. In particular, $c(1) = 3^{16}$.*

Proof. Same as for Proposition 3.8, only with $\mathrm{h} \le 8X^{1/2}$. □

3.3. *Proofs of Propositions 3.1 and 3.2.*

Proof of Proposition 3.1. By Proposition 3.8 we may assume that $\boldsymbol{m}$ has sup-norm

$$\|\boldsymbol{m}\| \le 24(3^{16}kX^{1/2})^{k-1}. \tag{3-20}$$

Using Corollary 3.5, we obtain

$$0 = \sum_{i=1}^{k} m_i \log|x_i| = \pi|D|^{1/2}L + O_1(k\|\boldsymbol{m}\|e^{-3Y^{1/2}/A}), \tag{3-21}$$

where L is the left-hand side of (3-2). (Recall that D denotes the common fundamental discriminant of $x_1, \dots, x_k$.) Using (3-5) and (3-20), we deduce from this the estimate $|L| \le 0.5A^{-k}$. Since L is a rational number with denominator not exceeding A^k, we must have $L = 0$. □

Proof of Proposition 3.2. We will prove only (3-8), because (3-9) is analogous.

Using Corollary 3.5 and Proposition 3.6, we obtain

$$0 = \frac{1}{\pi|\Delta|^{1/2}} \sum_{i=1}^{k} m_i \log|x_i| \ge \sum_{\substack{1\le i\le k\\ a(x_i)<A}} \frac{m'_i}{a(x_i)} - \sum_{m_i<0} \frac{|m'_i|}{\min\{a(x_i), A\}} + O_1\biggl(k\|\boldsymbol{m}'\| \frac{3\log X + e^{-3|\Delta|^{1/2}/A}}{\pi|\Delta|^{1/2}}\biggr),$$

Using our hypothesis (3-7), we obtain

$$\frac{3k\log X}{\pi|\Delta|^{1/2}} \le \frac{3}{\pi}\varepsilon, \qquad \frac{ke^{-3|\Delta|^{1/2}/A}}{\pi|\Delta|^{1/2}} \le 0.01\varepsilon,$$

and the result follows. □

4. Proof of Theorem 1.2

Throughout this section, unless the contrary is stated explicitly, x and y are distinct singular moduli and m, n are nonzero integers such that $\mathbb{Q}(x^m y^n) \ne \mathbb{Q}(x, y)$.

The proof of Theorem 1.2 is organized as follows. Assuming that

$$\max\{|\Delta_x|, |\Delta_y|\} \ge 10^6, \tag{4-1}$$

we show that one of the following two conditions is satisfied: either

$$\Delta_x = \Delta_y, \quad m = n, \quad [\mathbb{Q}(x, y) : \mathbb{Q}(x^m y^m)] = 2, \quad x \text{ and } y \text{ are conjugate over } \mathbb{Q}(x^m y^m) \tag{4-2}$$

(as wanted), or

$$\{\Delta_x, \Delta_y\} = \{\Delta, 4\Delta\} \quad \text{for some } \Delta \equiv 1 \bmod 8. \tag{4-3}$$

Unfortunately, we cannot rule out (4-3) assuming merely (4-1), but we show that (4-3) is impossible under the stronger hypothesis

$$\max\{|\Delta_x|, |\Delta_y|\} \geq 10^8, \tag{4-4}$$

completing thereby the proof.

We assume that x, y have the same fundamental discriminant (in the opposite case, the argument is much simpler: see Section 4.5.) We denote by K the common CM field of x, y, and we set $L = K(x, y)$. We also set

$$\alpha = x^m y^n, \quad F = \mathbb{Q}(\alpha), \quad G = \mathrm{Gal}(L/F).$$

Since F is a proper subfield of $\mathbb{Q}(x, y)$, there exists $\sigma \in G$ such that $x^\sigma \neq x$ or $y^\sigma \neq y$. We claim that

$$x^\sigma \neq x \quad \text{and} \quad y^\sigma \neq y. \tag{4-5}$$

Indeed, if, say, $y^\sigma = y$ then $(x^\sigma)^m = x^m$, which implies $x^\sigma = x$ by Lemma 2.22.

Remark 4.1. In the course of the argument we will study multiplicative relations

$$x^m y^n (x^\sigma)^{-m} (y^\sigma)^{-n} = 1, \tag{4-6}$$

with various choices of $\sigma \in G$ satisfying (4-5), and we will use Propositions 3.1 and 3.2 with parameters satisfying the following restrictions:

$$k \leq 4; \quad X = \max\{|\Delta_x|, |\Delta_y|\} \geq \begin{Bmatrix} 10^6 \\ 10^8 \end{Bmatrix}; \quad Y \geq \tfrac{1}{4}X; \quad A \leq 9; \quad \varepsilon = \begin{Bmatrix} 0.16 \\ 0.016 \end{Bmatrix}. \tag{4-7}$$

(The top/bottom alternation will be explained momentarily.) It is easy to verify that the conditions in (4-7) ensure that (3-5) and (3-7) are satisfied, so using the propositions is justified.

From here on through Section 4.4.1 we assume (4-1), and we use Proposition 3.2 with $X \geq 10^6$ and $\varepsilon = 0.16$. Starting from Section 4.4.2 we have (4-3) and assume (4-4), which will allow us to use Proposition 3.2 with $X \geq 10^8$ and $\varepsilon = 0.016$.

4.1. ***A special case.*** In this subsection we study the special case

$$m = -n \quad \text{and} \quad \Delta_x = \Delta_y. \tag{4-8}$$

We will need the result from this case to treat the general case. It will also be a good illustration of how our method works in a simple setup.

Let $\sigma \in G$ be such that (4-5) holds. We will apply Proposition 3.2 to the multiplicative relations

$$x^m y^{-m} (x^\sigma)^{-m} (y^\sigma)^m = 1, \tag{4-9}$$

$$x^m y^{-m} (x^{\sigma^{-1}})^{-m} (y^{\sigma^{-1}})^m = 1. \tag{4-10}$$

We may assume, up to Galois conjugation, that x is dominant. Then none of $y, x^\sigma, x^{\sigma^{-1}}$ is

$$a(x)=1, \quad a(y),\ a(x^\sigma),\ a(x^{\sigma^{-1}}) \ge 2.$$

If one of $a(y), a(x^\sigma)$ is ≥ 3, then, applying Proposition 3.2 to (4-9) with $A=3$ and $\varepsilon=0.16$, we obtain

$$m \le \left(\frac{1}{\min\{3, a(y)\}} + \frac{1}{\min\{3, a(x^\sigma)\}} + 0.16\right) m \le \left(\tfrac{1}{2}+\tfrac{1}{3}+0.16\right)m,$$

a contradiction.

Thus, we must have $a(y)=a(x^\sigma)=2$. This means that y and x^σ are either equal or 4-isogenous; see Proposition 2.21:3. Hence so are $y^{\sigma^{-1}}$ and x. Corollary 2.20 now implies that $a(y^{\sigma^{-1}}) \le 4$.

Applying Proposition 3.2 to (4-10) with $A=5$ and $\varepsilon=0.16$, we obtain

$$m + \frac{m}{a(y^{\sigma^{-1}})} \le m\left(\frac{1}{\min\{5, a(y)\}} + \frac{1}{\min\{5, a(x^{\sigma^{-1}})\}} + 0.16\right) \le m\left(\tfrac{1}{2}+\tfrac{1}{2}+0.16\right).$$

Since $a(y^{\sigma^{-1}}) \le 4$, we get a contradiction. We have proved that (4-8) is impossible.

4.2. ***The general case: preparations.*** Now we are ready to treat the general case. Pick $\sigma \in G$ satisfying (4-5). We have $(x/x^\sigma)^m = (y^\sigma/y)^n$, and, in particular, $\mathbb{Q}((x/x^\sigma)^m) = \mathbb{Q}((y/y^\sigma)^n)$. The special case of Theorem 1.2 treated in Section 4.1 implies that

$$\mathbb{Q}((x/x^\sigma)^m) = \mathbb{Q}(x, x^\sigma), \quad \mathbb{Q}((y/y^\sigma)^n) = \mathbb{Q}(y, y^\sigma).$$

Lemma 2.26:2 now implies that $K(x)=K(y)=L$, and Lemma 2.25 implies that there exists a discriminant Δ such that

$$\Delta_x = e_x^2\Delta, \quad \Delta_y = e_y^2\Delta, \quad (e_x, e_y) \in \{(1,1), (2,1), (1,2)\},$$

and, moreover,

$$\text{if } (e_x, e_y) \ne (1,1) \text{ then } \Delta \equiv 1 \bmod 8. \tag{4-11}$$

We may and will assume in the sequel that

$$m>0, \quad e_x m \ge e_y|n|, \quad a_x = 1. \tag{4-12}$$

If $(e_x, e_y) \ne (1,1)$ then x and y are not conjugate over $\mathbb{Q}$, and (4-5) becomes

$$x^\sigma, y^\sigma \notin \{x, y\}. \tag{4-13}$$

When $(e_x, e_y) = (1,1)$, Lemma 2.24 implies that σ can be redefined to satisfy either (4-13) or one of the following:

$$x^\sigma = y, \quad y^\sigma = x, \qquad [\mathbb{Q}(x,y):F]=2; \tag{4-14}$$

$$x^\sigma = y, \quad y^\sigma = z, \quad z^\sigma = x, \quad L=\mathbb{Q}(x,y), \quad [L:F]=3. \tag{4-15}$$

Case (4-14) is easy: relation (4-6) becomes $x^{m-n} = y^{m-n}$, and Lemma 2.22 implies that $m = n$, which means that we have (4-2).

We have to show that the other two cases are impossible. For (4-15) this is done in Section 4.3. Case (4-13) is much harder to dispose of; we deal with it in Section 4.4.

4.3. *Case* (4-15). We have

$$x^m y^{n-m} z^{-n} = 1.$$

Recall from (4-12) that $m \geq |n|$ and $a_x = 1$. This implies $a_y, a_z \geq 2$.

Assume first that $n > 0$. Then

$$\max\bigl\{m, |n-m|, |-n|\bigr\} = m.$$

Using Proposition 3.2 with $\varepsilon = 0.16$ and $A = 5$, we obtain

$$m \leq \frac{m-n}{\min\{5, a_y\}} + \frac{m}{\min\{5, a_z\}} n + 0.16m \leq \tfrac{1}{2}(m-n) + \tfrac{1}{2}n + 0.16m,$$

a contradiction.

Now assume that $n < 0$. Then

$$\max\bigl\{m, |n-m|, |-n|\bigr\} \leq 2m.$$

If $a_y \geq 3$ then Proposition 3.2 with $\varepsilon = 0.16$ and $A = 3$ implies that

$$m \leq \tfrac{1}{3} \cdot 2m + 0.16 \cdot 2m,$$

a contradiction. If $a_y = 2$ then x and y are 2-isogenous (see Proposition 2.21:1), and so are $y = x^\sigma$ and $z = y^\sigma$. This implies that $a_z \in \{1, 4\}$, see Corollary 2.20. But $a_z \geq 2$, and so $a_z = 4$. Proposition 3.1 now implies that

$$m + \frac{n-m}{2} - \frac{n}{4} = 0,$$

yielding $n = -2m$, again a contradiction. Thus, (4-15) is impossible.

4.4. *Case* (4-13). We will use the notation

$$m' = e_x m, \quad n' = e_y n.$$

Recall that $m' \geq |n'|$ and $a(x) = 1$; see (4-12). This implies, in particular, that $a(x^\sigma) \geq 2$.

4.4.1. *One of y, y^σ is dominant.* We start by showing that either y or y^σ is dominant.

Proposition 4.2. *If $n > 0$ then $a(y^\sigma) = 1$ and $\sigma^2 = 1$. If $n < 0$ then $a(y) = 1$. In both cases we have $(e_x, e_y) \neq (1, 1)$ and $\Delta \equiv 1 \bmod 8$.*

Proof. We treat separately $n > 0$ and $n < 0$.

Assume first that $n > 0$, but $a(y^\sigma) \ge 2$. We know already that $a(x^\sigma) \ge 2$. If one of $a(x^\sigma), a(y^\sigma)$ is ≥ 3 then, applying Proposition 3.2 with $A = 3$ and $\varepsilon = 0.16$ to the relation $x^m y^n (x^\sigma)^{-m} (y^\sigma)^{-n} = 1$, we obtain

$$m' \le \frac{m}{\min\{3, a(x^\sigma)\}} + \frac{n}{\min\{3, a(y^\sigma)\}} + 0.16m \le \bigl(\tfrac{1}{2} + \tfrac{1}{3} + 0.16\bigr)m,$$

a contradiction.

Thus, $a(x^\sigma) = a(y^\sigma) = 2$. This implies that $e_x = e_y = 1$, for otherwise $\Delta = 1 \bmod 8$ by (4-11), and one of Δ_x, Δ_y, being 4 mod 32, cannot admit singular moduli with denominator 2 (see Proposition 2.17:4).

Since $a(x^\sigma) = a(y^\sigma) = 2$ but $x^\sigma \ne y^\sigma$, the singular moduli x^σ and y^σ must be 4-isogenous; see Proposition 2.21:3. Hence so are x and y. Corollary 2.20 now implies that $a_y = 4$, and Proposition 3.1 yields

$$m' + \frac{n'}{4} - \frac{m'}{2} - \frac{n'}{2} = 0.$$

Hence $n' = 2m'$, again a contradiction. Thus, $n > 0$ implies that $a(y^\sigma) = 1$.

Note that if (4-13) holds for some $\sigma \in G$ then it also holds with σ replaced by σ^{-1}. Hence $n > 0$ implies that $a(y^{\sigma^{-1}}) = 1$ as well. Since there can be only one dominant singular modulus of a given discriminant, we must have $y^\sigma = y^{\sigma^{-1}}$. Hence $\sigma^2 = 1$ by Lemma 2.23.

Now assume that $n < 0$ but $a(y) \ge 2$. The same argument as above shows that $a(x^\sigma) = a(y) = 2$ and $e_x = e_y = 1$.

The singular moduli x^σ and y must be 4-isogenous. Hence so are x and $y^{\sigma^{-1}}$, which implies that $a(y^{\sigma^{-1}}) = 4$. Applying Proposition 3.2 with $A = 5$ and $\varepsilon = 0.16$ to

$$x^m y^{-|n|} (x^{\sigma^{-1}})^{-m} (y^{\sigma^{-1}})^{|n|} = 1,$$

we obtain

$$m' + \frac{|n'|}{4} \le \frac{|n'|}{2} + \frac{m'}{\min\{5, a(x^{\sigma^{-1}})\}} + 0.16m'.$$

Since $|n'| \le m'$ and $a(x^{\sigma^{-1}}) \ge 2$, this is impossible. Thus, we proved that $n < 0$ implies that $a(y) = 1$.

Finally, $(e_x, e_y) \ne (1, 1)$, because there cannot be two distinct dominant singular moduli of the same discriminant. Hence $\Delta \equiv 1 \bmod 8$ by (4-11). The proposition is proved. □

4.4.2. *Controlling the four denominators.* Thus, we know that two of the singular moduli x, y, x^σ, y^σ are dominant. Unfortunately, we have no control over the denominators of the other two.

We will now show that, with a suitably chosen Galois morphism θ, we can control the denominators of all four of $x^\theta, y^\theta, x^{\sigma\theta}, y^{\sigma\theta}$.

So far, we have assumed that $\max\{|\Delta_x|, |\Delta_y|\} \ge 10^6$ and used Proposition 3.2 with $\varepsilon = 0.16$. However, now we know that

$$\{\Delta_x, \Delta_y\} = \{\Delta, 4\Delta\} \quad \text{with } \Delta \equiv 1 \bmod 8,$$

which allows us (see Remark 4.1) to assume that $\max\{|\Delta_x|, |\Delta_y|\} \ge 10^8$ and to use Proposition 3.2 with $\varepsilon = 0.016$.

Proposition 4.3. *There exists $\theta \in \mathrm{Gal}(L/K)$ such that, when $(e_x, e_y) = (1, 2)$, we have*

$$a(x^\theta) = a(x^{\sigma\theta}) = 2, \quad a(y^\theta) = a(y^{\sigma\theta}) = 8, \tag{4-16}$$

and when $(e_x, e_y) = (2, 1)$, we have (4-16) *with x and y switched.*

To prove this proposition, we need to bound $|n'|$ from below.

Lemma 4.4. *Assume that $(e_x, e_y) = (2, 1)$. Then $|n'| \geq 0.85m'$.*

A similar estimate can be proved when $(e_x, e_y) = (1, 2)$, but we do not need this.

Proof. Assume first that $n < 0$. Then $a(x) = a(y) = 1$. In particular, x and y are 2-isogenous, and so are x^σ, y^σ. Write

$$x^m y^{-|n|} (x^\sigma)^{-m} (y^\sigma)^{|n|} = 1.$$

When $a(x^\sigma) \geq 8$ we use Proposition 3.2 with $A = 8$ and $\varepsilon = 0.016$ to obtain

$$m' \leq |n'| + \frac{m'}{8} + 0.016m',$$

which implies that $|n'| \geq 0.85m'$.

When $a(x^\sigma) \leq 7$, we must have $a(x^\sigma) \in \{3, 5, 7\}$ by Proposition 2.17:4, because $\Delta_x = 4\Delta \equiv 4 \bmod 32$. Since x^σ and y^σ are 2-isogenous, Corollary 2.20 implies that $a(y^\sigma) \in \{a(x^\sigma), a(x^\sigma)/4\}$, and we must have $a(y^\sigma) = a(x^\sigma)$. Using Proposition 3.1 with $A = 7$ we obtain

$$m' - |n'| - \frac{m'}{a(x^\sigma)} + \frac{|n'|}{a(x^\sigma)} = 0,$$

which shows that $|n'| = m'$. This proves the lemma in the case $n < 0$.

Now assume that $n > 0$. Then $a(x) = a(y^\sigma) = 1$ and $\sigma^2 = 1$. In particular, x and y^σ are 2-isogenous, and so are $x^{\sigma^{-1}} = x^\sigma$ and y. Arguing as above, we obtain that either $a(x^\sigma) \geq 8$ and $n' \geq 0.85m'$, or $a(x^\sigma) \in \{3, 5, 7\}$ and $n' = m'$. The lemma is proved. □

Proof of Proposition 4.3. Let us assume first that $n < 0$ and $(e_x, e_y) = (1, 2)$.

We have

$$\Delta_x = \Delta \equiv 1 \bmod 8, \quad \Delta_y = 4\Delta \equiv 4 \bmod 32. \tag{4-17}$$

By Proposition 2.17:1, there exist two distinct morphisms $\theta \in \mathrm{Gal}(L/K)$ such that $a(x^\theta) = 2$. Of the two, there can be at most one with the property $a(x^{\sigma\theta}) = 1$. Hence we may find θ satisfying

$$a(x^\theta) = 2, \quad a(x^{\sigma\theta}) \geq 2.$$

Since $n < 0$, we have $a(y) = 1$ by Proposition 4.2. Hence x and y are 2-isogenous, and so are x^θ and y^θ. It follows that $a(y^\theta) \in \{2, 8\}$. But $a(y^\theta) \neq 2$ by Proposition 2.17:4. Hence $a(y^\theta) = 8$.

Proposition 3.2, applied to

$$(x^\theta)^m (y^\theta)^{-|n|} (x^{\sigma\theta})^{-m} (y^{\sigma\theta})^{|n|} = 1$$

with $A = 9$ and $\varepsilon = 0.016$, implies that

$$\frac{m'}{2} \le \frac{|n'|}{8} + \frac{m'}{\min\{9, a(x^{\sigma\theta})\}} + 0.016m'.$$

If $a(x^{\sigma\theta}) \ge 3$ then this implies that

$$m'\left(\tfrac{1}{2} - \tfrac{1}{3}\right) \le \tfrac{1}{8}|n'| + 0.016m',$$

which is impossible because $m' \ge |n'|$. Hence $a(x^{\sigma\theta}) = 2$, and, as above, this implies that $a(y^{\sigma\theta}) = 8$.

Now assume that $n < 0$ and $(e_x, e_y) = (2, 1)$. We again have (4-17), but with x and y switched. Arguing as before, we find $\theta \in \mathrm{Gal}(L/K)$ such that

$$a(y^\theta) = 2, \quad a(y^{\sigma\theta}) \ge 2, \quad a(x^\theta) = 8.$$

As before, in the case $a(y^{\sigma\theta}) \ge 3$ we apply Proposition 3.2 to

$$(x^\theta)^{-m}(y^\theta)^{|n|}(x^{\sigma\theta})^m(y^{\sigma\theta})^{-|n|} = 1$$

and obtain

$$|n'|\left(\tfrac{1}{2} - \tfrac{1}{3}\right) \le \tfrac{1}{8}m' + 0.016m',$$

which is impossible because $|n'| \ge 0.85m'$. Hence $a(y^{\sigma\theta}) = 2$, and, as above, this implies that $a(x^{\sigma\theta}) = 8$.

Finally, let us assume that $n > 0$. Then $a(y^\sigma) = 1$ and $\sigma^2 = 1$. In particular, x and y^σ are 2-isogenous, and so are x^σ and $y^{\sigma^2} = y$. Now, writing

$$x^m(y^\sigma)^{-n}(x^\sigma)^{-m}y^n = 1,$$

we repeat the previous argument with y, y^σ switched, and with n replaced by $-n$. The proposition is proved. □

4.4.3. *Completing the proof.* Now we are ready to rule (4-13) out by deriving a contradiction. Let us summarize what we have. After renaming, we have distinct singular moduli x_1, x_2 of discriminant Δ and y_1, y_2 of discriminant 4Δ such that

$$a(x_1) = a(x_2) = 2, \quad a(y_1) = a(y_2) = 8,$$

and

$$x_1^{m_1}y_1^{n_1}x_2^{-m_1}y_2^{-n_1} = 1, \tag{4-18}$$

where m_1, n_1 is a permutation of m, n. We want to show that this impossible.

Proposition 3.10 implies that we may assume

$$\max\{|m_1|, |n_1|\} \le 10^{10}|\Delta|^{1/2}. \tag{4-19}$$

Note also that

$$|\Delta| \ge 10^7 \tag{4-20}$$

by the assumption (4-4).

Proposition 2.17 implies that, after possible renumbering, we have

$$\tau(x_1)=\frac{1+\sqrt{\Delta}}{4},\quad \tau(x_2)=\frac{-1+\sqrt{\Delta}}{4},\quad \tau(y_1)=\frac{b+\sqrt{\Delta}}{8},\quad \tau(y_2)=\frac{-b+\sqrt{\Delta}}{8},$$

where $b\in\{\pm1,\pm3\}$. Set $t=e^{-\pi|\Delta|^{1/2}/4}$ and $\xi=e^{b\pi i/4}$. Then

$$e^{2\pi i\tau(x_1)}=it^2,\quad e^{2\pi i\tau(x_2)}=-it^2,\quad e^{2\pi i\tau(y_1)}=\xi t,\quad e^{2\pi i\tau(y_2)}=\bar{\xi}t.$$

We deduce from (4-18) that

$$m_1\log(it^2x_1)-m_1\log(-it^2x_2)+n_1\log(\xi ty_1)-n_1\log(\bar{\xi}ty_2)\in\tfrac{1}{4}\pi i\mathbb{Z}. \tag{4-21}$$

Corollary 3.5:3 implies that

$$\begin{aligned}\log(it^2x_1)&=744it^2+O_1(5\cdot10^5t^4), &\qquad \log(-it^2x_2)&=-744it^2+O_1(5\cdot10^5t^4),\\ \log(\xi ty_1)&=744\xi t+O_1(5\cdot10^5t^2), &\qquad \log(\bar{\xi}ty_2)&=\ \ 744\bar{\xi}t+O_1(5\cdot10^5t^2).\end{aligned}$$

Transforming the left-hand side of (4-21) using these expansions, we obtain

$$744(\xi-\bar{\xi})tn_1+O_1\big(10^7t^2\max\{|m_1|,|n_1|\}\big)=\tfrac{1}{4}\pi ik$$

for some $k\in\mathbb{Z}$. An easy estimate using (4-19) and (4-20) shows that the left-hand side does not exceed 10^{-1000} in absolute value. Hence $k=0$, and we obtain, again using (4-19) and (4-20), that

$$744\,|\xi-\bar{\xi}||n_1|\le10^{17}|\Delta|^{1/2}e^{-\pi|\Delta|^{1/2}/4}<10^{-900}.$$

Hence $n_1=0$, a contradiction.

This proves the impossibility of (4-13), completing the proof of Theorem 1.2 in the case of equal fundamental discriminants.

4.5. ***Distinct fundamental discriminants.*** In this subsection $D_x\ne D_y$. Arguing as in the beginning of Section 4.2, we find a Galois morphism σ such that $\mathbb{Q}(x,x^\sigma)=\mathbb{Q}(y,y^\sigma)$. Lemma 2.26:1 implies that $\mathbb{Q}(x)=\mathbb{Q}(y)$. Corollary 2.13 implies that $h(\Delta_x)=h(\Delta_y)\le16$, and our hypothesis $\max\{|\Delta_x|,\Delta_y|\}\ge10^6$ contradicts Proposition 2.1. This concludes the proof of Theorem 1.2. □

5. Proof of Theorem 1.1

Throughout this section, unless stated otherwise, x,y,z are distinct singular moduli satisfying

$$\max\{|\Delta_x|,|\Delta_y|,|\Delta_z|\}\ge10^{10} \tag{5-1}$$

and such that there exist nonzero integers m,n,r with the property $x^my^nz^r\in\mathbb{Q}^\times$.

We assume that x,y,z have the same fundamental discriminant D; if this is not the case, then the argument is much simpler — see Section 5.5.

We denote by $K = \mathbb{Q}(\sqrt{D})$ the common CM field of x, y, z, and we set

$$L = K(x, y, z), \quad G = \mathrm{Gal}(L/K).$$

We set

$$f = \gcd(f_x, f_y, f_z), \quad e_x = \frac{f_x}{f}, \quad e_y = \frac{f_y}{f}, \quad e_z = \frac{f_z}{f}, \quad \Delta = Df^2.$$

Then $\gcd(e_x, e_y, e_z) = 1$ and

$$\Delta_x = e_x^2 \Delta, \quad \Delta_y = e_y^2 \Delta, \quad \Delta_z = e_z^2 \Delta. \tag{5-2}$$

5.1. ***The discriminants.*** The following property, showing that the ring class fields $K(x)$, $K(y)$ and $K(z)$ are closely related, is the basis for everything.

Proposition 5.1. (1) *We have*

$$L = K(x, y) = K(x, z) = K(y, z). \tag{5-3}$$

(2) *Each of the fields $K(x)$, $K(y)$, $K(z)$ is a subfield of L of degree at most* 2.

(3) *Up to permuting x, y, z (and correspondingly m, n, r) we have one of the cases from Table 2.*

Proof. By the assumption, $K(x^m) = K(y^n z^r) \subset K(y, z)$. Lemma 2.22 implies that $K(x) = K(y^n z^r)$, and Hence $x \in K(y, z)$. Hence $L = K(y, z)$. By symmetry, we obtain (5-3). This proves item (1).

From (5-1) we may assume that, for instance, $|\Delta_z| \geq 10^{10}$. Theorem 1.2 implies that the field $K(x) = K(y^n z^r)$ is a subfield of L of degree at most 2, and the same holds true for the fields $K(y)$. Unfortunately, we cannot make the same conclusion for $K(z)$, because, *a priori*, we cannot guarantee that $\max\{|\Delta_x|, |\Delta_y|\} \geq 10^8$, which is needed to apply Theorem 1.2 in this case. So a lengthy extra argument is required to prove that $[L : K(z)] \leq 2$. We split it into two cases.

Assume first that

$$\Delta \neq -3, -4 \quad \text{or} \quad \gcd(e_x, e_z) > 1. \tag{5-4}$$

In this case, setting $\ell = \mathrm{lcm}(e_x, e_y, e_z)$, Proposition 2.5 implies that $L = K[\ell f]$, the ring class field of K

e_x	$[L : K(x)]$	e_y	$[L : K(y)]$	e_z	$[L : K(z)]$	remarks
1	1	1	1	1	1	
1	1	1	1	2	1	$\Delta \equiv 1 \bmod 8$
1	1	2	1	2	1	$\Delta \equiv 1 \bmod 8$
1	2	2	1	2	1	$\Delta \equiv 0 \bmod 4$, $n = r$
1	2	3	1	3	1	$\Delta \equiv 1 \bmod 3$, $n = r$
2	2	3	1	3	1	$\Delta \equiv 1 \bmod 24$, $n = r$
1	2	4	1	4	1	$\Delta \equiv 1 \bmod 8$, $n = r$
1	2	6	1	6	1	$\Delta \equiv 1 \bmod 24$, $n = r$

Table 2. Data for Proposition 5.1.

of conductor ℓf. The class number formula (2-4) implies that

$$2 \geq [L : K(x)] = \Psi(\ell/e_x, \Delta_x), \tag{5-5}$$

which results in one of the following six cases:

$$\begin{aligned}
&\ell = e_x, && L = K(x); && \\
&\ell = 2e_x, && L = K(x), && \Delta_x \equiv 1 \bmod 8; \\
&\ell = 2e_x, && [L : K(x)] = 2, && \Delta_x \equiv 0 \bmod 4; \\
&\ell = 4e_x, && [L : K(x)] = 2, && \Delta_x \equiv 1 \bmod 8; \\
&\ell = 3e_x, && [L : K(x)] = 2, && \Delta_x \equiv 1 \bmod 3; \\
&\ell = 6e_x, && [L : K(x)] = 2, && \Delta_x \equiv 1 \bmod 24.
\end{aligned} \tag{5-6}$$

In particular, $e_x \geq \ell/6 \geq e_z/6$, which implies that $|\Delta_x| \geq |\Delta_z|/36 \geq 10^8$. Hence we may again apply Theorem 1.2 to conclude that $[L : K(z)] \leq 2$ in case (5-4).

We are left with the case

$$\Delta \in \{-3, -4\} \quad \text{and} \quad \gcd(e_x, e_z) = 1. \tag{5-7}$$

We want to show that it is impossible. We claim that in this case $\varphi(e_z) \leq 6$. Indeed, if $x \in K$ then

$$2 \geq [K(x, z) : K(x)] = [K(z) : K] = \frac{\Psi(e_z, \Delta)}{[\mathcal{O}_K^\times : \mathcal{O}_{K(z)}^\times]} \geq \frac{\Psi(e_z, \Delta)}{3},$$

which proves that $\varphi(e_z) \leq \Psi(e_z, \Delta) \leq 6$. And if $x \notin K$ then

$$\Psi(e_z, e_x^2\Delta) = \bigl[K[e_z e_x] : K[e_x]\bigr] = \bigl[K[e_z e_x] : K(x, z)\bigr] \cdot [K(x, z) : K(x)].$$

We have $\bigl[K[e_z e_x] : K(x, z)\bigr] \leq 3$ by Proposition 2.5, and

$$[K(x, z) : K(x)] = [L : K(x)] \leq 2,$$

as we have seen above. It follows that

$$\varphi(e_z) \leq \Psi(e_z, e_x^2\Delta) \leq 6.$$

From $\varphi(e_z) \leq 6$ we deduce $e_z \leq 18$. Hence $|\Delta_z| \leq 4 \cdot 18^2 < 10^{10}$, a contradiction. This shows the impossibility of (5-7), completing the proof of item (2).

We are left with part (3) of the proposition. As we have just seen, we have one of the six cases (5-6), and similarly with x replaced by y or by z. Since e_x, e_y, e_z are coprime, every prime number p does not divide one of them. If, say, $p \nmid e_x$, then

$$\nu_p(\ell) = \nu_p(\ell/e_x) \leq \begin{cases} 2 & \text{if } p = 2, \\ 1 & \text{if } p = 3, \\ 0 & \text{if } p \geq 5. \end{cases}$$

This proves that $\ell\,|\,12$. Moreover,

$$\text{if } 2|\ell \text{ then } \Delta\equiv 1 \bmod 8 \text{ or } \Delta \equiv 0 \bmod 4; \tag{5-8}$$

$$\text{if } 4|\ell \text{ then } \Delta\equiv 1 \bmod 8; \tag{5-9}$$

$$\text{if } 3|\ell \text{ then } \Delta\equiv 1 \bmod 3. \tag{5-10}$$

Indeed, assume that $2|\ell$ but $\Delta\equiv 5 \bmod 8$. Since e_x, e_y, e_z are coprime, one of them, say e_x, is not divisible by 2. Then $(\Delta_x/2)=-1$, and $\Psi(\ell/e_x, \Delta_x)$ must be divisible by 3, which contradicts (5-5). This proves (5-8).

In a similar fashion, one shows that $\Psi(\ell/e_x, \Delta_x)$ is divisible by 4 in each of the cases

$$4|\ell, \quad 2\nmid e_x, \quad \Delta\equiv 0 \bmod 4,$$
$$3|\ell, \quad 3\nmid e_x, \quad \Delta\equiv 2 \bmod 3,$$

and it is divisible by 3 in the case

$$3|\ell, \quad 3\nmid e_x, \quad \Delta\equiv 0 \bmod 3.$$

This proves (5-9) and (5-10).

It also follows from Theorem 1.2 that, when, say, $K(x)\neq L$, we must have $e_y=e_z$ and $n=r$.

A little PARI script (or verification by hand) shows that, up to permuting x, y, z, all possible cases are listed in Table 2. □

Recall that $f := \gcd(f_x, f_y, f_z)$ and $G := \mathrm{Gal}(L/K)$. Set $L_0 = K[f]$.

Corollary 5.2. 1. *Either*

$$L = L_0 = K(x) = K(y) = K(z)$$

or $[L : L_0] = 2$, *in which case exactly one of the fields* $K(x), K(y), K(z)$ *is* L_0 *and the other two are* L.

2. $[L_0 : K] \geq 101$.
3. *There exists* $\sigma\in G$ *such that* $a(x^\sigma), a(y^\sigma), a(z^\sigma) \geq 13$,
4. *There exists* $\sigma\in G$ *such that* $a(x^\sigma), a(y^\sigma) \geq 18$, *and the same statement is true for* x, z *and for* y, z.
5. *There exists* $\sigma\in G$ *such that* $a(x^\sigma)\geq 30$, *and the same statement holds true for* y *and for* z.

Proof. Item 1 is proved just by exploring Table 2. To prove item 2, note that, since $\max\{e_x, e_y, e_z\}\leq 6$, we have

$$|\Delta| \geq \max\{|\Delta_x|, |\Delta_y|, |\Delta_z|\}/36 \geq 10^8 \tag{5-11}$$

by (5-1). Hence $[L_0 : K] = h(\Delta) \geq 101$ by Proposition 2.1.

In proving item 3 we must distinguish the cases $L = L_0$ and $[L : L_0] = 2$. In the former case $L = K(x) = K(y) = K(z)$ and $|G| = [L : K] \geq 101$. Proposition 2.15 implies that there exist at most 32 elements $\sigma \in G$ such that $a(x^\sigma) < 13$, and the same for y and z. Since $|G| \ge 101 > 96$, we can find $\sigma \in G$ as wanted.

If $[L : L_0] = 2$ then, say,

$$K(x) = L_0, \quad K(y) = K(z) = L,$$

and $|G| = [L : K] \ge 202$. Again using Proposition 2.15, there exist at most 32 elements $\sigma \in G$ such that $a(y^\sigma) < 13$, the same for z, and at most 64 elements $\sigma \in G$ such that $a(x^\sigma) < 13$. Since $|G| \ge 202 > 128$, we again can find a σ as wanted. This proves item 3.

Item 4 is proved similarly. In the case $L = K(x) = K(y)$ there exist at most 48 elements $\sigma \in G$ such that $a(x^\sigma) < 18$, and the same for y. Since $|G| \ge 101 > 96$, we are done. In the case when one of $K(x)$, $K(y)$ is L, the other is L_0, and $[L : L_0] = 2$, we have $48 + 96 = 144$ unsuitable $\sigma \in G$; since $|G| \ge 202$, we are done again.

Item 5 is similar as well: there exist at most 99 unsuitable σ when $L = K(x)$, and at most 198 when $[L : K(x)] = 2$; in both cases we conclude as before. □

In the sequel we set

$$m' = me_x, \quad n' = ne_y, \quad r' = re_z.$$

We may and will assume that $m > 0$ and that

$$m' \ge \max\{|n'|, |r'|\}, \quad a_x = 1. \tag{5-12}$$

In the course of the argument we will study multiplicative relations

$$x^m y^n z^r (x^\sigma)^{-m} (y^\sigma)^{-n} (z^\sigma)^{-r} = 1, \tag{5-13}$$

with various choices of $\sigma \in G$, using Propositions 3.1 and 3.2. In our usage of Propositions 3.1 and 3.2 the parameters therein will satisfy the following restrictions:

$$\begin{aligned} &k = 6, \quad X = \max\{|\Delta_x|, \Delta_y|, |\Delta_z|\} \ge 10^{10}, \quad Y = |\Delta| \ge \tfrac{1}{36} X,\\ &A \le 162 \qquad\qquad\quad \text{for Proposition 3.1,}\\ &A \le 30, \quad \varepsilon = 0.01 \quad \text{for Proposition 3.2.} \end{aligned} \tag{5-14}$$

It is easy to verify that for any choice of parameters satisfying (5-14), conditions (3-5) and (3-7) are met, so using the propositions is justified.

5.2. *The denominators.* We already know that x is dominant, see (5-12). Our principal observation is that either one of y, z is dominant as well, or they both are subdominant. More precisely:

Proposition 5.3. *Up to interchanging y and z, one of the following conditions is satisfied*:

$$a_y = 1 \qquad \textit{and} \quad n < 0; \tag{5-15}$$

$$a_y = a_z = 2 \quad \textit{and} \quad n, r < 0. \tag{5-16}$$

Proof. With y and z possibly switched, we may assume that we are in one of the following cases:

$$n, r > 0; \tag{5-17}$$

$$n < 0 \quad \text{and} \quad r > 0; \tag{5-18}$$

$$n, r < 0. \tag{5-19}$$

We consider them separately.

Assume (5-17). Let σ be as in Corollary 5.2:3. Applying Proposition 3.2 to

$$x^m y^n z^r (x^\sigma)^{-m} (y^\sigma)^{-n} (z^\sigma)^{-r} = 1$$

with $A = 13$ and $\varepsilon = 0.01$, and using that $\max\{|m'|, |n'|, |r'|\} = m'$ by (5-12), we obtain

$$m' \le \frac{m'}{13} + \frac{n'}{13} + \frac{r'}{13} + 0.01m' \le m'\bigl(\tfrac{3}{13} + 0.01\bigr),$$

a contradiction. This shows that (5-17) is impossible.

Now assume (5-18). We want to show that $a_y = 1$ in this case. Thus, assume that $a_y \ge 2$. Using Corollary 5.2:4, we find $\sigma \in G$ such that $a(x^\sigma), a(z^\sigma) \ge 18$. Applying Proposition 3.2 to

$$x^m y^{-|n|} z^r (x^\sigma)^{-m} (y^\sigma)^{|n|} (z^\sigma)^{-r} = 1$$

with $A = 18$ and $\varepsilon = 0.01$, we obtain

$$m' \le \frac{|n'|}{\min\{18, a(y)\}} + \frac{m'}{18} + \frac{r'}{18} + 0.01m' \le m'\bigl(\tfrac{1}{2} + \tfrac{2}{18} + 0.01\bigr),$$

a contradiction. This shows that in the case (5-18) we must have (5-15).

Finally, let us assume (5-19). If one of a_y, a_z is 1 then we have (5-15), possibly after switching. If $a_y = a_z = 2$ then we have (5-16). Now let us assume that none of these is the case; that is, both a_y, a_z are ≥ 2 and one of them is ≥ 3. Again using Corollary 5.2, we may find $\sigma \in G$ such that $a(x^\sigma) \ge 30$. Applying Proposition 3.2, we obtain, in the same fashion as in the previous cases, the inequality

$$m' \le m'\bigl(\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{30} + 0.01\bigr),$$

a contradiction. The proposition is proved. □

We study cases (5-15) and (5-16) in Sections 5.3 and 5.4, respectively.

5.3. ***The dominant case.*** In this subsection we assume (5-15). Thus, we have

$$m > 0, \quad n < 0, \quad m' \ge \max\{|n'|, |r'|\}, \quad a_x = a_y = 1.$$

Since both x and y are dominant, we must have $e_x \ne e_y$. Exploring Table 2, we find ourselves in one of the following cases:

$$\{e_x, e_y\} = \{1, 2\}, \quad e_z \in \{1, 2\}, \quad \Delta \equiv 1 \bmod 8, \tag{5-20}$$
$$\{e_x, e_y\} = \{1, 2\}, \quad e_z = 2, \quad \Delta \equiv 0 \bmod 4, \tag{5-21}$$
$$\{e_x, e_y\} = \{1, 3\}, \quad e_z = 3, \quad \Delta \equiv 1 \bmod 3, \tag{5-22}$$
$$\{e_x, e_y\} = \{2, 3\}, \quad e_z = 3, \quad \Delta \equiv 1 \bmod 24, \tag{5-23}$$
$$\{e_x, e_y\} = \{1, 4\}, \quad e_z = 4, \quad \Delta \equiv 1 \bmod 8, \tag{5-24}$$
$$\{e_x, e_y\} = \{1, 6\}, \quad e_z = 6, \quad \Delta \equiv 1 \bmod 24. \tag{5-25}$$

Remark 5.4. It is crucial that, in each of these cases, a nontrivial congruence condition is imposed on Δ. This allows us to use Propositions 2.16 and 2.17 to find Galois morphisms σ with well-controlled denominators of x^σ, y^σ, z^σ, which is needed for the strategy described in Section 5.3.1 to work.

Here are some more specific observations.

1. We have either $e_z = e_x$ or $e_z = e_y$, which implies that
$$a_z \neq 1. \tag{5-26}$$
2. In case (5-20) we have $K(x) = K(y) = K(z) = L$.
3. In cases (5-21)–(5-25) we have $K(z) = L$, and one of the fields $K(x)$ or $K(y)$ is L as well, while the other is a degree 2 subfield of L. More precisely:
   - If $e_x < e_y = e_z$ then $K(y) = L$ and $[L : K(x)] = 2$.
   - If $e_y < e_x = e_z$ then $K(x) = L$ and $[L : K(y)] = 2$.
4. Theorem 1.2 implies that in cases (5-21)–(5-25) we have $n = r$ when $e_x < e_y$, and $m = r$ when $e_x > e_y$.

5.3.1. *The strategy.* In each of cases (5-20)–(5-25) we apply the following strategy.

- Find possible values for a_z.
- Using Proposition 2.16 or 2.17, find several $\sigma \in G$ such that we can control the denominators
$$a(x^\sigma), \quad a(y^\sigma), \quad a(z^\sigma). \tag{5-27}$$
- For every such σ, and every possible choice of a_z and of denominators (5-27), Proposition 3.1 implies the linear equation
$$m' + n' + \frac{r'}{a_z} = \frac{m'}{a(x^\sigma)} + \frac{n'}{a(y^\sigma)} + \frac{r'}{a(z^\sigma)}.$$

  With sufficiently many choices of σ, we may hope to have enough equations to conclude that $m' = n' = r' = 0$, a contradiction.

Practical implementation of this strategy differs from case to case. For instance, in cases (5-21)–(5-25) we have $m' = r'$ or $n' = r'$, so we need only two independent equations to succeed, while in case (5-20) three independent equations are needed.

Case (5-21) is somewhat special, because we get only one equation. To complete the proof in that case, we need to use an argument similar to that of Section 4.4.3.

Below details for all the cases follow.

5.3.2. *Cases* (5-22)–(5-25). In these cases $K(z) = L$, and one of the fields $K(x)$, $K(y)$ is also L while the other is a degree 2 subfield of L. In this subsection we make no use of the assumption $m' \geq |n'|$. Hence we may assume that $e_x < e_y = e_z$, in which case we have

$$K(y) = K(z) = L, \quad [L : K(x)] = 2. \tag{5-28}$$

Theorem 1.2 implies that in this case $n = r$, and that y, z are conjugate over $K(x)$:

$$y^\theta = z, \quad z^\theta = y, \tag{5-29}$$

where θ is the nontrivial element of $\mathrm{Gal}(L/K(x))$.

Let us specify the general strategy described in Section 5.3.1 for the cases (5-22)–(5-25).

1. Proposition 2.21 implies that x and y are ℓ-isogenous, where $\ell = e_x e_y$. Hence $x = x^\theta$ and $z = y^\theta$ are ℓ-isogenous as well. Using Corollary 2.20, we may now shortlist possible values of the denominator a_z. Precisely,

$$a_z \in \left(\frac{e_z}{e_x}\mathcal{Q}(\ell)\right) \cap \mathbb{Z}_{\geq 2},$$

where we use the notation $\lambda S = \{\lambda s : s \in S\}$. For instance, in case (5-23) we have $\ell = 6$, and

$$a_z \in \left(\tfrac{3}{2}\left\{\tfrac{1}{6}, \tfrac{2}{3}, \tfrac{3}{2}, 6\right\}\right) \cap \mathbb{Z}_{\geq 2} = \{9\}.$$

2. Propositions 2.16:1 and 2.17:1 imply the existence of morphisms σ_1 and σ_2 such that the three denominators $a(x)$ (which is 1), $a(x^{\sigma_1})$ and $a(x^{\sigma_2})$ are distinct. Precisely:

- If $\Delta_x \equiv 1 \bmod 3$ then 3 and 9 are denominators for Δ_x.
- If $\Delta_x \equiv 1 \bmod 8$ then 2 and 4 are denominators for Δ_x.

For instance, in case (5-23) we may find σ_1 and σ_2 to have

$$a(x^{\sigma_1}) = 3, \quad a(x^{\sigma_2}) = 9.$$

3. Using again Corollary 2.20, we may now shortlist the denominators $a(y^{\sigma_i})$ and $a(z^{\sigma_i})$. Precisely,

$$a(y^{\sigma_i}), a(z^{\sigma_i}) \in \left(a(x^{\sigma_i})\frac{e_z}{e_x}\mathcal{Q}(\ell)\right) \cap \mathbb{Z}_{\geq 1}.$$

For instance, in case (5-23) we have

$$a(y^{\sigma_1}), a(z^{\sigma_1}) \in \left(3 \cdot \tfrac{3}{2}\left\{\tfrac{1}{6}, \tfrac{2}{3}, \tfrac{3}{2}, 6\right\}\right) \cap \mathbb{Z}_{\geq 1} = \{3, 27\},$$
$$a(y^{\sigma_2}), a(z^{\sigma_2}) \in \left(9 \cdot \tfrac{3}{2}\left\{\tfrac{1}{6}, \tfrac{2}{3}, \tfrac{3}{2}, 6\right\}\right) \cap \mathbb{Z}_{\geq 1} = \{9, 81\}.$$

4. Now, Proposition 3.1 implies the system of linear equations

$$m' + \left(1 + \frac{1}{a(z)}\right) n' = \frac{m'}{a(x^{\sigma_i})} + \left(\frac{1}{a(y^{\sigma_i})} + \frac{1}{a(z^{\sigma_i})}\right) n' \quad (i = 1, 2). \tag{5-30}$$

(Recall that $n' = r'$). Solving the system, we find that $m' = n' = 0$ in every instance, a contradiction. This shows the impossibility of cases (5-22)–(5-25).

For instance, in case (5-23), equations (5-30) become

$$m' + \left(1 + \tfrac{1}{9}\right) n' = \tfrac{1}{3} m' + \lambda n', \qquad \lambda \in \left\{\tfrac{2}{3}, \tfrac{1}{3} + \tfrac{1}{27}, \tfrac{2}{27}\right\},$$
$$m' + \left(1 + \tfrac{1}{9}\right) n' = \tfrac{1}{9} m' + \mu n', \qquad \mu \in \left\{\tfrac{2}{9}, \tfrac{1}{9} + \tfrac{1}{81}, \tfrac{2}{81}\right\},$$

so nine systems in total, each having $m' = n' = 0$ as its only solution.

The numerical data obtained following these steps can be found in Table 3. We have 390 linear systems to solve: nine systems in cases (5-22), (5-23), 72 systems in case (5-24), and 300 systems in case (5-25). Doing this by hand is impractical, and we used a PARI script for composing Table 3 and solving the systems.

Remark 5.5. Using Propositions 2.16 and 2.17, we can further refine the lists of possible denominators for z, y^{σ_i} and z^{σ_i}. For instance, if the discriminant $\Delta_y = \Delta_z$ is $\equiv 0 \bmod 9$ then it cannot have denominators divisible by 3 but not by 9. Thus, in case (5-23), number 3 cannot be the denominator of y^{σ_1} or of z^{σ_1}, and so we must have $a(y^{\sigma_1}) = a(z^{\sigma_1}) = 27$. Arguments of this kind, used systematically, allow one to decimate the number of systems to solve.

However, the computational time for solving our systems being insignificant, we prefer to disregard this observation.

5.3.3. *Case* (5-21). This case is similar to cases (5-22)–(5-25), but somewhat special. Let us reproduce our data for the reader's convenience:

$$\{e_x, e_y\} = \{1, 2\}, \qquad e_z = 2, \qquad \Delta \equiv 0 \bmod 4, \qquad a_x = a_y = 1.$$

We may again assume that $e_x < e_y$, which means now that $e_x = 1$ and $e_y = 2$, and we again have (5-28) and (5-29). Step 1 of the strategy described in Section 5.3.2 works here as well: we prove that each

case	$\Delta \equiv$	e_x	e_y	ℓ	a_z	$a(x^{\sigma_1})$	$a(y^{\sigma_1}), a(z^{\sigma_1})$	$a(x^{\sigma_2})$	$a(y^{\sigma_2}), a(z^{\sigma_2})$	no. of systems
(5-22)	1 mod 3	1	3	3	9	3	$\in \{3, 27\}$	9	$\in \{9, 81\}$	9
(5-23)	1 mod 24	2	3	6	9	3	$\in \{3, 27\}$	9	$\in \{9, 81\}$	9
(5-24)	1 mod 8	1	4	4	$\in \{4, 16\}$	2	$\in \{2, 8, 32\}$	4	$\in \{4, 16, 64\}$	72
(5-25)	1 mod 24	1	6	6	$\in \{4, 9, 36\}$	2	$\in \{2, 8, 18, 72\}$	3	$\in \{3, 12, 27, 108\}$	300
(5-21)	4 mod 32	1	2	2	4	8	$\in \{8, 32\}$	16	$\in \{16, 64\}$	9

Table 3. Cases (5-22)–(5-25) and case (5-21) with $\Delta \equiv 4 \bmod 32$.

of y, z is 2-isogenous to x, which allows us to determine that $a_z = 4$. For later use, let us note that

$$\tau_x = \frac{\sqrt{\Delta}}{2}, \quad \tau_y = \sqrt{\Delta}, \quad \tau_z = \frac{b' + \sqrt{\Delta}}{4}, \tag{5-31}$$

where

$$b' = \begin{cases} 0 & \text{if } \Delta \equiv 4 \bmod 8, \\ 2 & \text{if } \Delta \equiv 0 \bmod 8. \end{cases} \tag{5-32}$$

The rest of the argument splits into two subcases. If $\Delta \equiv 4 \bmod 32$ then we may proceed as in Section 5.3.2. Proposition 2.17 implies that there exist $\sigma_1, \sigma_2 \in G$ such that $a(x^{\sigma_1}) = 8$ and $a(x^{\sigma_2}) = 16$. As before, we can now determine possible denominators of $y^{\sigma_i}, z^{\sigma_i}$ (see the bottom line of Table 3) and solve the resulting systems (5-30), concluding that $m = n = 0$.

Now assume that $\Delta \not\equiv 4 \bmod 32$. In this case 2 or 4 is a denominator for Δ; see Proposition 2.17:6. Since $\Delta_x = \Delta$, there exists $\sigma \in G$ such that $a(x^\sigma) \in \{2, 4\}$. We claim that

$$y^\sigma, z^\sigma \notin \{y, z\}. \tag{5-33}$$

Indeed, if, say, $y^\sigma = y$ then $\sigma = \mathrm{id}$ because $L = K(y)$; but $x^\sigma \neq x$, a contradiction. For the same reason, $z^\sigma \neq z$. Now assume that $y^\sigma = z$. Theorem 1.2 implies that y and z are conjugate over $K(x)$. Hence there exists $\theta \in G$ such that

$$x^\theta = x, \quad y^\theta = z, \quad z^\theta = y.$$

Then $z^{\theta\sigma} = z$, and, as before, $\theta\sigma = \mathrm{id}$, which is again a contradiction because $x^{\theta\sigma} = x^\sigma \neq x$. Similarly one shows that $z^\sigma \neq y$. This proves (5-33).

The cases $a(x^\sigma) = 2$ and $a(x^\sigma) = 4$ are very similar, but each one has some peculiarities, so we consider them separately.

Assume that $a(x^\sigma) = 2$. Then $a(y^\sigma) = a(z^\sigma) = 8$. Proposition 3.1 gives

$$m' + n'\left(1 + \tfrac{1}{4}\right) = \tfrac{1}{2}m' + n'\left(\tfrac{1}{8} + \tfrac{1}{8}\right),$$

which is just $m' = -2n'$. Hence $m = -4n$. It follows that $(x/x^\sigma)^4(y^\sigma/y)(z^\sigma/z)$ is a root of unity. Since the roots of unity in L are of order dividing 24 (see Corollary 2.7), we obtain

$$\left(x^4(x^\sigma)^{-4}y^{-1}z^{-1}y^\sigma z^\sigma\right)^{24} = 1. \tag{5-34}$$

Now we are going to argue as in Section 4.4.3. This means:

- We give explicit expressions for the τ- and q-parameters of all the six singular moduli occurring in (5-34). The τ-parameters for x, y, z are already given in (5-31), so we need to determine them only for $x^\sigma, y^\sigma, z^\sigma$.
- Taking the logarithm of (5-34), we deduce that a certain linear combination of logarithms is a multiple of $\pi i/12$.
- Using the q-expansion from Corollary 3.5, we obtain a contradiction.

Note, however, that in Section 4.4.3 the first order expansion (3-18) was sufficient, while now we would need the second order expansion (3-19).

Since y, z, x^σ are distinct and 2-isogenous to x, we must have, in addition to (5-31), (5-32),

$$\tau(x^\sigma) = \frac{b_2+\sqrt{\Delta}}{4}, \quad \text{where} \quad b_2 = \begin{cases} 0 & \text{if } b'=2, \\ 2 & \text{if } b'=0. \end{cases} \tag{5-35}$$

Since x, y^σ, z^σ are distinct and 2-isogenous to x^σ, we must have

$$\{\tau(y^\sigma), \tau(z^\sigma)\} = \left\{\frac{b_2+\sqrt{\Delta}}{8}, \frac{b_2'+\sqrt{\Delta}}{8}\right\}, \quad b_2' \in \{b_2+4, b_2-4\}.$$

Set $t = e^{-\pi|\Delta|^{1/2}/4}$ and $\xi = e^{\pi i b_2/4}$. Note that $\xi \in \{1, i\}$, and that

$$e^{\pi i b'/2} = -\xi^2, \quad e^{\pi i b_2'/4} = -\xi.$$

We obtain

$$e^{2\pi i \tau_x} = t^4, \quad e^{2\pi i \tau_y} = t^8, \quad e^{2\pi i \tau_z} = -\xi^2 t^2, \quad e^{2\pi i \tau(x^\sigma)} = \xi^2 t^2, \quad \{e^{2\pi i \tau(y^\sigma)}, e^{2\pi i \tau(z^\sigma)}\} = \{\xi t, -\xi t\}.$$

Taking the logarithm of (5-34), we obtain

$$4\log(t^4 x) - 4\log(\xi^2 t^2 x^\sigma) - \log(t^8 y) - \log(-\xi^2 t^2 z) + \log(-\xi t \cdot \xi t \cdot y^\sigma \cdot z^\sigma) \in \frac{\pi i}{12}\mathbb{Z}.$$

The q-expansion (3-19) from Corollary 3.5 implies that for some $k \in \mathbb{Z}$ we have

$$\frac{\pi i}{12} k = -162000\xi^2 t^2 + O_1(10^{10} t^3).$$

This easily leads to a contradiction, exactly as in Section 4.4.3.

Now assume that $a(x^\sigma)=4$. Since x^σ is 2-isogenous to y^σ and to z^σ, Corollary 2.20 and Proposition 2.17 imply that $a(y^\sigma), a(z^\sigma) \in \{4, 16\}$. Note however that

$$\Delta_y = \Delta_z = 4\Delta \equiv 0 \bmod 16,$$

and Proposition 2.17:5 implies that there may be at most one singular modulus of this discriminant with denominator 4. But we already have $a_z = 4$, and so neither of $a(y^\sigma), a(z^\sigma)$ can equal 4 by (5-33).

Thus, $a(y^\sigma) = a(z^\sigma) = 16$. Proposition 3.1 gives

$$m' + n'\left(1+\tfrac{1}{4}\right) = \tfrac{1}{4}m' + n'\left(\tfrac{1}{16}+\tfrac{1}{16}\right),$$

which is $m = -3n$. Arguing as before, we obtain

$$\left(x^3 (x^\sigma)^{-3} y^{-1} z^{-1} y^\sigma z^\sigma\right)^{24} = 1. \tag{5-36}$$

We have

$$\tau(x^\sigma) = \frac{b_4+\sqrt{\Delta}}{8},$$

and we want to specify this b_4. Since $(b_4)^2 \equiv \Delta \bmod 16$, we must have

$$\Delta \equiv 0, 4 \bmod 16 \quad \text{and} \quad b_4 = \begin{cases} 0 & \text{if } \Delta \equiv 16 \bmod 32, \\ 4 & \text{if } \Delta \equiv 0 \bmod 32, \\ \pm 2 & \text{if } \Delta \equiv 4 \bmod 16. \end{cases}$$

Hence,

$$b' \in \{b_4 + 2, b_4 - 2\}.$$

Finally, since both y^σ, z^σ have denominators 16 and are 2-isogenous to x^σ, we have

$$\{\tau(y^\sigma), \tau(z^\sigma)\} = \left\{ \frac{b_4 + \sqrt{\Delta}}{16}, \frac{b_4' + \sqrt{\Delta}}{16} \right\} \quad \text{and} \quad b_4' \in \{b_4 + 8, b_4 - 8\}.$$

Set $t = e^{-\pi|\Delta|^{1/2}/8}$ and $\xi = e^{\pi i b_4/8}$. Note that $\xi \in \{1, i, e^{\pm\pi i/4}\}$, and that

$$e^{\pi i b'/4} = \pm i \xi^2, \quad e^{\pi i b_4'/8} = -\xi.$$

We obtain

$$e^{2\pi i \tau_x} = t^8, \quad e^{2\pi i \tau_y} = t^{16}, \quad e^{2\pi i \tau_z} = \varepsilon i \xi^2 t^4, \quad e^{2\pi i \tau(x^\sigma)} = \xi^2 t^2, \quad \left\{ e^{2\pi i \tau(y^\sigma)}, e^{2\pi i \tau(z^\sigma)} \right\} = \{\xi t, -\xi t\},$$

where $\varepsilon \in \{1, -1\}$. Taking the logarithm of (5-36), we obtain

$$3\log(t^8 x) - 3\log(\xi^2 t^2 x^\sigma) - \log(t^{16} y) - \log(-\varepsilon i \xi^2 t^4 z) + \log(-\xi t \cdot \xi t \cdot y^\sigma \cdot z^\sigma) \in \frac{\pi i}{12}\mathbb{Z}.$$

The q-expansion (3-19) implies that for some $k \in \mathbb{Z}$ we have

$$\frac{\pi i}{12} k = -162000 \xi^2 t^2 + O_1(10^{10} t^3),$$

which again leads to a contradiction.

5.3.4. *Case* (5-20). We want to adapt the procedure described in Section 5.3.2 to this case. We reproduce our data for the reader's convenience:

$$\{e_x, e_y\} = \{1, 2\}, \quad e_z \in \{1, 2\}, \quad \Delta \equiv 1 \bmod 8, \quad a_x = a_y = 1. \tag{5-37}$$

The singular moduli x and y are 2-isogenous by Proposition 2.21. However, now we have $K(x) = K(y) = K(z) = L$, which means that there does not exist $\theta \in G$ with the properties $x^\theta = x$ and $y^\theta = z$. Hence, a priori we have no control of the degree of isogeny between x and z. To gain such control we need to determine the denominator a_z.

Proposition 5.6. *Assume* (5-37). *Then* $(e_x, e_y) = (1, 2)$ *and*

$$\textit{either} \quad e_z = 1, \quad a_z = 4 \quad \textit{or} \quad e_z = 2, \quad a_z = 3. \tag{5-38}$$

The proof consists of several steps. To start with, we eliminate the subcase $(e_x, e_y) = (2, 1)$.

Proposition 5.7. *In case* (5-37) *we must have* $(e_x, e_y) = (1, 2)$.

Proof. Note that $a_z > 1$; see (5-26). We will assume that $(e_x, e_y) = (2, 1)$ and get a contradiction.

Since $\Delta_y = \Delta \equiv 1 \bmod 8$, Proposition 2.17 implies that there are two elements $\sigma \in G$ such that $a(y^\sigma) = 2$. Since $L = K(z)$, at most one of them may satisfy $a(z^\sigma) = 1$. Hence there exists $\sigma \in G$ such that $a(y^\sigma) = 2$ and $a(z^\sigma) \geq 2$.

Since x and y are 2-isogenous, we must have $a(x^\sigma) \in \{2, 8\}$. But 2 is not a denominator for $\Delta_x = 4\Delta$ by Proposition 2.17, which implies that $a(x^\sigma) = 8$. Thus, we have found σ such that

$$a(x^\sigma) = 8, \quad a(y^\sigma) = 2, \quad a(z^\sigma) \geq 2.$$

We now want to derive a contradiction in each of the following cases:

$$\text{One of } a(z), a(z^\sigma) \text{ is } 2. \tag{5-39}$$

$$\text{Both } a(z), a(z^\sigma) \text{ are at least } 3. \tag{5-40}$$

Assume (5-39). Then $e_z = 1$, again by the same reason: 2 is not a denominator for 4Δ. Hence there exists $\theta \in G$ such that $y^\theta = z$. Since y, y^σ are 2-isogenous, so are $z = y^\theta$ and $z^\sigma = y^{\theta\sigma} = y^{\sigma\theta}$. It follows that if one of the denominators $a(z), a(z^\sigma)$ is 2, the other must be 4. Proposition 3.1 now implies that

$$m' + n' + \frac{r'}{a'} = \frac{m'}{8} + \frac{n'}{2} + \frac{r'}{a''} \quad \text{and} \quad \{a', a''\} = \{2, 4\}.$$

Hence

$$\frac{7}{8}m' = \frac{|n'|}{2} + r'\left(\frac{1}{a''} - \frac{1}{a'}\right) \leq m'\left(\frac{1}{2} + \frac{1}{4}\right),$$

a contradiction. This eliminates (5-39).

In case (5-40) we use Proposition 3.2 with $A = 9$ to obtain

$$m' + \frac{|n'|}{2} \leq \frac{m'}{8} + |n'| + \frac{|r'|}{d} + 0.01m' \quad \text{and} \quad d = \begin{cases} \min\{9, a(z^\sigma)\}, & r > 0, \\ \min\{9, a(z)\}, & r < 0. \end{cases}$$

Since $d \geq 3$, we obtain

$$\left(\frac{7}{8} - 0.01\right)m' \leq \frac{|n'|}{2} + \frac{|r'|}{3} \leq m'\left(\frac{1}{2} + \frac{1}{3}\right),$$

a contradiction. This rules (5-40) out as well. The proposition is proved. □

Next, we show the impossibility of $a_z = 2$.

Proposition 5.8. *In case* (5-37) *we must have* $a_z \geq 3$.

Proof. We already know that $a_z \geq 2$ and that $(e_x, e_y) = (1, 2)$. We also note the statement is immediate for $e_z = 2$, because 2 is not a denominator for 4Δ; see Proposition 2.17. Thus, let us assume that

$$(e_x, e_y, e_z) = (1, 2, 1), \quad a_z = 2,$$

and show that this is impossible.

Arguing as in the proof of Proposition 5.7 but with the roles of x and y interchanged, we find σ

satisfying

$$a(x^\sigma)=2, \quad a(y^\sigma)=8, \quad a(z^\sigma)\ge 2.$$

Since x, z are 2-isogenous, we have $a(z^\sigma)\in\{1,4\}$, and we must have $a(z^\sigma)=4$ because $a(z^\sigma)\ge 2$.

Next, let $\theta\in G$ be defined by $z^\theta=x$. Since x, z are 2-isogenous, we must have $a(x^\theta)=2$, which implies that $a(y^\theta)=8$.

Applying Proposition 3.1 to the relation

$$(x^\sigma)^m(y^\sigma)^n(z^\sigma)^r=(x^\theta)^m(y^\theta)^n(z^\theta)^r,$$

we obtain

$$\frac{m'}{2}+\frac{n'}{8}+\frac{r'}{4}=\frac{m'}{2}+\frac{n'}{8}+\frac{r'}{1},$$

which implies $r=0$, a contradiction. □

We also need to know that $|n'|$ is not much smaller than m'.

Proposition 5.9. *When $r>0$ we have $|n'|>0.87m'$. When $r<0$ and $a_z\ge a$ we have $|n'|>\lambda(a)m'$, where*

$$\lambda(a)=0.956-\frac{1}{\min\{30,a\}}.$$

Here are lower bounds for $\lambda(a)$ for some values of a that will emerge below:

a	3	5	6	24	30
$\lambda(a)$	>0.62	>0.75	>0.78	>0.91	>0.92

Proof. When $r>0$ we use Corollary 5.2 to find σ such that $a(x^\sigma), a(z^\sigma)\ge 18$. Now Proposition 3.2 gives

$$m'\le|n'|+\frac{m'}{18}+\frac{r'}{18}+0.01m'\le|n'|+m'\Big(\frac{2}{18}+0.01\Big),$$

which implies $|n'|>0.87m'$.

When $r<0$ we use Corollary 5.2 to find σ such that $a(x^\sigma)\ge 30$. When $a_z\ge a$, we obtain

$$m'\le|n'|+\frac{m'}{30}+\frac{|r'|}{\min\{30,a\}}+0.01m'\le|n'|+m'\left(\frac{1}{30}+\frac{1}{\min\{30,a\}}+0.01\right),$$

which implies $|n'|>\lambda(a)m'$. □

Proof of Proposition 5.6. The proof is similar to that of Proposition 5.7, but with the roles of x and y interchanged. This means that, instead of the inequality $m'\ge|n'|$, we have to use weaker inequalities from Proposition 5.9. This is why we cannot rule out (5-38).

We already know that $a_z\ge 3$ and that $(e_x, e_y)=(1,2)$. We also note that 4 is not a denominator for 4Δ; see Proposition 2.17. Hence it suffices to show that each of the cases

$$e_z=1, \quad a_z\ge 3, \quad a_z\ne 4, \tag{5-41}$$

$$e_z=2, \quad a_z\ge 5, \tag{5-42}$$

leads to a contradiction. As in the proof of Proposition 5.8, we fix $\sigma \in G$ satisfying

$$a(x^\sigma) = 2, \quad a(y^\sigma) = 8, \quad a(z^\sigma) \geq 2.$$

Assume (5-41). As in the proof of Proposition 5.7, we show that z, z^σ are 2-isogenous. Hence $\{a(z), a(z^\sigma)\} = \{a', 2a'\}$, where $a' \geq 3$. If $a' \geq 6$ then, using Proposition 3.2, we obtain

$$\frac{m'}{2} + |n'| \leq m' + \frac{|n'|}{8} + \frac{|r'|}{6} + 0.01m',$$

which implies that

$$|n'| \leq \tfrac{8}{7} m' \left(\tfrac{1}{2} + \tfrac{1}{6} + 0.01\right) < 0.78m',$$

contradicting the lower bound $|n'| > 0.78m'$ from Proposition 5.9.

If $a' \in \{3, 4, 5\}$ then Proposition 3.1 gives

$$m' + n' + \frac{r'}{a(z)} = \frac{m'}{2} + \frac{n'}{8} + \frac{r'}{a(z^\sigma)}.$$

This can be rewritten as

$$|n'| = \frac{8}{7}\left(\frac{m'}{2} + r'\left(\frac{1}{a(z)} - \frac{1}{a(z^\sigma)}\right)\right), \tag{5-43}$$

which implies

$$|n'| \leq \tfrac{8}{7} m' \left(\tfrac{1}{2} + \tfrac{1}{6}\right) < 0.77m', \tag{5-44}$$

When $r > 0$ this contradicts the lower bound $|n'| > 0.87m'$ from Proposition 5.9. When $r < 0$ and $a_z = 2a'$ this contradicts the lower bound $|n'| > 0.78m'$ from Proposition 5.9. Finally, when $r < 0$ and $a_z = a'$, we deduce from (5-43) the sharper upper bound $|n'| \leq \frac{4}{7}m'$, contradicting the lower bound $|n'| \geq 0.62m'$ from Proposition 5.9. This shows the impossibility of (5-41).

Now let us assume (5-42). Proposition 3.2 implies that

$$\frac{m'}{2} + |n'| \leq m' + \frac{|n'|}{8} + \frac{|r'|}{d} + 0.01m' \quad \text{and} \quad d = \begin{cases} \min\{9, a(z)\} & \text{if } r > 0, \\ \min\{9, a(z^\sigma)\} & \text{if } r < 0. \end{cases}$$

If $r > 0$ then $d \geq 5$, and we obtain

$$|n'| \leq \tfrac{8}{7} m' \left(\tfrac{1}{2} + \tfrac{1}{5} + 0.01\right) < 0.82m',$$

contradicting the lower bound $|n'| > 0.87m'$ from Proposition 5.9.

If $r < 0$ and $d \geq 8$ then

$$|n'| \leq \tfrac{8}{7} m' \left(\tfrac{1}{2} + \tfrac{1}{8} + 0.01\right) < 0.73m',$$

contradicting the lower bound $|n'| > 0.75m'$ from Proposition 5.9.

Thus,

$$r < 0, \quad 3 \leq a(z^\sigma) \leq 7.$$

Since $e_z = 2$, we must have $a(z^\sigma) = p \in \{3, 5, 7\}$. Hence y^σ and z^σ are $8p$-isogenous, and so are y and z. It follows that $a_z = 8p \geq 24$, and Proposition 5.9 implies the lower bound $|n'| > 0.91m'$. On the other

i	$a(x^{\sigma_i})$	$a(y^{\sigma_i})$	$a(z^{\sigma_i})$
1	2	8	$\in \{2, 8\}$
2	4	16	1
3	$\in \{2, 8\}$	$\in \{8, 32\}$	2

i	$a(x^{\sigma_i})$	$a(y^{\sigma_i})$	$a(z^{\sigma_i})$
1	2	8	24
2	3	3	1
3	$\in \{6, 24\}$	24	8

Table 4. Denominators for case (5-20). The table on the left refers to $(e_x, e_y, e_z) = (1, 2, 1)$ and $a_z = 4$; there are 8 systems in total. On the right, $(e_x, e_y, e_z) = (1, 2, 2)$ and $a_z = 3$, with 2 systems in total.

hand, Proposition 3.1 implies that

$$m' + n' + \frac{r'}{8p} = \frac{m'}{2} + \frac{n'}{8} + \frac{r'}{p},$$

which yields

$$|n'| = \frac{4m'}{7} + \frac{|r'|}{p} < 0.91m',$$

a contradiction. This shows impossibility of (5-42). The proposition is proved. □

Now it is easy to dispose of case (5-20), alias (5-37). We define σ_1 as σ from the proof of Proposition 5.6; that is, $a(x^{\sigma_1}) = 2$ and $a(z^{\sigma_1}) \neq 1$. Next, we define σ_2 from

$$z^{\sigma_2} = \begin{cases} x & \text{if } e_z = 1, \\ y & \text{if } e_z = 2. \end{cases}$$

Finally, we set $\sigma_3 = \sigma_2\sigma_1$.

Using Proposition 2.17 and Corollary 2.20, we calculate the possible denominators; see Table 4. A verification shows that, in each case, the system of 3 linear equations

$$m' + n' + \frac{r'}{a_z} = \frac{m'}{a(x^{\sigma_i})} + \frac{n'}{a(y^{\sigma_i})} + \frac{r'}{a(z^{\sigma_i})} \quad (i = 1, 2, 3) \tag{5-45}$$

has only the trivial solution.

5.4. ***The subdominant case.*** In this subsection we assume (5-16). Thus, we have

$$m > 0, \quad n, r < 0, \quad m' \geq \max\{|n'|, |r'|\}, \quad a_x = 1, \quad a_y = a_z = 2.$$

To start with, note that

$$e_y = e_z \quad \text{and} \quad \Delta_y = \Delta_z \equiv 1 \bmod 8. \tag{5-46}$$

Indeed, among the three numbers e_x, e_y, e_z there are only two distinct integers, see Table 2. If $e_y \neq e_z$ then, switching, if necessary, x and y, we may assume that $e_x = e_y$. Hence $K(x) = K(y)$, and we have one of two possibilities:

$$K(x) = K(y) = K(z) = L \tag{5-47}$$

or

$$K(x) = K(y) = L, \quad [L : K(z)] = 2.$$

In the latter case we must have $m = n$ by Theorem 1.2, which is impossible because $m > 0$ and $n < 0$. Thus, we have (5-47). Lemma 2.25 now implies that $e_z = 2$ and $\Delta_x = \Delta_y = \Delta \equiv 1 \bmod 8$. But then $\Delta_z \equiv 4 \bmod 32$, and we cannot have $a_z = 2$ by Proposition 2.17:4. Thus, $e_x = e_y$ is impossible, which proves that $e_y = e_z$. Now Proposition 2.17:2 implies that $\Delta_y = \Delta_z \equiv 1 \bmod 8$, which completes the proof of (5-46).

Exploring Table 2 and taking note of (5-46), we end up with one of the following cases:

$$e_x \in \{1, 2\}, \quad e_y = e_z = 1, \quad \Delta \equiv 1 \bmod 8, \quad L = K(x) = K(y) = K(z); \tag{5-48}$$

$$e_x \in \{1, 2\}, \quad e_y = e_z = 3, \quad \Delta \equiv 1 \bmod 24, \quad [L : K(x)] = 2, \quad n = r. \tag{5-49}$$

We cannot have $e_y = e_z = 4$ or $e_y = e_z = 6$, because in these cases $\Delta_y = \Delta_z \equiv 0 \bmod 4$, contradicting (5-46).

Each of the cases (5-48) and (5-49) can be disposed of using the strategy described in Section 5.3.1; moreover, the very first step of that strategy can be skipped, because a_z is already known.

Case (5-48) is analogous to case (5-20), but is much simpler, because, as indicated above, we already know a_z. We define $\sigma_1, \sigma_2, \sigma_3$ by

$$a(y^{\sigma_1}) = 1, \quad a(z^{\sigma_2}) = 1, \quad a(y^{\sigma_3}) = 8.$$

There can be several candidates for σ_3, we just pick one of them. The possible denominators, determined using Corollary 2.20 and Proposition 2.17:4, are given in Table 5. A verification with PARI shows that each of the 12 possible systems

$$m' + \frac{n'}{2} + \frac{r'}{2} = \frac{m'}{a(x^{\sigma_i})} + \frac{n'}{a(y^{\sigma_i})} + \frac{r'}{a(z^{\sigma_i})} \quad (i = 1, 2, 3)$$

has only the trivial solution $m' = n' = r' = 0$.

In case (5-49) we have $n = r$ (and also $n' = r'$), and so we need only σ_1 and σ_2. We do as in Section 5.3.2. Since $\Delta_x \equiv 1 \bmod 3$, we can find σ_1, σ_2 satisfying $a(x^{\sigma_1}) = 3$ and $a(y^{\sigma_2}) = 9$. Defining

$$\ell = \begin{cases} 6 & \text{when } e_x = 1, \\ 12 & \text{when } e_x = 2, \end{cases}$$

a quick verification shows that singular moduli x, y are ℓ-isogenous, and so are x, z. Using Corollary 2.20 and Proposition 2.16:2, we determine the possible denominators: in both cases $e_x = 1$ and $e_x = 2$ we find

$$a(y^{\sigma_1}), a(z^{\sigma_1}) = 54, \quad a(y^{\sigma_2}), a(z^{\sigma_2}) \in \{18, 162\}.$$

i	$a(x^{\sigma_i})$	$a(y^{\sigma_i})$	$a(z^{\sigma_i})$
1	2	1	4
2	2	4	1
3	$\in \{4, 16\}$	8	$\in \{2, 8, 32\}$

i	$a(x^{\sigma_i})$	$a(y^{\sigma_i})$	$a(z^{\sigma_i})$
1	8	1	4
2	8	4	1
3	$\in \{16, 64\}$	8	$\in \{2, 8, 32\}$

Table 5. Denominators for case (5-48), with (e_x, e_y, e_z) taking the values (1, 1, 1) (left) and (2, 1, 1) (right). In either case, there are 6 systems in total.

It follows that m', n' satisfy one of the three linear systems

$$\begin{cases} m'+n'\left(\frac{1}{2}+\frac{1}{2}\right)=\frac{1}{3}m'+\left(\frac{1}{54}+\frac{1}{54}\right)n', \\ m'+n'\left(\frac{1}{2}+\frac{1}{2}\right)=\frac{1}{9}m'+\lambda n', \end{cases} \qquad \text{where } \lambda\in\left\{\frac{1}{18}+\frac{1}{18},\ \frac{1}{18}+\frac{1}{162},\ \frac{1}{162}+\frac{1}{162}\right\}.$$

A verification shows that each of these systems has only the trivial solution $m'=n'=0$. This completes the proof of Theorem 1.1 for equal fundamental discriminants.

5.5. *Distinct fundamental discriminants.* We are left with the case when the fundamental discriminants D_x, D_y, D_z are not all equal. We may assume that $|\Delta_z|\ge|\Delta_y|\ge|\Delta_x|$. In particular, $|\Delta_z|\ge 10^{10}$ and $\Delta_z\ne\Delta_x$. Theorem 1.2 and Lemma 2.22 imply that $\mathbb{Q}(y)=\mathbb{Q}(y^n)=\mathbb{Q}(x^mz^r)=\mathbb{Q}(x,z)$, and so

$$\mathbb{Q}(x),\mathbb{Q}(z)\subset\mathbb{Q}(y). \tag{5-50}$$

Theorem 1.2 and Lemma 2.22 imply that $\mathbb{Q}(x)=\mathbb{Q}(x^m)=\mathbb{Q}(y^nz^r)$ is a subfield of $\mathbb{Q}(y,z)$ of degree at most 2. Since $z\in\mathbb{Q}(y)$, this implies that

$$[\mathbb{Q}(y):\mathbb{Q}(x)]\le 2.$$

Unfortunately, we cannot claim similarly that $[\mathbb{Q}(y):\mathbb{Q}(z)]\le 2$, because we do not know whether the singular moduli x, y satisfy the hypotheses of Theorem 1.2.

The rest of the proof splits into two cases: $D_x\ne D_y$ and $D_y\ne D_z$.

5.5.1. *The case $D_x\ne D_y$.* Since $[\mathbb{Q}(y):\mathbb{Q}(x)]\le 2$, Corollary 2.13 implies that $[\mathbb{Q}(y):\mathbb{Q}]\le 32$. It follows that $h(\Delta_z)=[\mathbb{Q}(z):\mathbb{Q}]\le 32$ as well. This contradicts Proposition 2.1 because $|\Delta_z|\ge 10^{10}$.

5.5.2. *The case $D_y\ne D_z$.* Since $z\in\mathbb{Q}(y)$, the field $\mathbb{Q}(z)$ must be 2-elementary by Proposition 2.8. Hence the proof in this case will be complete if we show either that

$$\rho_2(\Delta_z)\le 6 \tag{5-51}$$

or that

$$[\mathbb{Q}(y):\mathbb{Q}(z)]\le 2. \tag{5-52}$$

Recall that $\rho_2(\cdot)$ is the 2-rank; see Section 2.1.2.

Indeed, assume that (5-51) holds. Since $\mathbb{Q}(z)$ is 2-elementary, we have $h(\Delta_z)=2^{\rho_2(\Delta_z)}\le 64$, contradicting Proposition 2.1.

Similarly, if (5-52) holds then $h(\Delta_z)\le 16$ by Corollary 2.13, again contradicting Proposition 2.1.

We will show that, depending on the size of Δ_y, one of (5-51) or (5-52) holds.

If $|\Delta_y|\ge 10^8$, then Theorem 1.2 applies to the singular moduli x, y. It follows that $\mathbb{Q}(z)=\mathbb{Q}(z^r)=\mathbb{Q}(x^my^n)$ is a subfield of $\mathbb{Q}(x,y)$ of degree at most 2. Since $x\in\mathbb{Q}(y)$, we obtain $[\mathbb{Q}(y):\mathbb{Q}(z)]\le 2$, which is (5-52).

Next, let us assume that $10^6\le|\Delta_y|\le 10^8$. If $\Delta_y\not\equiv 4 \bmod 32$ then Theorem 1.2 again applies to the singular moduli x, y, and we may argue as above, obtaining (5-52).

If $\Delta_y \equiv 4 \bmod 32$, then $\rho_2(\Delta_y) = \omega(\Delta_y) - 2$; see Proposition 2.3. Since

$$10^8 < 4 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 = 446185740,$$

we have $\omega(\Delta_y) \leq 8$. Hence $\rho_2(\Delta_y) \leq 6$. Now let $K = \mathbb{Q}(\sqrt{\Delta_y})$ be the CM field of y. Since $\mathbb{Q}(z)$ is 2-elementary, K is not contained in $\mathbb{Q}(z)$ by Proposition 2.18. Since both K and $\mathbb{Q}(z)$ are Galois extensions of $\mathbb{Q}$, the group $\mathrm{Gal}(\mathbb{Q}(z)/\mathbb{Q})$ is isomorphic to $\mathrm{Gal}(K(z)/K)$, which is a quotient group of $\mathrm{Gal}(K(y)/K)$. In particular,

$$\rho_2(\mathrm{Gal}(\mathbb{Q}(z)/\mathbb{Q})) \leq \rho_2(\mathrm{Gal}(K(y)/K)) \leq 6,$$

which is (5-51).

Finally, let us assume that $|\Delta_y| \leq 10^6$. Since

$$10^6 < 4 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 = 1021020,$$

we have $\rho_2(\Delta_y) \leq \omega(\Delta) \leq 6$, again by Proposition 2.3. As we have seen, this implies (5-51). Theorem 1.1 is proved.

Acknowledgments

We thank Guy Fowler and the referee for many comments that helped us to correct inaccuracies and improve the presentation. We also thank Francesco Amoroso, Margaret Bilu, Igor Rapinchuk and Anatoly Vorobey for useful suggestions. Finally, we thank Bill Allombert and Amalia Pizarro, who allowed us to borrow Proposition 2.11 from [1].

All calculations were performed using PARI [19]. We thank Bill Allombert and Karim Belabas for the PARI tutorial. The reader may consult https://github.com/yuribilu/multiplicative to view the PARI scripts used for this article.

References

[1] B. Allombert, Y. Bilu, and A. Pizarro-Madariaga, "CM-points on straight lines", 2014. An early version of [2]. arXiv 1406.1274v1

[2] B. Allombert, Y. Bilu, and A. Pizarro-Madariaga, "CM-points on straight lines", pp. 1–18 in *Analytic number theory*, Springer, 2015. MR

[3] F. Amoroso and U. Zannier, "A uniform relative Dobrowolski's lower bound over abelian extensions", *Bull. Lond. Math. Soc.* **42**:3 (2010), 489–498. MR

[4] Y. André, "Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire", *J. Reine Angew. Math.* **505** (1998), 203–208. MR

[5] Y. Bilu, D. Masser, and U. Zannier, "An effective 'theorem of André' for CM-points on a plane curve", *Math. Proc. Cambridge Philos. Soc.* **154**:1 (2013), 145–152. MR

[6] Y. Bilu, F. Luca, and A. Pizarro-Madariaga, "Rational products of singular moduli", *J. Number Theory* **158** (2016), 397–410. MR

[7] Y. Bilu, F. Luca, and D. Masser, "Collinear CM-points", *Algebra Number Theory* **11**:5 (2017), 1047–1087. MR

[8] Y. Bilu, B. Faye, and H. Zhu, "Separating singular moduli and the primitive element problem", *Q. J. Math.* **71**:4 (2020), 1253–1280. MR

[9] Y. Bilu, F. Luca, and A. Pizarro-Madariaga, "Trinomials, singular moduli and Riffaut's conjecture", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **23**:4 (2022), 2003–2048. MR

[10] D. A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, 2nd ed., Wiley, Hoboken, NJ, 2013. MR

[11] B. Edixhoven, "Special points on the product of two modular curves", *Compositio Math.* **114**:3 (1998), 315–328. MR

[12] B. Faye and A. Riffaut, "Fields generated by sums and products of singular moduli", *J. Number Theory* **192** (2018), 37–46. MR

[13] G. Fowler, "Triples of singular moduli with rational product", *Int. J. Number Theory* **16**:10 (2020), 2149–2166. MR

[14] G. Fowler, "Equations in three singular moduli: the equal exponent case", *J. Number Theory* **243** (2023), 256–297. MR

[15] J. Klaise, *Orders in quadratic imaginary fields of small class number*, Master's thesis, University of Warwick, 2012, available at https://warwick.ac.uk/fac/cross_fac/complexity/people/students/dtc/students2013/klaise/janis_klaise_ug_report.pdf.

[16] L. Kühne, "An effective result of André–Oort type", *Ann. of Math.* (2) **176**:1 (2012), 651–671. MR

[17] L. Kühne, "An effective result of André–Oort type, II", *Acta Arith.* **161**:1 (2013), 1–19. MR

[18] D. W. Masser, "Linear relations on algebraic groups", pp. 248–262 in *New advances in transcendence theory* (Durham, 1986), Cambridge Univ. Press, 1988. MR

[19] "PARI/GP version 2.11.4", software, 2020, available at http://pari.math.u-bordeaux.fr.

[20] J. Pila and J. Tsimerman, "Multiplicative relations among singular moduli", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **17**:4 (2017), 1357–1382. MR

[21] A. Riffaut, "Equations with powers of singular moduli", *Int. J. Number Theory* **15**:3 (2019), 445–468. MR

[22] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR

[23] T. Tatuzawa, "On a theorem of Siegel", *Jpn. J. Math.* **21** (1951), 163–178. MR

[24] M. Watkins, "Class numbers of imaginary quadratic fields", *Math. Comp.* **73**:246 (2004), 907–938. MR

[25] P. J. Weinberger, "Exponents of the class groups of complex quadratic fields", *Acta Arith.* **22** (1973), 117–124. MR

Communicated by Jonathan Pila
Received 2022-08-03 Revised 2024-07-05 Accepted 2025-05-02

yuri@math.u-bordeaux.fr *Institut de Mathématiques de Bordeaux, Université de Bordeaux & CNRS, Talence, France*

sanoli@imsc.res.in *Institute of Mathematical Sciences, CIT Campus, Tharamani, Chennai, India*

emanuele.tron@math.u-bordeaux.fr *Institut de Mathématiques de Bordeaux, Université de Bordeaux & CNRS, Talence, France*

ALGEBRA AND NUMBER THEORY 20:6 (2026)
https://doi.org/10.2140/ant.2026.20.1125

The geometry of the unipotent component of the moduli space of Weil–Deligne representations

Daniel Funck

We study the moduli space of unipotent Weil–Deligne representations valued in a split reductive group G and characterise which irreducible components are smooth. We apply these smoothness results to show that a certain space of ordinary automorphic forms is a locally generically free module over the corresponding global deformation ring.

1. Introduction and overview

Let F be a local p-adic field and let G be a connected reductive algebraic group over F and $\hat{G}$ its Langlands dual. The local Langlands conjectures (proven for GL_n by Harris and Taylor in [HT01]) stipulate the existence of a natural map with finite fibres:

$$\frac{\{\text{smooth irreducible representations of } G(F)\}}{\{\text{isomorphism}\}} \to \frac{\{\text{L-parameters of } {}^LG\}}{\{\hat{G}\text{-conjugacy}\}}$$

Let $l \neq p$ be a prime. Let $L \subset \overline{\mathbb{Q}}_l$ be an l-adic field and $\mathcal{O}$ its ring of integers with residue field $\mathbb{F}$. In recent years, through the work in [BG19; Hel23; DHKM20; Zhu25; FS25], there has been great interest in studying the properties of a moduli space of L-parameters $\mathrm{Loc}_{\hat{G},\mathcal{O}}$ and a closely related space, the moduli space of framed L-parameters $\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}$.

The spaces $\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}$ and $\mathrm{Loc}_{\hat{G},\mathcal{O}}$ can be respectively defined as the scheme whose R-points are the set of L-parameters with R coefficients, and the algebraic stack obtained from the stack quotient of $\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}$ modulo the natural action of $\hat{G}$ via conjugation (that is, equivalence of representations):

$$\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}(R) = \{\text{L-parameters of } \hat{G} \text{ with } R\text{-coefficients}\}, \quad \mathrm{Loc}_{\hat{G},\mathcal{O}} = \big[\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}/\hat{G}\big].$$

MSC2020: primary 11F70, 11F80; secondary 11F33.
Keywords: number theory, moduli space of Langlands parameters, Galois representation, patching, automorphic form.

In addition to this interpretation, the scheme $\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}$ satisfies the secondary property that the (completions of) local rings of $\mathrm{Loc}^{\square}_{\hat{G},\mathcal{O}}$ can be interpreted as local Galois deformation rings. In this way, it is hoped to better understand these rings, which are crucial ingredients in the Taylor–Wiles–Kisin and Calegari–Geraghty patching methods.

Let W_F be the Weil group of the field F. In the case of a split reductive group G, one would want to define an L-parameter as a homomorphism $W_F \to \hat{G}(R)$ which satisfies some kind of continuity, but the problem with this naive approach is that the ring R has no topological structure in general. Historically, there have been multiple solutions to this issue, with varying degrees of usefulness. We are interested in the moduli spaces of Bellovin and Gee [BG19], defined via Weil–Deligne representations (defined below); and of Dat, Helm, Kurinczuk and Moss [DHKM20], defined through representations of a particular dense subgroup $W_F^0 \subseteq W_F$ (defined in Section 1.2 of the reference).

Definition 1.1. A Weil–Deligne representation valued in G with R-coefficients is a pair (r, N), where $r : W_F \to G(R)$ is a homomorphism with open kernel and N is an element of the nilpotent cone $\mathcal{N}_G(R) \subseteq \mathrm{Lie}(G)(R)$ such that $\mathrm{Ad}(g)N = |g|N$ for all $g \in W_F$, where $|\cdot| : W_F \to F^\times \to \mathbb{R}^{\geq 0}$ is the valuation of W_F coming from local class field theory.

The two moduli problems of [BG19] and [DHKM20] are representable by algebraic stacks $\mathrm{Loc}^{BG}_{G,\mathcal{O}}$ and $\mathrm{Loc}_{G,\mathcal{O}}$, respectively, and are disjoint unions of quotient stacks.

It is known, as in Proposition 2.7 of [DHKM20], that these definitions give isomorphic moduli spaces over fields of characteristic 0. In positive characteristic l (or mixed characteristic), only the latter moduli space is generally well-behaved, giving the deformation rings as the completions of its local rings. However, when l is greater than the Coxeter number h_G of G, the exponential and logarithm maps of Proposition 2.7 of [DHKM20] that arise from Grothendieck's l-adic monodromy are well defined polynomials, and we obtain an isomorphism between the two moduli spaces in this case too.

Let $\mathcal{O}$ be a discrete valuation ring (or field) of residue characteristic (resp. characteristic) $l > h_G$ or 0. Let $\mathcal{N}_G \subseteq \mathfrak{g}$ be the nilpotent cone inside the Lie algebra $\mathfrak{g}$. In this paper, we seek to understand the irreducible components of the scheme studied in [Hel23]. This is a reduced affine scheme of finite type $S_{G,\mathcal{O}}$, over the ring $\mathcal{O}$, whose R-points (R an $\mathcal{O}$-algebra) are given by

$$S_{G,\mathcal{O}}(R) = \{(\Phi, N) \in G(R) \times \mathcal{N}_G(R) | \mathrm{Ad}(\Phi)N = qN\},$$

where $q \in \mathcal{O}^\times$ is some prime power.

The scheme $S_{G,\mathcal{O}}$ is naturally the space of framed unipotent Weil–Deligne representations over $\mathcal{O}$ with values in G (following Definition 2.1.2 of [BG19]). We will be interested especially in the case where $\mathcal{O}$ is the ring of integers in a finite extension of $\mathbb{Q}_l$ because the $\mathfrak{m}_R$-adic completion of the local rings R of the closed points of this scheme can be interpreted as local Galois deformation rings for well behaved l (in fact, whenever the exponential and logarithm maps of Grothendieck's l-adic monodromy theorem give an isomorphism onto a connected component, as above). Since $S_{G,\mathcal{O}}$ is a connected component of the tame parameters $Z^1(W^0/P, G)_{\mathcal{O}}$, equation 4.5 of [DHKM20] extends the description of $S_{G,\overline{\mathbb{Q}}_l}$

for various groups G to a description of the geometry of many other connected components of $\mathrm{Loc}^{\square}_{G,\overline{\mathbb{Q}}_l}$ (those which correspond to the case where the action Ad_{φ} is trivial).

Outline. In Section 2, we collect some basic results for S_G in the mixed characteristic setting. The results of this paper depend strongly on the technical relationship between the residue characteristic l and the element $q \in \mathcal{O}$. Therefore, the section begins by defining the notions of G-banality and q-considerateness, and explaining how they are related. We then give a description of a decomposition of S_G into (unions of) irreducible components, generalising the decomposition of S_{GL_n} given in Proposition 2.1 of [Hel23], as follows: Let

$$p : S_{G,\mathcal{O}} \to \mathcal{N}_G$$

be the projection map onto the second factor. Let $C \subset \mathcal{N}_{G,L}$ be a G-conjugacy class inside $\mathcal{N}_{G,L}$. (In the case of GL_n, these can be characterised by partitions of n and in this situation we will denote the conjugacy class corresponding to λ by C_λ.) We note that because $S_{G,\mathcal{O}}$ is flat over $\mathcal{O}$, the irreducible components biject naturally with those of $S_{G,L}$. Then $\overline{p^{-1}(C)} \subseteq S_{G,\mathcal{O}}$ is a union of irreducible components of $S_{G,\mathcal{O}}$ (and, in the case of $G = \mathrm{GL}_n$, is itself irreducible). In Section 3, we expand on and generalise the results of Bellovin [Bel16, Section 7.2 and Proposition 7.10] by proving our Theorems 3.2 and 3.4, which together state:

Theorem 1.2. *Assume q is considerate towards $G_{\mathcal{O}}$* (see Definition 2.1).

(1) *Suppose $C \subseteq \mathcal{N}_{G,L}$ is a distinguished nilpotent orbit, or the zero orbit with corresponding component $X_C = \overline{p^{-1}(C)}$. Then X_C is a disjoint union of smooth connected components.*

(2) *Conversely, when $C \subseteq \mathcal{N}_{G,L}$ is a nontrivial nondistinguished orbit, the scheme X_C is singular.*

In Sections 4 and 5, we apply the smoothness result of Section 3 to Hida families of ordinary automorphic forms using the Taylor–Wiles–Kisin patching method in a situation very similar to that studied in [Ger19]. Let l be a prime and K be a finite extension of $\mathbb{Q}_l$ with ring of integers $\mathcal{O}$. Let F^+ be a totally real global number field, and consider an imaginary quadratic extension F of F^+. The Galois representations considered will correspond to certain Hida families of ordinary automorphic forms on a unitary algebraic group G_D/F^+ which is a unitary form of a unit group of a division algebra D/F^+. We will define a certain space of Hida families of ordinary automorphic forms $S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})_m$ for G_D with Hecke operators $\mathbb{T}$ and a corresponding deformation ring R_S^{univ}. We will then use the Taylor–Wiles patching method to deduce the following theorem.

Theorem 1.3 (Theorem 5.13). *Suppose $l > n$. The module $S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})_m^\vee[1/l]$ is a finite locally free $R_S^{\mathrm{univ}}[1/l]$-module.*

As a consequence, we can deduce that $R_S^{\mathrm{univ}}[1/l] \cong \mathbb{T}[1/l]$, and that the multiplicity of automorphic forms with a given characteristic zero Galois representation is constant along connected components of $R_S^{\mathrm{univ}}[1/l]$. One can then extend any such multiplicity results from the classical case to the case of nonclassical Hida families.

2. Considerateness and the relation to the stack of L-parameters

Let $\mathcal{O}$ be a discrete valuation ring or a field with residue field $\mathbb{F}$ of characteristic l or 0 and fraction field L. Let G be a connected reductive algebraic group over $\mathcal{O}$, let $\mathfrak{g}$ be its Lie algebra, and let h_G be its Coxeter number. Throughout the paper, we assume $l > h_G$ whenever the residue characteristic of $\mathcal{O}$ is finite.

Definition 2.1. Let h_G be the Coxeter number of G. Let $q \in \mathcal{O}^\times$ be an element of $\mathcal{O}$ such that $q^k - 1$ is invertible in $\mathcal{O}$ for all $k \le h_G$. When this occurs, we say that q is *considerate* towards G over $\mathcal{O}$.

In applications, the ring $\mathcal{O}$ will either be a field or the ring of integers in some field extension of $\mathbb{Q}_l$. Notice that, in this case, q-considerateness is equivalent to the condition that $1, q, q^2, \ldots, q^{h_G}$ are all distinct in the residue field $\mathbb{F}$ (in a sense, q "treads lightly" around G).

Definition 2.2. Let G be a split reductive group over a field L of characteristic l.

- l is called *G-banal* if l does not divide the order of the finite group $G(\mathbb{F}_q)$.
- l is called *geometrically G-banal* if, for any algebraically closed field E of characteristic l, any $\phi \in \mathrm{Loc}^{\square}_{G,E}$ can be *Frobenius twisted* by some $g \in C_G(\phi(I_F))$ (that is, the centraliser of the inertia subgroup) so that ϕ^g is a smooth point of $\mathrm{Loc}^{\square}_{G,E}$.

  The *Frobenius twist* of a representation $\phi : W_F \to G(L)$ by $g \in C_G(\phi(I_F))$ is the representation $\phi^g : W_F \to G(L)$ whose restriction to the inertia subgroup equals ϕ, and for which $\phi^g(\mathrm{Frob}) = \phi(\mathrm{Frob})g$.

Remark. Let $\hat{G}$ denote the Langlands dual group of G. We remark that the notion "l is geometrically $\hat{G}$-banal" is precisely the notion that "l is ${}^L G$-banal", as defined in Definition 5.27 of [DHKM20]. We introduce the notion simply to remove the extraneous Langlands dual, which is only required in the following Proposition.

Proposition 2.3. *Suppose that $\mathbb{F}$ is a field of positive characteristic $l > h_G$ and that G is a split reductive group. Then we have the following implications.*

- *If q is considerate towards $G_{/\mathbb{F}}$, then l is geometrically G-banal.*
- *If l is geometrically $\hat{G}$-banal (that is, ${}^L G$-banal), then l is G-banal.*
- *If $G = \mathrm{GL}_n$ or SL_n, then the condition "l is G-banal" implies that q is considerate towards $\hat{G}_{/\mathcal{O}}$. Thus, in these cases all three concepts are equivalent.*

Proof. By definition, q is considerate towards $G_{/\mathbb{F}}$ when the order of q within $\mathbb{F}$ is greater than the Coxeter number $h = h_G$. This is equivalent to $\prod_{n \le h} \Phi_n(q) \neq 0$ inside $\mathbb{F}$, where Φ_n is the n-th cyclotomic polynomial. This is the polynomial $\chi^*_{G,1}(q)$ of [DHKM20, Definition B.3]. Hence, by Theorems 5.6 and 5.7 of [DHKM20], this condition implies that l is ${}^L G$-banal.

That l being ${}^L G$-banal implies that l is G-banal is a consequence of the Chevalley–Steinberg formula

$$|G(\mathbb{F}_q)| = q^N \prod_d (q^d - 1)$$

(see [Ste16, Theorem 25a]), where d ranges over the fundamental degrees of the Weyl group of G. If l divides $\prod_d (q^d - 1)$, then l certainly divides $\prod_{n \le h} \Phi_n(q)$ as the Coxeter number is the highest fundamental degree. This shows the second statement by virtue of Theorem 5.7 in [DHKM20].

In the case $G = \mathrm{SL}_n$, we get $|\mathrm{SL}_n(\mathbb{F}_q)| = q^N \prod_{i=2}^{n} (q^i - 1)$. Hence, if l is G-banal, the order of q in $\mathbb{F}$ is at least n. Thus q is considerate towards $\hat{G}_{/\mathbb{F}}$. The case of GL_n is similar. □

Remark. It is worth noting that Corollary 5.27 of [DHKM20] gives the criterion that G-banal and ${}^L G$-banal are equivalent concepts whenever G is unramified and has no exceptional factors (where triality forms of type D_4 are also considered exceptional), but this does not hold in general (see, for example, Remark 5.22 of [DHKM20]). Outside the case of type A_n, considerateness is strictly weaker than geometric-G-banality. For example, the order of $G = \mathrm{Sp}_6(\mathbb{F}_q)$ is

$$|\mathrm{Sp}_6(\mathbb{F}_q)| = q^9 (q^2 - 1)(q^4 - 1)(q^6 - 1)$$

and the Coxeter number of $\hat{G} = \mathrm{SO}_7$ is $h = 6$. However, if $q^5 \equiv 1 \pmod{l}$, then l is G-banal, but q is not considerate towards $\hat{G}_{\mathcal{O}}$.

Definition 2.4. We define the affine scheme $S_{G,\mathcal{O}}$ over $\mathcal{O}$ as the scheme whose R points (for R an $\mathcal{O}$-algebra) are $\{(\Phi, N) \in G(R) \times \mathcal{N}_G(R) : \mathrm{Ad}(\Phi)N = qN\}$.

Corollary 5.4 of [Bel16] shows that this is a reduced scheme when $\mathcal{O}$ is a characteristic zero field, and hence, is a variety (i.e., a finite-type, separated, reduced scheme over a field). As discussed in the introduction, we may picture $S_{G,\mathcal{O}}$ as the moduli space of unipotent Weil–Deligne representations (r, N) valued in $G_{\mathcal{O}}$. In this context, "unipotent" means that the restriction of r to the inertia subgroup I_F is trivial (that is, $r(I_F) = 1$).

Proposition 2.5. (1) *Suppose q is considerate towards $G_{/\mathcal{O}}$. If $(\Phi, N) \in G \times \mathfrak{g}$ satisfies* $\mathrm{Ad}(\Phi)N = qN$, *then $N \in \mathcal{N}_G$. Hence,*

$$S_G(R) = \{(\Phi, N) \in G(R) \times \mathfrak{g}(R) : \mathrm{Ad}(\Phi)N = qN\},$$

and the requirement that $N \in \mathcal{N}_G(R)$ is redundant.

(2) *If G is split and $l > h_G$, then $S_{G,\mathcal{O}}$ is isomorphic to a closed subscheme of the moduli space of tame parameters $Z^1(W_F^0/P_F, G)_{\mathcal{O}}$* (see Section 1.2 of [DHKM20] for a definition of this space).

(3) *When l is geometrically G-banal, this space is a connected component of $Z^1(W_F^0/P_F, G)_{\mathcal{O}}$.*

Proof. Because $l > h_G$, the prime l is very good in the notation of Section 2.4 of [Cot22]. Hence, by Theorem 4.13 of [Cot22], we have an isomorphism of $\mathcal{O}$-algebras

$$\mathcal{O}[\mathfrak{g}]^G \to \mathcal{O}[\mathfrak{t}]^W$$

given by the restriction of functions on $\mathfrak{g}$ to $\mathfrak{t}$, where $\mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g}$ and W is the Weyl group.

By Chapter 3 of [Hum90] (see Table 1 of Section 3.7 and Table 2 of Section 3.18) the generators of $\mathcal{O}[\mathfrak{t}]^W$ are homogeneous of degree at most the Coxeter number h_G, and hence the same is true for $\mathcal{O}[\mathfrak{g}]^G$. (Although [Hum90] considers only the case of characteristic zero, the results extend to $\mathcal{O}$ because $|W|$ is invertible inside $\mathcal{O}$ and $\mathcal{O}[\mathfrak{t}]^W$ is a free $\mathcal{O}$-module.)

Let s be a generator of $\mathcal{O}[\mathfrak{g}]^G$ and suppose $(\Phi, N) \in G(R) \times \mathfrak{g}(R)$ satisfies $\mathrm{Ad}(\Phi)N = qN$. Then as s is G-invariant and homogeneous of degree at most the Coxeter number h_G, the condition $s(\mathrm{Ad}(\Phi)N) = s(qN)$ implies $s(N) = q^i s(N)$ for some $i \leq h_G$. As q is considerate towards $G_{/\mathcal{O}}$, we see that $q^i - 1$ is a nonzero divisor in $\mathcal{O}$, and hence that $s(N) = 0$. Thus the image of N in the GIT quotient $\mathfrak{g}//G$ is zero. Since l is very good, Theorem 4.12 of [Cot22] shows that N lies in the set of R-points of the nilpotent cone. Part (1) of the proposition follows.

Suppose G is a split group. As $l \neq p$, the space $Z^1 = Z^1(W_F^0/P_F, G)_{\mathcal{O}}$ has a model as a flat affine scheme over $\mathcal{O}$ with R-points equal to

$$Z^1(W_F^0/P_F, G)_{\mathcal{O}}(R) = \{(\phi, \sigma) \in G(R)^2 : \phi\sigma\phi^{-1} = \sigma^q\}.$$

Since $l > h_G$, the exponential and logarithm maps of Section 6 of [BDP17] are well defined polynomials, and thus we have an isomorphism between the nilpotent cone in $\mathcal{N}_G$ and unipotent cone $\mathcal{U}_G$. Hence, we have a map

$$S_{G,\mathcal{O}} \to Z^1(W_F^0/P_F, G)_{\mathcal{O}}, \quad (\Phi, N) \mapsto (\Phi, \exp N),$$

which is an isomorphism onto the closed subscheme of $Z^1(W_F^0/P_F, G)_{\mathcal{O}}$ given by those elements (ϕ, σ) with $\sigma \in \mathcal{U} \subseteq G$, where $\mathcal{U}$ is the unipotent cone. This shows (2).

For (3), suppose l is geometrically G-banal. Let $\mathcal{U}^+$ be the scheme-theoretic image of $Z^1(W_F^0/P_F, G)_{\mathcal{O}}$ through the second projection onto G. Proposition 2.6 of [DHKM20] tells us that the underlying reduced scheme of $\mathcal{U}^+$ is a subscheme of $\{\sigma \in G//G : \sigma^M = 1\}$ for some fixed $M \in \mathbb{N}$. Thus $\mathcal{U}^+$ is 0-dimensional over $\mathcal{O}$. As $Z^1(W_F^0/P_F, G)$ is flat and $\mathcal{O}$ is a discrete valuation ring, $\mathcal{U}^+$ is a finite flat $\mathcal{O}$-scheme.

We claim that no two distinct $\mathcal{O}$-points of $\mathcal{U}^+$ reduce to the same $\mathbb{F}$-point. The preimages of the L-points (resp. $\mathbb{F}$-points) of $\mathcal{U}^+$ in $Z^1(W_F^0/P_F, G)_L$ (resp. $Z^1(W_F^0/P_F, G)_{\mathbb{F}}$) are (unions of) connected components, thus it suffices to show that no two distinct points in $\mathcal{U}^+(L)$ reduce to the same point in $\mathcal{U}^+(\mathbb{F})$. This follows in turn from the statement that no two connected components of $Z^1(W_F^0/P_F, G)_L$ reduce to the same component of $Z^1(W_F^0/P_F, G)_{\mathbb{F}}$, which follows from Proposition 5.26 of [DHKM20], which states that $Z^1(W_F^0/P_F, G)_{\mathbb{F}}$ is reduced. We can conclude that the point $\operatorname{Spec} \mathcal{O} \hookrightarrow \mathcal{U}^+$ defined by $\sigma = 1$ is a connected component of $\mathcal{U}^+$. It follows that S_G, the preimage of this point, is a connected component of $Z^1(W_F^0/P_F, G)_{\mathcal{O}}$. □

We will also need the following results.

Proposition 2.6. (1) *The algebraic group G acts on S_G via simultaneous conjugation*:

$$g.(\Phi, N) = (g\Phi g^{-1}, \mathrm{Ad}(g)N).$$

Assume in addition that l is geometrically G-banal.

(2) *The scheme $S_{G,\mathcal{O}}$ is flat and a local complete intersection of relative dimension* $\dim G$ *over $\mathcal{O}$.*

(3) *Define the second projection map $p : S_G \to \mathcal{N}_G$ as before. If C is a $G_{/L}$-conjugacy class inside $\mathcal{N}_{G,L} \subseteq \mathcal{N}_G$, then the closed subscheme $X_C := \overline{p^{-1}(C)} \subset S_G$ is a union of irreducible components and $S_G = \bigcup_C X_C$.*

(4) *If in addition $G = \mathrm{GL}_n$, the X_C are irreducible components of $S_{n,\mathcal{O}} := S_{\mathrm{GL}_n,\mathcal{O}}$ and these can be naturally identified with partitions of n. For a partition p, we call the corresponding component X_p.*

(5) *The scheme $S_{G,\mathcal{O}}$ is reduced.*

Proof. (1) This is clear.

(2) This follows from Proposition 2.5(3) and Corollary 2.5 of [DHKM20].

(3) As $S_{G,\mathcal{O}}$ is flat over $\mathcal{O}$, the irreducible components of $S_{G,\mathcal{O}}$ are exactly those of the open subscheme $S_{G,L}$. This then follows from the proof of part 2, after noticing that $\mathcal{N}_{G,L} = \bigcup_C C_L$ as sets.

(4) Suppose $G = \mathrm{GL}_n$. Then C is a quotient of GL_n, and so is irreducible. Because centralisers inside GL_n are irreducible, the map $p^{-1}(C) \to C$ is flat with irreducible smooth fibres, and thus is smooth and open. By [Stacks, Lemma 004Z], it follows that $p^{-1}(C)$ is irreducible, and thus so is X_C. The final claim follows from the theory of Jordan normal forms.

(5) This follows from Proposition 2.8 of [DHKM20] and Proposition 2.5(3). □

2.1. ***Lemmas in commutative algebra and algebraic geometry.*** The remainder of this section proves some lemmas from algebraic geometry and commutative algebra that we will need later.

Lemma 2.7. *Let G be a smooth algebraic group over a scheme S, and let X be an S-scheme. Suppose that we have a morphism $m : G \times_S X \to X$ defining a group action of G on X. Then m is a smooth morphism.*

Proof. The morphism $p_X : G \times_S X \to X$ obtained by the base change of $G \to S$ is smooth. The automorphism ϕ of $G \times_S X$ given by $(g, x) \mapsto (g, g.x)$ is also smooth, since it is an isomorphism. As a composition of smooth morphisms, $m = p_X \circ \phi$ is smooth. □

Lemma 2.8. *Let $f : X \to Y$ be a smooth morphism of schemes. Let $p \in X$. Then Y is regular at $f(p)$ if and only if X is regular at p.*

Proof. After reducing the problem to local ring maps on stalks, this follows from Theorem 23.7 of [Mat86]. □

Lemma 2.9. *Assume that $\mathcal{O}$ is complete. Let R be a local $\mathcal{O}$-algebra, and assume that it is topologically of finite type with respect to the $\mathfrak{m}_R$-adic topology. Let x and $\bar{x}$ be prime ideals of R that give rise to the commutative diagram*

$$\begin{array}{ccccc} R & \xrightarrow{x} & \mathcal{O} & \hookrightarrow & L = \mathcal{O}[1/l] \\ & \searrow^{\bar{x}} & \downarrow & & \\ & & \mathbb{F} & & \end{array}$$

Then

$$R_{\bar{x}}^{\wedge}[1/l]_x^{\wedge} \cong R_x^{\wedge}.$$

Proof. Since $R \setminus x \supseteq R \setminus \bar{x} \cup \{1/l\}$, that $R_{\bar{x}}[1/l]_x \cong R_x$. As R is Noetherian, we have $\bigcap_n \bar{x}^n = 0$, and thus we have an injection $R_{\bar{x}} \hookrightarrow R_{\bar{x}}^{\wedge}$. This gives us a local homomorphism inclusion

$$R_x = R_{\bar{x}}[1/l]_x \hookrightarrow R_{\bar{x}}^{\wedge}[1/l]_x.$$

We notice that $R_x/x \cong L$, so that

$$\big[R_{\bar{x}}^{\wedge}[1/l]_x\big]/x \cong \big[\varprojlim_n (R/\bar{x}^n)/x\big][1/l] \cong \varprojlim_n \big(R/(x, l^n)\big)[1/l] \cong (\varprojlim \mathcal{O}/l^n)[1/l] = L,$$

the last equality arising because $\mathcal{O}$ is complete. Thus, by [Stacks, Lemma 0394], $R_{\bar{x}}^{\wedge}[1/l]_x^{\wedge}$ is the completion of $R_{\bar{x}}^{\wedge}[1/l]_x$ under the x-adic topology arising from R_x, and is a finite $R_x^{\wedge}$-module. It follows that the map

$$R_x^{\wedge} \to R_{\bar{x}}^{\wedge}[1/l]_x^{\wedge}$$

is an injection, and induces an isomorphism modulo x. If C is the cokernel (which we now know to be a finite $R_x^{\wedge}$-module) then we see $xC = C$, and so Nakayama's lemma shows us that $C = 0$, implying that the map is an isomorphism. □

3. Smoothness results for X_C

In this section we prove Theorem 1.2. Let G be a connected reductive group over $\mathcal{O}$ and $S_{G,\mathcal{O}}$ as before. As each map $X_C \to \operatorname{Spec}(\mathcal{O})$ is flat, we can (and do) reduce the problem to the case $\mathcal{O} = \mathbb{F}$ is a field of characteristic 0 or l (see [Stacks, Lemma 01V8]). Since smoothness is an fpqc-local property, we can assume that $\mathbb{F}$ is algebraically closed. We make these assumptions throughout this section.

3.1. ***Associated cocharacters.*** In what follows, we will require some setup, notation, and knowledge of Bala–Carter theory. Let G be a connected reductive group over an algebraically closed field $\mathbb{F}$ with Lie algebra $\mathfrak{g}$ and let $C \subseteq \mathcal{N}_G$ be a nilpotent orbit. In what follows, we restrict to the case where the derived subgroup of G is (almost) simple. When G is of adjoint type, we can do this because then $G = \prod_i G_i$ for G_i almost simple, adjoint, and $S_{G,\mathcal{O}} \cong \prod_i S_{G_i,\mathcal{O}}$. If G is not of adjoint type, then $S_{G,\mathcal{O}} \to S_{G^{ad},\mathcal{O}}$ is a $Z(G)$-torsor, and since $Z(G)$ is smooth (under our considerateness condition), any smoothness result translates between the cases for G and G^{ad}.

Let $\tilde{T}$ be a maximal torus of G and Π the set of roots. Let $e \in C \subseteq \mathfrak{g}$. Let $L \subseteq G$ be a Levi subgroup of G, minimal subject to $e \in \operatorname{Lie}(L)$. Let Z_L be the centre of L. Following Definition 2.8 of [FR08] (see also Section 2.3 of [Pre03]), the nilpotent element e is called distinguished in L if Z_L° is a maximal torus of the centraliser $C_L(e)$ of e in L. By Proposition 2.11 of [FR08], there is an associated cocharacter $\lambda : \mathbb{G}_m \to \tilde{T}$ such that $\operatorname{Ad}(\lambda(t)).e = t^2 e$ and $\operatorname{im}(\lambda) \subseteq [L, L]$.

The group $\mathbb{G}_m$ acts on $\operatorname{Lie}(G)$ through $\operatorname{Ad} \circ \lambda : \mathbb{G}_m \to \operatorname{Aut}(\mathfrak{g})$, which gives a decomposition

$$\mathfrak{g} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}(\lambda, i).$$

Through Lemma 5.6.5 of [Car93] and the preceding discussion, we choose a base of simple roots $\Delta \subseteq \Pi$ such that $\langle \alpha, \lambda \rangle \geq 0$ for all $\alpha \in \Delta$. Call the corresponding Borel subgroup B.

We define a parabolic subgroup $P_\lambda \subseteq L$ such that $\mathrm{Lie}(P_\lambda) = \bigoplus_{i \geq 0} \mathfrak{g}_L(\lambda, i)$, where each $\mathfrak{g}_L(\lambda, i) = \mathfrak{g}(\lambda, i) \cap \mathrm{Lie}(L)$. We note that P_λ has a Levi decomposition $P_\lambda = M_\lambda U_\lambda$ with $\mathrm{Lie}(U_\lambda) = \bigoplus_{i>0} \mathfrak{g}_L(\lambda, i)$ and $\mathrm{Lie}(M_\lambda) = \mathfrak{g}_L(\lambda, 0)$. We say P_λ is a *distinguished parabolic subgroup* in L if $\dim \mathfrak{g}_L(\lambda, 0) = \dim(\mathfrak{g}_L(\lambda, 2)) + \dim(Z_L)$, and by Proposition 2.5 of [Pre03] e is distinguished if and only if P_λ is distinguished.

The primary result of Bala–Carter theory is that there is a bijection between the adjoint orbits of $\mathcal{N}_G$ and pairs (M, P), where M is a Levi subgroup of G, and P is a distinguished parabolic subgroup of M.

3.2. ***The smoothness result for irreducible components corresponding to distinguished nilpotent orbits.*** This section contains the main result of the paper, concerning the smoothness of X_C (Theorem 3.2).

Lemma 3.1. *Assume $e \in C$ is a distinguished nilpotent element with λ an associated cocharacter. Let $C_G(e)$ be the centraliser of e in G with Levi decomposition $C_G(e) = MR$, where $R = R_u(C_G(e))$ is unipotent and M is reductive. Suppose that $t \in \mathbb{G}_m$ is sufficiently generic so that $\mathfrak{g}^{\mathrm{Ad}(\lambda(t))} = \mathfrak{g}(\lambda, 0)$. Then every element of $C_G(e)\lambda(t)$ is conjugate to an element of $M\lambda(t)$.*

Proof. Since e is distinguished, Theorem A of [Pre03] tells us that $M = M_\lambda \cap C_G(e)$ and $R = U_\lambda \cap C_G(e)$. Further, following Definition 2.8 of [FR08], the maximal torus of M is Z_G° so that M/Z_G° is a rank 0 reductive group; ergo finite. Thus, all unipotent elements of $C_G(e)$ lie in R. Let $g \in \lambda(t)C_G(e) \subseteq \tilde{T}C_G(e)$ have abstract Jordan decomposition $g = su$ with s semisimple and u unipotent. As s is semisimple, it lies in a maximal torus T' of $\tilde{T}C_G(e)$ and so there is some $x \in \tilde{T}C_G(e)$ such that $xsx^{-1} \in \tilde{T}$. In fact, we can assume without loss of generality that $x \in C_G(e)$, because $\tilde{T}$ is abelian. Then, because $\lambda(t)$ normalises $C_G(e)$, we see that $xgx^{-1} \in \lambda(t)C_G(e)$. Since u is unipotent in $\tilde{T}C_G(e)$ we obtain $u \in C_G(e)$, and thus we get

$$xsx^{-1} = xgx^{-1}(xux^{-1})^{-1} \in \lambda(t)C_G(e) \cap \tilde{T}.$$

Hence

$$xsx^{-1} \in \tilde{T} \cap \lambda(t)C_G(e) = \tilde{T} \cap \tilde{T}M \cap \lambda(t)C_G(e) = \lambda(t)\tilde{T} \cap \lambda(t)M = \lambda(t)[\tilde{T} \cap M].$$

Then $xux^{-1} \in C_G(xsx^{-1}) = C_G(\lambda(t)s')$ for some $s' \in \tilde{T} \cap M$. By the genericity condition on t, we see that $C_G(\lambda(t)) = M_\lambda$ and so, by Theorem 3.5.3 of [Car93], $C_G(\lambda(t)s') \subseteq C_G(\lambda(t)) = M_\lambda$. Hence $xux^{-1} \in M$ and $xux^{-1} = 1$. The result follows. □

Theorem 3.2. *Let $G_{/\mathcal{O}}$ be a connected reductive group with centre Z, let $\mathfrak{g}$ be the Lie algebra of G, and suppose $q \in \mathcal{O}$ is considerate towards G over $\mathcal{O}$. Suppose $C \subseteq \mathcal{N}_{G,L}$ is either 0 or a distinguished nilpotent adjoint orbit. Then X_C is smooth over $\mathcal{O}$ and there is a bijection between the connected components of X_C and the set of Φ_0-twisted conjugacy classes of the group $\pi_0(C_G(e))$.*

Proof. Consider first the case $C = 0$. Then $X_C = \{(\Phi, 0) \in S_{G,\mathcal{O}}\} \cong G$ via the map projecting to the Φ coordinate. Since G is smooth, this proves the theorem.

Keep the notation of before, with e a distinguished nilpotent element of $\mathfrak{g}$ with associated cocharacter $\lambda : \mathbb{G}_m \to T$. After making some choice of $\sqrt{q} \in \mathbb{F}$, set $\Phi_0 = \lambda(\sqrt{q})$. The centraliser $C_G(e)$ exhibits a Levi decomposition $C_G(e) = MR$ with M reductive and R the unipotent radical. Because e is a distinguished nilpotent element in $\mathfrak{g}$, we see that Z°, the connected component of the identity of the centre Z of G, is a maximal torus of M by Proposition 2.11(iii) of [FR08]. Since M/Z° is a split reductive group of rank 0, it is finite; because unipotent radicals are connected, it is isomorphic to the component group $A(e)$ of the centraliser $C_G(e)$. We choose a set of representatives S such that $M = \coprod_{s \in S} sZ^\circ$.

Define $Y = M\Phi \times \mathfrak{g}(\lambda, 2)$. Since the characteristic l exceeds h, we see that M is a smooth subgroup of G making Y a smooth $\mathbb{F}$-scheme. It is clear that Y is a closed subscheme of S_G, because if $(m\Phi_0, N) \in M\Phi_0 \times \mathfrak{g}(\lambda, 2)$ then

$$\mathrm{Ad}(m\Phi_0).N = \mathrm{Ad}(m)\mathrm{Ad}(\Phi_0).N = \mathrm{Ad}(m).qN = qN.$$

In fact, Y lies inside the irreducible component X_C. The distinguished element e lies in the unique open dense P_λ-orbit inside $\mathfrak{g}(2, \lambda)$ (See, for example, Proposition 5.8.7b of [Car93]) and thus $P_\lambda.[M\Phi_0 \times \{e\}]$ is a dense open subscheme of Y. As $M\Phi_0 \times \{e\} \subseteq p^{-1}(C)$, we obtain $P_\lambda.[M\Phi_0 \times \{e\}] \subseteq p^{-1}(C)$ and consequently $Y \subseteq X_C$. Define the morphism

$$f : G \times Y \to X_C, \quad (g, (\Phi, N)) \mapsto (g\Phi g^{-1}, \mathrm{Ad}(g)N).$$

As $G \times Y$ is a smooth variety, Lemma 2.8 implies the theorem provided we can show that f is a smooth surjective morphism.

Surjectivity is equivalent to the statement that every pair $(\Phi, N) \in X_C$ is conjugate to a pair in Y. To prove this, it suffices to show that Φ is conjugate to an element of $M\Phi_0$ whenever $(\Phi, N) \in X_C$. As there is some $\mathrm{Ad}(g).e \in \mathfrak{g}$ upon which Φ acts as ultiplication by q, we see that Φ is conjugate to some element of $C_G(e)\Phi_0$. By Lemma 3.1, then, Φ is conjugate to an element of $M\Phi_0$, proving surjectivity.

We now proceed to prove that f is smooth. Consider the commutative diagram

$$\begin{array}{ccc} G \times Y & \xrightarrow{f} & X_C \\ \downarrow & & \downarrow \\ G \times M\Phi_0 & \longrightarrow & G \end{array}$$

where the vertical maps come from the "forget N" projections $(g, m\Phi_0, N) \in G \times Y \mapsto (g, m\Phi_0) \in G \times M\Phi_0$ and $(\Phi, N) \in X_C \mapsto \Phi \in G$ and the horizontal maps are defined via the conjugation action of $g \in G$ on Y, so that the diagram commutes. Choosing a set of representatives S of M/Z°, the map $G \times M\Phi_0 \to G$ factors through

$$\begin{array}{ccc} G \times M\Phi_0 & \longrightarrow & G \\ \| & & \uparrow \\ \coprod_{s \in S} G \times sZ^\circ\Phi_0 & \xrightarrow{m} & \coprod_{s} Z^\circ G_{s\Phi_0} \end{array}$$

where $G_{s\Phi_0}$ denotes the conjugacy class of $s\Phi_0$ in G. Each $Z^\circ G_{s\Phi_0}$ defines a locally closed subvariety of G. If any two subschemes $Z^\circ G_{s\Phi_0}$ and $Z^\circ G_{t\Phi_0}$ intersect in G, say $x \in Z^\circ G_{s\Phi_0} \cap Z^\circ G_{t\Phi_0}$, then

$$z_1 g_1 s \Phi_0 g_1^{-1} = x = z_2 g_2 t \Phi_0 g_2^{-1}$$

implies that $t\Phi_0 = (z_1 z_2^{-1})(g_2^{-1} g_1) s \Phi_0 (g_2^{-1} g_1)^{-1}$; and thus leads us to $Z^\circ G_{s\Phi_0} = Z^\circ G_{t\Phi_0}$ as locally closed subschemes of G. Thus, by possibly restricting to a subset S' of the set of representatives S if necessary, we can view $\coprod_{s \in S'} Z^\circ G_{s\Phi_0}$ as a locally closed subscheme of G.

Because the map $f : G \times Y \to X_C$ is surjective, the map $X_C \to G$ also factors through $\coprod_{s \in S'} Z^\circ G_{s\Phi_0}$, giving us the commutative diagram

$$\begin{array}{ccc} G \times Y & \xrightarrow{f} & X_C \\ \downarrow & & \downarrow \\ G \times M\Phi_0 & \xrightarrow{m} & \coprod_{s \in S'} Z^\circ G_{s\Phi_0} \end{array} \tag{1}$$

We claim that this is a pullback square. Since e is distinguished, we see $\mathfrak{g}(\lambda, 1) = 0$ and thus each simple root α has its corresponding character eigenspace $\mathfrak{g}_\alpha$ either contained inside $\mathfrak{g}(\lambda, 0)$ or $\mathfrak{g}(\lambda, 2)$. Hence, as every positive root is the sum of at most $h - 1$ simple roots (where $h = h_G$ is the Coxeter number of G), we see that

$$\{i \in \mathbb{Z} : \mathfrak{g}(\lambda, i) \neq 0\} \subseteq 2\mathbb{Z} \cap [-2h + 2, 2h - 2].$$

Given that q is considerate towards $G_{/\mathbb{F}}$, it follows that the subspace of $\mathfrak{g}$ upon which $\Phi_0 = \lambda(\sqrt{q})$ acts as multiplication by q is precisely $\mathfrak{g}(\lambda, 2)$.

Thus, if $(g, \Phi) \in G \times M\Phi_0$ and $(\Phi', N') \in X_C$ are such that $g\Phi g^{-1} = \Phi'$, then $\mathrm{Ad}(\Phi)(\mathrm{Ad}(g^{-1})N') = q\mathrm{Ad}(g^{-1})N'$ and $\mathrm{Ad}(g^{-1})N' \in \mathfrak{g}(\lambda, 2)$ by the previous discussion. So the morphism

$$\big((g, \Phi), (\Phi', N')\big) \to \big(g, (\Phi, \mathrm{Ad}(g^{-1})N')\big)$$

gives an inverse to the natural morphism $G \times Y \to (G \times M\Phi_0) \times_{\coprod_s Z^\circ G_{s\Phi_0}} X_C$. This shows that the commutative diagram (1) is a pullback square.

By the theory of homogeneous spaces, the bottom map m is flat with fibres isomorphic to $\mathrm{Stab}_G(\Phi_0)$, which are smooth group schemes. This shows that m is smooth. Hence, since f is the base change of m to X_C, by Proposition 10.1 of [Har77] we see that f is smooth. We conclude X_C is smooth over $\mathbb{F}$.

The statement on the number of connected components is Theorem 2.5 of [Sho24], and is included for completeness. □

Remark. A question arises regarding the generality and necessity of the considerateness condition: When exactly is q-considerateness a necessary condition for smoothness? As one can see in the proof, we used considerateness to prove that $G \times Y$ was the pullback of the diagram

$$\begin{array}{ccc} G \times Y & \longrightarrow & X_C \\ \downarrow & & \downarrow \\ G \times M\Phi_0 & \longrightarrow & M.G_{\Phi_0} \end{array}$$

arising from the fact that $\{N \in \mathfrak{g} : \mathrm{Ad}(\Phi_0)N = qN\} = \mathfrak{g}(\lambda, 2)$. When C is the *regular* nilpotent orbit, $\mathfrak{g}(\lambda, i) \neq 0$ for every $i \in [-2h+2, 2h-2] \cap 2\mathbb{Z}$, so we see that q-considerateness is precisely the condition that $\{N \in \mathfrak{g} : \mathrm{Ad}(\Phi_0)N = qN\} = \mathfrak{g}(\lambda, 2)$. When C is distinguished and nonregular, we have $\mathfrak{g}(\lambda, 2h-2) = 0$. Hence, there is some $r < h$, depending on the distinguished orbit C, such that if $q \in \mathcal{O}$ has the property that $1, q, \dots, q^r$ are distinct, then X_C is smooth via the above proof. This r can always be taken to be

$$r = 1 + \max\{i : \mathfrak{g}(\lambda, 2i) \neq 0\}.$$

Remark. Let X_C^s be the image of f restricted to $G \times sZ^\circ\Phi_0$ (so that X_C^s is an irreducible component of X_C). Another way of interpreting the Cartesian diagram (1) is that $X_C^s \to Z^\circ G_{s\Phi_0}$ is the total space of the vector bundle on $Z^\circ G_{s\Phi_0}$ whose fibre above Φ is $\ker(\mathrm{Ad}(\Phi) - q) \subseteq \mathfrak{g}$. In particular, we obtain:

Corollary 3.3. *If C is a distinguished orbit and s, Φ_0 are as in the proof of Theorem 3.2, then X_C^s is described as the closed subscheme of $G \times \mathfrak{g}$ cut out by the equations*

- $\mathrm{Ad}(\Phi)N = qN$ *and*
- *any set of equations describing the closed orbit $Z^\circ G_{s\Phi_0} \subseteq G$.*

3.3. *The converse nonsmoothness result.* We now begin work towards the converse of Theorem 3.2. Consider the situation described at the beginning of the chapter with reductive group G over an algebraically closed field $\mathbb{F}$ with (almost) simple derived subgroup with maximal torus T, set of roots Π, a set of simple roots $\Delta \subseteq \Pi$, and a nilpotent element e with associated cocharacter λ. Let L be the smallest Levi subgroup of G with $e \in \mathrm{Lie}(L)$, so that e is distinguished for L. Let $\Delta_L \subseteq \Delta$ be the simple roots of L.

Theorem 3.4. *Let G be as before. Suppose $C \subseteq \mathcal{N}_G$ is a nilpotent adjoint orbit, distinguished in a proper and nontrivial Levi subgroup $T \neq L \subsetneq G$. Then $X_C \subseteq S_G$ is singular.*

Before proving this theorem, we need some terminology and a lemma. Let $D = (\Delta, \Sigma)$ be the Dynkin diagram of G, and $D_L = (\Delta_L, \Sigma_L) \subseteq D$ the maximal subdiagram containing exactly the vertices Δ_L. (Note that D is connected when its derived subgroup is (almost) simple but D_L may not be connected). Recall that, given a distinguished nilpotent e, one can attach to each vertex $\beta \in \Delta_L$ the number $\langle \alpha, \lambda \rangle \in \{0, 2\}$, and this is called the weighted Dynkin diagram $D_L(e)$.

We call a root $\alpha \in \Delta_L$ *exposed* if there is an edge in D connecting α to a root $\beta \in \Delta \setminus \Delta_L$.

Lemma 3.5. *Any exposed root $\alpha \in \Delta_L$ has $\langle \alpha, \lambda \rangle = 2$.*

Proof. In the case of types A, B, C and D, either all simple factors of the Levi subgroup are of type A, or exactly one of the almost simple factors of L is of type B, C, D, respectively. If all the simple factors of the Levi subgroup are type A, then all roots $\alpha \in \Delta_L$ have $\langle \alpha, \lambda \rangle = 2$ because all distinguished nilpotents orbits are regular in type A. In the case with one factor of type B, C or D, the only "exposed" root of this factor is on the end of the string, and one can see from the tables on pages 174 and 175 of [Car93]

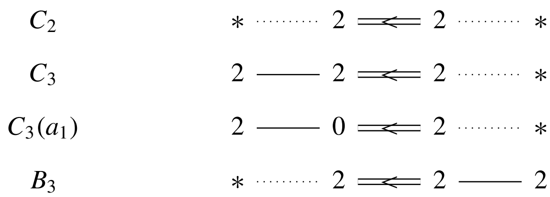


Figure 1. The distinguished weighted Dynkin diagrams of Levi subgroups of F_4.

that, independent of the choice of distinguished nilpotent, this exposed root α always has $\langle \alpha, \lambda \rangle = 2$. This proves the lemma in types A, B, C and D.

In type G_2, the only proper Levi subgroups are of type A_1, so all roots $\alpha \in \Delta_M$ have $\langle \alpha, \lambda \rangle = 2$.

In type F_4, there are three possibilities for a Levi factor not of type A, these being C_2, C_3 and B_3. The distinguished orbits of these Levi subgroups are described on pages 174 and 175 of [Car93] and are listed in Figure 1. From this we directly see that all exposed roots have $\langle \alpha, \lambda \rangle = 2$.

In type E_l, there are three possibilities for the Levi subgroup types. Either there are only factors of type A for which the result holds, or there is a unique factor of type D_n and $n \leq 7$, or there is a factor of type E_6 or E_7. In the case of a factor of type E_6 or E_7, the weighted Dynkin diagrams of distinguished parabolic subgroups (of which there are 3 of type E_6 and 6 of type E_7) are listed on page 176 of [Car93], from which it is clear that all exposed roots α have the desired property. See also the figures on page 1157.

This leaves only the case that L has a factor of type D_n. All distinguished orbits of D_n with $n \leq 7$ are listed on 1156, from whose table we see that all exposed roots have the desired property. □

Remark. In the description of nonregular distinguished parabolic subgroups in type B_l on page 175 of [Car93], one requires $k \geq 2$ (in the notation of the source) for the conditions to make sense, though this isn't explicit. In our application, the important fact is that for B_l with $l = 3$, the only distinguished orbit is the regular orbit. This can also be seen from the description of distinguished orbits in Theorem 8.2.14 of [CM93] via partitions of $2l + 1 = 7$ into distinct odd parts.

Proof of Theorem 3.4. Consider a point $P = (\Phi_0, 0) \in X_C$ with $\Phi_0 \in T$. Define four subvarieties of S_G that contain P as follows:

(1) Let $\mathcal{O} = G.P$ be the G-orbit of P.

(2) Let $\tilde{T}$ be the maximal torus of G seen as a closed subvariety of S_G via the inclusion $\Phi \mapsto (\Phi, 0)$.

(3) Let $\mathcal{N}_0 = \{N \in \mathfrak{g} : \mathrm{Ad}(\Phi_0)N = qN\}$ viewed as a closed subvariety of S_G via the inclusion $N \mapsto (\Phi_0, N)$.

(4) Let $U_0 = U^- \cap \mathrm{Stab}_G(P)$ where U^- is the opposite unipotent subgroup to $[B, B]$, viewed as a closed subvariety of S_G via the inclusion $u \mapsto (\Phi_0 u, 0)$.

Claim 1. *The tangent spaces $T_P\mathcal{O}$, $T_P\tilde{T}$, $T_P\mathcal{N}_0$ and T_PU_0 form a direct sum inside T_PS_G.*

Observe that $T_P(S_G) \subseteq T_P(G \times \mathfrak{g}) = \mathfrak{g} \times \mathfrak{g}$. We briefly describe $T_P\mathcal{O}$ as a subspace of $\mathfrak{g} \times \{0\}$ (which we will conflate with $\mathfrak{g}$ as this shouldn't cause confusion).

Consider the map $f : G \to G : g \mapsto g\Phi_0 g^{-1}\Phi_0^{-1}$, comprised of the conjugation action of g on Φ_0 followed by right multiplication by Φ_0^{-1} (to ensure the identity is sent to the identity). Then $\mathcal{O}$ is isomorphic to the set-theoretic image of f, which is a locally closed subscheme of G. The derivative of this map is $\mathrm{id} - \mathrm{Ad}(\Phi_0) : \mathfrak{g} \to \mathfrak{g}$, and it factors through $\mathfrak{g} \twoheadrightarrow T_P C \hookrightarrow \mathfrak{g}$. We hence see a natural identification of $T_P\mathcal{O}$ with $\mathrm{im}(\mathrm{id} - \mathrm{Ad}(\Phi_0))$.

We now proceed to prove the claim. Firstly, $T_P\tilde{T} \cap T_P U_0 = \{0\}$ because $\mathrm{Lie}(\tilde{T}) \cap \mathrm{Lie};\ (U^-) = \{0\}$. Next, consider $T_P\mathcal{O} \cap (T_P\tilde{T} \oplus T_P U_0)$. Note that $U_0, \tilde{T} \subseteq \mathrm{Stab}_G(\Phi) \subseteq G$ and $T_P\mathrm{Stab}_G(\Phi_0) = \ker(\mathrm{id} - \mathrm{Ad}(\Phi_0))$. Then because Φ_0 is semisimple, the intersection of $\mathrm{im}(\mathrm{id} - \mathrm{Ad}(\Phi_0))$ and $\ker(\mathrm{id} - \mathrm{Ad}(\Phi_0))$ is trivial (this is easily checked for GL_n, and extended to all G because G can always be embedded into some GL_N). Hence $T_P(\mathrm{Stab}_G(\Phi_0)) \cap T_P\mathcal{O} = 0$, and thus $T_P\mathcal{O} \cap (T_P\tilde{T} \oplus T_P U_0) = 0$.

To show that $T_P\mathcal{N}_0$ intersects $T_P\mathcal{O} + T_P\tilde{T} + T_P U_0$ at the origin, it suffices to notice that an element of $T_P C + T_P\tilde{T} + T_P U_0$ takes the form $P' = (\Phi', 0)$, while an element $P' \in T_P\mathcal{N}_0$ takes the form $P' = (\Phi_0, N) \in S_G(\mathbb{F}[\epsilon]/\epsilon^2)$. For these to be equal, we must have $\Phi' = \Phi_0$ and $N = 0$, so $P' = P$. This proves the claim.

Suppose that $C \subseteq \mathcal{N}_G$ is an adjoint orbit, neither zero nor a distinguished nilpotent orbit as in the hypothesis. Define $e \in C$, a choice of maximal torus and Borel $\tilde{T} \subset B$, the associated cocharacter λ and minimal Levi $L \subset G$, all as in the general setup. Set $D = \tilde{T} \cap X_C$ and $\mathcal{N}_1 = \mathcal{N}_0 \cap X_C$, and note that $U_0, \mathcal{O} \subseteq X_C$ already. Our aim is to show that there is a point $P \in X_C$ such that

$$\dim_{\mathbb{F}}(T_P\mathcal{O}) + \dim_{\mathbb{F}}(T_P D) + \dim_{\mathbb{F}}(T_P\mathcal{N}_1) + \dim_{\mathbb{F}}(T_P U_0) > \dim(X_C) = \dim(G).$$

It is clear that whenever $z \in Z_L$ and $\mu \in \mathbb{F}$, the point $(z\lambda(\sqrt{q}), \mu e)$ lies in X_C. We choose the point $P = (\Phi_0, 0) \in X_C$ with $\Phi_0 = z\lambda(\sqrt{q})$ for some $z \in Z_L$, which we will determine momentarily.

Regardless of the choice of z for now, recall the decomposition

$$\mathrm{Lie}(L) = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}_L(\lambda, i).$$

As $\mathfrak{g}_L(\lambda, 0)$ is a Levi subalgebra of $\mathrm{Lie}(L)$, there is a Levi subgroup $M_0 \subset L$ with $\mathrm{Lie}(M_0) = \mathfrak{g}_L(\lambda, 0)$. It is clear that $M_0 \subseteq \mathrm{Stab}_G(\Phi_0)$. Observe also that $\mathfrak{g}_L(\lambda, 2) \subseteq \mathcal{N}_1$ and $Z_L\Phi_0 \subseteq D$. We define positive integers

(1) $\epsilon_0 := \dim(\mathrm{Stab}_G(\Phi_0)) - \dim_{\mathbb{F}}(\mathfrak{g}_L(\lambda, 0)) = \dim(G) - \dim(\mathcal{O}) - \dim_{\mathbb{F}}(\mathfrak{g}_L(\lambda, 0))$,

(2) $\epsilon_1 := \dim_{\mathbb{F}}(T_P D) - \dim(Z_L)$,

(3) $\epsilon_2 := \dim_{\mathbb{F}}(T_P\mathcal{N}_1) - \dim_{\mathbb{F}}(\mathfrak{g}_L(\lambda, 2))$,

(4) $\epsilon_3 := \dim_{\mathbb{F}}(T_P U_0)$.

Putting this together, and using the fact that

$$\dim(\mathfrak{g}_L(\lambda, 0)) = \dim(\mathfrak{g}_L(\lambda, 2)) + \dim(Z_L)$$

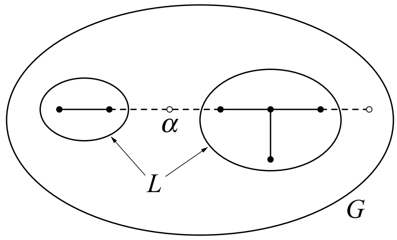


Figure 2. An example of α in the case where G is of type E_8 and L of type $D_4 \times A_2$.

because e is distinguished inside L (this follows from a generalisation of Lemma 8.2.1 in [CM93] to reductive groups in good characteristic), we see that

$$\begin{aligned}\dim(T_P X_C) &\geq \dim(T_P\mathcal{O}) + \dim(T_P D) + \dim(T_P\mathcal{N}_1) + \dim(T_P U_0)\\ &\geq [\dim(G) - \dim(\mathfrak{g}_L(\lambda, 0)) - \epsilon_0] + [\dim(Z_L) + \epsilon_1] + [\dim(\mathfrak{g}_L(\lambda, 2)) + \epsilon_2] + \epsilon_3\\ &= \dim(G) + \epsilon_1 + \epsilon_2 + \epsilon_3 - \epsilon_0.\end{aligned}$$

Thus, it is enough to find some choice of $z \in Z_L$ such that $\epsilon_1 + \epsilon_2 + \epsilon_3 - \epsilon_0 > 0$.

Fix some root $\alpha \in \Delta \setminus \Delta_L$ adjacent to a root in Δ_L in the Dynkin diagram. See Figure 2 for an example. There is a morphism of algebraic groups

$$B = \prod_{\beta\in\Delta} \beta : \tilde{T} \to \prod_{\beta\in\Delta} \mathbb{G}_m$$

which is surjective and has kernel $Z \subset \tilde{T}$. Because B is surjective, we can choose $z \in T$ such that $\alpha(z) = 1$ and $\beta(z) = 1$ whenever $\beta \in \Delta_L$ (so that $z \in Z_L$) and $\beta(z) = q \neq 1$ in all other cases. Consider $\Phi_0 = z\lambda(\sqrt{q})$. Whenever $\gamma \in \Pi_G^+$ is a positive root of G, it decomposes as a product of simple roots $\gamma = \prod_{\beta\in\Delta} \beta^{c_\beta} \in X(T)$ with $\sum_\beta c_\beta < h$. By design, each $\beta(\Phi_0)$ is either 1 or q, so $\gamma(\Phi_0) \in \{1, q, q^2, \dots, q^{h-1}\}$, and $\gamma(\Phi_0) = 1$ if and only if all the simple roots with $c_\beta \neq 0$ satisfy $\beta(\Phi_0) = 1$.

Claim 2. *If $\gamma \in \Pi_G^+ \setminus \Pi_L^+$ has $\gamma(\Phi_0) = 1$, then $\gamma = \alpha$.*

If γ is simple, then $\gamma = \alpha$ by our choice of Φ_0. Suppose for contradiction that γ is not simple. Then it contains at least two simple roots in its decomposition, and one of these must be α, as otherwise all simple roots are in Δ_L and $\gamma \in \Pi_L^+$. There must also be another root $\beta \in \Delta_L$ with $c_\beta \neq 0$, and for γ to be a root, there must be a path from α to β (in the Dynkin diagram) passing through vertexes β' with $c_{\beta'} \neq 0$, and each of these as such (since $\gamma(\Phi_0) = 1$) has $\beta'(\Phi_0) = 1$. But as $\alpha \in \Delta \setminus \Delta_L$ and $\beta \in \Delta_L$, at least one of the β' is an exposed root, and thus $\beta'(\Phi_0) = q$ by Lemma 3.5. This is a contradiction. We conclude that $\gamma(\Phi_0) = 1$ implies either $\gamma = \alpha$ or $\gamma \in \Pi_L^+$. This proves the claim.

When $\beta \in \Pi$, denote the root subgroup of β by $U_\beta \leq G$. By Theorem 3.5.3 of [Car93], the (connected) centraliser of Φ_0 is

$$C_G(\Phi_0)^\circ = \langle \tilde{T}, U_\beta, U_{-\beta} : \beta(\Phi_0) = 1 \rangle .$$

The subgroup generated by $\tilde{T}$ and all U_β with $\beta(\Phi_0) = 1$ and $\beta \in \Pi_L$ is simply M_0°, so we see that

$$C_G(\Phi_0)^\circ = \langle M_0^\circ, U_\alpha, U_{-\alpha} \rangle$$

and hence $\dim C_G(\Phi_0) = \dim(M_0) + 2$ (or, in other words, $\epsilon_0 = 2$).

The reflection $s_\alpha \in N(\tilde{T})/\tilde{T}$ acts on $\tilde{T}$ and stabilises Φ_0; thus it acts on $T_P X_C$. Further, this action preserves the subspaces $T_P D$ and $T_P \mathcal{N}_1$. However, since α is adjacent to a simple root of L, the reflection s_α does not preserve the Levi subgroup L, and hence preserves neither Z_L nor $\mathfrak{g}_L(\lambda, i)$, and we see that $s_\alpha(Z_L) \neq Z_L$ and $s_\alpha(\mathfrak{g}_L(\lambda, 2)) \neq \mathfrak{g}_L(\lambda, 2)$, so that $T_P D \supseteq s_\alpha(T_P Z_L) \cup T_P Z_L$ and $T_P \mathcal{N}_1 \supseteq s_\alpha(T_P \mathfrak{g}_L(\lambda, 2)) \cup T_P \mathfrak{g}_L(\lambda, 2)$, forcing $\epsilon_1 > 0$ and $\epsilon_2 > 0$ respectively.

For ϵ_3, consider any choice of isomorphism $u_{-\alpha} : \mathbb{G}_a \xrightarrow{\sim} U_{-\alpha}$ and the adjoint action of $u_{-\alpha}(a) \in U_{-\alpha}$ on $e = \sum_\beta e_\beta \in \mathrm{Lie}(L)$. Because α is a simple root in $\Delta_G \setminus \Delta_L$, we see that

$$\Big[e_{-\alpha}, \sum_\beta e_\beta\Big] = \sum_\beta [e_{-\alpha}, e_\beta] = 0$$

and thus that $\mathrm{Ad}(u_{-\alpha}(a))e = e$. Hence,

$$\big\{(\Phi_0 u_{-\alpha}(a), \mu e) : a \in \mathbb{G}_a, \mu \in \mathbb{G}_m\big\}$$

is a locally open subscheme of $p^{-1}(C)$, from which we see that $(\Phi_0 u_{-\alpha}(\epsilon), 0)$ is a deformation in $T_P U_0$, forcing $\epsilon_3 > 0$.

We then obtain the inequality $\epsilon_1 + \epsilon_2 + \epsilon_3 - \epsilon_0 \geq 3 - 2 = 1$, proving that $(\Phi_0, 0)$ is a singular point of X_C. □

Example 1. Consider the group

$$G = \mathrm{GSp}_4(R) = \big\{M \in \mathrm{GL}_4(R) : M\Omega M^{-1} = \lambda\Omega \text{ for some } \lambda \in \mathbb{G}_m(R)\big\}, \quad \Omega = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & -1 & & \\ -1 & & & \end{pmatrix},$$

where Ω is chosen so that a Borel subgroup can be given by the intersection of GSp_4 with the upper triangular matrices in GL_4. We let $L = \mathrm{GL}_2 \subseteq G$ be the Levi subgroup corresponding to the short root. Then

$$e = \begin{pmatrix} 0 & 1 & & \\ & 0 & 0 & \\ & & 0 & -1 \\ & & & 0 \end{pmatrix}$$

is distinguished in L and the associated cocharacter is $\lambda(t) = \mathrm{Diag}(t, t^{-1}, t, t^{-1})$. We choose $\Phi_0 = \mathrm{Diag}(q, 1, 1, q^{-1})$ and α to be the root corresponding to the one-parameter subgroup defined as

$$U_\alpha = \left\{\begin{pmatrix} 1 & & & \\ & 1 & * & \\ & & 1 & \\ & & & 1 \end{pmatrix}\right\}.$$

Explicitly, we now see that $\mathrm{Stab}_G(P) = \mathbb{G}_m \times \mathrm{GL}_2$. We can also describe the subvarieties:

$$Z_L = \left\{\left(\begin{pmatrix} ab & & & \\ & ab & & \\ & & ab^{-1} & \\ & & & ab^{-1} \end{pmatrix}, 0\right) : a, b \in \mathbb{G}_m\right\},$$

$$s_\alpha(Z_L) = \left\{ \left(\begin{pmatrix} ab & & & \\ & ab^{-1} & & \\ & & ab & \\ & & & ab^{-1} \end{pmatrix}, 0 \right) : a, b \in \mathbb{G}_m \right\},$$

$$\mathfrak{g}_L(\lambda, 2) = \left\{ \left(\Phi_0, \begin{pmatrix} 0 & a & & \\ & 0 & & \\ & & 0 & -a \\ & & & 0 \end{pmatrix} \right) : a \in \mathbb{G}_a \right\},$$

$$s_\alpha(\mathfrak{g}_L(\lambda, 2)) = \left\{ \left(\Phi_0, \begin{pmatrix} 0 & & a & \\ & 0 & & a \\ & & 0 & \\ & & & 0 \end{pmatrix} \right) : a \in \mathbb{G}_a \right\},$$

$$U_0 = \left\{ \left(\begin{pmatrix} q & & & \\ & 1 & & \\ & a & 1 & \\ & & & q^{-1} \end{pmatrix}, 0 \right) : a \in \mathbb{G}_a \right\}.$$

When we put all this together, we see the contribution from $D, \mathcal{N}_1$ and U_0 is 6-dimensional, and thus

$$\dim T_P X_C \geq \dim(\mathrm{GSp}_4) - \dim(\mathrm{Stab}_G(P)) + 6 = \dim(\mathrm{GSp}_4) + 1.$$

We can piece Theorems 3.2 and 3.4 together:

Corollary 3.6. *Let G be a connected reductive group over a field $\mathbb{F}$ of characteristic 0 or l, and suppose $G^{\mathrm{ad}} = \prod_i G_i$ where each G_i has a (almost) simple derived subgroup. Suppose q is considerate towards each $G_{i,\mathbb{F}}$. Then the smooth irreducible components of $S_{G^{\mathrm{ad}}}$ are precisely those of the form $\prod_i X_i$ where each $X_i \subseteq S_{G_i}$ is a smooth irreducible component of S_{G_i}. That is, each X_i corresponds to a distinguished nilpotent orbit of G_i or the zero orbit. The smooth components of S_G are precisely preimages the smooth components of $S_{G^{\mathrm{ad}}}$ under the obvious map $S_G \to S_{G^{\mathrm{ad}}}$.*

3.4. ***Distinguished orbits in type D and E.*** For the convenience of the reader in understanding Lemma 3.5, we give in the Appendix a list of weighted Dynkin diagrams for all distinguished orbits in types D_n and E_n with $n \leq 7$.

4. Automorphic forms for unitary groups

We now turn to an application of the smoothness result found in Section 3. In this section, we define the space of ordinary automorphic forms and the Hecke algebra attached to it. We then state a freeness result, which we will prove in Section 5.4 (Theorem 5.13).

Let l be a prime. Suppose F^+ is a totally real number field with an imaginary quadratic extension F such that all primes v of F^+ above l split in F. Let S_l be the set of all primes of F^+ that lie above l. Let L be a finite extension of $\mathbb{Q}_l$ with ring of integers $\mathcal{O}$ and residue field $\mathbb{F}$. Let $\bar{L}$ be a choice of algebraic closure. We will assume that L is large enough that all embeddings $F \hookrightarrow \bar{L}$ lie inside L. Let $c \in \mathrm{Gal}(F/F^+)$ be the unique nontrivial element, given by complex conjugation. For $a \in F$, we will denote $c(a)$ by $\bar{a}$ when convenient.

4.1. ***Unitary groups.*** Consider a central simple algebra D/F of F-dimension n^2, and let S_D be a finite set of primes of F^+ that split in F. Suppose that

(1) D splits at all places w of F that do not lie above any place in S_D;

(2) there is an isomorphism $D^{\mathrm{op}} \cong D \otimes_{F,c} F$ of F-algebras;

(3) the intersection $S_D \cap S_l$ is empty;

(4) D_w is a division algebra at all places w of F above a place in S_D;

(5) either n is odd, or n is even and $\frac{1}{2}n[F^+ : \mathbb{Q}] + \#S_D \equiv 0 \pmod 2$.

Because of condition (1) (which ensures that all places in S_D split), together with (2) and (5), we can find an involution of the second kind on D by [CHT08, Section 3.3, p. 95]. That is, we may construct a map

$$^* : D \to D$$

such that

- * is an F^+ linear anti-automorphism of D,
- $(a^*)^* = a$ for all $a \in D$, and
- the involution * coincides with complex conjugation when restricted to F.

In addition, we assume that this involution of the second kind is positive. That is, for any $\gamma \in D \setminus \{0\}$,

$$\mathrm{tr}_{F:\mathbb{Q}}[\mathrm{tr}_{D/F}(\gamma\gamma^*)] > 0.$$

Such an involution gives rise to a positive Hermitian form $\langle , \rangle : D \times D \to D$ given by $\langle x, y\rangle = x^* y$.

Let $\mathcal{O}_D$ be an order in D such that $\mathcal{O}_D^* = \mathcal{O}_D$ and such that $\mathcal{O}_{D,v}$ is a maximal order of D_v for any split prime v of F^+, as in Section 3.3 of [CHT08]. Define the unitary group over $\mathcal{O}_{F^+}$ whose R-points (where R is an $\mathcal{O}_{F^+}$-algebra) are given by $G_D(R) = \{g \in (\mathcal{O}_D \otimes_{\mathcal{O}_{F^+}} R)^\times : g^* g = 1\}$. Then G_D is an algebraic group over $\mathcal{O}_{F^+}$. By the positivity condition, we have $G_{D,v} \cong U(n)$ at each infinite place v of F^+.

For each prime v of F^+ that splits in F, choose a prime $\tilde{v}$ of F lying above v. This choice allows us to give an isomorphism $i_{\tilde{v}} : G_D(F_v^+) \to (D \otimes_F F_{\tilde{v}})^\times$ which restricts to an isomorphism $G_D(\mathcal{O}_{F^+,v}) \cong (\mathcal{O}_{D,\tilde{v}})^\times$ as in Section 3.3 of [CHT08]. Note that when $v \notin S_D$ is split in F, then G_D is split, so that $G_D(F_v^+) \cong (D \otimes_{F^+} F_v^+)^\times \cong (D \otimes_F F_{\tilde{v}})^\times \cong \mathrm{GL}_n(F_{\tilde{v}})$. If T is a set of primes of F^+ that splits in F, set $\tilde{T} = \{\tilde{v} : v \in T\}$.

4.2. *Automorphic forms of G_D.* We define the automorphic forms for G_D as in [Gro99] and [CHT08]. To do this, we recall some facts from the representation theory of reductive groups.

Let G be a split reductive group defined over L, and let $T \subseteq B \subseteq G$ be a choice of split maximal torus and Borel subgroup of G. Recall that finite-dimensional simple modules of G are uniquely determined by their highest weight in the character group of the torus $X^\bullet(T) := \mathrm{Hom}(T, \mathbb{G}_m)$, and that such a representation exists if and only if this highest weight ν lies in a dominant Weyl chamber.

In the case of GL_n and the standard upper Borel subgroup and maximal torus (defined over L), the set of weights naturally corresponds to $\mathbb{Z}^n$, and the set of dominant weights is $\mathbb{Z}^n_+ := \{\nu = (\nu_1, \dots, \nu_n) \in \mathbb{Z}^n : \nu_i \geq \nu_{i+1} \forall i\}$. We set the L-vector space W_ν to be the irreducible representation of highest weight ν. We will need to choose a $\mathcal{O}$ lattice of W_ν, which we do in the style of [Ger19, Section 1.1] as follows. Note that GL_n, B, and T are defined over $\mathcal{O}$. For a dominant weight ν, set ξ_ν to be the induced representation $\mathrm{Ind}_B^{\mathrm{GL}_n}(w_0(\nu))_{/\mathcal{O}}$ of the algebraic group $\mathrm{GL}_{n/\mathcal{O}}$ defined as the functor whose R points are

$$\mathrm{Ind}_B^G(w_0(\nu)) := \{ f \in R[\mathrm{GL}_n] : f(bg) = w_0(\nu)(b).f(g)\, \forall g \in \mathrm{GL}(R), b \in B(R) \},$$

where w_0 the longest element of the Weyl group. By Proposition II.2.2 and Corollary II.5.6 of [Jan03], the representation ξ_ν is irreducible of highest weight ν. We denote by M_ν the representation given by the $\mathcal{O}$-points of ξ_ν, so that $M_\nu \otimes_{\mathcal{O}} L \cong \xi_\nu(L) \cong W_\nu$.

Remark. The presence of w_0 is due to our convention that chooses B as the Borel of *upper* triangular matrices, whereas Jantzen induces from the Borel of *lower* triangular matrices. These two choices of Borel subgroup are related by w_0.

The finite-dimensional algebraic representations in L vector spaces of $G_{D,F_l^+} \cong \prod_{w \in \tilde{S}_l} \mathrm{GL}_{n,F_w}$ are characterised by the sequence of dominant weights, one for each embedding corresponding to $w \in \tilde{S}_l$. We define the set as $W = (\mathbb{Z}_+^n)^{\mathrm{Hom}(F^+,L)}$. For each $\mu \in W$, we can now define the algebraic representation of $G_{D/\mathcal{O}_{F^+}}$ with highest weight μ by $M_\mu = \bigotimes_{\tau \in \mathrm{Hom}(F^+,L),\mathcal{O}} M_{\nu_\tau}$, and $W_\mu = M_\mu \otimes_{\mathcal{O}} L$.

For each $v \in S_D$, choose a finite-free $\mathcal{O}$-module representation $\rho_v : G_D(\mathcal{O}_{F^+,v}) \to \mathrm{GL}(M_v)$ with open kernel such that $M_v \otimes \bar{L}$ is irreducible. Set $M_{\{\rho_v\}} = \bigotimes_{v \in S_D} M_v$. We set $M_{\mu,\{\rho_v\}} = M_\mu \otimes M_{\{\rho_v\}}$.

Definition 4.1. Let $\lambda = (\mu, \{\rho_v\})$ be as above. We define the space of automorphic forms for G_D of weight λ with A-coefficients $S_\lambda(A)$, where A is an $\mathcal{O}$-module, as the space of functions

$$f : G_D(F^+)\backslash G_D(\mathbb{A}_{F^+}^\infty) \to M_\lambda \otimes_{\mathcal{O}} A$$

such that there is an open compact subgroup

$$U \subset G_D(\mathbb{A}_{F^+}^{\infty,S_l}) \times G_D(\mathcal{O}_{F^+,l})$$

with

$$u|_{S_l \cup S_D} f(gu) = f(g)$$

for all $g \in G_D(\mathbb{A}_{F^+}^\infty)$ and $u \in U$ where $u|_{S_l \cup S_D}$ denotes the action of u on M_λ factoring through $\prod_{v \in S_D \cup S_l} G_D(F_v^+)$.

Notice that $S_\lambda(A)$ is a smooth representation of $G_D(\mathbb{A}_{F^+}^\infty)$ under the action

$$(h.f)(g) = h|_{S_l \cup S_D} f(gh).$$

(Again, $h|_{S_l \cup S_D}$ acts through the representation of $G_D(F_l^+) \times \prod_{v \in S_D} G_D(F_v^+)$ on M_λ.) We denote by $S_\lambda(U, A) = S_\lambda(A)^U$ the invariants under this action.

4.3. ***Hecke operators.*** For much of what remains, the argument will be a slight adaptation of that in [Ger19], the important details of which can be found in Sections 2 and 4. Let T be a finite set of places of F^+ containing $S_D \cup S_l$ such that all places in T split in F, and let $\tilde{T}$ be a set of primes of F above those in T as defined before. Fix an open compact subgroup $U = \prod_v U_v$ of $G_D(\mathbb{A}_{F^+}^\infty)$ such that U_v is hyperspecial at all places v outside T. Suppose further that U is sufficiently small; that is, there is a place

v such that U_v contains no elements of finite order other than the identity. We define the Hecke operators on the subspace $S_\lambda(U, A)$.

Hecke operators at unramified places. Let v be a place of F^+ split in F and $\tilde{v}$ be a place in F over v. Let $\varpi_{\tilde{v}}$ be a uniformiser. We can define the Hecke operators as the double coset operators

$$T_v^{(i)} = \left[i_v^{-1} \left(\mathrm{GL}_n(\mathcal{O}_{F,\tilde{v}}) \begin{pmatrix} \varpi_{\tilde{v}} I_i & 0 \\ 0 & I_{n-i} \end{pmatrix} \mathrm{GL}_n(\mathcal{O}_{F,\tilde{v}}) \right) \times U^v \right]$$

Hecke operators at places dividing l. At places dividing the residual characteristic of $\mathcal{O}$, we set

$$\alpha_{\tilde{v}}^{(i)} = \begin{pmatrix} \varpi_{\tilde{v}} I_i & 0 \\ 0 & I_{n-i} \end{pmatrix} \quad \text{and} \quad U_{\mu,\tilde{v}}^{(i)} = (w_0\mu_v)(\alpha_{\tilde{v}}^{(i)})^{-1}[U\alpha_{\tilde{v}}^{(i)}U],$$

where w_0 is the longest element of the Weyl group of GL_n and $\mu = (\mu_v) \in W$ with μ_v the dominant weight for the corresponding embedding $F^+ \hookrightarrow L$.

We make the following adjustment to the group U.

Definition 4.2. For v a place of F^+ above l and b a positive integer, let $I^b(\tilde{v})$ be the set of matrices in $\mathrm{GL}_n(\mathcal{O}_{F_{\tilde{v}}})$ which are upper-triangular and unipotent mod $\tilde{v}^b$. Define $U(l^b) = \prod_{v \in S_l} i_v^{-1}(I^b(\tilde{v})) \times U^l \subseteq G_D(\mathbb{A}_{F^+}^\infty)$ where U^l denotes the product $\prod_{v \nmid l} U_v$.

In the case with the group $U(l^b)$, we further define the following diamond operators:

Definition 4.3. Let T_n be the maximal torus inside GL_n as before. For $v \in S_l$, and $u \in T_n(\mathcal{O}_{F_{\tilde{v}}})$, define $\langle u \rangle$ as the operator

$$[U(l^b)uU(l^b)]$$

on $S_\lambda(U(l^b), A)$. For $u \in T_n(\mathcal{O}_{F^+,l}) = \prod_{v \in S_l} T_n(\mathcal{O}_{F_v}) \cong \prod_{v \in S_l} T_n(\mathcal{O}_{F_{\tilde{v}}})$, define $\langle u \rangle = \prod_{v \in S_l} \langle u_{\tilde{v}} \rangle$.

Let A be an $\mathcal{O}$-algebra and M an A-module. Define the Hecke algebra $\mathbb{T}^T = \mathbb{T}^T(U(l^b), M)$ as the A-subalgebra of $\mathrm{End}(S_\lambda(U(l^b), M))$ generated by all the operators

$$\left\{ \left(T_{\tilde{v}}^{(i)}, (T_{\tilde{v}}^{(n)})^{-1}\right) : v \notin T \text{ split in } F \right\} \cup \left\{ U_{\mu,\tilde{v}}^{(i)} : v \in S_l \right\} \cup \left\{ \langle u \rangle : u \in T_n(\mathcal{O}_{F^+,l}) \right\}.$$

Notice that the map $u \mapsto \langle u \rangle$ defines a group homomorphism

$$T_n(\mathcal{O}_{F^+,l}) \to \mathbb{T}^T(U(l^b), M)^\times$$

which factors through $T_n(\mathcal{O}_{F^+,l}/l^b) = \prod_{v \in S_l} T_n(\mathcal{O}_{F^+,v}/v^b)$.

4.4. ***Big ordinary Hecke algebras and the action of Λ.*** From this point on, we wish to focus on the cases where $A \in \mathcal{M}od_{\mathcal{O}}$ is one of $\mathcal{O}$, $L/\mathcal{O}$ or a finite module $\mathcal{O}/\pi^n\mathcal{O}$.

Recall from Hida theory, as fully explained in Section 2.4 of [Ger19], that for any place $v \in S_l$ and any i, the operator $e_v^{(i)} := \lim_{n\to\infty}(U_{\mu,\tilde{v}}^{(i)})^{n!}$ is a projection on $S_\lambda(U, A)$. We can further define the projection $e = \prod_{v,i} e_v^{(i)}$. We define the ordinary submodule $S_\lambda^{\mathrm{ord}}(U, A) := e.S_\lambda(U, A)$ as the image of this projection. Since all the Hecke operators commute, this is a Hecke invariant submodule. We also define $\mathbb{T}^{T,\mathrm{ord}}(U(l^b), A) = e\mathbb{T}^T(U(l^b), A)$.

Definition 4.4. Let T_n be the maximal torus of GL_n as before. For $b \geq 1$, let $T_n(l^b)$ be the kernel of $T_n(\mathcal{O}_{F^+,l}) \to T_n(\mathcal{O}/l^b)$. We define Λ as the algebra

$$\Lambda = \mathcal{O}[\![T_n(l)]\!] = \varprojlim_{b' \geq 1} \mathcal{O}[T_n(l)/T_n(l^{b'})].$$

We denote by a_N the kernel of the map $\Lambda \to \mathcal{O}[T_n(l)/T_n(l^N)]$. Since U is sufficiently small, we see $S_\lambda^{\mathrm{ord}}(U(l^{b,c}), A)$ is a free Λ/a_b-module, through the action of $T_n(\mathcal{O}_{F^+,l})$, and hence we have an inclusion $\Lambda/a_b \hookrightarrow \mathbb{T}(U(l^b), L/\mathcal{O})$ by Proposition 2.20 of [Ger19].

4.4.1. *Infinite level.* We need to consider the big ordinary Hecke algebra. Set

$$\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), A) = \varprojlim_{b>0} \mathbb{T}^{T,\mathrm{ord}}(U(l^b), A) \quad \text{and} \quad S^{\mathrm{ord}}(U(l^\infty), A) = \varinjlim_{b>0} S^{\mathrm{ord}}(U(l^b), A).$$

Because of the inclusions $\Lambda/a_b \hookrightarrow \mathbb{T}^{T,\mathrm{ord}}(U(l^{b,c}), L/\mathcal{O})$, we get an inclusion $\Lambda \hookrightarrow \mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), L/\mathcal{O})$, and we see that $S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})$ is a discrete Λ-module, so its Pontryagin dual is a compact Λ-module. (and in fact is finite free, by Proposition 2.20 of [Ger19] since we assume $U(l)$ is sufficiently small.)

We can now give a statement of a theorem that we prove by the application of Theorem 3.2. Under certain hypotheses, to be determined in Section 5, we have:

Theorem 5.13. *The $\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), L/\mathcal{O})$-module $S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})^\vee$ is locally free over the generic fibre $\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), L/\mathcal{O})[1/l]$.*

Therefore the multiplicity of $S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})^\vee$ is the same at every point of $\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), L/\mathcal{O})$ of characteristic zero, and thus we expect the multiplicity of nonclassical points (those corresponding to Hida families of ordinary automorphic forms) is the same as at classical automorphic forms in $S_\lambda(U, A)$.

5. Galois representations and deformation rings

5.1. ***Local deformation rings.*** In this section, we let G_{F^+} and G_F be the absolute Galois groups of F^+ and F, and $G_{F^+,v}$, $G_{F,w}$ be the decomposition groups at the places v, w of F^+ and F.

We now define a deformation problem. Let $v \in S_D$ with residue field of size q_v, and let $X_{\mathrm{St}} \subseteq S_{\mathrm{GL}_n}$ be the irreducible component corresponding to the regular nilpotent orbit. We say that an n-dimensional representation $\rho : G_{F^+,v} \to \mathrm{GL}_n(A)$ is Steinberg if the representation ρ lies in the A-points of this irreducible component X_{St}. When A is a field of characteristic zero and $WD(\rho) = (r, N)$ is the Weil–Deligne representation obtained from ρ, then this condition is equivalent to the condition r being unramified and the eigenvalues of $r(\mathrm{Frob}_{q_v})$ are in the ratio $q_v^{n-1} : q_v^{n-2} : \cdots : q_v : 1$.

Let $C_\mathcal{O}$ be the category of local Artinian $\mathcal{O}$-algebras with residue field $\mathbb{F}$, (that is, the category of coefficient rings as defined in Mazur's article in [CSS97]). For each $v \in S_D$ and Steinberg representation $\bar{\rho}_v : G_{F,\tilde{v}} \to \mathrm{GL}_n(\mathbb{F})$ define a functor

$$D_{\bar{\rho}_v}^{n,\square} : \mathcal{C}_\mathcal{O} \to \mathfrak{Set}, \quad A \mapsto \{\text{Steinberg liftings of } \bar{\rho}_v \text{ to } A\}.$$

This functor is pro-representable by the complete Noetherian local ring $R_v^{\square,\mathrm{st}} := \mathcal{O}^\wedge_{X_{\mathrm{St}},\bar{\rho}}$. When we view X_{St} as a scheme over L, Theorem 3.2 tells us, since q is not a root of unity in L, and is therefore considerate towards G_L, that any localisation of $R_v^{\square,\mathrm{st}}[1/l]$ is a regular ring and thus that $R_v^{\square,\mathrm{st}}[1/l]$ is regular.

We recall the definition of $\tilde{r}$-discrete series representations found in Section 2.4.5 in [CHT08].

Definition 5.1. Let $\tilde{r}_v : G_{F,\tilde{v}} \to \mathrm{GL}_d(\mathcal{O})$ be a representation with the following properties:

(1) $\tilde{r}_v \otimes \mathbb{F}$ is absolutely irreducible ($\mathbb{F}$ the residue field of $\mathcal{O}$).

(2) Every irreducible subquotient of $(\tilde{r}_v \otimes \mathbb{F})|_{I_{\tilde{v}}}$ is absolutely irreducible.

(3) $\tilde{r} \otimes \mathbb{F} \ncong \tilde{r} \otimes \mathbb{F}(i)$ for each $i = 0, \dots, m$, where $\_(i)$ denotes the twist by the unramified character sending Frob to q^i.

Whenever R is an $\mathcal{O}$ algebra, we say that $\rho : G_{F,\tilde{v}} \to \mathrm{GL}_{\mathrm{md}}(R)$ is an *$\tilde{r}$-discrete series representation* if there is a decreasing filtration $\{\mathrm{Fil}^i\}$ of ρ by R-direct summands such that

$$\mathrm{gr}^i \rho \cong \mathrm{gr}^0 \rho(i) \quad \text{for } i = 0, \dots, m-1 \qquad \text{and} \qquad \mathrm{gr}^0 \rho|_{I,\tilde{v}} \cong \tilde{r}|_{I,\tilde{v}} \otimes_{\mathcal{O}} R.$$

Proposition 5.2. *Suppose $l > h_G$. Let $\tilde{r}$ be a rank-d representation as above, and let n be an integer with $d|n$. Let $X_{\tilde{r},n}$ be the moduli space of framed $\tilde{r}$-discrete series representations of rank n, defined over $\mathcal{O}$. Then the base change $(X_{\tilde{r},n})_L$ is smooth over L.*

Proof. Let $S_{\tilde{r}}$ be the moduli stack over $\mathcal{O}$ of n-dimensional $\tilde{r}$-discrete representations, so that $S_{\tilde{r}} \cong [X_{\tilde{r}}/\mathrm{GL}_n]$, and let $S_{\mathbb{1}}$ be the stack of $m := n/d$-dimensional $\mathbb{1}$-discrete series representations. Let $S_{\tilde{r}}^{\mathrm{WD}}$ be the stack over L whose groupoid over R consists of objects (ρ', N), where ρ' is $\tilde{r}$-discrete series representation of rank $n = dm$ with open kernel, and N is an element of $\mathrm{End}_R(R^n)$ such that $\rho' N \rho'^{-1} = q^v N$. Define $S_{\mathbb{1}}^{\mathrm{WD}}$ analogously. Let t_l be the homomorphism $t_l : I \to \mathbb{Z}_l$ sending any lift of the topological generator of tame inertia to $1 \in \mathbb{Z}_l$. Recall that there is a morphism $S_{\tilde{r}}^{\mathrm{WD}} \to S_{\tilde{r}}$ given by (ρ', N) is sent to the unique representation ρ given by $\rho(g) = \rho'(g)\exp(t_l(g)N)$ for $g \in I$ and $\rho(\mathrm{Frob}) = \rho'(\mathrm{Frob})$. Recall that this is an isomorphism on the base change to L.

Then we have an morphism of algebraic stacks $S_{\mathbb{1}}^{\mathrm{WD}} \to S_{\tilde{r}}^{\mathrm{WD}}$ given by $(\rho', N) \mapsto (\rho', N) \otimes \tilde{r}$. We claim that this is an isomorphism. By an exercise in Clifford theory and by assumptions on $\tilde{r}$, the restriction $\tilde{r}|_I$ can be written as a direct sum of pairwise nonisomorphic absolutely irreducible I-representations $\tau \oplus \tau^{\mathrm{Frob}} \oplus \cdots \oplus \tau^{\mathrm{Frob}^{k-1}}$ for some $k \in \mathbb{N}$. As ρ' is $\mathbb{1}$-discrete series in characteristic zero, we see that $(\rho' \otimes \tilde{r})|_I \cong m(\tau \oplus \tau^{\mathrm{Frob}} \oplus \cdots \oplus \tau^{\mathrm{Frob}^{k-1}})$. Let $V_{\tilde{r}}(R) = \mathrm{End}_{R[I]}(\tilde{r}^m)$ be the space of I-equivariant maps of any representation in $S^{\mathrm{WD}_{\tilde{r}}}(R)$, and define $V_{\mathbb{1}}(R) = \mathrm{End}_{R[I]}(\mathbb{1}^m)$ similarly. The map

$$V_{\mathbb{1}}(R) \to V_{\tilde{r}}(R), \quad N \mapsto N \otimes \mathrm{id}_{\tilde{r}}, \tag{2}$$

is injective, and hence is isomorphic onto its image. We claim that if $(\rho, N) \in S_{\tilde{r}}^{\mathrm{WD}}(R)$, then N is in the image of this map.

First, note that N is I-equivariant. We calculate using Schur's lemma that $V_{\tilde{r}}(R) \cong M_m(R)^k$, since each τ^{Frob^i} is absolutely irreducible, and we see that the above map corresponds to the diagonal map $\Delta : M_m(R) \to M_m(R)^k$.

The space $V_{\tilde{r}}(R)$ has a natural action of Frobenius on it, and under this action $N = (N_1, \dots, N_k) \in M_m(R)^k$ has $\mathrm{Frob}.(N_1, \dots, N_k) = q(N_1, \dots, N_k)$. Notice that Frob induces an isomorphism of the underlying spaces $\tau^m \to (\tau^{\mathrm{Frob}})^m$, which gives us a commutative diagram

$$\begin{array}{ccc} \tau^m & \xrightarrow{\mathrm{Frob}} & (\tau^{\mathrm{Frob}})^m \\ \big\downarrow N_1 & & \big\downarrow qN_2 \\ \tau^m & \xrightarrow{\mathrm{Frob}} & (\tau^{\mathrm{Frob}})^m \end{array} \tag{3}$$

Hence, we see $(qN_2, \dots, qN_k, qN_1) = q(N_1, \dots, N_{k-1}, N_k)$, and thus N lies in the image of the diagonal map. This proves the claim.

Let $\chi_{\tilde{r}} = \mathrm{Hom}_I(\tau, \tilde{r})$. This is an unramified character. We claim that $(\mathrm{Hom}_I(\tau, \_) \otimes \chi_{\tilde{r}}^{-1}, \Delta^{-1}) : S_{\tilde{r}}^{\mathrm{WD}} \to S_{\mathbb{1}}^{\mathrm{WD}}$ is an inverse defining the equivalence.

We first show that the composition $S_{\tilde{r}}^{\mathrm{WD}}(R) \to S_{\mathbb{1}}^{\mathrm{WD}}(R) \to S_{\tilde{r}}^{\mathrm{WD}}(R)$ is the identity. For $(\Theta, N) \in S_{\tilde{r}}^{\mathrm{WD}}(R)$, the previous claim gives us an isomorphism on the N-part of the stacks $S_{\tilde{r}}^{\mathrm{WD}}(R)$, so we focus on the representation part. Since I acts through a finite quotient, and R is an algebra over a field of characteristic zero, we see that Θ is semisimple and hence we get a sequence of I-representation isomorphisms:

$$\begin{aligned} \Theta &\cong \bigoplus_{i=0}^{k-1} \mathrm{Hom}_I(\tau^{\mathrm{Frob}^i}, \Theta) \otimes \chi_{\tilde{r}}^{-1} \otimes \tau^{\mathrm{Frob}^i} \\ &\cong \mathrm{Hom}_I(\tau, \Theta) \otimes \chi_{\tilde{r}}^{-1} \otimes \bigoplus_{i=0}^{k-1} \tau^{\mathrm{Frob}^i} \cong \mathrm{Hom}_I(\tau, \Theta) \otimes \chi_{\tilde{r}}^{-1} \otimes \tilde{r}. \end{aligned}$$

To show this isomorphism also respects the W_F-action, we observe that each graded part of Θ has $\mathrm{gr}^i(\Theta) \cong \tilde{r} \otimes \chi(i)$, where χ is some unramified character. Then we obtain

$$\mathrm{Hom}_{L[I]}(\tau, \mathrm{gr}^i(\Theta)) \cong \mathrm{Hom}_{L[I]}(\tau, \tilde{r} \otimes \chi(i)) \cong \tilde{r} \otimes \chi(i),$$

so that both sides of the isomorphism are naturally $\tilde{r} \otimes \chi(i)$ as W_F-representations. Hence, the composition $S_{\tilde{r}}^{\mathrm{WD}}(R) \to S_{\mathbb{1}}^{\mathrm{WD}}(R) \to S_{\tilde{r}}^{\mathrm{WD}}(R)$ is the identity.

To show that $S_{\mathbb{1}}^{\mathrm{WD}}(R) \to S_{\tilde{r}}^{\mathrm{WD}}(R) \to S_{\mathbb{1}}^{\mathrm{WD}}(R)$ is the identity, let $\rho \in S_{\mathbb{1}}(R)$. Then the natural map

$$\rho \to \mathrm{Hom}_I(\tau, \rho \otimes \tilde{r}), \quad v \mapsto \{w \mapsto v \otimes w\}, \tag{4}$$

defines an I isomorphism. So we need only check that $\rho \otimes \chi_{\tilde{r}}$ and $\mathrm{Hom}_I(\tau, \rho \otimes \tilde{r})$ have the same action of Frobenius. This can be checked again, by looking at the character $\mathrm{gr}^i(\rho)$. Hence, we have exhibited an equivalence of categories $S_{\mathbb{1}} \leftrightarrow S_{\tilde{r}}$.

Given a choice of Frobenius Frob and a topological generator s of the tame inertia group we can explicitly write an isomorphism of stacks

$$\begin{aligned} S_{\mathbb{1}} &\cong [X_{\mathrm{St}}/\mathrm{GL}_m], \\ \rho &\mapsto \big(\rho(\mathrm{Frob}), \log(\rho(s))\big), \\ \rho_\Phi(\mathrm{Frob}^n x) = \Phi^n \exp(N t_l(x)) &\leftarrow\!\!\mid (\Phi, N). \end{aligned}$$

As $(X_{\mathrm{St}})_L$ is a smooth scheme by Theorem 3.2, this shows that $S_{\mathbb{1}}[1/l]$ is a smooth stack, and thus that $S_{\tilde{r}}[1/l]$ and $(X_{\tilde{r},n})_L$ are smooth. □

In light of this proposition, if $\bar{\rho} : G_{F,\tilde{v}} \to \mathrm{GL}_n(\mathbb{F})$ is an $\tilde{r}$-discrete series representation, we let $R_v^{\square,\tilde{r}}$ be the universal lifting ring of $\tilde{r}$-discrete series representations. By the proposition, the ring $R_v^{\square,\tilde{r}}[1/l]$ is regular at every maximal ideal.

5.1.1. *Deformation rings at primes above l.* For $v \in S_l$, let $\bar{I}_{\tilde{v}}$ be the inertia subgroup of $G_{F,\tilde{v}}^{\mathrm{ab}}$, let $\bar{I}_{\tilde{v}}(l)$ be the pro-l part, and let $\Lambda_{\tilde{v}} := \mathcal{O}[[\bar{I}_{\tilde{v}}(l)^n]]$, which we can identify with the universal lifting algebra of an ordered set of inertial characters $\{\bar{\chi}_i : I_{\tilde{v}} \to \mathbb{F}^\times\}_{i=1,\dots,n}$. Following chapter 3 of [Ger19] we can define a lifting $\Lambda_{\tilde{v}}$-algebra $R_v^{\triangle}$ as follows.

Take the universal lifting ring $R_v^{\square,\Lambda}$, so that a morphism $r : R_v^{\square,\Lambda} \to A$ corresponds to a pair $(\rho, \{\chi_i\}_{i=1,\dots,n})$ consisting of a representation $\rho : G_v \to \mathrm{GL}_n(A)$ lifting $\bar{\rho}$ and a sequence of characters $\chi_i : I_{\tilde{v}} \to A^\times$. Let $\mathcal{Flag}$ be the flag variety defined over $\mathcal{O}$. There is a subscheme $\mathcal{G}$ of $\mathcal{Flag} \times_{\mathcal{O}} \operatorname{Spec} R_v^{\square,\Lambda}$ whose A-points are the triples $(\mathrm{Fil}, \rho, \{\chi_i\}) \in (\mathcal{Flag} \times_{\mathcal{O}} \operatorname{Spec} R^{\square})(A)$ such that $\rho : G_v \to \mathrm{GL}_n(A)$ preserves the filtration Fil on A^n, and such that the action of $I_{\tilde{v}}$ on the graded part $\mathrm{Fil}_j/\mathrm{Fil}_{j-1}$ is χ_j. Then $R_v^{\triangle}$ is defined as the image of the natural morphism $R_v^{\square,\Lambda} \to \Gamma(\mathcal{G}, \mathcal{O}_{\mathcal{G}})$.

By Lemma 3.3 of [Ger19], the morphism $R_v^{\square} \to \mathcal{O}$ corresponding to a representation $\rho : G_v \to \mathrm{GL}_n(\mathcal{O})$ factors through $R_v^{\triangle}$ if and only if ρ is $\mathrm{GL}_n(\mathcal{O})$-conjugate to an upper triangular representation with diagonal characters equal to $\chi_1, \dots, \chi_n$ when restricted to inertia.

Definition 5.3. If A is a $\mathbb{Z}_l$-algebra and $v \in S_l$, we call a representation $\rho : G_v \to \mathrm{GL}_n(A)$ *ordinary* if it is $\mathrm{GL}_n(A)$-conjugate to an upper triangular matrix. Likewise, if $\rho : \mathrm{Gal}(\bar{F} : F) \to \mathrm{GL}_n(A)$ is a global Galois representation, we say ρ is ordinary if $\rho|_{G_v}$ is ordinary at all places $v \in S_l$.

In this terminology, the fact above can be restated by saying that a point $x : R_v^{\square} \to \mathcal{O}$ factors though $R_v^{\triangle}$ if and only if the corresponding representation ρ_x is ordinary.

Lemma 5.4. *Suppose that $\bar{\rho}_v : G_{F,\tilde{v}} \to \mathrm{GL}_n(\mathbb{F})$ is an ordinary Galois representation with diagonal characters $\bar{\chi}_1, \bar{\chi}_2, \dots, \bar{\chi}_n$ such that no pair $i < j$ has $\chi_i = \varepsilon\chi_j$, where ε is the cyclotomic character. Then $R_v^{\triangle}[1/l]$ is formally smooth of dimension $[F_{\tilde{v}} : \mathbb{Q}_l]\frac{1}{2}n(n+1) + n^2$ over L.*

Proof. This follows from Lemmas 3.17 and 3.7 of [Ger19]. (To apply Lemma 3.17 as stated there, one must note that $\mathcal{G}^{ar}$ is a union of irreducible components of $\mathcal{G}$ and $\bar{\rho}_v$ lies in the open subset of $\mathcal{G}[1/l]$ whose closure is defined to be $\mathcal{G}^{ar}$. Thus $R_v^{\triangle,ar} = R_v^{\triangle}$.) □

5.2. ***Local-global compatibility.*** We start by introducing the group $\mathcal{G}_n$ from [CHT08], defined as the group scheme that is the semidirect product of $\mathrm{GL}_n \times \mathrm{GL}_1$ with $C_2 = \{1, j\}$, where j acts as

$$j(g, \mu)j^{-1} = (\mu(g^{-1})^{\mathrm{T}}, \mu).$$

By Lemma 2.1.1 of [CHT08], representations $r : G_{F^+} \to \mathcal{G}_n(R)$ such that $r^{-1}(\mathrm{GL}_n(R) \times \mathrm{GL}_1(R)) = G_F$ are in correspondence with pairs (ρ, χ), where ρ is an n-dimensional representation of G_F and χ is a character of G_{F^+} such that $\rho^c \cong \chi\rho^\vee$ and $c \in G_{F^+}$ is sent to j.

For brevity, whenever we have a homomorphism $r : G_{F^+} \to \mathcal{G}_n(R)$ and a subgroup $H \subset G_{F^+}$, we use $r|_H$ to mean restriction to H, followed by projection to GL_n. Typically, H will be the subgroup G_F or its localisations $G_{F,w}$.

Proposition 5.5. *Suppose that* $\mathfrak{m} \trianglelefteq \mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), \mathcal{O})$ *is a maximal ideal with residue field* $\mathbb{F}$. *Then there is a unique continuous semisimple representation*

$$\bar{r}_\mathfrak{m} : G_F \to \mathrm{GL}_n(\mathbb{F})$$

satisfying the following conditions:

(1) $\bar{r}_\mathfrak{m}^c \cong \bar{r}_\mathfrak{m}^\vee \otimes \varepsilon^{1-n}$.

(2) $\bar{r}_\mathfrak{m}|_w$ *is unramified at all places* v *of* F^+ *outside* T.

(3) *If* v *additionally splits as* $v = ww^c$ *in* F, *then the characteristic polynomial of* $\bar{r}_\mathfrak{m}(\mathrm{Frob}_w)$ *is*

$$X^n - T_w^{(1)} X^{n-1} + \cdots + (-1)^j N(w)^{\frac{1}{2}j(j-1)} T_w^{(j)} X^{n-j} + \cdots + (-1)^n N(w)^{\frac{1}{2}n(n-1)} T_w^{(n)}$$

modulo $\mathfrak{m}$.

(4) *Let* $\tilde{r}_{\tilde{v}} : G_F \to \mathrm{GL}_{m_v}(\mathcal{O})$ *be constructed from the smooth representation* $\rho_v : G_D(F_v^+) \to \mathrm{GL}(M_v)$ *via the Jacquet–Langlands and local Langlands correspondences*, *as in* [CHT08, Section 3.3, p. 97]. *If* $v \in S_D$ *and* $U_v = G_D(\mathcal{O}_{F^+,v})$, *then* $\bar{r}_\mathfrak{m}|_{G_{F,v}}$ *is* $\tilde{r}_{\tilde{v}}$*-discrete series.*

Proof. We prove only (4), the other statements amounting to Proposition 2.28 in [Ger19]. By the argument in the proof of that same proposition, the maximal ideals of $\mathbb{T}$ are in bijection with those of $\mathbb{T}/m_\Lambda$. Hence, (4) follows from the classical situation (that is, usual automorphic forms for G_D rather than Hida families of ordinary automorphic forms). The proof of this can be found in Proposition 3.4.2 of [CHT08]. □

Proposition 5.6. *If* $\mathfrak{m}$ *is non-Eisenstein* (*that is, if* $\bar{r}_\mathfrak{m}$ *is irreducible*), *then* $\bar{r}_\mathfrak{m}$ *can be extended to a representation* $\bar{r}_\mathfrak{m} : G_{F^+} \to \mathcal{G}_n(\mathbb{F})$, *which in turn can be lifted to a representation*

$$r_\mathfrak{m} : G_{F^+} \to \mathcal{G}_n\big(\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), \mathcal{O})_\mathfrak{m}\big)$$

with the following properties:

(1) *If* $\nu : \mathcal{G}_n \to \mathrm{GL}_1$ *is the second projection, then* $\nu \circ r_\mathfrak{m} = \varepsilon^{1-n} \delta_{F/F^+}^{\mu_\mathfrak{m}}$, *where* ε *is the cyclotomic character,* δ_{F/F^+} *is the nontrivial character of* G_{F^+}/G_F, *and* $\mu_m \in \mathbb{Z}/2$.

(2) $\bar{r}_\mathfrak{m}|_{\tilde{v}}$ *is unramified at all places* $v \notin T$.

(3) *If* v *in addition splits as* $v = ww^c$ *in* F, *then the characteristic polynomial of* $\bar{r}_\mathfrak{m}(\mathrm{Frob}_w)$ *is*

$$X^n - T_w^{(1)} X^{n-1} + \cdots + (-1)^j N(w)^{\frac{1}{2}j(j-1)} T_w^{(j)} X^{n-j} + \cdots + (-1)^n N(w)^{\frac{1}{2}n(n-1)} T_w^{(n)}.$$

(4) *If* $v \in S_D$, *then* $r_\mathfrak{m}|_{G_{F,\tilde{v}}}$ *is* $\tilde{r}_{\tilde{v}}$*-discrete series.*

Proof. As with the previous proposition, this is Proposition 2.29 in [Ger19] along with the additional (4), which we prove. By the proof of Proposition 2.29 of [Ger19], we may find a sequence of maximal ideals $\mathfrak{m}_b \subset \mathbb{T}^{T,\mathrm{ord}}(U(l^b), \mathcal{O})$ such that $\mathbb{T}_{\mathfrak{m}} = \varprojlim_b \mathbb{T}^{T,\mathrm{ord}}(U(l^b), \mathcal{O})_{\mathfrak{m}_b}$, and we define $r_{\mathfrak{m}} = \varprojlim_b r_{\mathfrak{m}_b}$. By Lemma 3.4.4 of [CHT08], each $r_{\mathfrak{m}_b}|_{G_{F,\tilde{v}}}$ is $\tilde{r}_{\tilde{v}}$-discrete series, and so now it remains to show that $r_m|_{G_{F,\tilde{v}}}$ is, too. Since

$$r_{\mathfrak{m}_b} \otimes \mathbb{T}^{T,\mathrm{ord}}(U(l^{c,c}), \mathcal{O})_{\mathfrak{m}_c} = r_{\mathfrak{m}_c}$$

whenever $b > c$, the filtration Fil^i_b on $r_{\mathfrak{m}_b}$ descends to a filtration $\mathrm{Fil}^i_b \otimes \mathbb{T}^{T,\mathrm{ord}}(U(l^{c,c}), \mathcal{O})_{\mathfrak{m}_c}$ on $r_{\mathfrak{m}_c}$, and the graded parts have

$$[\mathrm{gr}^i(r_{\mathfrak{m},b})] \otimes \mathbb{T}^{T,\mathrm{ord}}(U(l^{c,c}), \mathcal{O})_{\mathfrak{m}_c} \cong \mathrm{gr}^i[r_{\mathfrak{m},b} \otimes \mathbb{T}^{T,\mathrm{ord}}(U(l^{c,c}), \mathcal{O})_{\mathfrak{m}_c}].$$

It follows that $\mathrm{Fil}^i_b \otimes \mathbb{T}^{T,\mathrm{ord}}(U(l^{c,c}), \mathcal{O})_{\mathfrak{m}_c}$ is a defining filtration on $r_{\mathfrak{m}_c}$. From Lemma 2.4.25 of [CHT08], such a filtration is unique, so we have a compatible system of filtrations on the $r_{\mathfrak{m}_b}$ which lift to a filtration on $r_{\mathfrak{m}}|_{G_{F,\tilde{v}}}$. We see from this compatibility that $\mathrm{gr}^i(r_{\mathfrak{m}}) = \varprojlim_b \mathrm{gr}^i(r_{\mathfrak{m}_b})$, and so $r_{\mathfrak{m}}|_{G_{F,\tilde{v}}}$ is $\tilde{r}_{\tilde{v}}$-discrete series. □

To complete the results we need for local-global compatibility, we need the following lemma:

Lemma 5.7. *Let $\tilde{v} \in \tilde{S}_l$, and let $R^{\triangle}_{\tilde{v}}$ be as before. Then there is a map $R^{\triangle}_{\tilde{v}} \to \mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), \mathcal{O})_{\mathfrak{m}}$ such that*

$$\begin{array}{ccc} G_{F,\tilde{v}} & \xrightarrow{\rho^{\triangle}} & \mathcal{G}_n(R^{\triangle}_{\tilde{v}}) \\ & \searrow^{r_{\mathfrak{m}}} & \downarrow \\ & & \mathcal{G}_n\big(\mathbb{T}^{T,\mathrm{ord}}(U(l^\infty))\big) \end{array}$$

commutes.

Proof. This follows directly from Corollary 4.3 of [Ger19]. □

5.3. ***Global deformation rings.*** Let F/F^+ and let $\bar{\rho} : G_F \to \mathrm{GL}_n(\mathbb{F})$ be a representation with local representations $\rho_w = \bar{\rho}|_{G_{F,w}}$, where w is a place of F. Let R be the set of places v of F^+ such that v splits and there is a place w of F above v where ρ ramifies. Set $T = S_l \sqcup S_D \sqcup R$, and define $\tilde{T}$ as before. We make the following assumptions:

- The representation $\bar{\rho}$ is an irreducible automorphic representation. That is, there is a non-Eisenstein maximal ideal $\mathfrak{m} \trianglelefteq \mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), \mathcal{O})$ such that $\bar{\rho} \cong \bar{r}_{\mathfrak{m}}$.
- The subgroup $\rho(G_{F^+(\zeta_l)}) \subseteq \mathcal{G}_n(\mathbb{F})$ is adequate in the sense of Definition 2.3 of [Tho12].
- The representation $\bar{\rho}$ is unramified outside $\tilde{T}$.
- At any place $v \in R$, any lift of $\bar{\rho}_v$ to $\overline{\mathbb{Q}}_l$ is nondegenerate in the sense of Section 3.3 of [Sho18]; in particular they lie on a single irreducible component of $\mathrm{Loc}^{\square}_{\mathrm{GL}_n,\mathbb{Q}_l}$.
- For each $v \in S_l$, we have $\mathrm{Hom}_{G_{F,\tilde{v}}}(\bar{\rho}_{\tilde{v}}, \bar{\rho}_{\tilde{v}}\varepsilon) = 0$, for ε the cyclotomic character.

As $\bar{\rho} \cong \bar{r}_{\mathfrak{m}}$ is irreducible, it can be extended to a representation $\bar{\rho} : G_{F^+} \to \mathcal{G}_n(\mathbb{F})$ such that $\nu \circ \bar{\rho} = \varepsilon^{1-n}\delta_{F/F^+}^{\mu_{\mathfrak{m}}}$ via Proposition 5.6. We fix such an extension.

For each $v \in T$, define $R_v^{\square}$ as the framed deformation ring for $\bar{\rho}_{\tilde{v}}$. Set

$$R^{\mathrm{loc}} := \Big(\widehat{\bigotimes_{v\in S_l}}{}_{\mathcal{O}} R_v^{\triangle}\Big) \hat{\otimes}_{\mathcal{O}} \Big(\widehat{\bigotimes_{v\in S_D}}{}_{\mathcal{O}} R_v^{\square,\tilde{r}_{\tilde{v}}}\Big) \hat{\otimes}_{\mathcal{O}} \Big(\widehat{\bigotimes_{v\in R}}{}_{\mathcal{O}} R_v^{\square}\Big)$$

as the local deformation ring for $\bar{\rho}$. Our first observation is that, since each $R_v^{\triangle}$ is a $\Lambda_{\tilde{v}}$-module, the algebra R^{loc} inherits the structure of a $\widehat{\bigotimes}_{v\in S_l} \Lambda_{\tilde{v}} \cong \Lambda$-module. The isomorphism $\widehat{\bigotimes}_{v\in S_l} \Lambda_{\tilde{v}} \cong \Lambda$ is inherited from the group isomorphisms

$$T_n(\mathrm{fl}) \cong \prod_{v\in S_l} T_n\mathcal{O}_{F^+,v}(l) \cong \prod_{v\in S_l} T_n\mathcal{O}_{F,\tilde{v}}(l) \cong \prod_{v\in S_l} \bar{I}_{\tilde{v}}(l)^n$$

where the final isomorphism is given by the Artin map of local class field theory.

Lemma 5.8. *The ring $R^{\mathrm{loc}}[1/l]$ is regular.*

Proof. Recall the construction of the complete ring $R_v^{\triangle}$ as the image of $R_v^{\square,\Lambda} = R_v^{\square} \hat{\otimes}_{\mathcal{O}} \Lambda_v$ in the global sections of $\mathcal{G} \subseteq \mathcal{F}lag \times_{\mathcal{O}} R_v^{\square,\Lambda}$. We "decomplete" $R_v^{\triangle}$ as follows: $R_v^{\square}$ is the completion of a local ring $\tilde{R}_v^{\square}$ at a closed point on a finite-type scheme over $\mathcal{O}$. The ring $\Lambda_v = \mathcal{O}[[\bar{I}_{\tilde{v}}(l)^n]]$ is a completed group algebra of a group which is topologically finitely generated, generated by a fixed choice of generators $\{s_i\}$. So we can choose a subring $\tilde{\Lambda}_v = \mathcal{O}[s_i^{\pm 1}]/\langle\text{relations}\rangle \subseteq \Lambda_v$ of finite type over $\mathcal{O}$ which is dense in Λ_v. Thus, there is a ring $\tilde{R}_v^{\square,\Lambda}$ of finite type over $\mathcal{O}$ whose completion is $R_v^{\square,\Lambda}$.

We can define a closed subscheme $\tilde{\mathcal{G}} \subseteq \mathcal{F}lag \times_{\mathcal{O}} \mathrm{Spec}(\tilde{R}_v^{\square,\Lambda})$ cut out by the same equations for $\mathcal{G}$ as in the definition of $R_v^{\triangle}$. Of course, there is a natural commutative diagram

$$\begin{array}{ccc} \tilde{R}_v^{\square,\Lambda} & \xrightarrow{\tilde{\phi}} & \mathcal{O}(\tilde{\mathcal{G}}) \\ \big\downarrow & & \big\downarrow \\ R_v^{\square,\Lambda} & \xrightarrow[\phi]{} & \mathcal{O}(\mathcal{G}) \end{array}$$

We set $\tilde{R}_v^{\triangle}$ as the image of $\tilde{\phi}$. It is a finite-type ring over $\mathcal{O}$ and, since the equations defining $\mathcal{G} \subseteq \mathcal{F}lag \times \mathrm{Spec}(R_v^{\square,\Lambda})$ are rational (that is, the defining ideal $\mathcal{I}$ equals $\tilde{\mathcal{I}}\mathcal{O}_{\mathcal{F}lag\times\mathrm{Spec}(R_v^{\square,\Lambda})}$ for some ideal $\tilde{\mathcal{I}} \subseteq \mathcal{F}lag \times \mathrm{Spec}(\tilde{R}_v^{\square,\Lambda})$), the image $\mathrm{im}(\phi) = R_v^{\triangle}$ is a completion of $\tilde{R}_v^{\triangle}$.

It follows that each of $R_v^{\square}$, $R_v^{\square,\tilde{r}_{\tilde{v}}}$ and $R_v^{\triangle}$ (for $v \in R$, S_D, S_l respectively) is a completion of a local ring at a closed point P_v on a finite-type $\mathcal{O}$-scheme X_v. By the last two hypotheses on $\bar{\rho}$ listed above, Lemma 5.4 and Proposition 3.6 of [Sho18], we see that $R_v^{\triangle}[1/l]$ (for $v \in S_l$) is regular and $R_v^{\square}[1/l]$ (for $v \in R$) is formally smooth. Thus, the closed points P_v on X_v (where $v \in S_l \cup R$) lie on an open subscheme $U_v \subseteq X_v$ whose generic fibre $U_v[1/l]$ is smooth over L. By Proposition 5.2, the same is true for $R_v^{\square,\tilde{r}_{\tilde{v}}}$ for $v \in S_D$. Set

$$\tilde{R}^{\mathrm{loc}} := \Big(\bigotimes_{\mathcal{O},v\in S_l} \tilde{R}_v^{\triangle}\Big) \otimes_{\mathcal{O}} \Big(\bigotimes_{\mathcal{O},v\in S_D} \tilde{R}_v^{\square,\tilde{r}_{\tilde{v}}}\Big) \otimes_{\mathcal{O}} \Big(\bigotimes_{\mathcal{O},v\in R} \tilde{R}_v^{\square}\Big).$$

Then $\tilde{R}^{\mathrm{loc}}$ is of finite type over $\mathcal{O}$ and has a maximal ideal m, corresponding to the closed point $(P_v)_v$, with respect to which R^{loc} is the m-adic completion. In addition, $\tilde{R}^{\mathrm{loc}}[1/l]$ is a regular L-algebra. To show that $R^{\mathrm{loc}}[1/l]$ is a regular ring is now a simple application of Lemma 2.9 and [Stacks, 07NY]. □

In fact, the same argument shows that $R_\infty[1/l]$ is a regular ring whenever R_∞ is a power series ring in a finite number of variables with coefficients in R^{loc}.

Let $\mathcal{S}$ be the tuple

$$\mathcal{S} = \big(F/F^+,\ T,\ \tilde{T},\ \varepsilon^{1-n}\delta_{F/F^+}^{\mu_m},\ \{R_v^{\triangle,ar} : v \in S_l\},\ \{R_v^{\square,st} : v \in S_D\},\ \{R_v^{\square} : v \in R\}\big)$$

and say that $\rho : G_{F^+} \to \mathcal{G}(A)$ is a lifting of $\bar{\rho}$ to $A \in \mathcal{C}_\Lambda$ of type $\mathcal{S}$ if it has the following properties:

(1) $\rho|_{G_F}$ lifts $\bar{r}_m$.

(2) ρ is unramified outside T.

(3) For $v \in S_D$, the local representation ρ_v is $\tilde{r}$-discrete series and gives rise to the morphism $R_v^{\square} \to A$ which factors through $R_v^{\square,\tilde{r}}$.

(4) For $v \in S_l$, the restriction ρ_v and the Λ-structure on A give a morphism $R_v^{\square} \otimes \Lambda \to A$ which factors through $R_v^{\triangle}$.

(5) $\nu \circ \rho = \varepsilon^{1-n}\delta_{F/F^+}^{\mu_m}$.

By Proposition 2.2.9 of [CHT08], we can construct the universal deformation ring $R_{\mathcal{S}}^{\mathrm{univ}}$ and the universal lifting ring $R_{\mathcal{S}}^{\square}$.

Let $h_0 = [F^+ : \mathbb{Q}] \cdot \frac{1}{2}n(n-1) + [F^+ : \mathbb{Q}] \cdot \frac{1}{2}n(1-(-1)^{\mu_{\mathrm{m}-1}})$ and let h be an integer larger than both h_0 and $\dim\big[H^1_{\mathcal{L}^\perp}(G_{F^+,T}, \operatorname{ad}\bar{\rho}(1))\big]$. (Here, the space $H^1_{\mathcal{L}^\perp}(G_{F^+,T}, \operatorname{ad}\bar{\rho}(1))$ is a particular subspace of the cohomology group $H^1(G_{F^+,T}, \operatorname{ad}\bar{\rho}(1))$ of the Galois group $G_{F^+,T}$ of the maximal extension of F^+ unramified outside of T, defined in Proposition 4.4 of [Tho12].)

After Thorne [Tho12], we will call a triple $(Q, \tilde{Q}, \{\bar{\psi}_v\}_{v\in Q})$ a Taylor–Wiles triple if

(1) Q is a set of primes of F^+ which split in F,

(2) $l | \mathrm{Nm}_{F^+}(v) - 1$ for each $v \in Q$,

(3) $|Q| = h$,

(4) $\tilde{Q}$ is the set $\{\tilde{v} | v \in Q\}$, and

(5) for each $v \in Q$, the representation $\bar{\rho}|_{G_v}$ splits as a direct sum into $\bar{s}_v \oplus \bar{\psi}_v$ where $\bar{\psi}_v$ is a generalised eigenspace with eigenvalue $\bar{\alpha}_v \in \mathbb{F}$ of dimension d_v.

For any Taylor–Wiles set Q we can define a deformation problem $\mathcal{S}(Q)$ that is the same as $\mathcal{S}$, but now we allow $\rho_{\tilde{v}}$, for $v \in Q$, to ramify in the following way: $\rho_{\tilde{v}}$ splits as a direct sum $s \oplus \psi$, and the two summands lift to $\bar{s}$ and $\bar{\psi}$ in such a way that s is unramified and $\psi|_{I_v} : I_v \to \mathrm{GL}_{d_v}$ factors through the scalar action on the underlying representation space. Using Proposition 2.2.9 in [CHT08] again, we can now take the universal deformation ring $R_{\mathcal{S}(Q)}^{\mathrm{univ}}$. Because stipulating that the local deformations at

Taylor–Wiles primes are unramified is a closed condition, this presents us with a surjection $R^{\mathrm{univ}}_{\mathcal{S}(Q)} \twoheadrightarrow R^{\mathrm{univ}}_{\mathcal{S}}$. We also have a natural map $R^{\mathrm{loc}} \to R^{\mathrm{univ}}_{\mathcal{S}(Q)}$ given by restrictions to the local subgroups at the level of functors.

Proposition 5.9. *For each $N \in \mathbb{N}$, we can find a Taylor–Wiles triple $(Q_N, \tilde{Q}_N, \{\bar{\psi}_v\}_{v\in Q})$ such that $l^N || \mathrm{Nm}_F(v) - 1$ for all $v \in Q_N$ and the global deformation ring $R^{\mathrm{univ}}_{\mathcal{S}(Q)}$ can be topologically generated over R^{loc} by $h - h_0$ generators.*

Proof. This follows from Lemma 4.4 of [Tho12] applied in the case of Theorem 8.6. □

In light of this proposition, set $R_\infty = R^{\mathrm{loc}}[[X_1, \dots, X_h]]$, set $R_N = R^{\mathrm{univ}}_{\mathcal{S}(Q_N)}$, and set $R_0 = R^{\mathrm{univ}}_{\mathcal{S}}$ so that we have surjections $R_\infty \twoheadrightarrow R_N$ and $R_N \twoheadrightarrow R_0$.

We now define some important subgroups of $G_D(\mathbb{A}^\infty_{F^+})$.

Definition 5.10. For $v \in Q_N$, suppose that $\bar{r}|_v = \bar{s} \oplus \bar{\psi}$ as before, with $\bar{\psi}$ a d_v-dimensional semisimple unramified representation with all Frobenius eigenvalues equal. We take the group $U_i(\tilde{v})$ to be the subgroup of $U_v \subseteq G_D(F_v^+)$ (identified with $\mathrm{GL}_n(F_{\tilde{v}})$ via the isomorphism $i_{\tilde{v}}$) of elements that take the form

$$\begin{pmatrix} \varpi_{\tilde{v}} * & * \\ 0 & aI_{d_v} \end{pmatrix}$$

modulo $\tilde{v}$ with $a \equiv 1 \mod \tilde{v}$ when $i = 1$, and arbitrary when $i = 0$. Set $U_i(Q) = U^Q \times \prod_{v\in Q} U_i(\tilde{v}) \subseteq G_D(\mathbb{A}^\infty_{F^+})$.

Let Δ_N be the maximal l-power quotient of $U_0(Q_N)/U_1(Q_N) \cong \prod_{v\in Q_N} k(\tilde{v})^\times$. We may view Δ_N as the maximal l-quotient of $\prod_{v\in Q_N} k(\tilde{v})^\times \cong (\mathbb{Z}/l^N)^q$. We claim there is an action of Δ_N on the ring $R^{\mathrm{univ}}_{\mathcal{S}(Q)}$. The map $\det \circ r^{\mathrm{univ}}_N : I_{F,\tilde{v}} \to (R^{\mathrm{univ}}_{\mathcal{S}(Q)})^\times$ given by the determinant of the universal deformation $r^{\mathrm{univ}}_N := r^{\mathrm{univ}}_{\mathcal{S}(Q_N),\bar{\rho}}$ factors through the kernel of $(R^{\mathrm{univ}}_{\mathcal{S}(Q)})^\times \to \mathbb{F}^\times$, which is an abelian l-power group. By local class field theory, there is an isomorphism $I_{F^{\mathrm{ab}},\tilde{v}} \to \mathcal{O}^\times_{F,\tilde{v}}$, and the l-power quotient of this group is the l-power quotient of $k(\tilde{v})^\times$. Hence we see that there is a map $\Delta_N \to (R^{\mathrm{univ}}_{\mathcal{S}(Q_N)})^\times$ and thus a ring map $\Lambda[\Delta_N] \to R^{\mathrm{univ}}_{\mathcal{S}(Q)}$, so that $R^{\mathrm{univ}}_{\mathcal{S}(Q_N)}$ inherits the structure of a finitely generated $\Lambda[\Delta_N]$-algebra. If $\mathfrak{a}_N$ is the augmentation ideal of $\Lambda[\Delta_N]$, then $R^{\mathrm{univ}}_{\mathcal{S}(Q_N)}/\mathfrak{a}_N$ is the ring of the universal deformation ring which parametrises Galois deformations of type $\mathcal{S}$. (These deformations are required to be unramified at places above Q_N.) Note that $\Delta_N \cong (\mathbb{Z}/l^n\mathbb{Z})^h$ by our choice of Q_N.

As in Sections 4.3 and 4.4, we can construct the Hecke operators

$$\mathbb{T}^{T\cup Q_N,\mathrm{ord}}\big(U_1(Q_N)(l^\infty), \mathcal{O}\big)$$

and, through a map $\mathbb{T}^{T\cup Q_N,\mathrm{ord}}\big(U_1(Q_N)(l^\infty), \mathcal{O}\big) \to \mathbb{T}^{T,\mathrm{ord}}(U(l^\infty), \mathcal{O})$, we can lift our choice of maximal ideal $\mathfrak{m}$ to a maximal ideal $\mathfrak{m}_N \subset \mathbb{T}^{T\cup Q_N,\mathrm{ord}}(U_1(Q_N)(l^\infty), \mathcal{O})$. Set $\mathbb{T}_{N,1} := \mathbb{T}^{T\cup Q_N,\mathrm{ord}}(U_1(Q_N)(l^\infty), \mathcal{O})^\wedge_{\mathfrak{m}_N}$ as the $\mathfrak{m}_N$-adic completion. As in Proposition 5.6, we can construct a representation $r_{\mathfrak{m}_N} : G_{F^+} \to \mathcal{G}_n(\mathbb{T}_{N,1})$ which by the proof of Theorem 6.8 of [Tho12] gives us an $\mathcal{S}(Q_N)$-lifting of $\bar{\rho}$. Hence, we get a surjection $R^{\mathrm{univ}}_{\mathcal{S}(Q)} \twoheadrightarrow \mathbb{T}_{N,1}$ for each N.

5.4. *Patching.* We now define a module H_N over $\mathbb{T}^{T\cup Q_N,\mathrm{ord}}(U_1(Q_N)(l^\infty),\mathcal{O})_m$ for each set Q_N, and quote a patching theorem that will allow us to construct the patched "limit" module H_∞, which we use to prove our local freeness result.

Define the space of automorphic forms $S^{\mathrm{ord}}(U_i(Q_N)(l^\infty), L/\mathcal{O})_m$ as before and set

$$H_0 = S^{\mathrm{ord}}(U(l^\infty), L/\mathcal{O})_m^\vee.$$

In Proposition 5.9 of [Tho12], Thorne describes a projection Pr_v on $S^{\mathrm{ord}}(U_i(Q_N)(l^\infty), L/\mathcal{O})_m$ and, later on, modules

$$H_{i,N} := \prod_{v\in Q_N} \mathrm{Pr}_v\big[S^{\mathrm{ord}}(U_i(Q_N)(l^\infty), L/\mathcal{O})_m\big]^\vee$$

with the following properties:

Proposition 5.11 [Tho12, part of proof of Theorem 6.8]. (1) *H_{1,Q_N} is a free $\Lambda[\Delta_{Q_N}]$-module and restriction to $S^{\mathrm{ord}}\big(U_0(Q_N)(l^\infty), L/\mathcal{O}\big)_m$ gives an isomorphism $H_{1,Q_N}/a_N \cong H_{0,Q_N}$.*

(2) *The map*

$$\Big(\prod_{v\in Q_N} \Pr_{\tilde{v}}\Big)^\vee : H_{0,Q_N} \to H_0$$

is an isomorphism.

Theorem 5.12 (patching). *Let $R \twoheadrightarrow \mathbb{T}$ be a surjective Λ-algebra homomorphism with $\mathbb{T}$ a finite Λ-algebra. Define $S_N = \Lambda[(\mathbb{Z}/l^n\mathbb{Z})^h] \cong \Lambda[\Delta_{Q_N}]$ with augmentation ideal $\mathfrak{a}_N$ and define the inverse limit $S'_\infty := \varprojlim \Lambda[\Delta_{Q_N}] \cong \Lambda[[Y_1,\dots,Y_h]]$. Set $S_\infty = S'_\infty \hat{\otimes}_{\mathcal{O}} \mathcal{T}$, where $\mathcal{T} = \mathcal{O}[[X_1,\dots,X_{|T|n^2}]]$. Suppose we have the following data*:

(1) *integers $t, h \geq 1$,*

(2) *a finite $\mathbb{T}$-module H,*

(3) *for each $N \geq 1$,*

  (a) *an S_N-algebra homomorphism $R_N \twoheadrightarrow \mathbb{T}_N$ that gets reduced to $R \twoheadrightarrow \mathbb{T}$ under reduction modulo $\mathfrak{a}_N$, and*

  (b) *a finite $\mathbb{T}_N$-module H_N, which is finite and free over S_N and whose S_N-rank is independent of N,*

  and

(4) *an S_∞-algebra R_∞ such that $R_\infty \twoheadrightarrow R_N$ with kernel* $\ker(S_\infty \to S_N)R_\infty$.

Then there is an $R_\infty \otimes S_\infty$-module H_∞, such that $H_\infty/aH_\infty \cong H$, H_∞ is a finite free S_∞-module, and the action of S_∞ on H_∞ factors through that of R_∞.

Proof. The details of the Taylor–Wiles–Kisin patching method used here are as in Chapter 4.3 of [Ger19]. They can also be found in Chapter 8 of [Tho12], under the heading "another patching argument". □

Theorem 5.13. *The module $H_0[1/l]$ is a finite locally free $R_S^{\mathrm{univ}}[1/l]$-module.*

Proof. We calculate that

$$\dim(S_\infty) = \dim(\Lambda) + h + |T|n^2 = n[F^+ : \mathbb{Q}]n + h + |T|n^2$$

and that

$$\begin{aligned}\dim(R_\infty) &= 1 + \sum_{v \in S_l} \left([F_{\tilde{v}} : \mathbb{Q}_l] \cdot \tfrac{1}{2}n(n+1) + n^2\right) + n^2|S_D \cup R| + h - h_0 \\ &= [F^+ : \mathbb{Q}] \cdot \tfrac{1}{2}n(n+1) + |T|n^2 + h - h_0 \\ &= [F^+ : \mathbb{Q}]n + |T|n^2 + h - [F^+ : \mathbb{Q}] \cdot \tfrac{1}{2}n(1 - (-1)^{\mu_{\mathrm{m}} - n})\end{aligned}$$

Consider the module $H_\infty^\square$. Since $H_\infty^\square$ is a finite free S_∞ module, and since the action of S_∞ factors through R_∞, we see that

$$\dim(S_\infty) = \mathrm{depth}_{S_\infty}(H_\infty^\square) \leq \mathrm{depth}_{R_\infty}(H_\infty^\square) \leq \dim(R_\infty)$$

and thus, the only possible way for this inequality to hold is if equality holds throughout. This implies $\mu_m \equiv n \mod 2$ and $H_\infty^\square$ is a maximal Cohen–Macaulay R_∞ module.

Now, consider the generic fibre. Let $m \subseteq R_\infty[1/l]$ be a maximal ideal. Lemma 5.8 shows that $R_\infty[1/l]_m$ is a regular local ring. Thus, any finitely generated maximal Cohen–Macaulay $R_\infty[1/l]_m$-module has finite projective dimension, and hence any maximal Cohen–Macaulay module is projective by the Auslander–Buchsbaum formula. This shows that $H_\infty^\square[1/l]_m$ is a free $R_\infty[1/l]_m$-module, this shows that $H_\infty^\square[1/l]$ is a locally finite free $R_\infty[1/l]$-module. It follows that $H_0[1/l]$ is a locally finite free $R_S^{\mathrm{univ}}[1/l]$-module. □

Corollary 5.14. $R_S^{\mathrm{univ}}[1/l] = \mathbb{T}[1/l].$

Proof. Let I be the kernel of the surjection $R_S^{\mathrm{univ}}[1/l] \to \mathbb{T}[1/l]$. Choose any maximal ideal m of $R_S^{\mathrm{univ}}[1/l]$. Since localisation is an exact functor, we get a short exact sequence

$$0 \to I_m \to R_S^{\mathrm{univ}}[1/l]_m \to \mathbb{T}[1/l]_m \to 0.$$

Note that the action of $R_S^{\mathrm{univ}}[1/l]_m$ on $H_0[1/l]_m$ factors through $\mathbb{T}[1/l]_m$, so that I_m annihilates all of $H_0[1/l]_m$. Since this is a free module, this shows that I_m is trivial. Since this is true for every m, this shows that $\mathrm{Supp}(I) = \varnothing$ and hence $I = 0$. Hence the surjection above is an isomorphism

$$R_S^{\mathrm{univ}}[1/l] \cong \mathbb{T}[1/l].$$ □

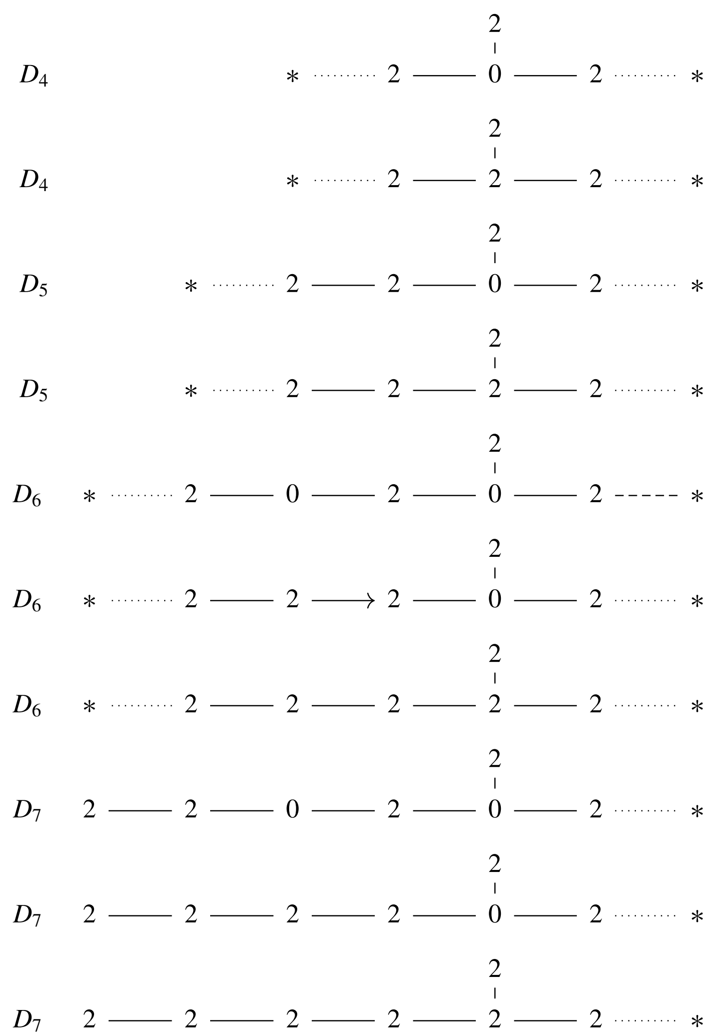
Remark. As an application of Theorem 5.13, whenever M is a locally free coherent sheaf on a connected space X, the rank function

$$X \to \mathbb{N} \cup \{0\}, \quad x \mapsto \mathrm{Rank}_x(M),$$

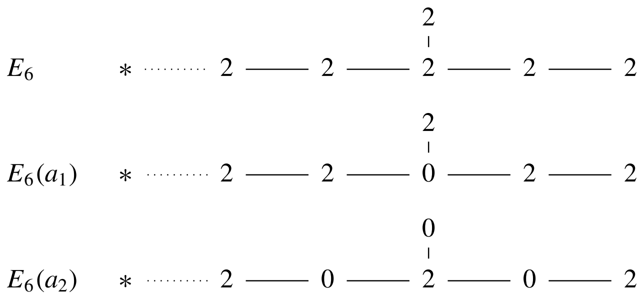
is locally constant. Therefore, the rank of a geometrically connected component can be calculated by calculating the rank at any special point $x \in X$. In our special case, the rank of the module $H_0[1/l]$ can be interpreted as the number of distinct automorphic forms with a given set of Hecke eigenvalues, which can be interpreted as the multiplicity of the Galois representation determined by said Hecke eigenvalues inside the space of automorphic forms. We have shown that for these automorphic forms, the multiplicity is determined only by the connected component of $R_\infty[1/l]$ on which the representation $\rho_{\mathfrak{m}}$ lies. By Lemma 4.2 of [Ger19], we see that the minimal primes of $R_\infty[1/l]$ biject with the minimal primes of Λ. Thus, if one could show that for each component of Spec Λ, there is an automorphic form of some classical weight had multiplicity 1, then all the Hida families of forms would also have multiplicity 1.

Appendix: Weighted Dynkin diagrams for distinguished orbits in types D_n and E_n with $n \leq 7$.

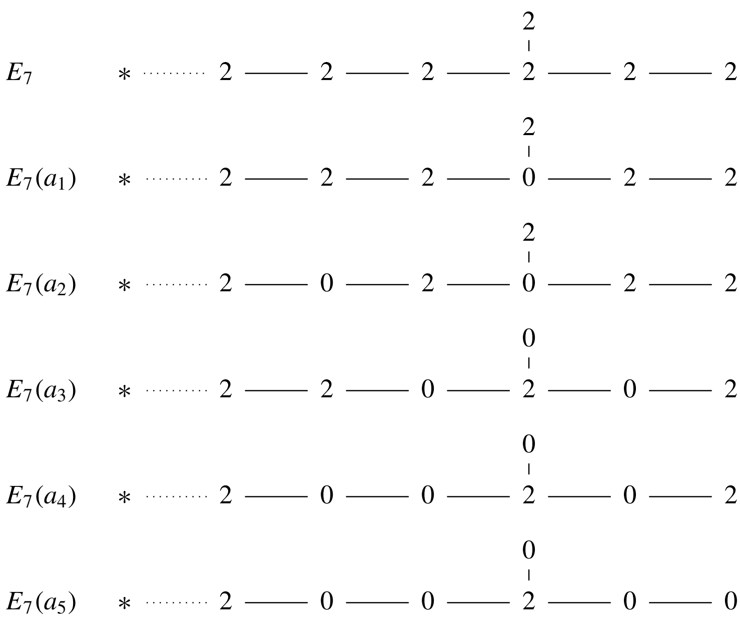
Weighted Dynkin diagrams of distinguished orbits of type D_n:



Weighted Dynkin diagrams of distinguished orbits of type E_6:



Weighted Dynkin diagrams of distinguished orbits of type E_7:



Acknowledgements

The research in this work is part of the author's PhD thesis, supported by the Engineering and Physical Sciences Research Council (EPSRC). I would like to thank my supervisor, Jack Shotton, for the advice and support given me throughout this project. I also thank reviewer for the detailed comments and suggestions given.

References

[BDP17] V. Balaji, P. Deligne, and A. J. Parameswaran, "On complete reducibility in characteristic p", *Épijournal Géom. Algébrique* **1** (2017), art. id. 3, 27 pp. MR

[Bel16] R. Bellovin, "Generic smoothness for G-valued potentially semi-stable deformation rings", *Ann. Inst. Fourier* (*Grenoble*) **66**:6 (2016), 2565–2620. MR

[BG19] R. Bellovin and T. Gee, "G-valued local deformation rings and global lifts", *Algebra Number Theory* **13**:2 (2019), 333–378. MR

[Car93] R. W. Carter, *Finite groups of Lie type*: *conjugacy classes and complex characters*, Wiley, New York, 1985. MR

[CHT08] L. Clozel, M. Harris, and R. Taylor, "Automorphy for some l-adic lifts of automorphic mod l Galois representations", *Publ. Math. Inst. Hautes Études Sci.* 108 (2008), 1–181. MR

[CM93] D. H. Collingwood and W. M. McGovern, *Nilpotent orbits in semisimple Lie algebras*, Van Nostrand, New York, 1993. MR

[Cot22] S. Cotner, "Springer isomorphisms over a general base scheme", preprint, 2022. arXiv 2211.08383

[CSS97] G. Cornell, J. H. Silverman, and G. Stevens (editors), *Modular forms and Fermat's last theorem: papers from the Instructional Conference on Number Theory and Arithmetic Geometry* (Boston, 1995), Springer, 1997. MR

[DHKM20] J.-F. Dat, D. Helm, R. Kurinczuk, and G. Moss, "Moduli of Langlands parameters", *J. Eur. Math. Soc.* (*JEMS*) **27**:5 (2025), 1827–1927. MR

[FR08] R. Fowler and G. Röhrle, "On cocharacters associated to nilpotent elements of reductive groups", *Nagoya Math. J.* **190** (2008), 105–128. MR

[FS25] L. Fargues, "Geometrization of the local Langlands correspondence: an overview", *J. Math. Sci. Univ. Tokyo* **32**:2 (2025), 157–240. MR

[Ger19] D. Geraghty, "Modularity lifting theorems for ordinary Galois representations", *Math. Ann.* **373**:3-4 (2019), 1341–1427. MR

[Gro99] B. H. Gross, "Algebraic modular forms", *Israel J. Math.* **113** (1999), 61–93. MR

[Har77] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR

[Hel23] E. Hellmann, "On the derived category of the Iwahori–Hecke algebra", *Compos. Math.* **159**:5 (2023), 1042–1110. MR

[HT01] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies **151**, Princeton University Press, 2001. MR

[Hum90] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge University Press, 1990. MR

[Jan03] J. C. Jantzen, *Representations of algebraic groups*, 2nd ed., Mathematical Surveys and Monographs **107**, Amer. Math. Soc., 2003. MR

[Mat86] H. Matsumura, *Commutative ring theory*, Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, Cambridge, 1986. MR

[Pre03] A. Premet, "Nilpotent orbits in good characteristic and the Kempf–Rousseau theory", *J. Algebra* **260**:1 (2003), 338–366. MR

[Sho18] J. Shotton, "The Breuil–Mézard conjecture when $l \neq p$", *Duke Math. J.* **167**:4 (2018), 603–678. MR

[Sho24] J. Shotton, "Irreducible components of the moduli space of Langlands parameters", *Int. Math. Res. Not.* **2024**:11 (2024), 9020–9035. MR

[Stacks] "The Stacks project", electronic reference, 2023, available at http://stacks.math.columbia.edu.

[Ste16] R. Steinberg, *Lectures on Chevalley groups*, corrected ed., University Lecture Series **66**, Amer. Math. Soc., 2016. MR

[Tho12] J. Thorne, "On the automorphy of l-adic Galois representations with small residual image", *J. Inst. Math. Jussieu* **11**:4 (2012), 855–920. MR

[Zhu25] X. Zhu, "Coherent sheaves on the stack of Langlands parameters", pp. 39–123 in *The Langlands program*, *II*: *Geometrization of the Langlands correspondence*, Proc. Sympos. Pure Math. **112.2**, Amer. Math. Soc., 2025. MR

Communicated by Frank Calegari
Received 2023-07-24 Revised 2025-03-25 Accepted 2025-06-27

danielfunckmaths@gmail.com *University of Tübingen, 72076 Tüebingen, Germany*

ALGEBRA AND NUMBER THEORY 20:6 (2026)
https://doi.org/10.2140/ant.2026.20.1159

Smoothness of stabilisers in generic characteristic

Ben Martin, David I. Stewart and Lewis Topley

Let R be a commutative unital ring. Given a finitely presented affine R-group scheme G acting on a finitely presented separated scheme X over R, we show that there is a prime p_0 such that for any R-algebra k that is a field of characteristic $p \geq p_0$, the centraliser in G_k of any closed subscheme of X_k is smooth. When X is not necessarily separated we show similarly that for any closed finitely presented subscheme $Y \subseteq X$ there is a p_1 depending on Y such that when k has characteristic $p \geq p_1$, the normaliser of Y_k in G_k is smooth. For the proof, we may assume k is algebraically closed, whence we prove these results using the Lefschetz principle together with careful application of Gröbner basis techniques, and using a suitable notion of the complexity of an action.

We apply our results to demonstrate that the Kostant–Kirillov–Souriau theorem holds for Lie algebras of algebraic groups in large positive characteristics: the coadjoint module of every such Lie algebra decomposes as a disjoint union of symplectic varieties, each of which is a coadjoint orbit.

1. Introduction

Let R be a commutative unital ring. By an *R-field* we mean an R-algebra that is also a field and by an *algebraic R-group* we mean a finitely presented affine R-group scheme. We prove the following.

Theorem 1.1. *Let G be an algebraic R-group and let X be a finitely presented separated G-scheme over R. Then there exists $p_0 \in \mathbb{N}$ such that whenever k is an R-field of characteristic $p \geq p_0$, the centraliser $C_{G_k}(Y)$ is smooth for every closed subscheme Y of X_k.*

Theorem 1.2. *Let G be an algebraic R-group and let X be a G-scheme of finite type over R. Let Y be a finitely presented closed subscheme of X. Then there exists $p_1 \in \mathbb{N}$ such that whenever k is an R-field of characteristic $p \geq p_1$, the normaliser $N_{G_k}(Y_k)$ is smooth.*

If we assume that R is noetherian then we can remove some of the hypotheses on X and Y: see Remark 4.15.

Theorem 1.1 uses the hypothesis that X is separated in order to infer that the centralisers are closed subschemes of G (see [Jan03, I.2.6]). Note also that the lower bound of p_0 in Theorem 1.1 for centralisers does not depend on Y; the bound in Theorem 1.2 for normalisers does, however, depend on Y, and this dependence cannot be removed: see Remark 4.12.

MSC2020: primary 03C60; secondary 20G07.
Keywords: algebraic group, smooth normaliser, smooth centraliser, smoothness, Gröbner basis, Lefschetz principle.

Natural examples of algebraic groups over rings abound, of course. The split reductive groups are all $\mathbb{Z}$-defined [Dem63]; and so too are the subgroups of reductive groups normalised by a split maximal torus — so-called *subsystem subgroups*. This class includes all parabolic subgroups of reductive groups.

There are several known special cases of the theorem. One of the most influential in Lie theory occurs when G is split reductive and X is either G itself or its Lie algebra $\mathfrak{g}$, on which G acts by the relevant adjoint action. Then it is well known that the group G, the algebra $\mathfrak{g}$, and the adjoint action are defined over $\mathbb{Z}$ (see [Jan03, II.1.1]) and the centralisers of single elements of $X(\bar{k})$ are smooth whenever p is a very good prime for G. This follows (see [BMRT10, Theorem 1.3(a)]) from work of Richardson [Ric67], who used the notion of a reductive pair to give an elegant proof for very good p that the number of unipotent and nilpotent orbits of $G_{\bar{k}}$ is finite. This smoothness result was generalised in [BMRT10, Theorem 1.2] to cover arbitrary subgroups of $G_{\bar{k}}$ and subalgebras of $\mathrm{Lie}(G_{\bar{k}})$. The hypotheses were further weakened in [Her13]. Normalisers, while much less well-behaved, were thoroughly considered in [HS16], where it was shown that (necessarily large) bounds on the characteristic exist, depending on the root system, which ensure the normalisers of subspaces of the Lie algebras of reductive groups are smooth. Through the classification of nilpotent and unipotent orbits, these results have found applications in developing the subgroup structure of simple algebraic groups and subalgebra structure of their Lie algebras; recent examples include [LT18] and [PS19].

One new feature of our work is that we move beyond the affine case: our results apply not just to affine varieties but to quasi-projective varieties and other varieties of finite type. Here is an application. The second author computed explicitly in [Ste16] the orbits of exceptional groups on their Lie algebras, determining when centralisers and stabilisers of lines are smooth for minimal induced (or dual-Weyl) modules; non-smoothness occurs only in characteristics 2 or 3. This motivated [op. cit., Question 1.4], to which our theorem provides the following strong answer.

Corollary 1.3. *Let G be an algebraic R-group and let V be a G-module which is finitely generated and projective as an R-module. Then there is a prime p_V such that whenever k is an R-field of characteristic $p \geq p_V$, the centraliser $C_{G_k}(W)$ and normaliser $N_{G_k}(W)$ are smooth for any k-subspace W of V_k.*

The smoothness of centralisers follows immediately from Theorem 1.1. To prove smoothness of normalisers one can apply Theorem 1.2. Alternatively, one can apply Theorem 1.1 to $X = \mathrm{Gr}_m(V)$, the Grassmannian of m-generated submodules of V: the idea is that the stabiliser of a subspace is the centraliser of the corresponding point in the Grassmannian — for details, see Section 4.6. The same conclusion does not hold if we weaken the hypothesis and consider all G_k-modules V of bounded dimension (see Remark 4.13), but it does if G_k is reductive and we bound the weights of the G_k-module V in an appropriate sense: see Proposition 4.17.

In the final section of this paper we apply our main theorem to prove a modular analogue of the Kostant–Kirillov–Souriau (KKS) theorem from symplectic geometry. Their theorem states that if G is a complex algebraic group with Lie algebra $\mathfrak{g}$ then the symplectic leaves of the Poisson variety $\mathfrak{g}^*$ are precisely the coadjoint orbits. When G is semisimple, this result leads to a classification of symplectic homogeneous G-varieties, since they all arise as finite covers of coadjoint orbits. We refer the reader to [GS77, § IV.7] for a detailed background of the theory. When k is an algebraically closed field of characteristic $p > 0$ one can ask whether the Poisson variety $\mathfrak{g}^*$ decomposes into a disjoint union of symplectic G-homogeneous subvarieties. In general the answer is negative, however we show that the KKS theorem holds whenever the characteristic is sufficiently large (Theorem 5.2).

Let us say some words on the proofs of our main results. The central idea is to combine the Lefschetz principle from first-order model theory with Gröbner basis techniques. This approach was suggested in [Sch00] as a method to solve certain problems in algebraic geometry, but to our knowledge the current paper is the first place it has been carried out in practice. Roughly speaking, the Lefschetz principle says that a first-order property that holds for algebraically closed fields k of characteristic 0 also holds for algebraically closed fields of large enough characteristic. It is well known that an algebraic k-group in characteristic 0 is smooth; we show that smoothness can be expressed in a first-order way, then use the Lefschetz principle to deduce smoothness for algebraic k-groups in sufficiently large characteristic. In order to do this, one needs suitable notions of complexity for schemes, morphisms and group actions. This allows us to work with honest first-order sentences without needing parameters from the field: rather than considering a fixed G and a fixed X, we quantify over all G and X of bounded complexity (a similar trick was used in [MST19]). Two crucial tools are Lemma 3.9, which allows us to give a bound on complexities arising from certain ideal membership problems, and Lemma 3.8. We find, surprisingly, that although certain complexities need to be constrained, others do not: for example, we prove variations of Theorems 1.1 and 1.2 which hold for schemes X that are not of finite type (see Corollaries 4.4 and 4.8). See also Remarks 4.15 and 4.18.

Outline. In Section 2 we recall the language of Hopf algebras and the Lefschetz principle from model theory, which we use to pass information from characteristic zero to large positive characteristics. Section 3 is the heart of the paper. Here we recall the theory of Gröbner bases which allows us to express a criterion for smoothness in terms of first-order sentences in the language of rings. We recall the definition of complexity for schemes and morphisms (Section 3.2). We introduce the notion of a d-bounded Hopf quadruple, and in Theorem 3.19 we prove that such Hopf algebras correspond to smooth algebraic groups in large characteristics $p > p_0(d)$. In Section 4 we provide the proofs of the main theorems, using the functor-theoretic descriptions of centralisers and normalisers and the notions of G-complexity (Definition 4.1) and (G, Δ)-complexity (Definition 4.5). Finally in Section 5 we provide a short proof of our modular version of the KKS theorem from symplectic geometry.

2. Preliminaries

Throughout, we consider a fixed commutative unital ring R. The closure of a field k is denoted by $\bar{k}$.

2.1. ***Schemes, group schemes and Hopf algebras.*** We take the functorial approach to schemes, as per [DG70] and [Jan03]. Thus for an R-algebra Q we think of $\mathrm{Spec}_R(Q)$ as the functor $\mathrm{Hom}_{R\text{-Alg}}(Q, -) : \underline{R\text{-Alg}} \to \underline{\mathrm{Set}}$. An R-*functor* is a functor $X : \underline{R\text{-Alg}} \to \underline{\mathrm{Set}}$; we call an R-functor X *affine* if it is isomorphic to $\mathrm{Spec}_R(R[X])$ for some R-algebra $R[X]$. If Q is an R-algebra then X_Q denotes the Q-functor obtained from X by base change. We say a subfunctor Y of an R-functor X is *open* if for every R-algebra Q and natural transformation $\beta : \mathrm{Spec}_R Q \to X$, the subfunctor $\beta^{-1}(Y)$ of $\mathrm{Spec}_R Q$ is an open subfunctor of $\mathrm{Spec}_R Q$, and is *closed* if for every R-algebra Q and natural transformation $\beta : \mathrm{Spec}_R Q \to X$, the subfunctor $\beta^{-1}(Y)$ of $\mathrm{Spec}_R Q$ is a closed subfunctor of $\mathrm{Spec}_R Q$. A closed subfunctor of an affine functor is affine. Then X is a *scheme* (or an *R-scheme*) if it is *local* in the sense of [Jan03, I.1.8] and admits a decomposition $X = \bigcup_{i\in\mathbb{I}} X_i$ for some indexing set $\mathbb{I}$, where the X_i are open affine subfunctors of X.

If Q is an R-algebra and X is an R-scheme then X_Q is a Q-scheme [Jan03, I.1.10]. We say X is of *finite type* if $\mathbb{I}$ is finite and each $R[X_i]$ is finitely generated over R; in this case we say X is *finitely presented* if each $R[X_i]$ is a finitely presented R-algebra. If R is Noetherian (e.g., a field) then any finitely generated R-algebra is finitely presented, so in this case any R-scheme of finite type is finitely presented. We do not assume that all schemes are separated; recall that a scheme is *separated* if the diagonal map $D : X \to X \times X$ is an embedding. By an embedding of schemes we mean a closed immersion.

An *affine R-group scheme G* is a functor from $\underline{R\text{-Alg}}$ to $\underline{\mathrm{Grp}}$ which, as a functor to $\underline{\mathrm{Set}}$, is naturally equivalent to one of the form $\mathrm{Spec}_R(R[G])$ for some finitely generated R-algebra $R[G]$. We consider only the case where $R[G]$ is finitely presented; in keeping with [Jan03, I.2.1], we then call G an *algebraic R-group*. We do not assume that algebraic R-groups are smooth. There is a natural notion of a closed subgroup of G (loc. cit.). The archetypal example of an algebraic R-group is GL_d, which is also an example of a split reductive group.

A Hopf R-algebra consists of data $(R[G], \Delta, \sigma, \epsilon)$ where $R[G]$ is an R-algebra, and there are R-algebra homomorphisms $\Delta : R[G] \to R[G] \otimes_R R[G]$, $\sigma : R[G] \to R[G]$ and $\epsilon : R[G] \to R$ satisfying the dual of the group axioms [Jan03, I.2.3(1–3)]:

$$(\Delta \otimes \mathrm{id}) \circ \Delta = (\mathrm{id} \otimes \Delta) \circ \Delta, \tag{2-1}$$

$$(\epsilon \mathbin{\overline{\otimes}} \mathrm{id}) \circ \Delta = (\mathrm{id} \mathbin{\overline{\otimes}} \epsilon) \circ \Delta, \tag{2-2}$$

$$(\sigma \mathbin{\overline{\otimes}} \mathrm{id}) \circ \Delta = \bar{\epsilon} = (\mathrm{id} \mathbin{\overline{\otimes}} \sigma) \circ \Delta. \tag{2-3}$$

Here the symbol $\varphi \mathbin{\overline{\otimes}} \psi$ denotes the tensor product of the maps φ, ψ followed by the natural multiplication map $R[G] \otimes R[G] \to R[G]$, and $\bar{\epsilon}$ denotes ϵ followed by the inclusion of R in $R[G]$. Hence by definition, the category of algebraic R-groups is the opposite category to the category of finitely presented Hopf algebras over R.

The Lie algebra $\mathrm{Lie}(G)$ of an algebraic R-group G is defined to be the R-module of all R-linear maps $I/I^2 \to R$ where $I = \mathrm{Ker}(\epsilon)$; in other words it is $\ker\big(G(R[\varepsilon]/(\varepsilon^2)) \to G(R)\big)$ where $R[\varepsilon]/(\varepsilon^2)$ is the algebra of dual numbers and the map takes ϵ to 0. Following [Jan03, I.7.7(3)] one obtains a natural R-linear Lie bracket on $\mathrm{Lie}(G)$ induced by the comultiplication Δ. Every morphism of algebraic R-groups induces a natural R-linear homomorphism of their Lie algebras.

When G is an algebraic R-group and k is any R-algebra, we can consider the base change G_k, which is an algebraic k-group. To see this concretely, suppose $R[G] \cong R[x_1, \ldots, x_n]/(g_1, \ldots, g_m)$. We have an obvious map $\omega : R[x_1, \dots, x_n] \to k[x_1, \dots, x_n]$ and we see that

$$G_k \cong \operatorname{Spec}_k\big(k[x_1, \dots, x_n]/(\omega(g_1), \dots, \omega(g_m))\big). \tag{2-4}$$

An action of G on an R-scheme X is a natural transformation $\alpha : G \times X \to X$ with the property that $\alpha(Q) : G(Q) \times X(Q) \to X(Q)$ is a group action for all R-algebras Q. If X is an R-module (see [Jan03, I.2.2]) and $G(Q)$ acts through Q-linear transformations of $X(Q)$, then we say X is a G-module. If X is affine, then we get a coaction map of R-algebras $\Delta_X : R[X] \to R[X] \otimes R[G]$. In case X is a G-module, this is the *comodule map* of [Jan03, I.2.8].

Let V be a G-module such that V is finitely generated and projective as an R-module. We may regard V as an affine G-scheme of finite type over R with co-ordinate ring the symmetric algebra $S(V^*)$ [Jan03, I.2.2 and I.2.7]. This construction commutes with base change.

2.2. ***Model theory and the Lefschetz principle.*** In this paper we use the Lefschetz principle to deduce statements about algebras over algebraically closed fields of large positive characteristic from the corresponding statements in characteristic zero. In doing so we pursue a theme from our earlier work [MST19], in which we proved a version of the first Kac–Weisfeiler theorem for representations of modular Lie algebras using the Lefschetz principle. A detailed introduction to the principle can be read in [Mar02] for example, and a concise overview may be found in [MST19, §2.1].

For future reference we state the Lefschetz principle, in a version taken from [MST19, Corollary 2.4]. The *language of rings* $\mathcal{L}_{\text{ring}}$ is the collection of first-order formulas that can be built from the symbols $\{\forall, \exists, \vee, \wedge, \neg, +, -, \times, 0, 1, =\}$ along with an arbitrary choice of variables. A *sentence* is a formula containing no free variables. For example, the formula $(\exists y) y^2 = x$ is not a sentence because the variable x is free, but by quantifying over x we can obtain a sentence: e.g., $(\forall x)(\exists y) y^2 = x$. Given a ring R, a sentence is either true or false for that ring: for instance, the sentence $(\exists y_1)(\exists y_2)\big((y_1 \neq y_2) \wedge (y_1^2 = y_2^2)\big)$ is true in every field of characteristic $p > 2$, but is false in every reduced ring of characteristic 2.

A *theory* is a set of sentences. A ring $\mathfrak{R}$ is a *model* of a theory T if every sentence belonging to T is true in $\mathfrak{R}$: for instance, if p is either 0 or prime then AC_p is the set of sentences that are true in every algebraically closed field of characteristic p, and any algebraically closed field of characteristic p is a model of AC_p.

Theorem 2.1 (Lefschetz principle). *Let ϕ be a sentence in $\mathcal{L}_{\text{ring}}$.*

(1) *If ϕ is true in some model of AC_p, where $p \geq 0$, then ϕ is true in every model of AC_p.*

(2) *If ϕ is true in some model of AC_0, then there exists $p_0 \in \mathbb{N}$ such that ϕ is true in any model of AC_p for $p > p_0$.*

2.3. ***Ideals in tensor products of algebras.*** Suppose that k is a field and that A, B are finitely generated k-algebras. Let K be an ideal of $A \otimes_k B$ with generators $f_1, \dots, f_n$ (note that K is finitely generated as $A \otimes_k B$ is noetherian). Fix a basis $\Xi = \{c_\lambda \mid \lambda \in \Lambda\}$ for B over k. We can write $f_j = \sum_m a_{j,m} \otimes b_{j,m}$ for elements $a_{j,m} \in A$ and $b_{j,m} \in \Xi$. Without loss of generality we can assume the elements $\{b_{j,m} \mid m\}$

are distinct for each fixed j, and under this assumption the ideal $K' \subseteq A$ generated by $\{a_{j,m} \mid j, m\}$ is uniquely determined by K, as we see from the next lemma.

Lemma 2.2. *K' is the smallest ideal of A such that $K \subseteq K' \otimes_k B$.*

Proof. Certainly $K \subseteq K' \otimes_k B$ and so it suffices to take an ideal $L \subseteq A$ such that $K \subseteq L \otimes_k B$ and show that $K' \subseteq L$. Observe that $A \otimes_k B$ is a free A-module with basis $\{1 \otimes c_\lambda \mid \lambda \in \Lambda\}$. Furthermore $L \otimes_k B = \bigoplus_{\lambda \in \Lambda} L \otimes c_\lambda$ and so if $f_j \in L \otimes_k B$ then $a_{j,m} \in L$ for all m appearing in the expression $f_j = \sum_m a_{j,m} \otimes b_{j,m}$. It follows that $K' \subseteq L$ as required. □

3. Smoothness of centralisers and Gröbner bases

3.1. ***Bounded polynomials and Gröbner bases.*** Throughout this subsection we fix $n \in \mathbb{N}$ and k denotes a field. We will want to quantify over all k-algebras of bounded presentation, equipped with the structure of a Hopf algebra of bounded presentation. Here "bounded" means that the lengths and degrees of the polynomial expressions appearing in the defining ideal of the underlying affine algebra, together with the comultiplication, antipode and counit, are bounded. To do so, we need to formulate statements to say that the Hopf algebra axioms are satisfied. Our main tool to this end will be to quantify over all Gröbner bases of bounded degree. We refer to [Eis95, Chapter 15] for a hearty introduction to Gröbner bases, but for our purposes we collect a simplified version here.

The basic principle is to provide a process for reduction of elements of $S := k[x_1, \ldots x_n]$ by elements of an ideal, which will terminate in a finite number of steps. Hence one wants to know when the size of an expression is reduced by an operation, and for this one first needs to choose a total order on monomials. This order needs to be *admissible* in the sense that $m_1 > m_2$ if and only if $m_1 m_3 > m_2 m_3 > m_2$ for any monomials m_1, m_2, m_3 such that $m_3 \neq 1$.

We will demand of the order that for any monomial $m \in S$, there are only finitely many $m' < m$. For instance, we may use the *homogeneous* (or *graded*) *lexicographic ordering*, in which

$$m := x_1^{a_1} \ldots x_n^{a_n} > m' := x_1^{b_1} \ldots x_n^{b_n} \iff \deg m > \deg m' \text{ or if } \deg m = \deg m' \text{ then } a_i > b_i \text{ for the first index } i \text{ with } a_i \neq b_i.$$

Thus the set of monomials is isomorphic to $\mathbb{N}$ as a totally ordered set. We define m_r to be the *r-th monomial in S*. If m is a monomial then we define $k^* \cdot m$ to be $\{\lambda m \mid \lambda \in k^*\}$ and we call elements of this form *terms*; every polynomial can be written uniquely as a sum of terms. We extend $>$ to terms by defining $\lambda m_i > \mu m_j$ if $i > j$ and $0 \neq \lambda, \mu \in k$, and we define the *initial term* $\mathrm{in}(f)$ to be the greatest term appearing in f with respect to $>$ (taking $\mathrm{in}(0) = 0$). For an ideal $I \subseteq S$, we define $\mathrm{in}(I)$ to be the ideal generated by the elements $\mathrm{in}(f)$ for all $f \in I$. If g and h are terms then there is a unique monomial m such that m divides g and h, and any other monomial dividing g and h also divides m: we define $\gcd(g, h) = m$.

Definition 3.1. A *Gröbner basis*[1] with respect to $>$ is an ordered list of elements $(g_1, \ldots, g_t) \in S^t$ for some t such that if I is the ideal of S generated by $g_1, \ldots, g_t$, then $\mathrm{in}(g_1), \ldots, \mathrm{in}(g_t)$ generate $\mathrm{in}(I)$. Note that the g_i need not be distinct and that although we work with ordered lists, the property of being a Gröbner basis does not depend on the ordering of the g_i.

Fix $d \in \mathbb{N}$. We wish to view a polynomial in $S = k[x_1, \ldots, x_n]$ as a finite list of its coefficients, where we will ultimately be quantifying over all possible lists of those coefficients. We define the degree $\deg(f)$ of a polynomial $f \in S$ to be the total degree of $\mathrm{in}(f)$. According to our chosen (homogeneous) monomial order, $\deg(f)$ is the highest total degree of any term in f. Let $\mathcal{X}_d \subset S$ be the set of monomials of degree at most d, and let $\ell_d := |\mathcal{X}_d|$. By homogeneity of the monomial order again, this means $\mathcal{X}_d = \{\mathsf{m}_1, \ldots, \mathsf{m}_{\ell_d}\}$. Furthermore, we may identify the set S_d of polynomials of degree at most d with the Cartesian product k^{ℓ_d}: the polynomial $\sum_{i=1}^{\ell_d} \lambda_i \mathsf{m}_i$ corresponds to $(\lambda_1, \lambda_2, \ldots, \lambda_{\ell_d}) \in k^{\ell_d}$.

Definition 3.2. Let $S = R[x_1, \ldots, x_n]$ be a polynomial ring over a ring R. Let S_d denote the polynomials in S of degree at most d. We say that an ordered list $\mathscr{B}$ of polynomials in R is *d-bounded* if $|\mathscr{B}| = \ell_d$ and $\mathscr{B}$ consists of elements of S_d.

If $R = k$ and $\mathscr{B}$ is also a Gröbner basis, we say $\mathscr{B}$ is a *d-bounded Gröbner basis.*

We identify the set of d-bounded lists of elements of S with $S_d^{\ell_d} = k^{\ell_d^2}$.

Remarks 3.3. (i) Any Gröbner basis of length greater than ℓ_d consisting of polynomials of degree at most d can be reduced to a Gröbner basis of cardinality at most ℓ_d. For if there are at least $\ell_d + 1$ elements then two, f and g say, must have the same leading monomial. So for some $\lambda \in k$, $g - \lambda f$ has a lower leading monomial and replacing g by $g - \lambda f$ we still have a Gröbner basis, directly from Definition 3.1. Inductively we may assume g is zero, thus it can be removed to produce a smaller Gröbner basis.

(ii) Conversely, any finite list of polynomials (resp. Gröbner basis) can be embedded into a d-bounded list of polynomials (resp. d-bounded Gröbner basis) for some d by appending an appropriate number of zero polynomials to the end of the list.

Lemma 3.4. *Let $d \in \mathbb{N}$ and let $1 \le e \le \ell_d$. Then there is a first-order formula $\phi_{e,d}$ in the language $\mathcal{L}_{\mathrm{ring}}$ of rings with ℓ_d free variables such that for any polynomial $f \in S$ of degree at most d,*

$$\phi_{e,d}(f) \text{ holds} \iff \mathrm{in}(f) \in k^* \cdot \mathsf{m}_e.$$

Proof. After identifying the set S_d with the space k^{ℓ_d}, so that the polynomial $f = \sum_{i=1}^{\ell_d} \lambda_i \mathsf{m}_i$ identifies with $(\lambda_1, \ldots, \lambda_{\ell_d})$, the required formula is

$$(\lambda_e \neq 0) \wedge (\lambda_{e+1} = 0) \wedge (\lambda_{e+2} = 0) \wedge \cdots \wedge (\lambda_{\ell_d} = 0). \qquad \square$$

Given a d-bounded list of polynomials, we need to check with a first-order formula that it forms a Gröbner basis. For this, we appeal to Buchberger's criterion [Eis95, Theorem 15.8], which we reproduce here.

[1] In contrast to [BW93], but consistently with [Eis95], we allow elements of Gröbner bases to be zero.

Let c be any integer and let $\mathscr{B} = (g_1, \dots, g_c) \in S^c$. For each pair of indices $1 \le i, j \le c$, we define

$$m_{ij} = \frac{\mathrm{in}(g_i)}{\gcd(\mathrm{in}(g_i), \mathrm{in}(g_j))} \in S.$$

(We interpret this as 0 if $g_i = 0$ or $g_j = 0$.) Then it follows from the division algorithm [Eis95, Proposition 15.6] that there exist $f_u^{(ij)} \in S$ with $\mathrm{in}(m_{ji} g_i - m_{ij} g_j) \ge \mathrm{in}(f_u^{(ij)} g_u)$ for each $1 \le u \le c$, and remainders $h_{ij} \in S$, none of whose terms is in $(\mathrm{in}(g_1), \dots, \mathrm{in}(g_c))$, such that

$$m_{ji} g_i - m_{ij} g_j = \left(\sum_u f_u^{(ij)} g_u \right) + h_{ij}. \tag{3-1}$$

We call an expression (3-1) a *standard expression* for $m_{ji} g_i - m_{ij} g_j$.

Theorem 3.5 (Buchberger's criterion). *The set $\mathscr{B}$ is a Gröbner basis if and only if there exist standard expressions* (3-1) *such that $h_{ij} = 0$ for all $1 \le i, j \le c$.*

Lemma 3.6. *If g_i and g_j have no variables in common, there is a standard expression for $m_{ji} g_i - m_{ij} g_j$ such that $h_{ij} = 0$.*

Proof. Since $m_{ij} = \mathrm{in}(g_i)$ and $m_{ji} = \mathrm{in}(g_j)$, (3-1) becomes

$$m_{ji} g_i - m_{ij} g_j = \underbrace{-(g_j - \mathrm{in}(g_j))}_{f_i^{(ij)}} g_i + \underbrace{(g_i - \mathrm{in}(g_i))}_{f_j^{(ij)}} g_j.$$

Set $f_u^{(ij)} = 0$ for $u \ne i, j$. Write $g_i = \mathrm{in}(g_i) + \widetilde{\mathrm{in}(g_i)} + g_i'$ with $\widetilde{\mathrm{in}(g_i)}$ the initial term of $g_i - \mathrm{in}(g_i)$, and likewise for g_j. Then $m_{ji} g_i - m_{ij} g_j = \mathrm{in}(g_j)\,\widetilde{\mathrm{in}(g_i)} + \mathrm{in}(g_j) g_i' - \mathrm{in}(g_i)\,\widetilde{\mathrm{in}(g_j)} - \mathrm{in}(g_i) g_j'$ has as initial term whichever is the larger of $-\mathrm{in}(g_j)\,\widetilde{\mathrm{in}(g_i)}$ and $\mathrm{in}(g_i)\,\widetilde{\mathrm{in}(g_j)}$ (these two terms cannot cancel each other as g_i and g_j have no variables in common). But the initial terms of $f_i^{(ij)}$ and $f_j^{(ij)}$ are $-\widetilde{\mathrm{in}(g_j)}\,\mathrm{in}(g_i)$ and $\widetilde{\mathrm{in}(g_i)}\,\mathrm{in}(g_j)$, respectively, so $\mathrm{in}(m_{ji} g_i - m_{ij} g_j) \ge \mathrm{in}(f_u^{(ij)} g_u)$ for $u = i$ and $u = j$. Hence we get $h_{ij} = 0$, as required. □

Lemma 3.7. *Let $d \in \mathbb{N}$. There is a first-order formula β_d in the language of rings with ℓ_d^2 free variables such that if $\mathscr{B}$ is a d-bounded list of elements of S, then*

$$\beta_d(\mathscr{B}) \text{ holds} \iff \mathscr{B} \text{ is a Gröbner basis.}$$

Proof. Suppose $\mathscr{B} = (g_1, \dots, g_{\ell_d}) \in S_d^{\ell_d}$. We will produce a first-order formula that detects whether there exist expressions (3-1) for each pair (g_i, g_j), with $h_{i,j} = 0$. Suppose $\mathrm{in}(g_i) \in k^* \cdot \mathsf{m}_a$ and $\mathrm{in}(g_j) \in k^* \cdot \mathsf{m}_b$ and that there is a formula $\chi_{a,b}$ such that $\chi_{a,b}(g_i, g_j)$ is true if and only if there exist expressions (3-1) such that $h_{ij} = 0$. Then using Lemma 3.4 we set $\beta_d(\mathscr{B})$ to be the formula

$$\bigwedge_{1 \le i, j \le \ell_d} \Bigl(\bigvee_{1 \le a, b \le \ell_d} \bigl(\chi_{a,b}(g_i, g_j) \wedge \phi_{a,d}(g_i) \wedge \phi_{b,d}(g_j) \bigr) \Bigr),$$

and we see that $\beta_d(\mathscr{B})$ is true if and only if $\mathscr{B}$ satisfies the necessary and sufficient criterion of Buchberger's criterion to deduce that $\mathscr{B}$ is a Gröbner basis.

Thus we have reduced the problem, without loss of generality, to showing the existence of $\chi_{a,b}(g_1, g_2)$. For fixed a and b, $\mathsf{m}_{e'} := \gcd(\mathsf{m}_a, \mathsf{m}_b)$ is also fixed, depending just on the bijection between $\mathbb{N}$ and the monomials in S, hence so are the monomials m_{12} and m_{21}. Now, the highest monomial appearing in the left-hand side of (3-1) is at most the d'-th, where d' is given by $\mathsf{m}_{d'+1} = (\mathsf{m}_a \mathsf{m}_b / \mathsf{m}_{e'})$. Suppose $\operatorname{in}(m_{ji} g_i - m_{ij} g_j) = \mathsf{m}_e$. Then there is a finite set of pairs $P = \{(g_{a_b}, \mathsf{m}_{a_b})\}_{1 \le b \le p}$ such that $\operatorname{in}(g_{a_b} \mathsf{m}_{a_b}) \le \mathsf{m}_e$. Hence, setting $\chi_{e,a,b}(g_1, g_2)$ to be the formula

$$(\exists \lambda_1) \dots (\exists \lambda_p)(m_{21} g_1 - m_{12} g_2 - \sum_{1 \le b \le p} \lambda_b g_{a_b} \mathsf{m}_{a_b} = 0),$$

we see that $\chi_{e,a,b}(g_1, g_2)$ holds if and only if there is an expression of the form (3-1) for $m_{21} g_1 - m_{12} g_2$ with $h_{i,j} = 0$ (given that $\operatorname{in}(m_{ji} g_i - m_{ij} g_j) = \mathsf{m}_e$). Lastly, let $\chi_{a,b}(g_1, g_2)$ be the formula

$$\bigvee_{e=1}^{d'} \big(\phi_{e,d}(m_{2,1} g_1 - m_{1,2} g_2) \wedge \chi_{e,a,b}(g_1, g_2)\big);$$

this will do. □

Another important thing we need to be able to encode with a first-order statement is the dimension $\dim(I) = \dim(\operatorname{Spec}_k(S/I))$ of the scheme determined by an ideal $I \subseteq S = k[x_1, \dots, x_n]$. If $I = (g_1, \dots, g_{\ell_d})$ then in general it is not easy to read off $\dim I$ from the elements $\{g_1, \dots, g_{\ell_d}\}$. However, when $\{g_1, \dots, g_{\ell_d}\}$ form a Gröbner basis for I there is a simple method: the dimension is the maximal size of a subset $X \subseteq \{x_1, \dots, x_n\}$ such that $\operatorname{in}(g_1), \dots, \operatorname{in}(g_n)$ depend only on the elements of $\{x_1, \dots, x_n\} \setminus X$ [BW93, Definition 9.22 and Corollary 9.28]. Using this fact along with Lemma 3.4, we can determine dimension with a first-order formula.

Lemma 3.8. *Let $d \in \mathbb{N}$ and $0 \le e \le n$. Then there is a first-order formula $\delta_{e,d}$ in the language $\mathcal{L}_{\mathrm{ring}}$ of rings, with ℓ_d^2 free variables, such that if $\mathscr{B}$ is any d-bounded Gröbner basis with $I = (\mathscr{B})$, then*

$$\delta_{e,d}(\mathscr{B}) \text{ holds} \iff \dim(I) = e.$$

Proof. There is obviously a finite collection of lists of monomials that could play the role of initial terms of the elements of a d-bounded Gröbner basis defining an ideal of dimension e. More formally, there is a set $\mathcal{X}_e = \{X_j \mid j \in T\}$, where T is some finite index set, and where each X_j is a d-bounded list of monomials in S satisfying: (i) there are distinct $i_1, \dots, i_e \in \{1, \dots, , n\}$ such that each $m \in X_j$ does not involve $x_{i_1}, \dots, x_{i_e}$; (ii) for any distinct $i_1, \dots, i_{e+1} \in \{1, \dots, n\}$, there is $m \in X_j$ depending on x_{i_k} for some $1 \le k \le e+1$. For convenience we assume that the X_j are ordered sets and identify the monomials with their ordinal via the bijection of monomials of S with $\mathbb{N}$. Then we may set $\delta_{e,d}(\mathscr{B})$ to be the formula

$$\bigvee_{X_j = (a_1, \dots, a_{\ell_d}) \in \mathcal{X}_e} \big(\phi_{a_1,d}(g_1) \wedge \phi_{a_2,d}(g_2) \wedge \dots \wedge \phi_{a_{\ell_d},d}(g_{\ell_d})\big). \qquad □$$

The next lemma uses the ideal membership algorithm for Gröbner bases to write a first-order formula whose truth determines whether an element is in an ideal. If $\mathscr{B}$ is a d-bounded Gröbner basis and $f \in S_d$ then we may identify $(\mathscr{B}, f)$ with an element of $k^{{\ell_d}^2+\ell_d}$ in the usual manner.

Lemma 3.9. *Let $d \in \mathbb{N}$. Then there is a first-order formula ι_d in $\mathcal{L}_{\mathrm{ring}}$ with $\ell_d^2 + \ell_d$ free variables, such that for any $f \in S_d$ and d-bounded Gröbner basis $\mathscr{B}$ with $I := (\mathscr{B})$, we have*

$$\iota_d(\mathscr{B}, f) \text{ holds} \iff f \in I.$$

Proof. Let $(g_1, \dots, g_{\ell_d})$ be a d-bounded Gröbner basis and let $f \in S_d$. Since the elements of $\mathscr{B}$ have bounded total degree d, and $<$ is a homogeneous order, there are only finitely many monomials m such that $\mathrm{in}(mg_i) \le \mathrm{in}(f)$ for some $1 \le i \le \ell_d$, where this number depends only on d. Let $\mathsf{m}_{d'}$ be the greatest such monomial. Thus we set $\iota_d(\mathscr{B}, f)$ to be the formula

$$(\exists \lambda_{i,j})_{1 \le i \le d', 1 \le j \le d} \quad f = g_1\left(\sum_{i=0}^{d'} \lambda_{i,1}\mathsf{m}_i\right) + g_2\left(\sum_{i=0}^{d'} \lambda_{i,2}\mathsf{m}_i\right) + \cdots + g_d\left(\sum_{i=0}^{d'} \lambda_{i,d}\mathsf{m}_i\right). \qquad (\dagger)$$

We claim that $\iota_d(\mathscr{B}, f)$ is true if and only if $f \in I$. This follows by induction on e where $\mathrm{in}(f) = \mathsf{m}_e$: since $\mathscr{B}$ is a Gröbner basis, by [BW93, 5.35(vii)], f is *top-reducible* by some g_i or is not in I. In the former case, this means that there is a term m such that $\mathrm{in}(f - g_i m) < \mathrm{in}(f)$. By the inductive hypothesis, $\iota_d(\mathscr{B}, f - g_i m)$ is true whenever $f - g_i m \in I$, which is the case if and only if $f \in I$. If $f - g_i m \in I$ this says that there is an expression of the form (†) with f replaced by $f - g_i m$; moving $g_i m$ to the other side of the equation, this says that there is also one for f. □

3.2. ***Complexity of schemes and their morphisms.*** We recall some terminology, now reasonably common in the literature, to describe the boundedness of affine schemes; see [Sch00, Definition 4.1] for example. It is closely related to the notion of d-boundedness above.

Definition 3.10. (a) For $n \in \mathbb{N}$, we say that an ideal I of $S = R[x_1, \dots, x_n]$ *has complexity at most d* if I can be generated by polynomials of degree at most d; in this case we also say that the affine R-scheme X defined by the vanishing locus of I has complexity at most d.

(b) If in addition, $m \in \mathbb{N}$ we say that a homomorphism of R-algebras $\varphi : S \to T := R[y_1, \dots, y_m]$ has *degree at most d* if $\varphi(x_i)$ is a polynomial in the y_j of degree at most d.

More generally, we say that a morphism of affine schemes $f : Y \to X$ has *complexity at most d* if there are embeddings $X \subseteq \mathbb{A}_R^n$ and $Y \subseteq \mathbb{A}_R^m$ determined by ideals I of $S = R[x_1, \dots, x_n]$ and J of $T = R[y_1, \dots, y_m]$, such that the comorphism $f^* : S \to T$ applied to each x_i is represented modulo J by a polynomial in the y_j of degree at most d.

In particular, when $m = n$, $I = 0$ and f is the embedding $Y \hookrightarrow X = \mathbb{A}^n$, then f^* is just the quotient of S by J and has complexity at most 1.

Note that our definition differs slightly from that in [Sch00]: we regard n as fixed but we do not require that $n \le d$, as this is not necessary for our purposes.

By an *affine embedding* of an R-scheme X, we mean an embedding of X in $\mathbb{A}^n_R$ for some $n \in \mathbb{N}$. Below when we speak of the complexity of an R-scheme, we mean it to be taken with respect to a fixed affine embedding, and likewise for morphisms of R-schemes; we will not mention the affine embedding explicitly unless it is necessary. If we are given an affine embedding of X then we use the same embedding for any closed subscheme of X. If G is an algebraic R-group then we pick an affine embedding arising from a Hopf quadruple in the sense of Definition 3.13 below. If we are given affine embeddings of R-schemes X_1 and X_2 then we take our affine embedding of $X_1 \times X_2$ to be the product embedding.

Remarks 3.11. (i) If I has complexity at most d then a generating set of polynomials of degree at most d can be transformed into a d-bounded list as per Remarks (ii)(i) and (ii). This allows us to apply the results from Section 3.1 replacing hypotheses involving boundedness with hypotheses involving bounded complexity.

(ii) Let X, X' and X'' be affine schemes corresponding to ideals I of S, I' of S' and I'' of S'', respectively. Let $f : X \to X'$ and $g : X' \to X''$ be maps of complexity at most r and s, respectively. It follows immediately from the definitions that $g \circ f$ has complexity at most rs. Similarly, if Y' is a closed subscheme of X' and Y' has complexity at most d then $f^{-1}(Y')$ has complexity at most dr — note that this bound does not depend on the complexity of X and X'.

(iii) Suppose $\{X_i \mid i \in I\}$ is a family of affine R-schemes given by ideals I_i of S. It is immediate that if each X_i has complexity at most d then $\bigcap_{i \in I} X_i$ also has complexity at most d.

(iv) The notion of complexity behaves well under base change in the following sense. Suppose X is an affine R-scheme with a given embedding in some affine space $\mathbb{A}^n_R$. Let k be an R-field. Then base change gives an affine k-scheme X_k with an embedding in $\mathbb{A}^n_k$. If $f_1, \ldots, f_t \in R[x_1, \ldots, x_n]$ generate the vanishing ideal of X in $\mathbb{A}^n_R$ then the images of the f_i in $k[x_1, \ldots, x_n]$ generate the vanishing ideal of X_k in $\mathbb{A}^n_k$. Hence if X has complexity at most d then X_k also has complexity at most d. The analogous result holds for morphisms.

(v) Suppose X and Y are both affine R-schemes, say

$$R[X] = R[x_1, \ldots, x_r]/I \quad \text{and} \quad R[Y] = R[y_1, \ldots, y_s]/J$$

with

$$I = (f_1, \ldots, f_t) \quad \text{and} \quad J = (g_1, \ldots, g_u).$$

Then

$$R[X \times Y] \cong R[X] \otimes R[Y] \cong R[x_1, \ldots, x_r, y_1, \ldots y_s]/K,$$

where K is the ideal generated by the concatenation of the f_i and g_i (after extending the domain of f_i and g_i to be trivial functions of the y_j and x_j respectively). Then one sees that if the complexity of X is d_1 and of Y is d_2, we get a presentation of $R[X \times Y]$ which shows its complexity is at most $\max\{d_1, d_2\}$. In particular, arguing by induction one shows that if S/I has complexity at most d, then $(S/I)^{\otimes r}$ has complexity at most d.

Definition 3.12. Let G be an algebraic R-group and let X be an affine G-scheme with action map $\alpha : G \times X \to X$. We say that *the action of G on X has complexity at most d* if α does.

Definition 3.13. Recall that $S = R[x_1, \dots, x_n]$. Let $H := (\mathscr{B}, \Delta, \sigma, \epsilon)$ be a quadruple with $(\mathscr{B}) = I \trianglelefteq S$; and $\Delta : S \to S^{\otimes 2}$, $\sigma : S \to S$ and $\epsilon : S \to R$ being R-algebra homomorphisms satisfying $\Delta(I) \subseteq I \otimes S + S \otimes I$, $S(I) \subseteq I$ and $\epsilon(I) = 0$. Then we say H is a *Hopf quadruple* if S/I equipped with the maps Δ, σ and ϵ forms a Hopf algebra. We say that a Hopf quadruple is of *complexity at most d* if $\mathscr{B}$ consists of polynomials of degree at most d and the complexity of Δ, σ and ϵ are at most d (this is automatic for ϵ, as ϵ has complexity 0). Dually, if an affine algebraic R-group G is described by the data $(\mathrm{Spec}(S/(\mathscr{B})), \Delta^*, \sigma^*, \epsilon^*)$, where $(\mathscr{B}, \Delta, \sigma, \epsilon)$ is a Hopf quadruple of complexity at most d, then we say that G is an *algebraic group of complexity at most d*. If G' is a closed subgroup of G then we may represent G' by the Hopf quadruple $(\mathscr{B}', \Delta, \sigma, \epsilon)$, where $\mathscr{B}' \supseteq \mathscr{B}$; if the polynomials in $\mathscr{B}'$ all have degree at most d then G' also has complexity at most d.

When R is a field, we call such a quadruple a *GroHo quadruple* if $\mathscr{B}$ happens to be a Gröbner basis.

We now work over a field k. The above definition invites us to consider the complexity of homomorphisms from S to $S' = S^{\otimes r}$ for various r. To that end, write $\ell_{d,r}$ for the total number of monomials of degree at most d in S'. Let $I = (\mathscr{B}) \subseteq S$ be an ideal generated by a d-bounded list $\mathscr{B} = \{f_1, \dots, f_{\ell_d}\}$ of elements of S. Write $f_i = f_i(x_1, \dots, x_n)$ and define $f_{i,j} = f_i(x_{j,1}, \dots, x_{j,n})$ for $1 \le j \le r$. Let $J_r = (\mathscr{B}_r)$, where $\mathscr{B}_r = \{f_{1,1}, \dots, f_{1,\ell_d}, \dots, f_{r,1}, \dots, f_{r,\ell_d}\}$. Then there is an isomorphism

$$\varphi_r : (S/I)^{\otimes r} \to k[x_{1,1}, \dots, x_{1,n}, \dots, x_{r,1}, \dots, x_{r,n}]/J_r. \tag{3-2}$$

It follows that the scheme corresponding to $(S/I)^{\otimes r}$ also has complexity at most d.

Lemma 3.14. *Let $d, r \in \mathbb{N}$. Then there is a first-order formula $\zeta_{d,r}$ in $\mathcal{L}_{\mathrm{ring}}$ with $n\ell_{d,r} + \ell_d^2$ free variables such that if $\Lambda : S \to S^{\otimes r}$ is a homomorphism of complexity at most d and $\mathscr{B} = \{f_1, \dots, f_{\ell_d}\}$ is any d-bounded Gröbner basis for an ideal I of S,*

$$\zeta_{d,r}(\mathscr{B}, \Lambda) \text{ holds} \iff \Lambda \text{ factors to a homomorphism } S/I \to (S/I)^{\otimes r}.$$

Proof. Recall that $S = k[x_1, \dots, x_n]$ and $I = (\mathscr{B})$. With reference to Remark 3.11(v), we may take a presentation of $(S/I)^{\otimes r}$ as

$$(S/I)^{\otimes r} \cong k[x_{i,j}]/K$$

where $1 \le i \le n$, $1 \le j \le r$ and K is the ideal generated by the disjoint union $\mathscr{B}_r := \{f_{i,j}\}$, with $f_{i,j}$ acting on $x_{l,m}$ as zero if $j \ne m$ and otherwise acting as f_i acts on the x_l.

To deploy our Gröbner basis formulae from earlier, we need to specify a monomial order on the $x_{i,j}$'s. Define first an order on the (i, j) for $1 \le i \le n$ and $1 \le j \le r$ by $(i, j) < (i', j')$ if $j < j'$ or $j = j'$ and $i < i'$. Now define an order on monomials by $m := \prod_{i,j} x_{i,j}^{a_{i,j}} < m' := \prod_{i,j} x_{i,j}^{b_{i,j}}$ if and only if $\deg m < \deg m'$ or if $\deg m = \deg m'$ and $a_{i,j} < b_{i,j}$ for the greatest pair (i, j) such that $a_{i,j} \ne b_{i,j}$. We claim that $\mathscr{B}_r$ as above is a Gröbner basis for the homogeneous lexicographic monomial order on the monomials in $x_{i,j}$. Since our order extends the monomial orders on the subalgebras $k[x_{1,i}, \dots, x_{n,i}]$ for any fixed i, we see that Buchberger's criterion (Theorem 3.5) holds for all pairs $(f_{i,j}, f_{i',j'})$; furthermore if $j \neq j'$ then $f_{i,j}$ and $f_{i',j'}$ have no variables in common, so Buchberger's criterion holds for $(f_{i,j}, f_{i',j'})$ by Lemma 3.6. This proves the claim.

Now since Λ has complexity at most d and $\mathscr{B}_r$ is d-bounded, we may appeal to Lemma 3.9 to get first-order formulas $\iota_{d,r}$ such that $\iota_{d,r}(\mathscr{B}_r, \varphi_r(\Lambda(x_i)))$ holds if and only if $\varphi_r(\Lambda(x_i)) \in J_r$. Hence we set $\zeta_{d,r}$ to be the formula

$$\bigwedge_{i=1}^{n} \iota_{d,r}(\mathscr{B}_r, \varphi_r(\Lambda(x_i))). \qquad \square$$

Recall the axioms of a Hopf algebra, listed in (2-1)–(2-3).

Lemma 3.15. *Let $d \in \mathbb{N}$. There is a formula $\eta_d \in \mathcal{L}_{\text{ring}}$ with $\ell_d^2 + n(\ell_{d,2} + \ell_d + 1)$ free variables such that if $\mathscr{B}$ is any d-bounded Gröbner basis, with $I = (\mathscr{B})$ and $\Delta : S \to S^{\otimes 2}$, $\sigma : S \to S$ and $\epsilon : S \to k$ any d-bounded homomorphisms, then*

$$\eta_d(\mathscr{B}, \Delta, \sigma, \epsilon) \text{ holds} \iff (S/I, \Delta, \sigma, \epsilon) \text{ is a Hopf algebra.}$$

Proof. Assume Δ, σ and ϵ factor as $S/I \to (S/I)^{\otimes r}$. We must find formulas $\eta_d^{(1)}$ (resp. $\eta_d^{(2)}, \eta_d^{(3)}$) which hold if and only if (2-1) (resp. (2-2), (2-3)) are satisfied. Since the constructions are almost identical for each formula, we give the details for $\eta_d^{(1)}$. To see that (2-1) holds, it clearly suffices to check that $(\Delta \otimes \text{id}) \circ \Delta(x_i + I) - (\text{id} \otimes \Delta) \circ \Delta(x_i + I) = 0 \in (S/I)^{\otimes 3} \cong S^{\otimes 3}/J_3$ for each $1 \le i \le n$, where $J_3 = (\mathscr{B}_3)$ as above. This amounts to checking that $(\Delta \otimes \text{id}) \circ \Delta(x_i) - (\text{id} \otimes \Delta) \circ \Delta(x_i) \in J$. Since Δ is d-bounded, $\varphi_2(\Delta(x_i))$ is a d-bounded polynomial in $S^{\otimes 2}$; similarly $f_i := \varphi_3((\Delta \otimes \text{id}) \circ (\Delta(x_i)))$ is a d^2-bounded polynomial in $S^{\otimes 3}$. Likewise $\mathscr{B}_3$ is d-bounded. Thus we may set $\eta_d^{(1)}(\mathscr{B}, \Delta, \sigma, \epsilon)$ to be the formula

$$\bigwedge_{i=1}^{n} \iota_{d^2}(\mathscr{B}_3, f_i).$$

Finally, we set $\eta_d(\mathscr{B}, \Delta, \sigma, \epsilon)$ to be the formula

$$\zeta_{d,2}(\mathscr{B}, \Delta) \wedge \zeta_{d,1}(\mathscr{B}, \sigma) \wedge \zeta_{d,0}(\mathscr{B}, \epsilon) \wedge \eta_d^{(1)}(\mathscr{B}, \Delta, \sigma, \epsilon) \wedge \eta_d^{(2)}(\mathscr{B}, \Delta, \sigma, \epsilon) \wedge \eta_d^{(3)}(\mathscr{B}, \Delta, \sigma, \epsilon),$$

where the $\zeta_{d,r}$ are as in Lemma 3.14. $\square$

3.3. ***Generic smoothness of algebraic groups of bounded complexity.*** Here we invoke the Lefschetz principle and our first-order formulas to show that algebraic groups of bounded complexity are generically smooth. We use the fact that when k is a field, an algebraic k-group G is smooth if and only if $\dim(G) = \dim(\text{Lie}(G))$ [Jan03, I.7.17]. Note also that if k'/k is a field extension then G is smooth if and only if $G_{k'}$ is smooth.

As a k-vector space, the Lie algebra $\text{Lie}(G)$ is the tangent space $T_{\epsilon^*}(G)$, where ϵ^* is the identity element of G. Thus its dimension is the nullity of the $\ell_d \times n$ matrix $\mathscr{J}$ where $\mathscr{J}_{kl} = \epsilon(\partial f_k / \partial x_l)$.

Lemma 3.16. *Let $d \in \mathbb{N}$, and $0 \le e \le d$. There is a first-order formula $\tau_{e,d}$ in $\mathcal{L}_{\text{ring}}$ with ℓ_d^2 free variables such that for any GroHo quadruple $(\mathscr{B}, \Delta, \sigma, \epsilon)$ of complexity at most d that describes the algebraic k-group G we have*

$$\tau_{e,d}(\mathscr{B}) \text{ holds} \iff \dim \operatorname{Lie}(G) = e.$$

Proof. As we identify each $f_i \in \mathscr{B}$ with the set of its ℓ_d coefficients λ_{ij}, partial differentiation by $\partial/\partial x_i$ gives a linear map $k^{\ell_d} \to k^{\ell_d}$. Composing with ϵ is then a linear map $k^{\ell_d} \to k$. Hence each $\mathscr{J}_{kl}$ is a fixed linear combination of the λ_{ij}'s. The statement that the nullity of $\mathscr{J}$ is e is equivalent to the statement that there are e linearly independent vectors $v_1, \ldots, v_e \in k^{\ell_d}$ satisfying $\mathscr{J} \cdot v_i = 0$ and given any $v_{e+1} \in k^{\ell_d}$ such that $v_1, \ldots, v_{e+1} \in k^{\ell_d}$ is linearly independent, there exists $v \in \langle v_1, \ldots, v_{e+1}\rangle$ such that $\mathscr{J} \cdot v \neq 0$. This statement can be given as a formula in $\mathcal{L}_{\text{ring}}$ in an obvious way (see [MST19, Example 2.1(i)] for example). □

Lemma 3.17. *Let $d \in \mathbb{N}$. Then there is a first-order formula θ_d with ℓ_d^2 free variables such that for any GroHo quadruple $H := (\mathscr{B}, \Delta, \sigma, \epsilon)$ of complexity at most d,*

$$\theta_d(\mathscr{B}) \text{ holds} \iff H \text{ describes a smooth } k\text{-group}.$$

Proof. The k-group G described by H is a subscheme of $\operatorname{Spec}(S) \cong \mathbb{A}^n$, so $0 \le \dim G \le n$. Then invoking Lemmas 3.16 and 3.8 we may set $\theta_d(\mathscr{B}, \Delta, \sigma, \epsilon)$ to be the following formula:

$$\bigvee_{e=0}^{n} \left(\delta_{e,d}(\mathscr{B}) \wedge \tau_{e,d}(\mathscr{B})\right). \qquad \square$$

We wish to apply the above results to obtain statements about the set of algebraic groups of complexity at most d. This amounts to statements about the Hopf quadruples of complexity at most d. The following theorem of Dubé guarantees that any algebraic group of complexity at most d can in fact be described by a GroHo quadruple of complexity at most D, where D depends just on d and n.

Theorem 3.18 [Dub90]. *If $\mathscr{B}' \subset S$ is a set of polynomials of degree at most d, then $(\mathscr{B}')$ has a Gröbner basis $\mathscr{B}$ consisting of polynomials of degree at most*

$$2\left(\frac{d^2}{2} + d\right)^{2^{n-1}}.$$

Theorem 3.19. *Let $d \in \mathbb{N}$. Then there is a prime $p_0 = p_0(n, d)$ such that whenever $\operatorname{char}(k) \ge p_0$, any algebraic k-group of complexity at most d is smooth.*

Proof. By Theorem 3.18, each affine algebraic group of complexity at most d is described by a GroHo quadruple of complexity at most D, where D depends just on d and n; here n has been fixed. So it suffices to prove that there is a $p_0(D)$ such that any GroHo quadruple of complexity at most D over a field k of characteristic $p \ge p_0$ describes a smooth algebraic k-group.

First suppose that k is algebraically closed. Let $H = (\mathscr{B}, \Delta, \sigma, \epsilon)$ be a GroHo quadruple of complexity at most D. Recall we identify H with a string of $\ell_D^2 + n(\ell_{D,2} + 2\ell_D + 1)$ coefficients in k, which we write

$(\lambda_i)_{i=1}^{\ell_D^2+n(\ell_{D,2}+2\ell_D+1)}$.Then invoking Lemmas 3.7, 3.15 and 3.17, the following formula Φ_D is a *sentence* in $\mathcal{L}_{\text{ring}}$ that is true if and only if all GroHo quadruples of complexity at most D describe smooth algebraic groups:

$$(\forall\lambda_1)\cdots(\forall\lambda_{\ell_D^2+n(\ell_{D,2}+2\ell_D+1)})(\beta_D(\mathscr{B})\wedge\eta_D(\mathscr{B},\Delta,\sigma,\epsilon)\wedge\theta_D(\mathscr{B})).$$

By Cartier's theorem [Jan03, I.7.17(2)], Φ is true for all fields of characteristic 0. Therefore the Lefschetz principle (Theorem 2.1) guarantees the existence of a prime p_0 such that the same is true for all algebraically closed fields of characteristic $p \geq p_0$.

Now let k be arbitrary and let G be an algebraic k-group of complexity at most d. Suppose $p \geq p_0$. We can choose a GroHo quadruple $H = (\mathscr{B}, \Delta, \sigma, \epsilon)$ of complexity at most D such that G is the corresponding algebraic k-group. By changing base from k to $\bar{k}$, we may regard H as a GroHo quadruple that defines $G_{\bar{k}}$. Recall that complexity does not increase under base change, so the complexity of this GroHo quadruple is still at most d. So $G_{\bar{k}}$ is smooth by the algebraically closed case, which implies that G is smooth as well. This proves the theorem. □

4. Normalisers and centralisers

We now prove our main results Theorems 1.1 and 1.2. We will work with non-affine schemes, so we need some preliminary material. We use the definitions and terminology of [Jan03, Chapter I]. Although we work mainly over k, we need to consider arbitrary R and study the behaviour of some of the constructions below under base change to k. Note that every R-scheme is locally free in the sense of [Jan03, I.1.15] if R is a field. We fix an R-scheme X and an open covering $\{X_i \mid i \in \mathbb{I}\}$ of X such that each X_i is affine. For each i, we fix an affine embedding of X_i in some $\mathbb{A}_R^{n_i}$. This allows us to talk about the complexity of certain schemes and maps involving the X_i.

We also fix a (not necessarily smooth) algebraic R-group G described by a Hopf quadruple $(S/I, \Delta, \sigma, \epsilon)$. We fix an action $\alpha : G \times X \to X$ of G on X.

4.1. *The functor* $\mathfrak{Mor}(-,-)$. We review some basic constructions of algebraic geometry: see [Jan03, I.1.15, I.2.6]. Recall that for R-schemes Z and W, we get an R-functor

$$\mathfrak{Mor}(Z, W) : \underline{R\text{-Alg}} \to \underline{\text{Set}}$$

given by $\mathfrak{Mor}(Z, W)(Q) = \text{Mor}(Z_Q, W_Q)$ for any R-algebra Q; if $\varphi : Q \to Q'$ is a homomorphism of R-algebras then $\mathfrak{Mor}(Z, W)(\varphi) : \text{Mor}(Z_Q, W_Q) \to \text{Mor}(Z_{Q'}, W_{Q'})$ is given by base change. Given another R-scheme C and a morphism $\alpha : C \times Z \to W$ of R-schemes, we get a map of R-functors $\nu : C \to \mathfrak{Mor}(Z, W)$ such that for an R-algebra Q and $c \in C(Q)$, ν maps c to $\alpha(c, -) \in \text{Mor}(Z_Q, W_Q)$.

Now let W' be a closed subscheme of W. If Z is locally free as an R-scheme then by [Jan03, I.1.15] we may regard $\mathfrak{Mor}(Z, W')$ as a closed R-subfunctor of $\mathfrak{Mor}(Z, W)$. So by [Jan03, I.1.15(3)] we obtain a closed subscheme $\nu^{-1}(\mathfrak{Mor}(Z, W'))$ of C.

4.2. ***G-complexity.*** Next we describe the precise boundedness condition that we need. With the notation of the last subsection, let Y be a closed subscheme of X. Set $Y_i = Y \cap X_i$ for each i. Recall $A := R[G] \cong S/I$ for $S = R[x_1, \ldots, x_n]$ and let $B_i = R[Y_i]$. Let $\alpha_i : G \times Y_i \to X$ be the restriction of α to $G \times Y_i$. Then $\alpha_i^{-1}(Y)$ is a closed subscheme of $G \times Y_i$ [Jan03, I.1.12(2)], so it corresponds to an ideal K_i of $R[G \times Y_i] = A \otimes_R B_i$. We can write $K_i = (f_j^{(i)} \mid j \in \mathbb{J}_i)$, where $\mathbb{J}_i$ is some (possibly infinite) indexing set, and where each $f_j^{(i)}$ has the form

$$f_j^{(i)} = \sum_m a_{mj}^{(i)} \otimes b_{mj}^{(i)} \tag{4-1}$$

for some $a_{mj}^{(i)} \in A$ and $b_{mj}^{(i)} \in B_i$.

Definition 4.1. Let G, X and Y be as above. We say that Y *has G-complexity at most d* if there exist $f_j^{(i)}$ as above such that each $a_{mj}^{(i)}$ has a representative $\tilde{a}_{mj}^{(i)}$ in S of degree at most d.

Note that we do not place any restrictions on the degrees of the $b_{mj}^{(i)}$ in the definition, and we do not require the K_i to be finitely generated.

Remark 4.2. Let G, X and Y be as above. The G-complexity condition in Definition 4.1 can be hard to verify, but here is a useful special case. Let $X \subseteq \mathbb{A}^s$ be affine; let J be an ideal of $R[\mathbb{A}^s] = R[t_1, \ldots, t_s]$ of complexity at most e defining a closed subscheme Y of X. Furthermore, recalling Definition 3.12, suppose the action $\alpha : G \times X \to X$ has complexity at most e'. Then we claim Y has G-complexity at most ee'.

To see this, let $\tilde{\alpha}$ be the restriction of α to $G \times Y$, and note that $\tilde{\alpha}^{-1}(Y)$ is the closed subscheme in $R[G] \otimes_R R[Y]$ defined by the vanishing of $\tilde{\alpha}^*(J)$. So since J is generated by polynomials in the t_i of degree at most e, then applying the algebra homomorphism $\tilde{\alpha}^*$ to them leads to polynomials of degree at most ee' (see Remark 3.11(ii)).

Much of the time we can do better than this; if the image of $\tilde{\alpha}$ is Y — so that G normalises Y — then $\tilde{\alpha}^{-1}(Y) = G \times Y$ is the closed subscheme of $G \times Y$ defined by the zero ideal. In that case, Y has G-complexity 0.

Now let Q be an R-algebra. Change of base yields an algebraic Q-group G_Q acting on a Q-scheme X_Q with an open covering by affine schemes $(X_i)_Q$. If Y is a closed subscheme of X then Y_Q is a closed subscheme of X_Q. The various constructions in Section 4.2 are well-behaved with respect to base change, so we see that if Y has G-complexity at most d then Y_Q has G_Q-complexity at most d.

4.3. ***Normalisers.*** We start by recalling the scheme-theoretic definition of the normaliser (see [Jan03, I.2.6] for details). In this subsection we assume the ground ring is a field k. This implies that the local freeness condition used in Section 4.1 holds. Let Y be a closed subscheme of X. Then the *normaliser of Y* (denoted $N_G(Y)$) is the k-subgroup functor of G given by

$$N_G(Y)(Q) = \{g \in G(Q) \mid \alpha(g,h) \in Y(Q') \text{ for all } h \in Y(Q') \text{ and all } Q\text{-algebras } Q'\}.$$

It is clear from the definition that if k' is a k-algebra then $N_{G_{k'}}(Y_{k'}) = (N_G(Y))_{k'}$. We will need the explicit description of $N_G(Y)$ afforded by the k-functor $\mathfrak{Mor}(-,-)$. The action $\alpha : G \times X \to X$ gives rise to a map $\nu : G \to \mathfrak{Mor}(X, X)$ as described in Section 4.1. We have a map $\mathfrak{Mor}(X, X) \to \mathfrak{Mor}(Y, X)$ given by restriction, and we let $\gamma : G \to \mathfrak{Mor}(Y, X)$ be the composition with ν. Then $N_G(Y) = \gamma^{-1}(\mathfrak{Mor}(Y, Y)) \cap i_G(\gamma^{-1}(\mathfrak{Mor}(Y, Y)))$, where $i_G : G \to G$ is the inversion map. Since Y is locally free, $\mathfrak{Mor}(Y, Y)$ is closed in $\mathfrak{Mor}(Y, X)$ and it follows that $N_G(Y)$ is a closed subgroup functor of G. Now $N_G(Y)$ is an algebraic k-group since k is noetherian.

Theorem 4.3. *Let $d \in \mathbb{N}$ and let G, X and Y be as above. Suppose G has complexity at most d and Y has G-complexity at most d. Then $N_G(Y)$ is a closed subscheme of G of complexity at most d^2.*

Proof. We have $N_G(Y) = \gamma^{-1}(\mathfrak{Mor}(Y, Y)) \cap i_G(\gamma^{-1}(\mathfrak{Mor}(Y, Y)))$. By Remark 3.11(ii), it is enough to prove that $\gamma^{-1}(\mathfrak{Mor}(Y, Y))$ has complexity at most d (since i_G has complexity at most d).

Each Y_i is a closed subscheme of X_i [Jan03, I.1.13, Lemma], and the Y_i form an open cover of Y. Let γ_i be the composition $G \to \gamma\mathfrak{Mor}(Y, X) \to \mathfrak{Mor}(Y_i, X)$, where the second map is given by restriction. By [Jan03, I.1.15(2)], $\gamma^{-1}(\mathfrak{Mor}(Y, Y)) = \bigcap_{i\in\mathbb{I}} \gamma_i^{-1}(\mathfrak{Mor}(Y_i, Y))$. Each $\gamma_i^{-1}(\mathfrak{Mor}(Y_i, Y))$ is a closed subscheme of G by an argument similar to the one for $\gamma^{-1}(\mathfrak{Mor}(Y, Y))$. In the proof of [Jan03, I.1.15(3)], one considers a map f from $\mathrm{Spec}_k(R)$ to $\mathfrak{Mor}(Y_i, X)$, where R is an arbitrary k-algebra; one obtains a map f' from $\mathrm{Spec}_k(R) \times Y_i$ to X. We apply this construction when $R = A = k[G]$, $B_i = k[Y_i]$ and $f = \gamma_i$; then f' is just the map $\alpha_i : G \times Y_i \to X$. Let K_i be the ideal of $k[G \times Y_i] = A \otimes_k B_i$ corresponding to $\alpha_i^{-1}(Y)$ as before, and let K_i' be the ideal of A corresponding to $\gamma_i^{-1}(\mathfrak{Mor}(Y_i, Y))$. By the argument of loc. cit., K_i' is the smallest ideal of A such that $K_i' \otimes B_i$ contains K_i.

Fix $i \in \mathbb{I}$. Choose $f_j^{(i)}$, $a_{mj}^{(i)}$, $\widetilde{a}_{mj}^{(i)}$ and $b_{mj}^{(i)}$ as in Definition 4.1, where each $\widetilde{a}_{mj}^{(i)}$ has degree at most d. Rewriting and expanding as necessary, we may assume each $b_{mj}^{(i)}$ is a member of a fixed k-basis of B_i; evidently this does not affect the bound on the degree of the $\widetilde{a}_{mj}^{(i)}$. Thanks to Lemma 2.2 we see that K_i' is generated by the $a_{mj}^{(i)}$, thus has complexity at most d.

Let $I_0 = I + \sum_{i\in\mathbb{I}} K_i'$ be the ideal of S corresponding to $\gamma^{-1}(\mathfrak{Mor}(Y, Y))$. Since each K_i' has complexity at most d, so does I_0. The result now follows. □

Corollary 4.4. *Let $d, n \in \mathbb{N}$. Then there is a prime $p_1 = p_1(n, d)$ such that if:*

- *k is any field of characteristic $p \geq p_1$;*
- *G is any affine algebraic k-group of complexity at most d;*
- *X is any k-scheme on which G acts;*
- *Y is any closed subscheme of X of G-complexity at most d;*

then $N_G(Y)$ is a smooth closed subscheme of G.

Proof. This follows from Theorem 3.19 and Theorem 4.3. We need to be working with a fixed n as well as a fixed d in order to apply Theorem 3.19; but this is the case here as we are considering subgroups of a fixed G. □

4.4. ***(G, Δ)-complexity.*** Once again we work over an arbitrary ring R. We need a slightly different boundedness condition for Theorem 4.7. Here we assume again that $X = \bigcup_{i\in\mathbb{I}} X_i$ is an R-scheme with Y a closed subscheme. This time, we also assume that X is separated; this means that the diagonal Δ_X is closed in $X \times X$. Set $Y_i = Y \cap X_i$ for each i as before. Let $A = R[G]$ and let $B_i = R[Y_i]$ for each i.

Define $\beta : G \times X \to X \times X$ by $\beta = \alpha \times \mathrm{pr}_2$, where $\mathrm{pr}_2 : G \times X \to X$ is the projection; let β_i, β_Y and β_{Y_i} be its restriction to $G \times X_i$, $G \times Y$ and $G \times Y_i$, respectively. Thanks to [Jan03, I.1.12(2)] we see that $\beta_{Y_i}^{-1}(\Delta_X)$ is a closed subscheme of $G \times Y_i$, so it corresponds to an ideal K_i of $R[G \times Y_i] = A \otimes_R B_i$. We can write $K_i = (f_j^{(i)} \mid j \in \mathbb{J}_i)$, where each $f_j^{(i)}$ has the form

$$f_j^{(i)} = \sum_m a_{mj}^{(i)} \otimes b_{mj}^{(i)} \tag{4-2}$$

for some $a_{mj}^{(i)} \in A$ and some $b_{mj}^{(i)} \in B_i$.

Definition 4.5. Let G, X and Y be as above. We say that Y *has (G, Δ)-complexity at most d* if there exist $f_j^{(i)}$ as above such that each $a_{mj}^{(i)}$ has a representative $\tilde{a}_{mj}^{(i)}$ in S of degree at most d.

Lemma 4.6. *Let G, X and Y be as above. Suppose X has (G, Δ)-complexity at most d. Then Y has (G, Δ)-complexity at most d.*

Proof. Fix $i \in \mathbb{I}$. By hypothesis, $\beta_i^{-1}(\Delta_X)$ is the closed subscheme of $G \times X_i$ defined by elements of the form $f_j^{(i)} = \sum_m a_{mj}^{(i)} \otimes b_{mj}^{(i)}$ for j in some indexing set $\mathbb{J}_i$, where each $a_{mj}^{(i)} \in A$ has a representative $\tilde{a}_{mj}^{(i)}$ in S of degree at most d and each $b_{mj}^{(i)} \in B_i$. Let $L_i = (h_\ell^{(i)} \mid \ell \in \mathbb{L}_i)$ be the ideal cutting out Y_i in X_i, where $\mathbb{L}$ is a (possibly infinite) indexing set. Now, we have $\beta_{Y_i}^{-1}(\Delta_X) = \beta_i^{-1}(\Delta_X) \cap (G \times Y_i)$, so the ideal cutting out $\beta_{Y_i}^{-1}(\Delta_X)$ is generated by the $f_j^{(i)}$ for $j \in \mathbb{J}_i$ together with the elements $1 \otimes h_\ell^{(i)}$ for $\ell \in \mathbb{L}$. As the constant polynomial 1 has complexity 0, the result now follows. □

Now let Q be an R-algebra. We see that if X has (G, Δ)-complexity at most d then X_Q has (G_Q, Δ_Q)-complexity at most d.

4.5. ***Centralisers.*** The argument is similar to the one for normalisers, and we recommence with the notation from Section 4.3. In particular, our ground ring is a field k. We assume additionally that the k-scheme X is separated. Recall the definition of the centraliser $C_G(Y)$ from [Jan03, I.2.6]: it is the k-subgroup functor of G whose Q-points are

$$C_G(Y)(Q) = \{g \in G(Q) \mid \alpha(g, y) = y \text{ for all } y \in Y(Q') \text{ and all } Q\text{-algebras } Q'\},$$

where Q is any k-algebra. It is clear from the definition that if k' is a k-algebra then $C_{G_{k'}}(Y_{k'}) = (C_G(Y))_{k'}$.

The maps β and β_Y from Section 4.2 give rise to maps

$$\delta : G \to \mathfrak{Mor}(X, X \times X) \quad \text{and} \quad \delta_Y : G \to \mathfrak{Mor}(Y, X \times X),$$

as described in Section 4.1. We have a map $\mathfrak{Mor}(X, X \times X) \to \mathfrak{Mor}(Y, X \times X)$ given by restriction, and it is easily seen that δ_Y is the composition $G \to \delta\mathfrak{Mor}(X, X \times X) \to \mathfrak{Mor}(Y, X \times X)$.

Since X is separated, Δ_X is closed in $X \times X$. Then $C_G(Y) = \delta^{-1}(\mathfrak{Mor}(Y, \Delta_X))$, and this is a closed subgroup functor of G; in particular, $C_G(Y)$ is an algebraic k-group.

Theorem 4.7. *Let $d \in \mathbb{N}$ and let G, X and Y be as above. Suppose G has complexity at most d and X has (G, Δ)-complexity at most d. Then $C_G(Y)$ is a closed subscheme of G of complexity at most d.*

Proof. Let δ_i be the composition $G \to \gamma\mathfrak{Mor}(Y, X \times X) \to \mathfrak{Mor}(Y_i, X \times X)$, where the second map is given by restriction. We have $C_G(Y) = \delta^{-1}(\mathfrak{Mor}(Y, \Delta_X))$. By [Jan03, I.1.15(2)], $\delta^{-1}(\mathfrak{Mor}(Y, \Delta_X)) = \bigcap_{i \in \mathbb{I}} \delta_i^{-1}(\mathfrak{Mor}(Y_i, \Delta_X))$. Each $\delta_i^{-1}(\mathfrak{Mor}(Y_i, \Delta_X))$ is a closed subscheme of G. In the proof of [Jan03, I.1.15(3)], one considers a map f from $\mathrm{Spec}_k(R)$ to $\mathfrak{Mor}(Y_i, X)$, where R is an arbitrary k-algebra; one obtains a map f' from $\mathrm{Spec}_k(R) \times Y_i$ to X. We apply this construction when $R = A = k[G]$, $B_i = k[Y_i]$ and $f = \delta_i$; then f' is just the map $\beta_i : G \times Y_i \to X \times X$. Let K_i be the ideal of $k[G \times Y_i] = A \otimes_k B_i$ corresponding to $\beta_{Y_i}^{-1}(\Delta_X)$, and K_i' the ideal of A corresponding to $\delta_i^{-1}(\mathfrak{Mor}(Y_i, \Delta_X))$. By the argument of loc. cit., K_i' is the smallest ideal of A such that $K_i' \otimes B_i$ contains K_i.

Since Y has (G, Δ)-complexity at most d by Lemma 4.6, K_i is generated by elements of the form $f_j^{(i)} = \sum_m a_{mj}^{(i)} \otimes b_{mj}^{(i)}$ for some $a_{mj}^{(i)} \in A$ and some $b_{mj}^{(i)} \in B_i$, where each $a_{mj}^{(i)}$ has a representative $\widetilde{a}_{mj}^{(i)}$ in S of degree at most d. As in the proof of Theorem 4.3, we know K_i' is generated by the elements $a_{mj}^{(i)}$, and has complexity at most d. Hence $C_G(Y)$ has complexity at most d, and we are done. □

Corollary 4.8. *Let $d, n \in \mathbb{N}$. Then there is a prime $p_0 = p_0(n, d)$ such that if:*

- *k is any field of characteristic $p \geq p_0$;*
- *G is any affine algebraic k-group of complexity at most d;*
- *X is any separated k-scheme of (G, Δ)-complexity at most d;*

then for any closed subscheme Y of X, the centraliser $C_G(Y)$ is a smooth closed subscheme of G.

Proof. This follows from Theorem 3.19 and Theorem 4.7. As in Corollary 4.4, we are working with subgroups of a fixed G, so there is no problem with applying Theorem 3.19. □

Remark 4.9. Because of Lemma 4.6, the complexity hypotheses in Corollary 4.8 do not involve Y; therefore we get better smoothness results for centralisers than for normalisers (cf. Remark 4.12). The complexity hypothesis on X is difficult to check in general, but it clearly holds if X is of finite type, since then we can take I to be finite. Here is a useful special case. Suppose X is finitely presented and affine and suppose the action $\alpha : G \times X \to X$ has complexity at most d. By assumption, X corresponds to an ideal I of $R[x_1, \ldots, x_n]$, where I is generated by polynomials $f_1, \ldots, f_t$ of $R[x_1, \ldots, x_n]$ for some $n, t \in \mathbb{N}$. Then X has complexity at most d', where d' is the maximum of the degrees of the f_i. We claim that X has (G, Δ)-complexity at most dd'. To see this, observe that Δ_X corresponds to the ideal J of $R[x_1, \ldots, x_n] \otimes_R R[x_1, \ldots, x_n]$ generated by the polynomials $f_i \otimes 1 - 1 \otimes f_i$ for $1 \leq i \leq t$, so Δ_X has complexity at most d'. The map $\beta : G \times X \to X \times X$ has complexity at most d, since α has complexity at most d and pr_2 has complexity 1. Hence $\beta^{-1}(\Delta_X)$ has complexity at most dd' by Remark 3.11(ii), and the claim follows.

Remark 4.10. Let G, X and Y be as above (defined over arbitrary R). Suppose G has complexity at most d. Then for any R-algebra Q and any $y \in Y_Q(Q)$, the singleton $\{y\}$ has G_Q-complexity at most d. In case $Q = k$ is algebraically closed, suppose Y_k is reduced; thus its k-points $Y_k(k)$ are dense. It follows that $C_{G_k}(Y_k)$ is equal to the closed subgroup $\bigcap_{y \in Y(k)} N_{G_k}(\{y\})$. Each $N_{G_k}(\{y\})$ has G-complexity at most d^2 by Theorem 4.3, so $C_{G_k}(Y_k)$ has complexity at most d^2. Hence by Theorem 3.19 there exists $p_0 = p_0(n, d)$ such that if $\mathrm{char}(k) \geq p_0$ then we have that $C_{G_k}(Y_k)$ is smooth. Note that by translating the problem to normalisers of points, we do need to assume that X is separated.

4.6. *Proof of main theorems and examples.* We can now give a quick proof of our main results.

Proof of Theorems 1.1 and 1.2. Let G and X be as in the statement of Theorem 1.1. Since G is affine and finitely presented, there exists $d_1 \in \mathbb{N}$ such that G has complexity at most d_1. Since X is finitely presented, there exists $d_2 \in \mathbb{N}$ such that X has (G, Δ)-complexity at most d_2; note that there are only finitely many ideals K_i that we need consider in the definition of (G, Δ)-complexity for X, and each K_i is finitely generated. Hence for any R-field k, we see from Remark 3.11(iv) that G_k has complexity at most d and X_k has (G_k, Δ)-complexity at most d, where $d = \max\{d_1, d_2\}$. Now Theorem 1.1 follows from Corollary 4.8.

Now let G, X and Y be as in the statement of Theorem 1.2. Since X is of finite type and Y is finitely presented, we can choose a finite cover of X by open affine subsets X_i such that $Y_i := Y \cap X_i$ is finitely presented for each i. There exists $d_2 \in \mathbb{N}$ such that Y has G-complexity at most d_2 (we need consider only finitely many finitely generated ideals K_i in the definition of G-complexity). We finish the proof as we did in the centraliser case, using Corollary 4.4 in place of Corollary 4.8. □

Proof of Corollary 1.3. The G-module V is a finitely presented affine G-scheme over R. The proof for centralisers of subspaces follows immediately from Theorem 1.1 applied to G and V.

For normalisers we will use Corollary 4.4. The dual of a finitely generated projective module is finitely generated and projective, so we can choose generators $\alpha_1, \ldots, \alpha_n$ for V^* as an R-module. The map $\alpha_1 \times \cdots \times \alpha_n$ gives an R-linear embedding of V as a subspace of $\mathbb{A}^n_R$. There exists $d \in \mathbb{N}$ such that G has complexity at most d and the map $G \times V \to V$ given by the action has complexity at most d. Now let k be an R-field. Then G_k has complexity at most d and the map $G_k \times V_k \to V_k$ given by the action has complexity at most d, by Remark 3.11(iv). Let W be a subspace of V_k. Then W has complexity at most 1 with respect to the embedding of V_k in $\mathbb{A}^n_k$, so W has G-complexity at most d by Remark 4.2. The result now follows from Corollary 4.4. □

Remark 4.11. We sketch another proof of the normaliser part of Corollary 1.3. Let k be an R-field and let W be a subspace of V_k. Set $r = \dim(W)$. Then $N_{G_k}(W) = C_{G_k}(x)$, where x is the element of $\mathbb{P}(\Lambda^r(V_k))$ corresponding to the line $\Lambda^r(W)$ in the exterior power $\Lambda^r(V_k)$, so we can deduce the result by applying Theorem 1.1 to the G-module $\mathbb{P}(\Lambda^r(V))$. We leave the details to the reader.

Remark 4.12. Any hope to extend Theorem 1.2 to deal with normalisers of arbitrary closed subschemes of X will fail without first imposing some further hypotheses. For instance, [HS16, Lemma 11.11] gives for each prime p a smooth subgroup H_p of GL_3 over an algebraically closed field of characteristic p such that the normaliser of H_p is nonsmooth. Here we can take $G = X = \mathrm{GL}_3$ over $R = \mathbb{Z}$ with G acting on X by conjugation; Theorem 1.2 does not apply because our closed subschemes H_p are not of the form Y_k for any fixed closed subscheme Y of X.

Remark 4.13. Likewise, Theorem 1.1 fails without some kind of complexity hypothesis. For example, let G be a split simple and simply connected group over $\mathbb{Z}$ and $N = V_G(\lambda)$ the Weyl module for G with minuscule highest weight λ. Then N_k is irreducible for each algebraically closed field k. When $\operatorname{char} k = p > 0$, let $M_k = (N_k)^{[1]}$ be the Frobenius twist of N_k through $F : G_k \to G_k^{(1)}$; as $G_k \cong F(G_k) = G_k^{(1)}$ we have M_k irreducible too. By irreducibility, $C_{G_k}(m) \subsetneq G_k$ for any $0 \neq m \in M_k$. The k-group G being connected and smooth it follows that $\dim_k(C_{G_k}(m)) < \dim G_k$, yet $\mathrm{Lie}(G_k)$ is in the kernel of the action on M_k. Thus $\dim_k \mathrm{Lie}(C_{G_k}(m)) = \dim G_k$; it follows that $C_{G_k}(m)$ is not smooth. Note that X is of finite type here and G and X are fixed, but the action is not.

Remark 4.14. Here is a closely related example of the limits of Theorem 1.1; this time X, G and the G-action on X are fixed but X is not of finite type. Let $R = \mathbb{Z}$, $G = \mathrm{SL}_2$ and X be the G-scheme not of finite type which is the disjoint union of the G-modules $H^0(p) = \mathrm{Ind}_B^G(p)$ for every prime p — here B is a Borel subgroup of G and the integer p denotes a free $\mathbb{Z}$-module of rank 1 on which B acts with weight p through the quotient map to a maximal torus. When $\operatorname{char} k = p$ and $k = \bar{k}$, the simple socle of the G_k-module $H^0(p)_k$ is isomorphic to a Frobenius twist $L(1)_k^{[1]}$ of the natural 2-dimensional G_k-module $L(1)_k$. If v is a point in the socle of $H^0(p)_k$ then its centraliser is a proper subgroup of G_k, but v is centralised by the whole Lie algebra. We conclude that the centraliser of v is not smooth. In this instance, the action map $\alpha : G \times X \to X$ is not d-bounded for any d.

Remark 4.15. If we assume that R is noetherian then we obtain some variations on Theorems 1.1 and 1.2, as follows. First consider normalisers. Let R be noetherian, let G be an algebraic R-group and let X be a G-scheme. Let Y be a closed subscheme of X. We claim that there exists $p_1 \in \mathbb{N}$ such that whenever k is an R-field of characteristic $p \geq p_1$, $N_{G_k}(Y_k)$ is smooth.

To see this, choose an open covering of X by affine schemes X_i for i in some indexing set $\mathbb{I}$. Set $Y_i = Y \cap X_i$, let $\alpha_i : G \times Y_i \to X$ be the restriction of the action and let K_i be the ideal of $A \otimes_R B_i$ corresponding to $\alpha_i^{-1}(Y_i)$, where $A := R[G]$ and $B_i := R[Y_i]$. Choose generators $f_j^{(i)}$ for K_i, where j runs over some indexing set $\mathbb{J}_i$. For each $i \in \mathbb{I}$ and each $j \in \mathbb{J}_i$, we can write $f_j^{(i)} = \sum_m a_{mj}^{(i)} \otimes b_{mj}^{(i)}$, where each $a_{mj}^{(i)}$ belongs to A and each $b_{mj}^{(i)}$ belongs to B_i. Now let K_i' be the ideal of A generated by the $a_{mj}^{(i)}$, and let $K' = \sum_{i \in \mathbb{I}} K_i'$. Since G is algebraic, A is finitely generated. It follows that K' is finitely generated as R is noetherian: so some finite subset F of the $a_{mj}^{(i)}$ generates K'.

Now let k be an R-field. We obtain X_k, Y_k, $(Y_i)_k$, $(\alpha_i)_k$ by changing base. The ideal of $(\alpha_i)_k^{-1}(Y_k)$ is $K_i \otimes_R k$, and we have $\sum_{i \in \mathbb{I}} K_i' \otimes_R k = K' \otimes_R k$. But $K' \otimes_R k$ is generated by the elements $a_{mj}^{(i)} \otimes 1$, where the $a_{mj}^{(i)}$ run over the elements of F. It follows that $K' \otimes_R k$ has complexity bounded by some d which is independent of k. Applying Theorem 3.19 yields the result.

By a very similar argument we obtain the following result for centralisers. Let R be noetherian, let G be an algebraic R-group and let X be a separated G-scheme over R. Let Y be a closed subscheme of X. Then there exists $p_0 \in \mathbb{N}$ such that whenever k is an R-field of characteristic $p \geq p_0$, $C_{G_k}(Y_k)$ is smooth.

Note that in these results we did not need X or Y to be finitely presented, or even locally of finite type. On the other hand, Remark 4.14 shows that Theorem 1.1 becomes false if we remove the assumption that X is of finite type, even when R is noetherian.

4.7. ***Modules for reductive groups.*** Let k be a field of characteristic $p > 0$ and let G be a split reductive k-group (which by convention means that it is connected). We recall some of the basic representation theory of G as found in [Jan03, II.1, II.2]. Let B be a Borel subgroup of G, containing a split maximal torus T of G. This choice defines a set of simple roots of the (reduced) root lattice Φ of G, and a subset of dominant weights $X(T)^+$ of the character lattice $X(T) = \mathrm{Hom}(T, \mathbb{G}_m)$. Moreover, there is a 1–1 correspondence between the dominant weights $\lambda \in X(T)^+$ and the simple G-modules $L(\lambda)$. Let k_λ denote the 1-dimensional k-module on which T acts with weight λ; this is also a B-module via the canonical projection $B \to T$. The induced representation $H^0(\lambda) := \mathrm{Ind}_B^G(k_\lambda)$ is finite-dimensional and contains $L(\lambda)$ as its unique simple submodule; i.e., as its socle.

There is a natural pairing of $X(T)$ with the cocharacter lattice $Y(T) = \mathrm{Hom}(\mathbb{G}_m, T)$ denoted

$$\langle\ ,\ \rangle : X(T) \times Y(T) \to \mathbb{Z};\ (\lambda, \varphi) \mapsto \lambda \circ \phi.$$

One can show that any root $\alpha \in \Phi \subseteq X(T)$ gives rise to a coroot $\alpha^\vee \in Y(T)$. The next result follows from [Jan03, V.5.6].

Lemma 4.16. *Let ρ denote half the sum of the positive roots.*

(i) *Let λ be a dominant weight. If $\langle \lambda + \rho, \alpha^\vee \rangle \leq p$ for all $\alpha \in \Phi^+$ then $H^0(\lambda) = L(\lambda)$.*

(ii) *If V is a G-module such that $\langle \lambda + \rho, \alpha^\vee \rangle \leq p$ for all $\alpha \in \Phi^+$ and all dominant weights λ in V, then V is semisimple.*

Say a dominant weight λ is *d-bounded* if $\langle \lambda + \rho, \alpha^\vee \rangle \leq d$ for all $\alpha \in \Phi^+$.

Proposition 4.17. *Let Φ be a (reduced) root system and fix $d \in \mathbb{N}$. There are primes p_2 and p_3 with the following properties. If k is a field of characteristic $p \geq p_2$ (resp. $p \geq p_3$), if G is any connected reductive k-group such that $G_{\bar{k}}$ has root system Φ, and if V is a finite-dimensional G-module such that the dominant weights of $V_{\bar{k}}$ are all d-bounded, then the centralisers of all closed subschemes of V in G (resp., the normalisers of all closed subschemes of V of complexity at most d) are smooth.*

Proof. Since centralisers, normalisers and smoothness are well-behaved under field extensions, we may assume without loss that $k = \bar{k}$. We may assume $p_2, p_3 \geq d$. Then by Lemma 4.16 and local finiteness [Jan03, I.2.13–14],

$$V \cong \bigoplus_{\lambda \in \Lambda} L(\lambda) \cong \bigoplus_{\lambda \in \Lambda} H^0(\lambda),$$

for some indexing set Λ of dominant weights. Recall that the group G is defined over $\mathbb{Z}$ in the sense that there is an algebraic $\mathbb{Z}$-group $\boldsymbol{G}$ such that $\boldsymbol{G}_k \cong G$ [Jan03, II.1.17]; furthermore $\boldsymbol{G}$ has a Borel subgroup $\boldsymbol{B}$ containing a split maximal torus $\boldsymbol{T}$ and such that $\boldsymbol{B}_k$ is a Borel subgroup of G and $\boldsymbol{T}_k$ is a split maximal torus of G. All Borel subgroups of G are $G(k)$-conjugate and all split maximal tori of B are $B(k)$-conjugate, so without loss of generality, $B = \boldsymbol{B}_k$ and $T = \boldsymbol{T}_k$. Since $\boldsymbol{G}$ is split, there is a character $\boldsymbol{\lambda}$ of $\boldsymbol{T}$ such that $\lambda = \boldsymbol{\lambda}_k$. Now induction commutes with flat base change [Jan03, I.3.5(3)], which implies $H^0(\lambda) = \operatorname{Ind}_{\boldsymbol{B}}^{\boldsymbol{G}}(\mathbb{Z}_{\boldsymbol{\lambda}})_k$. Thus V is isomorphic to the base change to k of a $\boldsymbol{G}$-module $\boldsymbol{V}$, which is free as a $\mathbb{Z}$-module. Moreover as there are only finitely many possible isomorphism classes of the direct summands of $\boldsymbol{V}$ — there are only finitely many d-bounded weights — the complexity of $\boldsymbol{G}$, the complexity of $\boldsymbol{V}$ and the complexity of the action map of $\boldsymbol{G}$ on $\boldsymbol{V}$ are bounded by a function $f = f(\Phi, d)$ depending just on Φ and d, being that of the maximal summand. (To obtain the bound on the complexity of $\boldsymbol{V}$ we used Remark 3.11(v).) Hence by Remark 3.11(iv) the complexity of G, the complexity of V and the complexity of the action map of G on V are bounded by a function $f = f(\Phi, d)$. It follows from Remark 4.9 that V has (G, Δ)-complexity at most $f(\Phi, d)^2$, and it follows from Remark 4.2 that if Y is a closed subscheme of V of complexity at most d then Y has G-complexity at most $df(\Phi, d)$. Now we are done by an application of Corollaries 4.8 and 4.4. □

Remark 4.18. Note that we do not need any bound on $\dim(V)$ in Proposition 4.17.

5. Application: the Kostant–Kirillov–Souriau theorem in characteristic *p*

A foundational result in the theory of smooth complex Poisson varieties in Weinstein's symplectic foliation theorem, which states that every such variety decomposes into a disjoint union of its symplectic leaves. In general these are not complex submanifolds and, even when they are, they usually fail to be locally closed for the Zariski topology (see [BG03, Remark 3.6(1)] for example). Nevertheless there is a large class of Poisson varieties which admit locally closed symplectic leaves. If G is a complex algebraic group with Lie algebra $\mathfrak{g}$ then $\mathfrak{g}^*$ carries a natural Poisson structure. A fundamental result in symplectic geometry states that the coadjoint orbits coincide with the symplectic leaves of $\mathfrak{g}^*$; this is known as the Kostant–Kirillov–Souriau theorem. For semisimple groups this construction is exhaustive in a precise sense: every symplectic homogeneous space is a finite covering of such an orbit (see [GS77, §IV.7]).

The theory of symplectic varieties in positive characteristic is still in its early stages, although the foundations have been carefully lain in the landmark work of Bezrukavnikov and Kaledin [BK08], where they classified Frobenius constant quantisations of smooth symplectic varieties. Another notable result is the proof of the formal version of Weinstein's splitting theorem [Tik18], which decomposes a restricted Poisson variety into a product of a symplectic subvariety and a transverse Poisson slice, in a formal neighbourhood of a point.

Although one cannot define symplectic leaves in positive characteristic, many Poisson varieties decompose into a disjoint union of quasi-affine symplectic subvarieties. This seems to be the most natural replacement for the symplectic foliation in this setting.

When G is an algebraic group over an algebraically closed field of positive characteristic we can ask whether the coadjoint orbits are symplectic subvarieties of $\mathfrak{g}^*$. In general the answer is negative, and the failure can be traced back to the fact that the quotient map from a group to an orbit is not always separable. We now demonstrate that the Kostant–Kirillov–Souriau theorem holds for algebraic groups over algebraically closed fields of sufficiently large positive characteristics, using Theorem 1.1. First, we require a preparatory lemma. If G is an algebraic R-group and k is an algebraically closed R-field then we write $\mathfrak{g}_k$ for $\mathrm{Lie}(G_k)$.

Proposition 5.1. *Let G be an algebraic R-group. There exists a prime $p_4 \in \mathbb{N}$ with the property that if k is any algebraically closed field of characteristic $p \geq p_4$, then for all $x \in \mathfrak{g}_k$ and all $\chi \in \mathfrak{g}_k^*$, the subgroups $C_{G_k}(x)$ and $C_{G_k}(\chi)$ are smooth.*

Proof. Since G is finitely presented, we have $A = R[G] = R[x_1, \dots, x_n]/I$ with $I = (f_1, \dots, f_r)$, and suppose G has complexity at most d — which implies we can choose the f_i to have degree at most d. We may assume G is nontrivial and so $d \geq 1$. The vanishing ideal I_1 at the identity is then finitely generated of complexity at most d and so the Lie algebra $\mathrm{Lie}(G) \cong (I_1/I_1^2)^*$ and the adjoint action $G \times \mathrm{Lie}(G) \to \mathrm{Lie}(G)$ have bounded complexity as a function just of d; see [Jan03, I.2.4(8)] for a formula. (Certainly d^4 will suffice.) Moreover, if Q is any R-algebra then $Q[G_Q] \cong R[G] \otimes_R Q$ has complexity at most d and the same formula implies the adjoint action has complexity at most d^4 also. A similar argument yields the bounded complexity of the co-adjoint action. The result now follows from Remark 4.9 and Corollary 4.8. □

Let G be an algebraic R-group and let p_4 be as Proposition 5.1. Pick an algebraically closed R-field k of characteristic p. The following is a version of the Kostant–Kirillov–Souriau theorem.

Theorem 5.2. *If $p \geq p_4$ then the induced Poisson structure on coadjoint orbits in $\mathfrak{g}_k^*$ is symplectic. Hence $\mathfrak{g}_k^*$ decomposes into a disjoint union of locally closed symplectic subvarieties.*

Proof. Since $p \geq p_4$ it follows from Proposition 5.1 that the coadjoint stabilisers in $\mathfrak{g}_k$ are all smooth. For the proof we fix $\chi \in \mathfrak{g}_k^*$ and write Ω for the coadjoint orbit of χ, write ad^* for the coadjoint representation of $\mathfrak{g}_k$ on $\mathfrak{g}_k^*$ and write $\mathfrak{g}_k^\chi$ for the stabiliser of χ. Thanks to [Jan04, 2.1], the natural bijective morphism $G_k/C_{G_k}(\chi) \to \Omega$ is separable and so we have isomorphisms

$$T_\chi \Omega \xrightarrow{\sim} \mathrm{ad}^*(\mathfrak{g}_k)\chi \xrightarrow{\sim} \mathfrak{g}_k/\mathfrak{g}_k^\chi. \tag{5-1}$$

Let $I \subseteq k[\mathfrak{g}_k]$ be the defining ideal of $\overline{\Omega}$. Since Ω is G-stable, I is G-stable and hence $\mathfrak{g}_k$-stable. Hence I is a Poisson ideal and $k[\overline{\Omega}]$ inherits a Poisson structure. Let $x_1, \dots, x_n$ be a basis for $\mathfrak{g}_k$. The rank of the Poisson structure at χ is the rank of the matrix π^χ such that $\pi^\chi_{i,j} = \chi([x_i, x_j])$. However π^χ is nothing other than the matrix of the linear form $\bigwedge^2 \mathfrak{g}_k \to k$ given by $(x, y) \mapsto \chi([x, y])$. The radical of this form is $\mathfrak{g}_k^\chi$ and so we conclude that the rank of the Poisson structure on $\overline{\Omega}$ at χ is $\dim \mathfrak{g}_k/\mathfrak{g}_k^\chi$. It follows from (5-1) that the rank coincides with $\dim \Omega$. Since G acts by Poisson automorphisms and Ω is homogeneous we conclude that the Poisson structure on $\overline{\Omega}$ has full rank at every point of Ω, as desired. □

Acknowledgements

Stewart's research is funded by Leverhulme Trust Research Project Grant number RPG-2021-080. Topley is grateful for the support of EPSRC grant EP/N034449/1, as well as funding from the UKRI, grant number MR/S032657/1, MR/S032657/2, MR/S032657/3. We are grateful to Sylvy Anscombe and Charlotte Kestner for helpful discussions. We thank the referees for their careful reading of the paper and many useful comments.

References

[BG03] K. A. Brown and I. Gordon, "Poisson orders, symplectic reflection algebras and representation theory", *J. Reine Angew. Math.* **559** (2003), 193–216. MR

[BK08] R. Bezrukavnikov and D. Kaledin, "Fedosov quantization in positive characteristic", *J. Amer. Math. Soc.* **21**:2 (2008), 409–438. MR

[BMRT10] M. Bate, B. Martin, G. Röhrle, and R. Tange, "Complete reducibility and separability", *Trans. Amer. Math. Soc.* **362**:8 (2010), 4283–4311. MR

[BW93] T. Becker and V. Weispfenning, *Gröbner bases: a computational approach to commutative algebra*, Graduate Texts in Mathematics **141**, Springer, 1993. MR

[Dem63] M. Demazure, "Fibrés tangents, algèbres de Lie", fascicle 1, exposé 2, 40 pp. in *Schémas en groupes* (IHES, 1963), Inst. Hautes Études Sci., Paris, 1963. MR

[DG70] M. Demazure and P. Gabriel, *Groupes algébriques*, *I*: *Géométrie algébrique*, *généralités*, *groupes commutatifs*, Masson, Paris, 1970. MR

[Dub90] T. W. Dubé, "The structure of polynomial ideals and Gröbner bases", *SIAM J. Comput.* **19**:4 (1990), 750–775. MR

[Eis95] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR

[GS77] V. Guillemin and S. Sternberg, *Geometric asymptotics*, Mathematical Surveys **14**, Amer. Math. Soc., 1977. MR

[Her13] S. Herpel, "On the smoothness of centralizers in reductive groups", *Trans. Amer. Math. Soc.* **365**:7 (2013), 3753–3774. MR

[HS16] S. Herpel and D. I. Stewart, "On the smoothness of normalisers, the subalgebra structure of modular Lie algebras, and the cohomology of small representations", *Doc. Math.* **21** (2016), 1–37. MR

[Jan03] J. C. Jantzen, *Representations of algebraic groups*, 2nd ed., Mathematical Surveys and Monographs **107**, Amer. Math. Soc., 2003. MR

[Jan04] J. C. Jantzen, "Nilpotent orbits in representation theory", pp. 1–211 in *Lie theory*, Progr. Math. **228**, Birkhäuser, Boston, 2004. MR

[LT18] A. J. Litterick and A. R. Thomas, "Complete reducibility in good characteristic", *Trans. Amer. Math. Soc.* **370**:8 (2018), 5279–5340. MR

[Mar02] D. Marker, *Model theory*: *an introduction*, Graduate Texts in Mathematics **217**, Springer, 2002. MR

[MST19] B. Martin, D. Stewart, and L. Topley, "A proof of the first Kac–Weisfeiler conjecture in large characteristics", *Represent. Theory* **23** (2019), 278–293. MR

[PS19] A. Premet and D. I. Stewart, "Classification of the maximal subalgebras of exceptional Lie algebras over fields of good characteristic", *J. Amer. Math. Soc.* **32**:4 (2019), 965–1008. MR

[Ric67] R. W. Richardson, Jr., "Conjugacy classes in Lie algebras and algebraic groups", *Ann. of Math.* (2) **86** (1967), 1–15. MR

[Sch00] H. Schoutens, "Uniform bounds in algebraic geometry and commutative algebra", pp. 43–93 in *Connections between model theory and algebraic and analytic geometry*, Quad. Mat. **6**, Dept. Math., Seconda Univ. Napoli, Caserta, 2000. MR

[Ste16] D. I. Stewart, "On the minimal modules for exceptional Lie algebras: Jordan blocks and stabilizers", *LMS J. Comput. Math.* **19**:1 (2016), 235–258. MR

[Tik18] A. Tikaradze, "On the local structure of quantizations in characteristic p", *Bull. Lond. Math. Soc.* **50**:1 (2018), 159–165. MR

Communicated by Anand Pillay
Received 2023-11-03 Revised 2024-12-06 Accepted 2025-06-27

b.martin@abdn.ac.uk *Department of Mathematics, University of Aberdeen, Aberdeen, United Kingdom*

david.i.stewart@manchester.ac.uk *Department of Mathematics, University of Manchester, Manchester, United Kingdom*

lt803@bath.ac.uk *Department of Mathematical Sciences, University of Bath, Bath, United Kingdom*

ALGEBRA AND NUMBER THEORY 20:6 (2026)
https://doi.org/10.2140/ant.2026.20.1185

Derived isogenies and isogenies for abelian surfaces

Zhiyuan Li and Haitao Zou

We study twisted Fourier–Mukai partners of abelian surfaces. Following Huybrechts (2019), we introduce twisted derived equivalences (also called derived isogenies) between abelian surfaces. We show that there is a twisted derived Torelli theorem for abelian surfaces over algebraically closed fields with characteristic $\neq 2, 3$.

Our approach involves extending to rational Hodge structures, ℓ-adic Tate modules and F-crystals a trick introduced by Shioda in the context of integral Hodge structures. Using this trick, we can confirm the Tate conjecture in a special case. Then we make use of Tate's isogeny theorem to give a characterization of derived isogenies between abelian surfaces via so-called principal isogenies. As a consequence, we show the two abelian surfaces are principally isogenous if and only if they are derived isogenous.

1. Introduction

1.1. ***Background.*** In the study of abelian varieties, a natural question is to classify their Fourier–Mukai partners. Thanks to Orlov and Polishchuk's *derived Torelli theorem* for abelian varieties [56; 58], there is a geometric/cohomological classification of derived equivalences between them. More generally, one can consider the notion of *twisted derived equivalence* or *derived isogeny* between abelian varieties, in the spirit of [31]:

Definition 1.1.1. Two abelian varieties X and Y are derived isogenous if they can be connected by derived equivalences between twisted abelian varieties, i.e., if there exist twisted abelian varieties (X_i, α_i) and

This project is supported by the NKRD Program of China (no. 2020YFA0713200), NSFC (nos. 12121001, 12171090 and 12425105) and Shanghai Pilot Program for Basic Research (no. 21TQ001). Z. Li is also a member of LMNS. H. Zou is also supported by the Deutsche Forschungsgemeinschaft, project ID 491392403 – TRR 358.
MSC2020: primary 14F08, 14K02; secondary 14G17.
Keywords: abelian surface, isogeny, derived category, twisted sheaf, Torelli theorem.

(X_i, β_i) such that there is a sequence of derived equivalences

$$\begin{aligned} \mathrm{D}^b(X,\alpha) &\xrightarrow{\simeq} \mathrm{D}^b(X_1,\beta_1), \\ \mathrm{D}^b(X_1,\alpha_2) &\xrightarrow{\simeq} \mathrm{D}^b(X_2,\beta_2), \\ &\vdots \\ \mathrm{D}^b(X_n,\alpha_{n+1}) &\xrightarrow{\simeq} \mathrm{D}^b(Y,\beta_n), \end{aligned} \tag{1.1.1}$$

where $\mathrm{D}^b(X,\alpha)$ is the bounded derived category of α-twisted coherent sheaves on X.

In [66] (see especially Theorem 1.2 there) Stellari proved that derived isogenous complex abelian surfaces are isogenous, using the Kuga–Satake varieties associated to their transcendental lattices. The converse is not true: there are isogenous abelian surfaces that are not derived isogenous [66, Remark 4.4(ii)]. In this paper we give a cohomological and geometric classification of derived isogenies between abelian surfaces over algebraically closed fields of arbitrary characteristic.

1.2. ***A twisted derived Torelli theorem for abelian surfaces in characteristic zero.*** We first classify derived isogenies between abelian surfaces in term of isogenies. For this purpose, we introduce a new kind of isogeny: two abelian surfaces X and Y are *principally isogenous* if there is an isogeny f from X to Y of square degree. For example, X and its dual abelian variety $\hat{X}$ are principally isogenous since any polarization $\mathcal{L}$ on X induces an isogeny $f_{\mathcal{L}} : X \to \hat{X}$ of degree $\chi(\mathcal{L})^2$.

The first main result is this:

Theorem 1.2.1. *Let X and Y be two abelian surfaces over $k = \bar{k}$ with* $\operatorname{char} k = 0$. *The following statements are equivalent.*

(i) *X and Y are derived isogenous.*

(ii) *X and Y are principally isogenous.*

A notable fact for abelian surfaces is that besides their first cohomology groups, their second cohomology groups also carry rich structures. In the untwisted case, Mukai and Orlov [49; 56] have shown that

$$\mathrm{D}^b(X) \cong \mathrm{D}^b(Y) \iff \widetilde{\mathrm{H}}(X,\mathbb{Z}) \cong_{\mathrm{Hdg}} \widetilde{\mathrm{H}}(Y,\mathbb{Z}) \iff \mathrm{T}(X) \cong_{\mathrm{Hdg}} \mathrm{T}(Y),$$

where $\widetilde{\mathrm{H}}(X,\mathbb{Z})$ and $\widetilde{\mathrm{H}}(Y,\mathbb{Z})$ are Mukai lattices, $\mathrm{T}(X) \subseteq \mathrm{H}^2(X,\mathbb{Z})$ and $\mathrm{T}(Y) \subseteq \mathrm{H}^2(Y,\mathbb{Z})$ are transcendental lattices, and $\cong_{\mathrm{Hdg}}$ stands for an integral Hodge isometry (see[12, Theorem 5.1]). The next statement can be viewed as a generalization of Mukai and Orlov's result.

Corollary 1.2.2. *Statements* (i) *and* (ii) *of Theorem 1.2.1 are also equivalent to the each of the following conditions*:

(iii) *The associated Kummer surfaces* $\mathrm{Km}(X)$ *and* $\mathrm{Km}(Y)$ *are derived isogenous.*

(iv) *The Chow motives are isomorphic as exterior algebras*: $\mathfrak{h}(X) \cong \mathfrak{h}(Y)$. *Their even degree parts are isomorphic as Frobenius algebras*: $\mathfrak{h}^{\mathrm{even}}(X) \cong \mathfrak{h}^{\mathrm{even}}(Y)$.

(v) *The even degree Chow motives are isomorphic as Frobenius algebras*: $\mathfrak{h}^{\mathrm{even}}(X) \cong \mathfrak{h}^{\mathrm{even}}(Y)$.

When $k = \mathbb{C}$, these conditions are also equivalent to each of the following:

(vi) $\mathrm{H}^2(X, \mathbb{Q}) \cong \mathrm{H}^2(Y, \mathbb{Q})$ *is a rational Hodge isometry.*

(vii) $\widetilde{\mathrm{H}}(X, \mathbb{Q}) \cong \widetilde{\mathrm{H}}(Y, \mathbb{Q})$ *is a rational Hodge isometry.*

(viii) $\mathrm{T}(X) \otimes \mathbb{Q} \cong \mathrm{T}(Y) \otimes \mathbb{Q}$ *is a rational Hodge isometry.*

Here, the motive $\mathfrak{h}(X)$ admits a canonical motivic decomposition

$$\mathfrak{h}(X) = \bigoplus_{i=0}^{4} \mathfrak{h}^i(X) \tag{1.2.1}$$

à la Deninger and Murre [18], such that $\mathrm{H}^*(\mathfrak{h}^i(X)) \cong \mathrm{H}^i(X)$ for any Weil cohomology $\mathrm{H}^*(-)$. It satisfies $\mathfrak{h}^i(X) = \bigwedge^i \mathfrak{h}^1(X)$ for all i, $\mathfrak{h}^4(X) \simeq \mathbb{1}(-4)$ and $\bigwedge^i \mathfrak{h}^1(X) = 0$ for $i > 4$ (see [37]). The motive $\mathfrak{h}(X)$ is an exterior algebra object in the category of Chow motives over k and the even-degree part

$$\mathfrak{h}^{\mathrm{even}}(X) = \bigoplus_{k=0}^{2} \textstyle\bigwedge^{2k} \mathfrak{h}^1(X) \tag{1.2.2}$$

forms a Frobenius algebra object in the sense of [23].

Equivalences (i) $\Longleftrightarrow$ (iv) $\Longleftrightarrow$ (v) are motivic realizations of derived isogenies between abelian surfaces, which can be viewed as an analogue of the motivic global Torelli theorem on K3 surfaces (compare [31, Conjecture 0.3] and [23, Theorem 1]). Equivalences (i) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (viii) can be viewed as a generalization of [66, Theorem 1.2]. The Hodge-theoretic realization (i) $\Longleftrightarrow$ (vi) follows a strategy similar to that of [31, Theorem 0.1], which makes use of Shioda's period map and the Cartan–Dieudonné decomposition of a rational isometry. Equivalences (vi) $\Longleftrightarrow$ (vii) $\Longleftrightarrow$ (viii) follow from the Witt cancellation theorem (see page 1226).

1.3. ***Shioda's trick.*** The proof of Theorem 1.2.1 involves a new ingredient, which we call the *rational Shioda's trick on abelian surfaces.* The original Shioda's trick in [63] plays a key role in the proof of Shioda's global Torelli theorem for abelian surfaces, which links the weight-1 integral Hodge structure of an abelian surface to its weight-2 integral Hodge structure. This is its generalization, proved in Section 4:

Theorem 1.3.1 (Shioda's trick). *Let X and Y be complex abelian surfaces. For any admissible Hodge isometry*

$$\psi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q}),$$

we can find an isogeny $f : Y \to X$ of degree d^2 such that $\psi = f^*/d$.

As an application, the generalized Shioda's trick gives the algebraicity of some cohomological cycles. For any integer d, one can consider a Hodge similitude of degree d,

$$\mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q}),$$

called a *Hodge isogeny of degree d*. From the Hodge conjecture on products of abelian surfaces, we know that every Hodge isogeny is algebraic. Our generalized Shioda's trick actually shows that Hodge isogenies between abelian surfaces are induced by certain isogenies. We also prove ℓ-adic and p-adic Shioda's tricks, which lead to a proof of the Tate conjecture for isometries between second cohomology groups (as either Galois modules or crystals) of abelian surfaces over finitely generated fields. See Corollary 4.6.3 for details.

1.4. *Results in positive characteristic.* The second part of this paper investigates the twisted derived Torelli theorem over positive characteristic fields. Because no satisfactory global Torelli theorem exists, one cannot follow the argument in characteristic zero directly. Instead, we need some input from p-adic Hodge theory:

Theorem 1.4.1. *Let X and Y be abelian surfaces over $k = \bar{k}$ with* $\operatorname{char} k = p > 3$. *The following statements are equivalent*:

(i′) *X and Y are prime-to-p derived isogenous.*

(ii′) *X and Y are prime-to-p principally isogenous.*

If X is supersingular, then Y is derived isogenous to X if and only if Y is supersingular.

Here, we say a derived isogeny as (1.1.1) is *prime-to-p* if its crystalline realization is integral (see Definition 3.1.1 for details), which is a somewhat technical condition. The main ingredient in the proof of Theorem 1.4.1 is the lifting-specialization technique, which works well for prime-to-p derived isogenies. Actually, our method shows that there is an implication (i′) ⇒ (ii′) for derived isogenies which are not necessarily prime-to-p (see Theorem 6.3.1). Conversely, we believe that the existence of quasiliftable isogenies will imply the existence of a derived isogeny (see Conjecture 6.3.2). The only obstruction is the existence of the specialization of non-prime-to-p derived isogenies between abelian surfaces. See Remark 6.2.2.

Another natural question is whether two abelian surfaces are derived isogenous if and only if their associated Kummer surfaces are derived isogenous over positive characteristic fields. We cannot prove this equivalence, but we provide a partial solution to the question in Theorem 6.4.1.

Similarly, one may ask whether such results also hold for K3 surfaces. Let $\mathbb{F}_q$ be a finite field, with $q = p^r$. Two K3 surfaces S and S' over $\mathbb{F}_q$ are (geometrically) isogenous in the sense of Yang [69] if there exists an algebraic correspondence Γ that induces an isometry of $\operatorname{Gal}(\bar{\mathbb{F}}_p/k)$-modules

$$\Gamma_\ell^* : \mathrm{H}^2_{\text{ét}}(S_{\bar{\mathbb{F}}_p}, \mathbb{Q}_\ell) \xrightarrow{\sim} \mathrm{H}^2_{\text{ét}}(S'_{\bar{\mathbb{F}}_p}, \mathbb{Q}_\ell)$$

for all $\ell \nmid p$ and an isometry of isocrystals

$$\Gamma_p^* : \mathrm{H}^2_{\text{crys}}(S_k/K) \xrightarrow{\sim} \mathrm{H}^2_{\text{crys}}(S'_k/K),$$

where $k/\mathbb{F}_q$ is a finite extension and K is the fraction field of the Cohen ring $W = W(k)$. More generally, we can take a finitely generated field k over $\mathbb{F}_p$ and a Cohen ring of k. Then we say that the isogeny is prime-to-p if the isometry Γ_p^* is integral, i.e., $\Gamma_p^*(\mathrm{H}^2_{\mathrm{crys}}(S_k/W)) = \mathrm{H}^2_{\mathrm{crys}}(S'_k/W)$. This leads us to a formulation of the twisted derived Torelli conjecture for K3 surfaces.

Conjecture 1.4.2. *For K3 surfaces S and S' over a finitely generated field k, the following statements are equivalent.*

(a) *There exists a derived isogeny* $\mathrm{D}^b(S) \sim \mathrm{D}^b(S')$.

(b) *There exists an isogeny between S and S'.*

The implication (a) $\Longrightarrow$ (b) is clear, while the converse remains open if $\operatorname{char} k > 0$. In the case of Kummer surfaces, our results provide some evidence of Conjecture 1.4.2. We mention that Bragg and Yang have studied derived isogenies between K3 surfaces over positive characteristic fields, proving a weaker version of Conjecture 1.4.2 (see [10, Theorem 1.2]).

Outline. The next two sections review some known constructions and facts; specifically, Section 2 contains computations for the Brauer group of abelian surfaces using the Kummer construction. This will allow us to prove the lifting lemma for twisted abelian surfaces of finite height.

In Section 3, we collect knowledge on derived isogenies between abelian surfaces and their cohomological realizations, which include the motivic realization, $\boldsymbol{B}$-field theory, twisted Mukai lattices, and a filtered Torelli theorem and its relation to the moduli space of twisted sheaves. At the end of the section, we follow Bragg and Lieblich's twistor line argument in [8] to conclude the supersingular case of Theorem 1.4.1.

In Section 4, we review Shioda's work and extend it to rational Hodge isogenies. This is the key ingredient for proving Theorem 1.2.1. Then, after introducing admissible ℓ-adic and p-adic bases, we prove the ℓ-adic and p-adic Shioda's tricks for admissible isometries on abelian surfaces. In an application, we prove the algebraicity of these isometries on abelian surfaces over finitely generated fields.

Sections 5 and 6 are devoted to proving Theorems 1.2.1 and 1.4.1, the first of which is restated as (essentially) Theorems 5.1.3 and 5.3.4. The proof of Theorem 1.4.1 is much more subtle. We establish the lifting and specialization theorem for prime-to-p derived isogeny. Then we can conclude (i′) $\Longleftrightarrow$ (ii′) from Theorem 1.2.1 for abelian surfaces of finite heights.

Notation and conventions.

(1) Throughout, k will denote a field. If k is a perfect field and $\operatorname{char} k = p > 0$, we write $W := W(k)$ for the ring of Witt vectors in k, which is equipped with a morphism $\sigma : W \to W$ induced by the Frobenius map on k. If k is not perfect, we consider the Cohen ring W with $W/pW = k$. Inside the ring of Witt vectors in a fixed algebraic closure $\bar{k}$ of k, we get a fixed Frobenius lift $\sigma : W \to W$ of k.

(2) Let X be a smooth projective variety over k. We denote by $\mathrm{H}^\bullet_{\text{ét}}(X_{\bar{k}}, \mathbb{Z}_\ell)$ the ℓ-adic étale cohomology group of $X_{\bar{k}}$. The $\mathbb{Z}_\ell$-module $\mathrm{H}^\bullet_{\text{ét}}(X_{\bar{k}}, \mathbb{Z}_\ell)$ is endowed with a canonical $G_k = \mathrm{Gal}(\bar{k}/k)$-action. We use $\mathrm{H}^i_{\mathrm{crys}}(X/W)$ to denote the i-th crystalline cohomology group of X over the p-adic base $W \twoheadrightarrow k$, which is a W-module.

(3) For any abelian group G and integer n, we denote by $G[n]$ the n-torsion subgroup of G and by $G\{n\}$ the union of all n-power torsion elements. For a lattice L in $\mathbb{Z}$ or $\mathbb{Q}$ and an integer n, we use $L(n)$ for the lattice twisted by n, that is, $L = L(n)$ as a $\mathbb{Z}$ or $\mathbb{Q}$-module, but with

$$\langle x, y\rangle_{L(n)} = n\langle x, y\rangle_L.$$

The reader should not confuse this with the Tate twist.

(4) Let X and Y be abelian surfaces. Here is a list of the various notions of isogeny between X and Y.

- An *isogeny* is a surjective homomorphism $X \to Y$ with finite kernel.
- A *quasi-isogeny* is a $\mathbb{Q}$-isogeny.
- A *prime-to-ℓ quasi-isogeny* is a $\mathbb{Z}_{(\ell)}$-isogeny.
- A *principal quasi-isogeny* is a quasi-isogeny whose degree is a square.
- A *derived isogeny* is a chain of twisted derived equivalences from X to Y.
- A *prime-to-ℓ derived isogeny* is a derived isogeny whose cohomological realization is prime-to-ℓ.

2. Twisted abelian surfaces

In this section, we give some preliminary results in the theory of twisted abelian surfaces, especially in positive characteristic. Many of them are well-known to experts.

2.1. ***Gerbes on abelian surfaces.*** Let X be a smooth projective variety over a field k and let $\mathscr{X} \to X$ be a μ_n-gerbe over X. This corresponds to a pair (X, α) for some $\alpha \in \mathrm{H}^2_{\mathrm{fl}}(X, \mu_n)$, where the cohomology group is with respect to the fppf topology. Since μ_n is commutative, there is a bijection of sets

$$\mathrm{H}^2_{\mathrm{fl}}(X, \mu_n) \xrightarrow{\sim} \{\mu_n\text{-gerbes on } X\}/\simeq,$$

where $\simeq$ is the μ_n-equivalence defined as in [25, IV.3.1.1]. We may write $\alpha = [\mathscr{X}]$. For any integer m, let $\mathscr{X}^{(m)}$ be the gerbe corresponding to the cohomological class $m[\mathscr{X}] \in \mathrm{H}^2_{\mathrm{fl}}(X, \mu_n)$.

The Kummer exact sequence induces a surjective map

$$\mathrm{H}^2_{\mathrm{fl}}(X, \mu_n) \to \mathrm{Br}(X)[n], \tag{2.1.1}$$

where the right-hand side is the *cohomological Brauer group* $\mathrm{Br}(X) := \mathrm{H}^2_{\text{ét}}(X, \mathbb{G}_m)$. There is an associated $\mathbb{G}_m$-gerbe on X via the map (2.1.1), denoted by $\mathscr{X}_{\mathbb{G}_m}$. Let $[\mathscr{X}_{\mathbb{G}_m}]$ denote the corresponding class in $\mathrm{Br}(X)[n]$. If $[\mathscr{X}_{\mathbb{G}_m}] = 0$, we will call $\mathscr{X}$ an *essentially trivial* μ_n-gerbe.

Following [39, §2], one can define twisted coherent sheaves and their twisted derived category in terms of gerbes.

Definition 2.1.1. Let $\mathscr{X} \to X$ be a μ_n-gerbe or $\mathbb{G}_m$-gerbe over X. Let $\mathtt{Coh}^{(m)}(\mathscr{X})$ be the abelian category of *$\mathscr{X}^{(m)}$-twisted coherent sheaves*, consisting of m-fold coherent sheaves on the stack $\mathscr{X}$. We define $\mathrm{D}^{(m)}(\mathscr{X})$ as the bounded derived category of $\mathtt{Coh}^{(m)}(\mathscr{X})$.

As shown in [39, Propositions 2.1.2.6 and 2.1.3.3], there are natural equivalences

$$\mathtt{Coh}^{(1)}(\mathscr{X}) \simeq \mathtt{Coh}^{(1)}(\mathscr{X}_{\mathbb{G}_m}) \simeq \mathtt{Coh}(X, [\mathscr{X}_{\mathbb{G}_m}]),$$

the last of which is the abelian category of twisted sheaves defined by Căldăraru [14]. Throughout this paper, we mainly use Lieblich's terminology.

For two G-gerbes $\mathscr{X} \to X$ and $\mathscr{Y} \to Y$, we denote by $\mathscr{X} \wedge_{i,j} \mathscr{Y}$ the G-gerbe on $X \times Y$ given by the image of $G \times G$-gerbe $\mathscr{X} \times \mathscr{Y}$ under the map

$$\mathrm{H}^2_{\mathrm{fl}}(X \times Y, G \times G) \to \mathrm{H}^2_{\mathrm{fl}}(X \times Y, G)$$

induced by the multiplication $G \times G \to G$, $(g_1, g_2) \mapsto (g_1^i g_2^j)$. There is an equivalence

$$\mathtt{Coh}^{(1)}(\mathscr{X} \wedge_{i,j} \mathscr{Y}) \xrightarrow{\sim} \mathtt{Coh}^{(i,j)}(\mathscr{X} \times \mathscr{Y}),$$

where the right side is the subcategory of (i, j)-fold coherent sheaves on $\mathscr{X} \times \mathscr{Y}$ [28, Corollary 2.3.2]. When $i = j = 1$, we simply write $\mathscr{X} \wedge \mathscr{Y}$ for $\mathscr{X} \wedge_{1,1} \mathscr{Y}$.

A *derived equivalence* means a k-linear exact equivalence between triangulated categories in the form

$$\Phi : \mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y}).$$

If Φ is of the form

$$\Phi^{\mathcal{P}}(\mathcal{E}) = \boldsymbol{R}q_*(p^*\mathcal{E} \otimes \mathcal{P}),$$

we call it a Fourier–Mukai transform with a kernel $\mathcal{P} \in \mathrm{D}^{(-1,1)}(\mathscr{X} \times \mathscr{Y})$ and projections $p : \mathscr{X} \times \mathscr{Y} \to \mathscr{X}$, $q : \mathscr{X} \times \mathscr{Y} \to \mathscr{Y}$, and $\mathscr{X}$, $\mathscr{Y}$ are called a pair of Fourier–Mukai partners. If these gerbes are (essentially) trivial, then by Orlov's result, any k-linear exact equivalence between bounded derived categories of smooth projective varieties is of this form.

Similarly to Orlov's theorem, Canonaco and Stellari showed that any twisted derived equivalence is also of Fourier–Mukai type:

Proposition 2.1.2 [15]. *Any derived equivalence* $\mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y})$ *can be written uniquely* (*up to isomorphism*) *as a Fourier–Mukai transform*

$$\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y}),$$

whose kernel $\mathcal{P}$ *is a perfect complex in* $\mathrm{D}^{(-1,1)}(\mathscr{X} \times \mathscr{Y})$.

2.2. *Kummer construction.* If k has characteristic $p \neq 2$, there is an associated Kummer surface $\mathrm{Km}(X)$ constructed as follows:

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{\tilde{\sigma}} & X \\ \downarrow{\scriptstyle \pi} & & \downarrow \\ \mathrm{Km}(X) & \xrightarrow{\sigma} & X/\iota \end{array} \tag{2.2.1}$$

where ι is the involution of X given by sending x to $-x$, σ is the crepant resolution of quotient singularities, $\tilde{\sigma}$ is the blow-up of X along the closed subscheme $X[2] \subset X$. Its birational inverse is denoted by $\tilde{\sigma}^{-1}$.

Let $E \subset \tilde{X}$ be the exceptional locus of $\tilde{\sigma}$. For a classical cohomology theory $\mathrm{H}^\bullet(-)$ (such as Betti, étale and crystalline) with coefficients in R, if 2 is invertible in R, we have a canonical decomposition

$$\mathrm{H}^2(\mathrm{Km}(X)) \cong \mathrm{H}^2(X) \oplus \pi_* \Sigma_X, \tag{2.2.2}$$

where Σ_X is the summand in $\mathrm{H}^2(\tilde{X})$ generated by irreducible components of E.

Moreover, we have a composition of the sequence of morphisms

$$(\tilde{\sigma}^{-1})^* : \mathrm{Br}(\tilde{X}) \to \mathrm{Br}(\tilde{X} \setminus E) \cong \mathrm{Br}(X \setminus X[2]) \cong \mathrm{Br}(X).$$

Here, the last isomorphism $\mathrm{Br}(X) \to \mathrm{Br}(X \setminus X[2])$ is due to Grothendieck's purity theorem (see [26; 16]).

Proposition 2.2.1. *When $k = \bar{k}$ and $p \neq 2$, the $(\tilde{\sigma}^{-1})^* \pi^*$ induces an isomorphism between cohomological Brauer groups*

$$\Theta : \mathrm{Br}(\mathrm{Km}(X)) \to \mathrm{Br}(X). \tag{2.2.3}$$

In particular, when X is supersingular over $\bar{k}$, then $\mathrm{Br}(X)$ *is isomorphic to the additive group $\bar{k}$.*

Proof. For torsions of (2.2.3) whose orders are coprime to p, the proof is essentially the same as that of [65, Proposition 1.3], by the Hochschild–Serre spectral sequence and the fact that $\mathrm{H}^2(\mathbb{Z}/2\mathbb{Z}, k^*) = 0$ as the characteristic p is not 2 [68, Proposition 6.1.10]. See also [66, Lemma 4.1] for the case $k = \mathbb{C}$. For p-primary torsion part, we have

$$\mathrm{Br}(\mathrm{Km}(X))\{p\} \cong \mathrm{Br}(X)^\iota\{p\}$$

from the Hochschild–Serre spectral sequence, where $\mathrm{Br}(X)^\iota$ is the ι-invariant subgroup. Hence, it suffices to prove that ι acts trivially on $\mathrm{Br}(X)$. This is well-known to experts and works for any abelian varieties over an algebraically closed field. (See the proof of [53, Lemma 8.1], for example.)

In fact, $\mathrm{H}^2_{\mathrm{fl}}(X, \mu_p)$ can be ι-equivariantly embedded to $\mathrm{H}^2_{\mathrm{dR}}(X/k)$ by de Rham–Witt theory (see [51, Proposition 1.2]). The action of ι on $\mathrm{H}^2_{\mathrm{dR}}(X/k) = \bigwedge^2 \mathrm{H}^1_{\mathrm{dR}}(X/k)$ is the identity, as its action on $\mathrm{H}^1_{\mathrm{dR}}(X/k)$ is given by $x \mapsto -x$. Thus the involution on $\mathrm{H}^2_{\mathrm{fl}}(X, \mu_p)$ is trivial. Then by the exact sequence

$$0 \to \mathrm{NS}(X) \otimes \mathbb{Z}/p \to \mathrm{H}^2_{\mathrm{fl}}(X, \mu_p) \to \mathrm{Br}(X)[p] \to 0,$$

we can deduce that $\mathrm{Br}(X)[p]$ is invariant under the involution. For p^n-torsions with $n \geq 2$, we can proceed by induction on n. Assume that all elements in $\mathrm{Br}(X)[p^d]$ are ι-invariant if $1 \leq d < n$. By abuse of notation, we still use ι to denote the induced map $\mathrm{Br}(X) \to \mathrm{Br}(X)$. For $\alpha \in \mathrm{Br}(X)[p^n]$, $p\alpha \in \mathrm{Br}(X)[p^{n-1}]$ is ι-invariant. This gives

$$p\alpha = \iota(p\alpha) = p\iota(\alpha),$$

which implies $\alpha - \iota(\alpha) \in \mathrm{Br}(X)[p]$. Applying ι to $\alpha - \iota(\alpha)$, we obtain

$$\alpha - \iota(\alpha) = \iota(\alpha) - \alpha.$$

This implies that $\alpha - \iota(\alpha)$ is also a 2-torsion element. Since p is coprime to 2, we conclude that $\alpha = \iota(\alpha)$.

If X is supersingular, then $\mathrm{Km}(X)$ is also supersingular. By [2], the Brauer group of a supersingular K3 surface is isomorphic to k. Thus, $\mathrm{Br}(X) \cong k$. □

Remark 2.2.2. In the case where X is supersingular, the method of [2] cannot be applied directly to show that $\mathrm{Br}(X) = k$ as $\mathrm{H}^1_{\mathrm{fl}}(X, \mu_{p^n})$ is not trivial in general for an abelian surface X.

2.3. ***A lifting lemma.*** In [7], Bragg showed that a twisted K3 surface can be lifted to characteristic 0. Though his method cannot be applied directly to twisted abelian surfaces, one can still obtain a lifting result for twisted abelian surfaces via the Kummer construction. The following result will be used frequently in this paper.

Lemma 2.3.1. *Let $\mathscr{X}_0 \to X_0$ be a $\mathbb{G}_m$-gerbe on an abelian surface X_0 over $k = \bar{k}$. Suppose* char $k > 2$ *and X has finite height. Then there exists a complete discrete valuation ring V with residue field k, fraction field K and the following properties*:

- *There is a smooth projective abelian scheme $\mathscr{X}_V \to X_V$ over* Spec V *whose special fiber is isomorphic to $\mathscr{X}_0 \to X_0$.*
- *There is a sequence of isomorphisms*

$$\mathrm{NS}(X_{\overline{K}}) \xleftarrow{\ \sim\ } \mathrm{NS}(X_V) \xrightarrow{\ \sim\ } \mathrm{NS}(X_0).$$

Here $\mathrm{NS}(X_V)$ *is the group or Cartier divisors on X_V modulo numerical equivalence over V, and the morphisms are given by pullback.*

Proof. The existence of such a lifting is ensured by [7, Theorem 7.10; 38, Lemma 3.9] and Proposition 2.2.1. Generally speaking, let $\mathscr{S}_0 \to \mathrm{Km}(X_0)$ be the associated twisted Kummer surface via the isomorphism (2.2.3). Then [7, Theorem 7.10] (by taking $\mathrm{Pic}(\mathrm{Km}(X_0))$ as a saturated sublattice of itself) asserts that there exists some discrete valuation ring V and a projective family of K3 surfaces

$$\begin{array}{ccc} \mathscr{S}_V & \longrightarrow & S_V \\ & \searrow & \downarrow \\ & & \mathrm{Spec}\, V \end{array}$$

such that the special fiber is $\mathscr{S}_0 \to \mathrm{Km}(X_0)$ and the specialization map $\mathrm{NS}(S_{\overline{K}}) \to \mathrm{NS}(\mathrm{Km}(X_0))$ of Néron–Severi lattices is an isomorphism, where $K = \mathrm{Frac}(V)$. Now we can apply [38, Lemma 3.9] to get a lifting $X_V \to \mathrm{Spec}\, V$ of X such that $\mathrm{Km}(X_V) \cong S_V$ over Spec V.
We have an isomorphism $\mathrm{NS}(X_V) \cong \mathrm{NS}(X_K)$ since X_V is regular. Consider the commutative diagram

$$\mathrm{NS}(X_{\overline{K}}) \longleftarrow \mathrm{NS}(X_K) \cong \mathrm{NS}(X_V) \longrightarrow \mathrm{NS}(X_0), \qquad \mathrm{sp}: \mathrm{NS}(X_{\overline{K}}) \to \mathrm{NS}(X_0) \tag{2.3.1}$$

(see [44, Proposition 3.3] and its proof). The morphism $\mathrm{NS}(X_V) \to \mathrm{NS}(X_0)$ is injective by Proposition 3.6 of [44] since $\mathrm{NS}(X_K)$ is torsion-free. The morphism $\mathrm{NS}(X_K) \to \mathrm{NS}(X_{\bar{K}})$ is a primitive embedding since $\mathrm{Br}(V) = 0$. Thus, it is sufficient to see that the specialization map sp is an isomorphism. The relative Kummer construction $\mathrm{Km}(X_V) \cong S_V$ canonically identifies $\mathrm{NS}(X_K)$ (resp. $\mathrm{NS}(X_0)$) with a sublattice of $\mathrm{NS}(S_{\bar{K}})$ (resp. $\mathrm{NS}(\mathrm{Km}(X_0))$) after division by 2 (see [51, Lemma 7.11] or [64, Proposition 3.1]). The identification is compatible under specialization. We conclude using the isomorphism $\mathrm{NS}(S_{\bar{K}}) \cong \mathrm{NS}(\mathrm{Km}(X_0))$.

Lifting the $\mathbb{G}_m$-gerbe $\mathscr{X}_0 \to X_0$ to Spec V is equivalent to finding find a Brauer class in $\mathrm{Br}(X_V)$ whose restriction to X_0 is $[\mathscr{X}_0]$. Analogously to the proof of Proposition 2.2.1, there is a canonical map between the cohomological Brauer groups

$$\Theta = (\tilde{\sigma}^{-1})^* \pi^* : \mathrm{Br}(\mathrm{Km}(X_V)) \to \mathrm{Br}(X_V)$$

as in (2.2.3). The image $\Theta([\mathscr{S}_V]) \in \mathrm{Br}(X_V)$ is the desired lifting of $[\mathscr{X}_0]$. □

2.4. ***Flat cohomology of abelian surfaces.*** Finally, we consider the representability of the flat cohomology of abelian surfaces. Let $f : X \to S$ be a flat and proper morphism of algebraic spaces of finite type over k. Consider the sheaf of the abelian groups $\mathrm{R}^i f_* \mu_p$ on the big fppf site $(\mathtt{Sch}/S)_{\mathrm{fl}}$, which can be expressed as the fppf sheafification of

$$S' \mapsto \mathrm{H}^i_{\mathrm{fl}}(X_{S'}, \mu_p)$$

for any S-scheme S'. The representability of $\mathrm{R}^i f_* \mu_p$ is difficult to determine due to the complexity of flat cohomology with p-torsion coefficients. We will prove it for abelian surfaces.

Proposition 2.4.1. *Let $f : X \to S$ be an abelian S-scheme of relative dimension* 2*. Then $\mathrm{R}^1 f_* \mu_p \cong \hat{X}[p]$ is a finite flat S-group scheme.*

Proof. It suffices to check the statement affine locally on the base. Assume S is an affine scheme of finite type over k. Taking the Stein factorization, we can further assume $f_* \mathcal{O}_X \cong \mathcal{O}_S$. Then $f_* \mu_p \cong \mu_p$ also holds universally. Under this assumption, we have an exact sequence of fppf-sheaves by Kummer theory:

$$0 \to \mathrm{R}^1 f_* \mu_p \to \mathrm{R}^1 f_* \mathbb{G}_m \to \mathrm{R}^1 f_* \mathbb{G}_m. \tag{2.4.1}$$

Since $\mathrm{R}^1 f_* \mathbb{G}_m$ computes the relative Picard scheme $\mathrm{Pic}_{X/S}$ and the Néron–Severi group of X is torsion-free, we get

$$\mathrm{R}^1 f_* \mu_p \cong \ker\big(\mathrm{Pic}_{X/S} \xrightarrow{\cdot p} \mathrm{Pic}_{X/S}\big) \cong \ker\big(\mathrm{Pic}^0_{X/S} \xrightarrow{\cdot p} \mathrm{Pic}^0_{X/S}\big).$$

But $\mathrm{Pic}^0_{X/S}$ is representable by the dual abelian S-scheme $\hat{X}$, by [50, Corollary 6.8]. Thus, $\mathrm{R}^1 f_* \mu_p$ is representable by the commutative finite group S-scheme $\hat{X}[p]$. □

Proposition 2.4.2. *Let $f : \mathcal{X} \to S$ be a proper smooth family of abelian surfaces over an algebraic space S. Then $\mathrm{R}^2 f_* \mu_p$ is representable by an algebraic space, which is separated and locally of finite presentation over S.*

Proof. This follows from [9, Corollary 1.11 and Example 11.5], because $\mathrm{R}^1 f_* \mu_p$ is representable by Proposition 2.4.1. □

Remark 2.4.3. The case in which $X \to S = \mathrm{Spec}(k)$ is a smooth surface for some field k is claimed by Artin in [2, Theorem 3.1] without proof. Bragg and Olsson provided a proof (Corollary 1.6 in [9]). For relative K3 surfaces, there is a moduli-theoretic proof given by Bragg and Lieblich using the stack of Azumaya algebras (see [8, Theorem 2.1.6]). Their proof cannot be used directly for relative abelian surfaces as the essential assumption $\mathrm{R}^1 f_* \mu_p = 0$ fails in the fppf site $(\mathtt{Sch}/S)_{\mathrm{fl}}$.

Remark 2.4.4. An alternative proof for Proposition 2.4.2 consists in applying Artin's representability criterion [1, Theorem 5.3]. The most technical part is to see the separatedness.

The following observation is essential in the construction of the twistor space of supersingular abelian or K3 surfaces.

Corollary 2.4.5 [8, Proposition 2.2.4]. *Suppose that each geometric fiber of $f : \mathcal{X} \to S$ is supersingular. The connected components of any geometric fiber of $\mathrm{R}^2 f_* \mu_p \to S$ are isomorphic to the additive group scheme $\mathbb{G}_a$.*

Proof. The completion of each geometric fiber of $\mathrm{R}^2 f_* \mu_p$ at $\bar{s} \in S$, along the identity section, is isomorphic to the formal Brauer group $\widehat{\mathrm{Br}}_{X_{\bar{s}}/k(\bar{s})}$, which is isomorphic to $\widehat{\mathbb{G}}_a$. The only smooth connected p-torsion group scheme at $k(\bar{s})$ with this property is $\mathbb{G}_a$. □

3. Cohomological realizations of derived isogeny

We next provide a summary of the derived isogenies on the cohomology groups of abelian surfaces and introduce the notion of prime-to-ℓ derived isogenies. This action can be described in two ways: via

(1) the motivic realization, which provides rational isomorphisms on the cohomology groups; or

(2) the realization on the integral twisted Mukai lattices.

Following [27; 40], we then extend the filtered Torelli theorem to twisted abelian surfaces over an algebraically closed field k with $\mathrm{char}\, k \neq 2$. As a corollary, we show in Theorem 3.5.3 that any Fourier–Mukai partner of a twisted abelian surface is isomorphic to a moduli space of stable twisted sheaves.

3.1. ***Motivic realization of derived isogeny on cohomology groups.*** It is known that (twisted) derived equivalent smooth projective surfaces over a field k have isomorphic Chow motives (see [30, §2.4] and [23, §1.2], for example). We record these results for convenience, focusing on abelian surfaces over k for concreteness.

For any algebraic surface X over a field k, one may consider idempotent correspondences $\pi^2_{\mathrm{alg},X}$ and $\pi^2_{\mathrm{tr},X}$ in $\mathrm{CH}^2(X \times X)_{\mathbb{Q}}$ defined as

$$\pi^2_{\mathrm{alg},X} := \sum_{i=1}^{\rho} \frac{1}{\deg(E_i \cdot E_i)} E_i \times E_i, \quad \pi^2_{\mathrm{tr},X} = \pi^2_X - \pi^2_{\mathrm{alg},X},$$

where π_X^2 is the idempotent correspondence given by the Chow–Künneth decomposition (1.2.1) and the E_i are divisors generating the Néron–Severi group $\mathrm{NS}(X_{k^s})$ such that $E_i \cdot E_i \neq 0$ and $E_i \cdot E_j = 0$ for any $i \neq j$. Consider the decomposition

$$\mathfrak{h}^2(X) = \mathfrak{h}^2_{\mathrm{alg}}(X) \oplus \mathfrak{h}^2_{\mathrm{tr}}(X)$$

given by $\pi^2_{\mathrm{alg},X}$ and $\pi^2_{\mathrm{tr},X}$. It is not hard to see that $\mathfrak{h}^2_{\mathrm{alg}}(X)$ is a Tate motive after base change to the separable closure k^s, whose Chow realization is

$$\mathrm{CH}^*_{\mathbb{Q}}(\mathfrak{h}^2_{\mathrm{alg}}(X_{k^s})) \cong \mathrm{NS}(X_{k^s})_{\mathbb{Q}}.$$

Let $\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y})$ be a derived equivalence between two twisted abelian surfaces over k. Consider the cycle class

$$\mathrm{ch}_{\mathscr{X}^{(-1)} \wedge \mathscr{Y}}(\mathcal{P}) \cdot \sqrt{\mathrm{td}_{X \times Y}} = \mathrm{ch}_{\mathscr{X}^{(-1)} \wedge \mathscr{Y}}(\mathcal{P}) \in \mathrm{CH}^*(X \times Y)_{\mathbb{Q}}. \tag{3.1.1}$$

Here $\mathrm{ch}_{\mathscr{X}^{(-1)} \wedge \mathscr{Y}}(-)$ is the twisted Chern character defined as in (3.3.2); this provides an isomorphism

$$\mathfrak{h}(X) \xrightarrow{\sim} \mathfrak{h}(Y),$$

which preserves the even-degree parts:

$$\mathfrak{h}^{\mathrm{even}}(-) := \bigoplus_{k=0}^{2} \mathfrak{h}^{2k}(-) \cong \bigoplus_{k=0}^{2} \textstyle\bigwedge^{2k} \mathfrak{h}^1(-)$$

(cf. [23, §§1.2.3]). For a Weil cohomology theory H, its cohomological realization

$$\mathrm{H}^{\mathrm{even}}(X) \xrightarrow{\sim} \mathrm{H}^{\mathrm{even}}(Y) \tag{3.1.2}$$

preserves the Mukai pairing. The cohomological realization (3.1.2) is not integral in general. We can introduce the prime-to-ℓ derived isogeny via integral cohomological realizations, which will be used in the rest of the paper.

Definition 3.1.1. Let ℓ be a prime and $\operatorname{char} k = p$. When $\ell \neq p$, the derived isogeny $\mathrm{D}^b(X) \sim \mathrm{D}^b(Y)$ given by (1.1.1) is called *prime-to-ℓ* if each cohomological realization in the sequence

$$\tilde{\varphi}_\ell : \mathrm{H}^{\mathrm{even}}_{\mathrm{\acute{e}t}}(X_{i-1,\bar{k}}, \mathbb{Q}_\ell) \xrightarrow{\sim} \mathrm{H}^{\mathrm{even}}_{\mathrm{\acute{e}t}}(X_{i,\bar{k}}, \mathbb{Q}_\ell)$$

is integral, i.e., $\tilde{\varphi}_\ell\left(\mathrm{H}^{\mathrm{even}}_{\mathrm{\acute{e}t}}(X_{\bar{k}}, \mathbb{Z}_\ell)\right) = \mathrm{H}^{\mathrm{even}}_{\mathrm{\acute{e}t}}(Y_{\bar{k}}, \mathbb{Z}_\ell)$. In the case $\ell = p$, it is called *prime-to-p* if each $\tilde{\varphi}_p : \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(X_{i-1}/K) \xrightarrow{\sim} \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(X_i/K)$ is integral.

Remark 3.1.2. The correspondence (3.1.1) does not necessarily preservecohomological degrees. However, it admits a modification that is an isomorphism between degree-two parts: indeed, the cycle class $[\Gamma_{\mathrm{tr}}] \in \mathrm{CH}^2(X \times Y)_{\mathbb{Q}}$ given by the codimension-two component of (3.1.1) induces an isomorphism of transcendental motives by a weight argument

$$[\Gamma_{\mathrm{tr}}]_2 := \pi^2_{\mathrm{tr},Y} \circ [\Gamma_{\mathrm{tr}}] \circ \pi^2_{\mathrm{tr},X} : \mathfrak{h}^2_{\mathrm{tr}}(X) \xrightarrow{\sim} \mathfrak{h}^2_{\mathrm{tr}}(Y).$$

It extends to an isomorphism $\mathfrak{h}^2(X) \xrightarrow{\sim} \mathfrak{h}^2(Y)$ since their algebraic parts are abstractly isomorphic, as X and Y have the same Picard number. This supports the implication (v) $\Longrightarrow$ (vii) in Corollary 1.2.2.

3.2. *Mukai lattices and B-fields.* Let k be an algebraically closed field with $\operatorname{char} k \neq 2$. Let X be an abelian surface over k. When $k = \mathbb{C}$, the *Mukai lattice* of X is defined as

$$\widetilde{\mathrm{H}}(X, \mathbb{Z}) := \mathrm{H}^0(X, \mathbb{Z}(-1)) \oplus \mathrm{H}^2(X, \mathbb{Z}) \oplus \mathrm{H}^4(X, \mathbb{Z}(1))$$

with the Mukai pairing

$$\langle (r_1, b_1, s_1), (r_2, b_2, s_2) \rangle := b_1 \cdot b_2 - r_1 s_2 - r_2 s_1 \tag{3.2.1}$$

and a pure $\mathbb{Z}$-Hodge structure of weight 2. In general, we have the following notion of Mukai lattices, taken from [40, §2]. (The definition there is for K3 surfaces, but in fact it works well for any smooth surface with trivial canonical bundle.)

- Let $\widetilde{\mathrm{N}}(X)$ be the *extended Néron–Severi lattice*, defined as $\widetilde{\mathrm{N}}(X) := \mathbb{Z} \oplus \mathrm{NS}(X) \oplus \mathbb{Z}$, with Mukai pairing
$$\langle (r_1, c_1, s_1), (r_2, c_2, s_2) \rangle = c_1 \cdot c_2 - r_1 s_2 - r_2 s_1.$$
The Chow realization of
$$\mathfrak{h}^0(X)(-1) \oplus \mathfrak{h}^2_{\mathrm{alg}}(X) \oplus \mathfrak{h}^4(X)(1)$$
can be identified with $\widetilde{\mathrm{N}}(X)_{\mathbb{Q}}$.
- If $\operatorname{char} k = 0$ or if $\operatorname{char} k = p > 0$ and $\ell \neq \operatorname{char} k$ is a prime, the ℓ-adic Mukai lattice is defined on the even degrees of the integral ℓ-adic cohomology of X by
$$\mathrm{H}^0_{\text{ét}}(X, \mathbb{Z}_\ell(-1)) \oplus \mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell) \oplus \mathrm{H}^4_{\text{ét}}(X, \mathbb{Z}_\ell(1)),$$
with $\widetilde{\mathrm{H}}(X, \mathbb{Z}_\ell)$ a Mukai pairing defined analogously to (3.2.1).
- If $\operatorname{char} k = p > 0$, the p-adic Mukai lattice $\widetilde{\mathrm{H}}(X, W)$ is defined on the even degrees of the crystalline cohomology of X with coefficients in $W(k)$ by
$$\mathrm{H}^0_{\mathrm{crys}}(X/W(k))(-1) \oplus \mathrm{H}^2_{\mathrm{crys}}(X/W(k)) \oplus \mathrm{H}^4_{\mathrm{crys}}(X/W(k))(1),$$
where the twist (i) is given by replacing the Frobenius by $F \mapsto p^{-i} F$, and the Mukai pairing is given similarly to formula (3.2.1).

Hodge B-fields. Assume $k = \mathbb{C}$. For any $\mathbb{G}_m$-gerbe $\mathscr{X} \to X$, one can find a lift $B \in \mathrm{H}^2(X, \mathbb{Q})$ of $[\mathscr{X}] \in \mathrm{Br}(X)$ from the exponential sequence. Such a B is called a ***B***-field lift of α. We define the *twisted Mukai lattice* of $\mathscr{X}$ as

$$\widetilde{\mathrm{H}}(X, \mathbb{Z}; B) := \exp(B) \cdot \widetilde{\mathrm{H}}(X, \mathbb{Z}) \subset \widetilde{\mathrm{H}}(X, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q},$$

which is isomorphic to $\widetilde{\mathrm{H}}(X, \mathbb{Z})$. For simplicity of notation, we still use (r, c, s) to denote the vector $\exp(B)(r, c, s)$. There is an induced pure Hodge structure of weight 2 on $\widetilde{\mathrm{H}}(X, \mathbb{Z}; B)$ given by

$$\widetilde{\mathrm{H}}^{0,2}(X; B) = \exp(B)\widetilde{\mathrm{H}}^{0,2}(X),$$

(see [32, Definition 2.3]). It is clear that a different choice of lift B' satisfies $B - B' \in \mathrm{H}^2(X, \mathbb{Z})$, and thus there is a Hodge isometry

$$\exp(B - B') : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B') \xrightarrow{\sim} \widetilde{\mathrm{H}}(X, \mathbb{Z}; B).$$

This means that, up to isomorphisms, $\widetilde{\mathrm{H}}(X, \mathbb{Z}; B)$ is independent of the choice of the $\boldsymbol{B}$-field lifting and can also be denoted by $\widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z})$.

As shown in [71, Corollary 4.4], for any derived equivalence $\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y})$ between two twisted abelian surfaces, the Fourier–Mukai kernel induces a Hodge isometry

$$\tilde{\varphi} = \varphi_{B,B'} : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B) \xrightarrow{\sim} \widetilde{\mathrm{H}}(Y, \mathbb{Z}; B') \tag{3.2.2}$$

for suitable $\boldsymbol{B}$-field lifts B, B'. It provides the cohomological realization as in (3.1.2) rationally.

ℓ-adic and crystalline B-fields. Let us recall the generalized notions of $\boldsymbol{B}$-fields in ℓ-adic cohomology [42, §3.2] and crystalline cohomology [6, §3], as analogues in Betti cohomology. Full considerations for the cases ℓ-adic and p-adic are given in [10, §2], and are applicable to both K3 and abelian surfaces. Therefore, we omit some technical details here.

For a prime $\ell \neq p$ and $n \in \mathbb{N}$, the Kummer sequence of étale sheaves

$$1 \to \mu_{\ell^n} \to \mathbb{G}_m \xrightarrow{(\cdot)^{\ell^n}} \mathbb{G}_m \to 1 \tag{3.2.3}$$

induces a long exact sequence

$$\cdots \to \mathrm{Pic}(X) \xrightarrow{\cdot l^n} \mathrm{Pic} X \to \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n}) \to \mathrm{Br}(X)[\ell^n] \to 0.$$

Taking the inverse limit over n, we get a map

$$\pi_\ell : \mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1)) = \varprojlim_n \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n}) \to \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n}) \twoheadrightarrow \mathrm{Br}(X)[\ell^n].$$

Lemma 3.2.1. *The map π_ℓ is surjective.*

Proof. By [45, Chapter V, Lemma 1.11], we have a short exact sequence

$$0 \to \mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1))/\ell^n \to \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n}) \to \mathrm{H}^3_{\text{ét}}(X, \mathbb{Z}_\ell(1))[\ell^n] \to 0.$$

Since $\mathrm{H}^3_{\text{ét}}(X, \mathbb{Z}_\ell(1))$ is torsion-free for any abelian surface X, we have an isomorphism

$$\mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1))/\ell^n \cong \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n}).$$

Therefore, the reduction morphism $\mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1)) \to \mathrm{H}^2_{\text{ét}}(X, \mu_{\ell^n})$ can be identified with

$$\mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1)) \twoheadrightarrow \mathrm{H}^2_{\text{ét}}(X, \mathbb{Z}_\ell(1))/\ell^n,$$

which is surjective. The assertion follows. □

For any $\alpha \in \mathrm{Br}(X)[\ell^n]$ such that $\ell \neq p$, let $B_\ell(\alpha) := \pi_\ell^{-1}(\alpha)$, which is nonempty by Lemma 3.2.1.

For a Brauer class $\alpha \in \mathrm{Br}(X)[p^n]$, we need the following commutative diagram arising from de Rham–Witt theory [34, I.3.2, II.5.1, théorème 5.14]:

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathrm{H}^2(X, \mathbb{Z}_p(1)) & \longrightarrow & \mathrm{H}^2_{\mathrm{crys}}(X/W) & \xrightarrow{p-F} & \mathrm{H}^2_{\mathrm{crys}}(X/W) \\ & & \downarrow & & \downarrow{\scriptstyle p_n := (\otimes W_n)} & & \\ & & \mathrm{H}^2_{\mathrm{fl}}(X, \mu_{p^n}) & \xrightarrow{d\log} & \mathrm{H}^2_{\mathrm{crys}}(X/W_n) & & \end{array} \tag{3.2.4}$$

Here $\mathrm{H}^2(X, \mathbb{Z}_p(1)) := \varprojlim_n \mathrm{H}^2_{\mathrm{fl}}(X, \mu_{p^n})$. The map $d\log$ is injective by flat duality [51, Proposition 1.2]. Since the crystalline cohomology groups of an abelian surface are torsion-free, the mod p^n reduction map p_n is surjective. Consider the canonical surjective map

$$\pi_p : \mathrm{H}^2_{\mathrm{fl}}(X, \mu_{p^n}) \twoheadrightarrow \mathrm{Br}(X)[p^n]$$

induced by the Kummer sequence. We set

$$B_p(\alpha) := \left\{ b \in \mathrm{H}^2_{\mathrm{crys}}(X/W) \mid p_n(b) = d\log(t) \text{ for some } t \in \mathrm{H}^2_{\mathrm{fl}}(X, \mu_{p^n}) \text{ such that } \pi_p(t) = \alpha \right\}.$$

Following [10, Definitions 2.16 and 2.17], we can introduce (mixed) $\boldsymbol{B}$-fields for twisted abelian surfaces.

Definition 3.2.2. Let $\mathscr{X} \to X$ be a μ_n-gerbe and $[\mathscr{X}_{\mathbb{G}_m}] \in \mathrm{Br}(X)[n]$.

- If $n = \ell^t$ for some prime ℓ, an *ℓ-adic $\boldsymbol{B}$-field lift* of $\mathscr{X} \to X$ is an element $B = b/\ell^t$, where $b \in B_\ell([\mathscr{X}_{\mathbb{G}_m}])$. When $\ell = p$, it is also called a *crystalline $\boldsymbol{B}$-field* lift.
- In general, a mixed $\boldsymbol{B}$-field lift of $\mathscr{X} \to X$ is a collection $B = \{B_\ell\}$ consisting of a choice of an ℓ-adic $\boldsymbol{B}$-field lift B_ℓ of $[\mathscr{X}_{\mathbb{G}_m}^{(n\ell^{-t_\ell})}]$ for all prime factors $\ell \mid n$, where t_ℓ is the ℓ-adic valuation of n.

Remark 3.2.3. Not all elements in $\mathrm{H}^2_{\mathrm{crys}}(X/W)\left[\frac{1}{p}\right]$ are crystalline $\boldsymbol{B}$-fields, since the map $d\log$ is not surjective. From the first row in the diagram (3.2.4), we can see $B \in \mathrm{H}^2_{\mathrm{crys}}(X/W)\left[\frac{1}{p}\right]$ is a $\boldsymbol{B}$-field lift of some Brauer class if and only if $F(B) = pB$.

3.3. ***Twisted Mukai lattice over arbitrary fields.*** Let $\pi : \mathscr{X} \to X$ be a μ_n-gerbe with $\mathrm{ord}([\mathscr{X}_{\mathbb{G}_m}]) = n$, and let $B = \{B_\ell\}$ be a mixed $\boldsymbol{B}$-field lift of $[\mathscr{X}_{\mathbb{G}_m}]$. We define the ℓ-adic twisted Mukai lattice as

$$\widetilde{\mathrm{H}}(X, B_\ell) = \begin{cases} \exp(B_\ell)\widetilde{\mathrm{H}}(X, \mathbb{Z}_\ell) & \text{if } \ell \neq p, \\ \exp(B_\ell)\widetilde{\mathrm{H}}(X, W) & \text{if } \ell = p, \end{cases} \tag{3.3.1}$$

endowed with the Mukai pairing (3.2.1), where $\exp(B_\ell) = 1 + B_\ell + \frac{1}{2}B_\ell^2$.

Up to isomorphisms, the twisted Mukai lattice $\widetilde{\mathrm{H}}(X, B_\ell)$ is independent of the choice of the $\boldsymbol{B}$-field lift. We may use $\widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell)$ or $\widetilde{\mathrm{H}}(\mathscr{X}, W)$ to denote the twisted Mukai lattices to highlight the coefficients, irrespective of the choice of the $\boldsymbol{B}$-field lift.

Definition 3.3.1. Let $\mathrm{K}_0^{(1)}(\mathscr{X})$ be the Grothendieck group of $\mathtt{Coh}^{(1)}(\mathscr{X})$. The *twisted Chern character map* is the unique additive group homomorphism

$$\mathrm{ch}_{\mathscr{X}} : \mathrm{K}_0^{(1)}(\mathscr{X}) \to \widetilde{\mathrm{N}}(X)_{\mathbb{Q}}$$

such that for any locally free $\mathscr{X}$-twisted sheaf $\mathcal{E}$ on $\mathscr{X}$ with positive rank we have

$$\mathrm{ch}_{\mathscr{X}}(\mathcal{E}) = \sqrt[n]{\pi_*(\mathcal{E}^{\otimes n})} \in \widetilde{\mathrm{N}}(X)_{\mathbb{Q}}, \tag{3.3.2}$$

where $\sqrt[n]{-}$ means a choice of n-th roots such that the 0-codimension component of $\mathrm{ch}_{\mathscr{X}}(\mathcal{E})$ is equal to rank $\mathcal{E}$.

Denote by $\widetilde{\mathrm{N}}(\mathscr{X})$ the image of $\mathrm{K}_0^{(1)}(\mathscr{X})$ in $\widetilde{\mathrm{N}}(X)_{\mathbb{Q}}$ under the twisted Chern character map; we call it the *extended twisted Néron–Severi lattice*. For $\mathcal{E} \in \mathrm{D}^{(1)}(\mathscr{X})$, we define $v(\mathcal{E}) = \mathrm{ch}_{\mathscr{X}}([\mathcal{E}]) \in \widetilde{\mathrm{N}}(\mathscr{X})$ to be the Mukai vector of $\mathcal{E}$.

One can also define the twisted Chern character map to a cohomological twisted Mukai lattice

$$\mathrm{ch}_B : \mathrm{K}_0^{(1)}(\mathscr{X}) \to \widetilde{\mathrm{H}}(X, B_\ell);$$

see [42, §3.3] and [6, Appendix A3] for the ℓ-adic and crystalline cases, respectively. For any mixed $\boldsymbol{B}$-field lift B of $[\mathscr{X}_{\mathbb{G}_m}]$, the twisted Chern character ch_B factors through $\widetilde{\mathrm{N}}(\mathscr{X})$:

$$\begin{array}{ccc} \mathrm{K}_0^{(1)}(\mathscr{X}) & \xrightarrow{\ \mathrm{ch}_{B_\ell}\ } & \widetilde{\mathrm{H}}(X, B_\ell) \\ & {\scriptstyle \mathrm{ch}_{\mathscr{X}}} \searrow \quad \nearrow {\scriptstyle \exp(B_\ell)\,\mathrm{cl}_{\mathrm{H}}} & \\ & \widetilde{\mathrm{N}}(\mathscr{X}) & \end{array}$$

where cl_{H} is the cycle class map to the cohomology theory $\mathrm{H}(-)$. The following result is essentially proved in [10].

Proposition 3.3.2 [10, Proposition 3.5]. *Let B be a mixed $\boldsymbol{B}$-field lift of $[\mathscr{X}_{\mathbb{G}_m}] \in \mathrm{Br}(X)$. Then*

$$\widetilde{\mathrm{N}}(\mathscr{X}) \cong \bigcap_{\ell} \left(\widetilde{\mathrm{N}}(X) \otimes \mathbb{Z}\left[\tfrac{1}{\ell}\right] \cap \widetilde{\mathrm{H}}(X, B_\ell)\right),$$

where the intersection $\widetilde{\mathrm{N}}(X) \otimes \mathbb{Z}\left[\tfrac{1}{\ell}\right] \cap \widetilde{\mathrm{H}}(X, B_\ell)$ is taken in $\widetilde{\mathrm{N}}(X) \otimes \mathbb{Q}_\ell$ and the intersection $\bigcap_\ell$ is taken in $\widetilde{\mathrm{N}}(\tilde{X}) \otimes \mathbb{Q}$. The lattice $\widetilde{\mathrm{N}}(\mathscr{X})$ only depends on the associated $\mathbb{G}_m$-gerbe $\mathscr{X}_{\mathbb{G}_m}$, up to a lattice isomorphism.

Similarly, one can define the relative extended twisted Mukai lattice on smooth projective families of twisted abelian surfaces.

3.4. ***A filtered Torelli theorem.*** In [40; 41], Lieblich and Olsson introduced the notion of filtered derived equivalence and demonstrated that K3 surfaces with such equivalence are isomorphic. We will present an analogous result for (twisted) abelian surfaces. The proof is simpler than for K3 surfaces, as the bounded derived category of a (twisted) abelian surface corresponds to a generic K3 category [33].

Let $\mathscr{X} \to X$ be a μ_n-gerbe. The rational numerical Chow ring $\mathrm{CH}^*_{\mathrm{num}}(\mathscr{X})_{\mathbb{Q}} \cong \mathrm{CH}^*_{\mathrm{num}}(X)_{\mathbb{Q}}$ is equipped with a codimension filtration

$$\mathrm{Fil}^i \mathrm{CH}^*_{\mathrm{num}}(\mathscr{X})_{\mathbb{Q}} := \bigoplus_{k \geq i} \mathrm{CH}^k_{\mathrm{num}}(\mathscr{X})_{\mathbb{Q}}.$$

Since X is a surface, we have a natural identification $\widetilde{\mathrm{N}}(\mathscr{X})_{\mathbb{Q}} \cong \mathrm{CH}^*_{\mathrm{num}}(\mathscr{X})_{\mathbb{Q}}$.

Definition 3.4.1. Let $\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \to \mathrm{D}^{(1)}(\mathscr{Y})$ be a Fourier–Mukai transform. The derived equivalence $\Phi^{\mathcal{P}}$ is called *filtered* if its induced isomorphism $\Phi^{\mathcal{P}}_{\mathrm{CH}} : \widetilde{\mathrm{N}}(\mathscr{X}) \xrightarrow{\sim} \widetilde{\mathrm{N}}(\mathscr{Y})$ preserves the induced codimension filtrations.

Since the isomorphism $\widetilde{\mathrm{N}}(\mathscr{X}) \xrightarrow{\sim} \widetilde{\mathrm{N}}(\mathscr{Y})$ preserves the Mukai pairing, it is not hard to see that $\Phi^{\mathcal{P}}$ is filtered if and only if it sends the Mukai vector $(0, 0, 1)$ to $(0, 0, \pm 1)$. At the cohomological level, the codimension filtration on $\widetilde{\mathrm{H}}(X)\left[\frac{1}{\ell}\right]$ (the prime ℓ depending on the choice of ℓ-adic or crystalline twisted Mukai lattice) is given by $F^i = \oplus_{r \geq i} \mathrm{H}^{2r}(X)\left[\frac{1}{\ell}\right]$. The filtration on $\widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell)$ is defined by

$$F^i \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell) = \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell) \cap F^i \widetilde{\mathrm{H}}(X, \mathbb{Z}_\ell)\left[\tfrac{1}{\ell}\right].$$

If we choose a $\boldsymbol{B}$-field lift B_ℓ, a direct computation shows that the graded pieces of $F^\bullet$ are

$$\begin{aligned}
\mathrm{Gr}^0_F \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell) &= \left\{\left(r, rB_\ell, \tfrac{1}{2} r B_\ell^2\right) \mid r \in \mathrm{H}^0(X, \mathbb{Z}_\ell(-1))\right\}, \\
\mathrm{Gr}^1_F \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell) &= \left\{(0, b, b \cdot B_\ell) \mid b \in \mathrm{H}^2(X, \mathbb{Z}_\ell)\right\} \cong \mathrm{H}^2(X, \mathbb{Z}_\ell), \\
\mathrm{Gr}^2_F \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell) &= \left\{(0, 0, s) \mid s \in \mathrm{H}^4(X, \mathbb{Z}_\ell(1))\right\} \cong \mathrm{H}^4(X, \mathbb{Z}_\ell(1)).
\end{aligned} \tag{3.4.1}$$

Lemma 3.4.2. *A Fourier–Mukai transform $\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \to \mathrm{D}^{(1)}(\mathscr{Y})$ is filtered if and only if its cohomological realization is filtered for all $\boldsymbol{B}$-field liftings.*

Proof. A Fourier–Mukai transform that is filtered is necessarily cohomologically filtered. This is because the map

$$\exp(B_\ell) \cdot \mathrm{cl}_H : \widetilde{\mathrm{N}}(\mathscr{X}) \to \widetilde{\mathrm{H}}(\mathscr{X}, \mathbb{Z}_\ell)$$

preserves the filtrations for any $\boldsymbol{B}$-field lift B of $[\mathscr{X}_{\mathbb{G}_m}]$.

For the converse, notice that $\Phi^{\mathcal{P}}$ is filtered if and only if the induced map $\Phi^{\mathcal{P}}_{\mathrm{CH}}$ takes the vector $(0, 0, 1)$ to $(0, 0, \pm 1)$. As $\Phi^{\mathcal{P}}$ is cohomologically filtered for B, the cohomological realization of $\Phi^{\mathcal{P}}$ preserves the graded piece Gr^2_F in (3.4.1). This implies that $\Phi^{\mathcal{P}}_{\mathrm{CH}}$ takes $(0, 0, 1)$ to $(0, 0, \pm 1)$. □

Proposition 3.4.3 (filtered Torelli theorem for twisted abelian surfaces). *Suppose $k = \bar{k}$ is such that $\mathrm{char}\, k \neq 2$. Let $\mathscr{X} \to X$ and $\mathscr{Y} \to Y$ be μ_n-gerbes on abelian surfaces. The following statements are equivalent.*

(1) *There is an isomorphism between the associated $\mathbb{G}_m$-gerbes $\mathscr{X}_{\mathbb{G}_m}$ and $\mathscr{Y}_{\mathbb{G}_m}$.*

(2) *There is a filtered Fourier–Mukai transform $\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{X}) \to \mathrm{D}^{(1)}(\mathscr{Y})$.*

Proof. For untwisted case, i.e., $\mathscr{X} = X$ and $\mathscr{Y} = Y$, this is exactly [27, Proposition 3.1]. Here we extend it to the twisted case. As one direction is obvious, it suffices to show that (2) implies (1).

Firstly, we claim that all semirigid objects in $\mathrm{D}^{(1)}(\mathscr{Y})$ are in $\mathtt{Coh}^{(1)}(\mathscr{Y})$ up to a shift. According to Remark 3.13 in [33], it is sufficient to show that there are no stable spherical sheaves in $\mathtt{Coh}(\mathscr{Y}^{(1)})$. If $\mathcal{E}$ is a spherical $\mathscr{Y}^{(1)}$-twisted sheaf with rank $\mathcal{E} = 0$, then $c_1(\mathcal{E})^2 = -\chi(\mathcal{E}, \mathcal{E}) = -2$, which is impossible for the abelian surface. Suppose that there is a stable spherical $\mathscr{Y}$-twisted sheaf $\mathcal{E}$ with Mukai vector $v = (r, c, s)$ such that $r > 0$. Choose a polarization $H \in \mathrm{Pic}(Y)$ so that $\mathcal{E}$ is H-semistable. Let $M_H(\mathscr{Y}, v)$ be the moduli space of H-semistable $\mathscr{Y}$-twisted sheaves on Y. Then $M_H(\mathscr{Y}, v)$ is nonempty. Consider the determinant morphism to the Picard stack of invertible $\mathscr{Y}^{(r)}$-twisted sheaves

$$\mathbf{det} : \mathscr{M}_H(\mathscr{Y}, v) \to \mathrm{Pic}(\mathscr{Y}^{(r)}).$$

For any $\mathcal{L} \in \mathrm{Pic}^0(Y)$ and $\mathcal{E} \in \mathscr{M}_H(\mathscr{Y}, v)$, the tensor product $\mathcal{E} \otimes \mathcal{L}$ is still a stable $\mathscr{Y}$-twisted sheaf with the Mukai vector v. Thus, the map **det** dominates the component of $\mathrm{Pic}(\mathscr{Y}^{(r)})$ containing $\det(\mathcal{E})$, which is of dimension 2. Therefore, the deformation theory of twisted coherent sheaf implies

$$\dim_k \mathrm{Ext}^1(\mathcal{E}, \mathcal{E}) \geq \dim \mathscr{M}_H(\mathscr{Y}, v) \geq 2,$$

contradicting the assumption that $\mathcal{E}$ is spherical.

Let $\Phi^{\mathcal{P}} : \mathrm{D}^b(\mathscr{X}^{(1)}) \to \mathrm{D}^b(\mathscr{Y}^{(1)})$ be a Fourier–Mukai transform. For a closed point $x \in X$, denote

$$\mathcal{P}_x := \Phi^{\mathcal{P}}(k(x)) = \mathcal{P}|_{\{x\}\times\mathscr{Y}},$$

by image of the skyscraper sheaf $k(x)$. Since $k(x)$ is semirigid, $\mathcal{P}_x$ is also semirigid. The previous discussion implies that there is an integer m such that $\mathcal{H}^i(\mathcal{P}_x) = 0$ for any $i \neq m$ and closed point $x \in \mathcal{Y}$. Therefore, there is a $\mathscr{X}^{(-1)} \wedge \mathscr{Y}$-twisted sheaf $\mathcal{E} \in \mathtt{Coh}(\mathscr{X}^{(-1)} \times \mathscr{Y})$ such that $\mathcal{P} \cong \mathcal{E}[m]$.

Suppose $\Phi^{\mathcal{P}}$ is filtered. Composing it with the shift functor $\mathcal{F} \mapsto \mathcal{F}[1]$ if necessary, we may assume that the cohomological realization of $\Phi^{\mathcal{P}}$ sends $(0, 0, 1)$ to $(0, 0, 1)$. In this case, $\mathcal{E}_x$ is just a skyscraper sheaf on $\{x\} \times Y$. The same argument as in [15, Corollary 5.3] or [29, Corollary 5.22, 5.23] shows that there is an isomorphism $f : X \to Y$ such that $f^*([\mathscr{Y}_{\mathbb{G}_m}]) = [\mathscr{X}_{\mathbb{G}_m}]$. □

3.5. ***Twisted FM partners via moduli space of twisted sheaves.*** In the rest of this section, we will assume that $k = \bar{k}$ and $\mathrm{char}\, k = p \neq 2$. Let $\mathscr{X} \to X$ be a twisted abelian surface over k.

Definition 3.5.1 [70, Definition 0.1]. Let $v = (r, c, s) \in \widetilde{\mathrm{N}}(\mathscr{X})$ be a primitive Mukai vector such that $v^2 = 0$. If

(1) $r > 0$, or

(2) $r = 0$, c is effective and $s \neq 0$, or

(3) $r = c = 0$ and $s > 0$,

then v is called *positive*.

We denote by $\mathscr{M}_H(\mathscr{X}, v)$ the moduli stack of H-semistable $\mathscr{X}$-twisted sheaves with the Mukai vector $v \in \widetilde{\mathrm{N}}(\mathscr{X})$, where H is a v-generic ample divisor on X. Here, we record a well-known nonemptiness criterion for $\mathscr{M}_H(\mathscr{X}, v)$ when X is not supersingular. We will extend this result to the supersingular case in Theorem 3.6.6, using the theory of supersingular twistor space.

Proposition 3.5.2 (Minamide, Yanagida and Yoshioka; Bragg and Lieblich). *Suppose X is an abelian surface over k that is not supersingular. If v is positive with $v^2 = 0$, then for any v-generic polarization H, the coarse moduli space $M_H(\mathscr{X}, v)$ is an abelian surface, and the moduli stack $\mathscr{M}_H(\mathscr{X}, v)$ is a $\mathbb{G}_m$-gerbe on $M_H(\mathscr{X}, v)$.*

Proof. In characteristic 0, this is [71, Theorem 3.16].

When $\operatorname{char} k = p > 2$, the nonemptiness can be seen through a lifting argument, as shown in [8, Proposition 4.1.20] and [46, Proposition A.2.1]. Since X is of finite height when $\operatorname{char} k = p > 0$, Lemma 2.3.1 implies there exists a DVR V with residue field k and a projective lifting $\mathscr{X}_V \to X_V$ of $\mathscr{X} \to X$ over Spec V, together with an extension $v_V \in \widetilde{\mathrm{N}}(\mathscr{X}_V)$ and a polarization $H_V \in \mathrm{NS}(X_V)$ such that $H_V|_{\operatorname{Spec} k} = H$. Consider the relative moduli space of twisted sheaves $\mathscr{M}_{H_V}(\mathscr{X}_V, v_V)$ over Spec V. Its (geometric) generic fiber is a moduli space of twisted sheaves with positive Mukai vector in characteristic zero, which is nonempty by Yoshioka's result. Thus its special fiber, which is isomorphic $\mathscr{M}_H(\mathscr{X}, v)$, is also nonempty by Langton's semistable reduction theorem. □

The following is an extension of [27, Theorem 1.2].

Theorem 3.5.3. *Assume $k = \bar{k}$ with $\operatorname{char} k \neq 2$. Let $\mathscr{X} \to X$ and $\mathscr{Y} \to Y$ be $\mathbb{G}_m$-gerbes over an abelian surface defined over k. Then $\mathrm{D}^{(1)}(\mathscr{X}) \simeq \mathrm{D}^{(1)}(\mathscr{Y})$ if and only if $\mathscr{Y}^{(-1)} \to Y$ is isomorphic to the moduli stack $\mathscr{M}_H(\mathscr{X}, v) \to M_H(\mathscr{X}, v)$ for some $v \in \widetilde{\mathrm{N}}(\mathscr{X})$ and v-generic polarization H.*

Proof. For the "if" part, note that the universal family of twisted sheaves on $\mathscr{M}_H(\mathscr{X}, v) \times \mathscr{X}$ induces a derived equivalence.

For the other direction, suppose $\mathrm{D}^{(1)}(\mathscr{X}) \simeq \mathrm{D}^{(1)}(\mathscr{Y})$ are equivalent. Let

$$\Phi^{\mathcal{P}} : \mathrm{D}^{(1)}(\mathscr{Y}) \to \mathrm{D}^{(1)}(\mathscr{X})$$

be a Fourier–Mukai transform. Let $v \in \widetilde{\mathrm{N}}(\mathscr{X})$ be the image of $(0, 0, 1) \in \widetilde{\mathrm{N}}(\mathscr{Y})$ under $\Phi^{\mathcal{P}}$. Up to a shift, we can assume that v is a positive vector. By Proposition 3.5.2, $M_H(\mathscr{X}, v)$ is an abelian surface and $\mathscr{M}_H(\mathscr{X}, v) \to M_H(\mathscr{X}, v)$ is a $\mathbb{G}_m$-gerbe over it.

Let $\mathcal{E}$ be a universal $\mathscr{X}$-twisted sheaf on $\mathscr{M}_H(\mathscr{X}, v) \times \mathscr{X}$ that is a $(1, 1)$-fold twisted sheaf and induces a derived equivalence

$$\Phi^{\mathcal{E}} : \mathrm{D}^{(-1)}(\mathscr{M}_H(\mathscr{X}, v)) \to \mathrm{D}^{(1)}(\mathscr{X}),$$

whose cohomological realization maps the Mukai vector $(0, 0, 1)$ to v. Composing $\Phi^{\mathcal{E}}$ with the derived equivalence

$$(\Phi^{\mathcal{P}})^{-1} \simeq \Phi^{\mathcal{P}^\vee}[2] : \mathrm{D}^{(1)}(\mathscr{X}) \to \mathrm{D}^{(1)}(\mathscr{Y}),$$

we obtain a filtered derived equivalence from $\mathscr{M}_H(\mathscr{X}, v)^{(-1)}$ to $\mathscr{Y}$, which induces an isomorphism from $\mathscr{M}_H(\mathscr{X}, v)^{(-1)}$ to $\mathscr{Y}$ by Proposition 3.4.3. □

3.6. *Supersingular twisted abelian surfaces.* Finally, we discuss the case of supersingular twisted abelian surfaces. In this part, we extend the construction the supersingular twistor space as [8] via the Ogus crystalline Torelli theorem for supersingular abelian surfaces (see [51, §2]).

Definition 3.6.1. Let p be a prime $\neq 2$. Let Λ be an indefinite p-elementary even lattice, meaning that $\operatorname{disc}(\Lambda \otimes \mathbb{Q}) = -1$ and $\Lambda^\vee/\Lambda$ is p-torsion. Then $|\Lambda^\vee/\Lambda| = p^{2\sigma_0(\Lambda)}$ for $1 \le \sigma_0(\Lambda) \le \frac{1}{2}n$ and the integer $\sigma_0(\Lambda)$ is called the *Artin invariant* of Λ. We define M_Λ to be the Ogus *moduli space of characteristic subspaces* of $p\Lambda^\vee/p\Lambda$.

When Λ has signature $(1, n-1), n \ge 2$, as shown in [61, Section 1], Λ is uniquely determined by its Artin invariant. When $n = 6$, we call it a *supersingular abelian surface lattice*, because for every supersingular abelian surface X, the Néron–Severi lattice $\operatorname{NS}(X)$ is a supersingular abelian surface lattice [52, (1.6)].

From now on, let us assume that Λ is a supersingular abelian surface lattice. For simplicity, denote by σ_0 the Artin invariant $\sigma_0(\Lambda)$. We set

$$\tilde{\Lambda} = \Lambda \oplus U(p),$$

where $U(p)$ is the twisted hyperbolic plane generated by the vectors e and f such that $e^2 = f^2 = 0$ and $e \cdot f = -p$. Let $M^{\langle e \rangle}_{\tilde{\Lambda}} \subseteq M_{\tilde{\Lambda}}$ be the moduli space of characteristic subspaces of $p\tilde{\Lambda}^\vee/p\tilde{\Lambda}$ that do not contain e.

Proposition 3.6.2 [8, §3]. *The moduli stacks $M^{\langle e \rangle}_{\tilde{\Lambda}}$ and M_Λ are representable by schemes over $\mathbb{F}_p$, which are smooth of dimensions σ_0 and $\sigma_0 - 1$, respectively. There is a smooth morphism*

$$\pi_e : M^{\langle e \rangle}_{\tilde{\Lambda}} \to M_\Lambda$$

whose fiber at a closed point is isomorphic to a group scheme with connected components $\mathbb{A}^1$.

Proof. The first assertion is given in [52, Proposition 4.6]. Let us sketch the construction of π_e. Given any $\widetilde{\mathcal{K}} \in M^{\langle e \rangle}_{\tilde{\Lambda}}(T)$ over an $\mathbb{F}_p$-scheme T, a characteristic subspace

$$\mathcal{K} \subseteq (p\Lambda^\vee/p\Lambda) \otimes \mathcal{O}_T$$

can be formed as the image of $\widetilde{\mathcal{K}} \cap (e^\perp \otimes \mathcal{O}_T)$ in $(e^\perp/e) \otimes \mathcal{O}_T$ (see [8, Lemma 3.1.9]). Consequently, the map $\widetilde{\mathcal{K}} \mapsto \mathcal{K}$ defines a morphism

$$\pi_e : M^{\langle e \rangle}_{\tilde{\Lambda}} \to M_\Lambda.$$

The rest of the assertion is a consequence of [8, Lemma 3.1.15]. □

Definition 3.6.3. The *twistor line* in $M_{\tilde{\Lambda}} \otimes_{\mathbb{F}_p} k$ is an affine line $\mathbb{A}^1 \subset M_{\tilde{\Lambda}} \otimes_{\mathbb{F}_p} k$ that is a connected component of a fiber of π_e over a k-point of $M_\Lambda(k)$ for some isotropic vector $e \in \tilde{\Lambda}$.

The moduli functor S_Λ of Λ-marked supersingular abelian surfaces is representable by a locally separated and smooth algebraic space of dimension $\sigma_0 - 1$ over k, by the crystalline Torelli theorem [51, Theorem 7.3] together with the argument in [52, Theorem 2.7]. Consider the universal family of supersingular abelian surfaces

$$u : \mathcal{X} \to S_\Lambda,$$

which is smooth with relative dimension 2. By Proposition 2.4.2, the higher direct image $\mathrm{R}^2 u_* \mu_p$ is representable by an algebraic group space over S_Λ, denoted by

$$\pi : \mathscr{S}_\Lambda \to S_\Lambda.$$

The connected component of the identity $\mathscr{S}^o_\Lambda \subset \mathscr{S}_\Lambda$ parameterizes the μ_p-gerbes which are not essentially trivial except the identity, on each Λ-marked supersingular abelian surface in $S_\Lambda(k)$. Then there are (twisted) period morphisms following the approach in [52, §3].

Proposition 3.6.4. *There are (twisted) period morphisms*

$$\rho : S_\Lambda \to \overline{M}_\Lambda := M_\Lambda \otimes_{\mathbb{F}_p} k \quad \text{and} \quad \tilde{\rho} : \mathscr{S}^o_\Lambda \to \overline{M}^{\langle e \rangle}_{\tilde{\Lambda}} := M^{\langle e \rangle}_{\tilde{\Lambda}} \otimes_{\mathbb{F}_p} k$$

such that the commutative diagram

$$\begin{array}{ccc} \mathscr{S}^o_\Lambda & \xrightarrow{\pi|_{\mathscr{S}^o_\Lambda}} & S_\Lambda \\ \big\downarrow{\scriptstyle \tilde{\rho}_\Lambda} & & \big\downarrow{\scriptstyle \rho_\Lambda} \\ \overline{M}^{\langle e \rangle}_{\tilde{\Lambda}} & \xrightarrow{\pi_e} & \overline{M}_\Lambda \end{array} \tag{3.6.1}$$

is Cartesian. Moreover, ρ and $\tilde{\rho}$ are étale surjective when $p > 2$.

Proof. This was proved by Bragg and Lieblich in the case of supersingular K3 surfaces [8, §3 and§5]. But everything works for supersingular abelian surfaces as well. We mention that one can also use the Kummer construction to deduce the statement from the K3 case.

For reference, let us sketch the construction of $\tilde{\rho}_\Lambda$ and ρ_Λ. Let (X, η) be a Λ-marked supersingular abelian surface. The K3-crystal $\mathrm{H}^2_{\mathrm{crys}}(X/W)$ determines a characteristic subspace

$$\mathcal{K}_{\mathrm{H}^2(X)} := \ker(\mathrm{NS}(X) \otimes k \to \mathrm{H}^2_{\mathrm{crys}}(X/W) \otimes k).$$

Then $\rho_\Lambda(X, \eta)$ is the characteristic subspace $\eta^{-1}(\mathcal{K}_{\mathrm{H}^2(X)})$ in $(p\Lambda^\vee/p\Lambda) \otimes_{\mathbb{F}_p} k$. Suppose $\mathscr{X} \to X$ is a μ_p-gerbe. We define

$$\mathcal{K}_{\widetilde{\mathrm{H}}(\mathscr{X})} := \ker(\widetilde{\mathrm{N}}(\mathscr{X}) \otimes k \to \widetilde{\mathrm{H}}(\mathscr{X}, W) \otimes_W k) \subset \frac{p\widetilde{\mathrm{N}}(\mathscr{X})^\vee}{p\widetilde{\mathrm{N}}(\mathscr{X})} \otimes k$$

as the strictly characteristic subspace of $\widetilde{\mathrm{H}}(\mathscr{X}, W)$. Note that there is an extended map of K3 crystals,

$$\tilde{\Lambda} \otimes \mathbb{Z}_p \xrightarrow{\tilde{\eta}} \widetilde{\mathrm{N}}(\mathscr{X}) \otimes \mathbb{Z}_p \to \widetilde{\mathrm{H}}(\mathscr{X}, W),$$

where $\tilde{\eta}$ is given by $e \mapsto (0,0,1)$, $f \mapsto (p,0,0)$ and $c \mapsto \eta(c)$ for all $c \in \Lambda$. Then $\tilde{\rho}_\Lambda(\mathscr{X}, \eta) = \tilde{\eta}^{-1}(\mathcal{K}_{\widetilde{\mathrm{H}}(\mathscr{X})})$ is the characteristic subspace of $(p\tilde{\Lambda}^\vee / p\tilde{\Lambda}) \otimes k$. □

Remark 3.6.5. In one view, the moduli space M_Λ is a crystalline analog of the classical period domain. Let H be a supersingular K3 crystal. The associated Tate module $\mathrm{T}_H \subseteq H$ is a supersingular K3 $\mathbb{Z}_p$-lattice in the sense of Ogus (see [51, 3.13]). According to [51, Theorem 3.20], the functor

$$H \rightsquigarrow (T_H, \mathcal{K}_H),$$

where $\mathcal{K}_H = \ker(T_H \otimes k \to H \otimes k)$, defines an equivalence between the category of supersingular K3 crystals and the category of strictly characteristic subspaces of a supersingular K3 $\mathbb{Z}_p$-lattice.

Using the twisted period map, we obtain:

Theorem 3.6.6. *Let $\mathscr{X} \to X$ be a μ_p-gerbe over a supersingular abelian surface X over k.*

(1) *If a primitive vector $v \in \widetilde{\mathrm{N}}(\mathscr{X})$ is positive and isotropic, the coarse moduli space $M_H(\mathscr{X}, v)$ is an abelian surface.*

(2) *If $\mathscr{Y} \to Y$ is another twisted abelian surface, we have $\mathrm{D}^{(1)}(\mathscr{X}) \simeq \mathrm{D}^{(1)}(\mathscr{Y})$ if and only if there is an isomorphism*

$$\widetilde{\mathrm{H}}(\mathscr{X}, W) \cong \widetilde{\mathrm{H}}(\mathscr{Y}, W)$$

of K3 crystals.

(3) *There is a derived equivalence*

$$\mathrm{D}^{(1)}(\mathscr{X}_0) \simeq \mathrm{D}^b(X),$$

where $\mathscr{X}_0 \to X_0$ is a μ_p-gerbe over the unique superspecial abelian surface X_0.

Proof. For (1), if $\mathscr{X} \to X$ is an essentially trivial μ_p-gerbe over a supersingular abelian surface X, this can be proved by a standard lifting argument (see also [22, Proposition 6.9]). When $\mathscr{X} \to X$ is nontrivial, we can take the universal family of μ_p-gerbes

$$f : \mathfrak{X} \to \mathbb{A}^1$$

on the connected component $\mathbb{A}^1 \subset \mathrm{R}^2 u_* \mu_p$ that contains $\mathscr{X}$ (see Corollary 2.4.5). The fibers of f contain $\mathscr{X} \to X$ and the trivial μ_p-gerbe over X. By taking the relative moduli space of twisted sheaves (with a suitable v-generic polarization) on $\mathfrak{X} \to \mathbb{A}^1$, one obtains the nonemptiness of $M_H(\mathscr{X}, v)$ from the case of essentially trivial gerbes.

For the proof of the forward direction of (2), we notice that by Remark 3.6.5, it is sufficient to find an isomorphism between pairs

$$(\widetilde{\mathrm{N}}(\mathscr{X}), \mathcal{K}_{\widetilde{\mathrm{H}}(\mathscr{X})}) \xrightarrow{\sim} (\widetilde{\mathrm{N}}(\mathscr{Y}), \mathcal{K}_{\widetilde{\mathrm{H}}(\mathscr{Y})});$$

this is provided by the de Rham realization of the derived equivalence $\mathrm{D}^{(1)}(\mathscr{X}) \simeq \mathrm{D}^{(1)}(\mathscr{Y})$. The other direction is handled like the case of K3 surfaces proved in [6, Theorem 3.5.5]. The key is that if

$\widetilde{\mathrm{H}}(\mathscr{X}, W) \cong \widetilde{\mathrm{H}}(\mathscr{Y}, W)$, then there exists $v \in \widetilde{\mathrm{N}}(\mathscr{X})$ such that the induced isomorphism

$$\widetilde{\mathrm{H}}(\mathscr{M}_H(\mathscr{X}, v)^{(-1)}, W) \cong \widetilde{\mathrm{H}}(\mathscr{Y}, W)$$

of K3 crystals sends (0, 0, 1) to (0, 0, 1). The assertion then essentially follows from the Ogus crystalline Torelli theorem for supersingular abelian surfaces (Theorem 7.3 of [51]), as in [6, Theorem 3.5.2]. We omit details.

For (3), due to (2), it suffices to find a μ_p-gerbe $\mathscr{X}_0 \to X_0$ such that there is a supersingular K3 crystal isomorphism

$$\widetilde{\mathrm{H}}(\mathscr{X}_0, W) \cong \widetilde{\mathrm{H}}(X, W).$$

By Remark 3.6.5, this is equivalent to finding $\mathscr{X}_0 \to X_0$ and an isometry $\widetilde{\mathrm{N}}(\mathscr{X}_0) \otimes \mathbb{Z}_p \cong \widetilde{\mathrm{N}}(X) \otimes \mathbb{Z}_p$ sending $\mathcal{K}_{\widetilde{\mathrm{H}}(\mathscr{X}_0)}$ to $\mathcal{K}_{\widetilde{\mathrm{H}}(X)}$.

Let us give an explicit construction of $\mathscr{X}_0 \to X_0$ via the twisted period map. If X is superspecial, no further proof is necessary. Suppose X is not superspecial. Then $\sigma_0(\mathrm{NS}(X)) = 2$ by Proposition 3.7 of [64]. Let Λ be the supersingular abelian surface lattice with Artin invariant 2 and let $\eta : \Lambda \xrightarrow{\sim} \mathrm{NS}(X)$ be a Λ-marking. As shown in [61, Section 2] (see also [22, Proposition 6.1]), $\Lambda = U(p) \oplus \Lambda'$ contains $U(p)$ as a direct summand and the image of $U(p)$ in $(p\Lambda^\vee/p\Lambda) \otimes k$ is not contained in the strictly characteristic subspace $\rho_\Lambda(X, \eta)$.

Note that the lattice $\Lambda_0 = U \oplus \Lambda'$ is a supersingular abelian lattice with Artin invariant 1. There is a natural isomorphism

$$\widetilde{\mathrm{N}}(X) \xrightarrow{\eta \oplus \mathrm{id}} \Lambda \oplus U \cong \Lambda_0 \oplus U(p) = \tilde{\Lambda}_0 \tag{3.6.2}$$

and we can identify $\mathcal{K}_{\widetilde{\mathrm{H}}(X)}$ with $\rho_\Lambda(X, \eta)$ via the isometry $\eta \oplus \mathrm{id}$. Let

$$\mathcal{K} \subseteq (p\tilde{\Lambda}_0^\vee / p\tilde{\Lambda}_0) \otimes k$$

be the image of $\mathcal{K}_{\widetilde{\mathrm{H}}(X)}$ through the map induced by (3.6.2). By our assumption, $\mathcal{K}$ does not contain the image of some isotropic vector $e \in U(p)$ and therefore can be viewed as a point in $M^{\langle e \rangle}_{\tilde{\Lambda}_0}(k)$. As $\tilde{\rho}_{\Lambda_0}$ is surjective, there is a Λ_0-marked supersingular abelian surface $(\mathscr{X}_0 \to X_0, \eta_0)$ such that $\tilde{\rho}_{\Lambda_0}(\mathscr{X}_0, \eta_0) = \mathcal{K}$. It is easy to see that $\mathscr{X}_0 \to X_0$ is as desired. □

4. Shioda's Torelli theorem for abelian surfaces

In [63], Shioda discovered that there is a way, now called Shioda's trick, to extract information about the first cohomology of a complex abelian surface from its second cohomology. This established a global Torelli theorem for complex abelian surfaces via second cohomology, which is also a key step in Piatetskii-Shapiro and Shafarevich's proof of the Torelli theorem for K3 surfaces (see [57, §5, Lemma 4 and Theorem 1]).

The aim of this section is to generalize Shioda's method to all fields and establish an isogeny theorem for abelian surfaces via the second cohomology. We will deal with Shioda's trick for Betti cohomology, étale cohomology and crystalline cohomology separately.

4.1. ***Recap of Shioda's trick for Hodge isometry.*** We first recall Shioda's construction. Suppose X is a complex abelian surface. Its singular cohomology ring $\mathrm{H}^\bullet(X, \mathbb{Z})$ is canonically isomorphic to the exterior algebra $\bigwedge^\bullet \mathrm{H}^1(X, \mathbb{Z})$. Let V be a free $\mathbb{Z}$-module of rank 4. We denote by Λ the lattice $(\bigwedge^2 \mathrm{V}, q)$, where $q : \bigwedge^2 \mathrm{V} \times \bigwedge^2 \mathrm{V} \to \bigwedge^4 \mathrm{V} \cong \mathbb{Z}$ is the wedge product. After choosing a $\mathbb{Z}$-basis $\{v_i\}_{1\le i\le 4}$ for $\mathrm{H}^1(X, \mathbb{Z})$, we have an isometry of $\mathbb{Z}$-lattices $\Lambda \xrightarrow{\sim} \mathrm{H}^2(X, \mathbb{Z})$. The set of vectors

$$\{v_{ij} := v_i \wedge v_j\}_{0\le i<j\le 4}$$

clearly forms a basis of $\mathrm{H}^2(X, \mathbb{Z})$, which will be called an *admissible basis* of A for its second singular cohomology. For another complex abelian surface Y, a Hodge isometry

$$\varphi : \mathrm{H}^2(Y, \mathbb{Z}) \xrightarrow{\sim} \mathrm{H}^2(X, \mathbb{Z})$$

will be called *admissible* if $\det(\varphi) = 1$ with respect to some admissible bases on X and Y. It is clear that the admissibility of a morphism is independent of the choice of admissible bases.

In terms of admissible bases, we can view φ as an element in $\mathrm{SO}(\Lambda)$. On the other hand, we have the exact sequence of groups

$$1 \to \{\pm 1\} \to \mathrm{SL}_4(\mathbb{Z}) \xrightarrow{\bigwedge^2} \mathrm{SO}(\Lambda). \tag{4.1.1}$$

Shioda observed that the image of $\mathrm{SL}_4(\mathbb{Z})$ in $\mathrm{SO}(\Lambda)$ is a subgroup of index two and does not contain $-\mathrm{id}_\Lambda$. From this, he proved:

Theorem 4.1.1 (Shioda [63, Theorem 1]). *For any admissible integral Hodge isometry ψ, there is an isomorphism of integral Hodge structures*

$$\psi : \mathrm{H}^1(Y, \mathbb{Z}) \xrightarrow{\sim} \mathrm{H}^1(X, \mathbb{Z})$$

such that $\bigwedge^2(\psi) = \varphi$ or $-\varphi$.

This is what we call "Shioda's trick". As we can assume that a Hodge isometry is admissible after possibly taking the dual abelian variety for one of them (see Example 4.2.3 below), we can obtain the Torelli theorem for complex abelian surfaces by using weight-two Hodge structures — that is, X is isomorphic to Y or its dual $\hat{Y}$ if and only if there is an integral Hodge isometry $\mathrm{H}^2(X, \mathbb{Z}) \cong \mathrm{H}^2(Y, \mathbb{Z})$ [63, Theorem 1].

4.2. ***Admissible bases.*** To extend Shioda's work to arbitrary fields, we must define admissibility for different cohomology theories (e.g., étale and crystalline cohomology).

Let k be a field with $\operatorname{char} k = p \ge 0$. Suppose X is an abelian surface over k and $\ell \nmid p$ is a prime. For simplicity of notations, we will denote $\mathrm{H}^\bullet(-)_R$ for one of the following cohomology theories:

(1) if $k \hookrightarrow \mathbb{C}$ and $R = \mathbb{Z}$ or any number field E, then $\mathrm{H}^\bullet(X)_R = \mathrm{H}^\bullet(X(\mathbb{C}), R)$ the singular cohomology.

(2) if $R = \mathbb{Z}_\ell$ or $\mathbb{Q}_\ell$, then $\mathrm{H}^\bullet(X)_R = \mathrm{H}^\bullet_{\text{ét}}(X_{\bar{k}}, R)$, the ℓ-adic étale cohomology.

(3) if $\operatorname{char} k = p > 0$, then we can take $R = W$ a Cohen ring of k or the fraction field K of W, then $\mathrm{H}^\bullet(X)_R = \mathrm{H}^\bullet_{\text{crys}}(X/W)$ or $\mathrm{H}^\bullet_{\text{crys}}(X/W) \otimes K$, the crystalline cohomology.

There is an isomorphism between the cohomology ring $\mathrm{H}^\bullet(X)_R$ and the exterior algebra $\bigwedge^\bullet \mathrm{H}^1(X)_R$. We denote by $\mathrm{tr}_X : \mathrm{H}^4(X)_R \xrightarrow{\sim} R$ the corresponding trace map. The Poincaré pairing $\langle - , - \rangle$ on $\mathrm{H}^2(X)_R$ can be realized as

$$\langle \alpha, \beta \rangle = \mathrm{tr}_X(\alpha \wedge \beta).$$

Analogously to Section 4.1, an R-basis $\{v_i\}$ of $\mathrm{H}^1(X)_R$ will be called a *d-admissible basis* if it satisfies

$$\mathrm{tr}_X(v_1 \wedge v_2 \wedge v_3 \wedge v_4) = d$$

for some $d \in R^*$. When $d = 1$, we speak of an *admissible basis*. For any d-admissible (resp. admissible) basis $\{v_i\}$, the associated R-basis $\{v_{ij} := v_i \wedge v_j\}_{i<j}$ of $\mathrm{H}^2(X)_R$ will also be called d-admissible (resp. admissible).

Example 4.2.1. Let $\{v_1, v_2, v_3, v_4\}$ be an R-linear basis of $\mathrm{H}^1(X)_R$. Suppose

$$\mathrm{tr}_X(v_1 \wedge v_2 \wedge v_3 \wedge v_4) = t \in R^*.$$

For any $d \in R^*$, there is a natural d-admissible R-linear basis $\{(d/t)v_1, v_2, v_3, v_4\}$.

Definition 4.2.2. Let X and Y be abelian surfaces over k.

- An R-linear isomorphism $\psi : \mathrm{H}^1(X)_R \to \mathrm{H}^1(Y)_R$ is d-admissible if it takes an admissible basis to a d-admissible basis.
- An R-linear isomorphism $\varphi : \mathrm{H}^2(X)_R \to \mathrm{H}^2(Y)_R$ is d-admissible if

$$\mathrm{tr}_Y \circ \textstyle\bigwedge^2(\varphi) = d\,\mathrm{tr}_X$$

  for some $d \in R^*$, or equivalently, if it sends an admissible basis to a d-admissible basis. When $d = 1$, it will also be called admissible.

The set of d-admissible isomorphisms is denoted by $\mathrm{Isom}^{\mathrm{ad},(d)}(\mathrm{H}^i(X)_R, \mathrm{H}^i(Y)_R)$ accordingly.

For any isomorphism $\varphi : \mathrm{H}^2(X)_R \xrightarrow{\sim} \mathrm{H}^2(Y)_R$, let $\det(\varphi)$ be the determinant of the matrix with respect to some admissible bases. It is not hard to see $\det(\varphi)$ is independent of the choice of admissible bases, and φ is admissible if and only if $\det(\varphi) = 1$.

Example 4.2.3. Let $\{v_i\}$ be an admissible basis of $\mathrm{H}^1(X)_R$. For the dual abelian surface $\hat{X}$, the dual basis $\{v_i^*\}$ with respect to the Poincaré pairing naturally forms an admissible basis of $\hat{X}$, under the identification $\mathrm{H}^1(X)_R^\vee \cong \mathrm{H}^1(\hat{X})_R$. Let

$$\varphi^{\mathcal{P}} : \mathrm{H}^2(X)_R \to \mathrm{H}^2(\hat{X})_R$$

be the isomorphism induced by the Poincaré bundle $\mathcal{P}$ on $X \times \hat{X}$. A direct computation (see Lemma 9.3 of [29], for instance) shows that $\varphi^{\mathcal{P}}$ is nothing but

$$-\mathrm{D} : \mathrm{H}^2(X)_R \xrightarrow{\sim} \mathrm{H}^2(X)_R^\vee \cong \mathrm{H}^2(\hat{X})_R,$$

where D is Poincaré duality. For an admissible basis $\{v_i\}$ of X, its R-linear dual $\{v_i^*\}$ with respect to Poincaré pairing forms an admissible basis of $\hat{X}$. By our construction, we see that

$$\mathrm{D}(v_{12}, v_{13}, v_{14}, v_{23}, v_{24}, v_{34}) = (v_{34}^*, -v_{24}^*, v_{23}^*, v_{14}^*, -v_{13}^*, v_{12}^*),$$

which implies that D is of determinant -1 under these admissible bases. Thus the determinant of $\varphi^{\mathcal{P}}$ is not admissible.

Example 4.2.4. Let $f : X \to Y$ be an isogeny of degree d between two abelian surfaces, for some positive integer d. If d is coprime to ℓ, then it will induce an isomorphism

$$f^* : \mathrm{H}^2(Y)_{\mathbb{Z}_\ell} \xrightarrow{\sim} \mathrm{H}^2(X)_{\mathbb{Z}_\ell},$$

that is d-admissible. If, in addition, $d = n^2$, then f^*/n will be an admissible $\mathbb{Z}_\ell$-integral isometry with respect to the Poincaré pairing. If $d = k^4$, then f/k will be a $\mathbb{Z}_{(\ell)}$-isogeny such that its pullback is admissible integral.

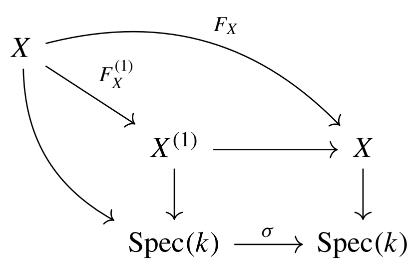
Example 4.2.5. Suppose X is an abelian surface over a perfect field k with $\operatorname{char} k = p > 0$. The F-crystal $\mathrm{H}^1(X)_W$, together with the trace map

$$\mathrm{tr}_X : \mathrm{H}^4(X)_W \xrightarrow{\sim} W,$$

forms an abelian crystal (of genus 2) in the sense of [51, §6]. We can see that an isomorphism of F-crystals $\mathrm{H}^1(X)_W \xrightarrow{\sim} \mathrm{H}^1(Y)_W$ is admissible if and only if it is an isomorphism between abelian crystals, i.e., it is compatible with trace maps.

4.3. ***More on admissible bases of F-crystals.*** In contrast to ℓ-adic étale cohomology, the semilinear structure on crystalline cohomology from its Frobenius is trickier to work with. Therefore, we spend some words on the interaction of Frobenius with admissible bases.

Suppose k is a perfect field with $\operatorname{char} k = p > 0$. We have the following Frobenius pullback diagram:



Via the natural identification $\mathrm{H}^1_{\mathrm{crys}}(X^{(1)}/W) \cong \mathrm{H}^1_{\mathrm{crys}}(X/W) \otimes_\sigma W$, the σ-linearization of Frobenius action on $\mathrm{H}^1_{\mathrm{crys}}(X/W)$ can be viewed as the injective W-linear map

$$F^{(1)} := (F_X^{(1)})^* : \mathrm{H}^1_{\mathrm{crys}}(X^{(1)}/W) \hookrightarrow \mathrm{H}^1_{\mathrm{crys}}(X/W).$$

If k is not perfect, then after passing to $W(\bar{k})$ or equivalently choosing a Frobenius lift on the Cohen ring W, we also get a Frobenius action on $\mathrm{H}^1_{\mathrm{crys}}(X/W)$, whose linearization is given by the relative Frobenius morphism.

There is a decomposition $\mathrm{H}^1_{\mathrm{crys}}(X/W) = H_0(X) \oplus H_1(X)$ such that

$$F^{(1)}\big(\mathrm{H}^1_{\mathrm{crys}}(X^{(1)}/W)\big) \cong H_0(X) \oplus pH_1(X), \tag{4.3.1}$$

and $\mathrm{rank}_W H_i = 2$ for $i = 0, 1$, which is related to the Hodge decomposition of the de Rham cohomology of X/k by Mazur's theorem; see [3, §8, Theorem 8.26].

The Frobenius map can be expressed in terms of admissible bases. We can choose an admissible basis $\{v_i\}$ of $\mathrm{H}^1_{\mathrm{crys}}(X/W)$ such that

$$v_1, v_2 \in H_0(X) \quad \text{and} \quad v_3, v_4 \in H_1(X).$$

Then $\{p^{\alpha_i} v_i\} := \{v_1, v_2, pv_3, pv_4\}$ forms an admissible basis of $\mathrm{H}^1_{\mathrm{crys}}(X^{(1)}/W)$ under the identification (4.3.1), since $\mathrm{tr}_{X^{(1)}} \circ \bigwedge^4 F^{(1)} = p^2 \sigma_W \circ \mathrm{tr}_X$. In term of these bases, the Frobenius map can be written as

$$F^{(1)}(p^{\alpha_i} v_i) = \sum_j c_{ij} p^{\alpha_j} v_j, \tag{4.3.2}$$

where the c_{ij} form an invertible 4×4-matrix $C_X = (c_{ij})$ with coefficients in W.

Suppose Y is another abelian surface over k, $\psi : \mathrm{H}^1_{\mathrm{crys}}(X/W) \to \mathrm{H}^1_{\mathrm{crys}}(Y/W)$ is an admissible map, and $\psi^{(1)}$ is the induced map $\psi \otimes_\sigma W : \mathrm{H}^1_{\mathrm{crys}}(X^{(1)}/W) \to \mathrm{H}^1_{\mathrm{crys}}(Y^{(1)}/W)$. Denote by M and M' the matrices of ψ and $\psi^{(1)}$, respectively, with respect to the chosen admissible bases.

Lemma 4.3.1. *The map ψ commutes with Frobenius if and only if $C_Y M' C_X^{-1} = M$.*

Proof. By definition, ψ commutes with Frobenius if and only if $(F_Y^{(1)})^* \circ \psi^{(1)} = \psi \circ (F_X^{(1)})^*$. The statement is then clear from (4.3.2). □

4.4. ***Generalizing Shioda's trick.*** Let us review some basic properties of the special orthogonal group scheme over an integral domain. Our main reference is [17, Appendix C].

Let Λ be an even $\mathbb{Z}$-lattice of rank $2n$. We can associate with it a vector bundle $\underline{\Lambda}$ on $\mathrm{Spec}(\mathbb{Z})$ with constant rank $2n$ equipped with a quadratic form q over $\mathrm{Spec}(\mathbb{Z})$ obtained from Λ. The functor

$$A \mapsto \{g \in \mathrm{GL}(\Lambda_A) \,|\, q_A(g \cdot x) = q_A(x) \text{ for all } x \in \Lambda_A\}$$

is representable by a $\mathbb{Z}$-subscheme of $\mathrm{GL}(\Lambda)$, denoted by $\mathrm{O}(\Lambda)$. There is a homomorphism of $\mathbb{Z}$-group schemes

$$D_\Lambda : \mathrm{O}(\Lambda) \to \underline{\mathbb{Z}/2\mathbb{Z}},$$

called the Dickson morphism (see [17, p. 313] for the definition). Roughly speaking,

$$D_\Lambda(g) = \begin{cases} 0 & \text{if } g \text{ is a product of an even number of reflections,} \\ 1 & \text{if } g \text{ is a product of an odd number of reflections,} \end{cases}$$

for a point $g \in \mathrm{O}(\Lambda)$ over a field in characteristic zero. The Dickson morphism is surjective since Λ is even, and its construction is compatible with any base change [17, Proposition C.2.8]. The *special orthogonal group scheme* over $\mathbb{Z}$ with respect to Λ is defined as the kernel of D_Λ and denoted by $\mathrm{SO}(\Lambda)$. We have

$$\mathrm{SO}(\Lambda)_{\mathbb{Z}[1/2]} \cong \ker\left(\det : \mathrm{O}(\Lambda) \to \mathbb{G}_m\right)_{\mathbb{Z}[1/2]}.$$

By [17, Proposition C.2.11], for example, $\mathrm{SO}(\Lambda) \to \mathrm{Spec}(\mathbb{Z})$ is smooth in relative dimension $\frac{1}{2}n(n-1)$ and has connected fibers.

For any ℓ, the special orthogonal group scheme

$$\mathrm{SO}(\Lambda_{\mathbb{Z}_\ell}) \cong \mathrm{SO}(\Lambda)_{\mathbb{Z}_\ell}$$

is smooth over $\mathbb{Z}_\ell$ with connected fibers, which implies that its generic fiber $\mathrm{SO}(\Lambda_{\mathbb{Q}_\ell})$ is connected. Thus, $\mathrm{SO}(\Lambda_{\mathbb{Z}_\ell})$ is connected as a group scheme over $\mathbb{Z}_\ell$, since $\mathrm{SO}(\Lambda_{\mathbb{Q}_\ell}) \subset \mathrm{SO}(\Lambda_{\mathbb{Z}_\ell})$ is dense.

The special orthogonal group scheme admits a universal covering

$$\mathrm{Spin}(\Lambda) \to \mathrm{SO}(\Lambda),$$

i.e., a simply connected central isogeny; see [17, Section C.4] for the construction.

Lemma 4.4.1. *Let* V *be free* $\mathbb{Z}$*-module of rank* 4 *and* $\Lambda = \bigwedge^2 \mathrm{V}$. *Let* R *be a ring of coefficients as listed in Section 4.2. There is an exact sequence of smooth* R*-group schemes*

$$1 \to \mu_{2,R} \to \mathrm{SL}(\mathrm{V})_R \xrightarrow{\bigwedge^2(-)_R} \mathrm{SO}(\Lambda)_R \to 1$$

(*as fppf-sheaves if* $\frac{1}{2} \notin R$). *Moreover, there is an exact sequence*

$$1 \to \{\pm \mathrm{id}_4\} \to \mathrm{SL}(\mathrm{V})(R) \xrightarrow{\bigwedge^2(-)_R} \mathrm{SO}(\Lambda)(R) \to R^*/(R^*)^2. \tag{4.4.1}$$

Proof. For the first statement, it suffices to assume $R = \mathrm{Spec}(\bar{k})$ for an algebraically closed field $\bar{k}$, where it is clear from a computation. Note that we have an exact sequence on rational points (see [25, Proposition 3.2.2])

$$1 \to \mu_2(R) \to \mathrm{SL}(\mathrm{V})(R) \to \mathrm{SO}(\Lambda)(R) \to \mathrm{H}^1(\mathrm{Spec}(R), \mu_2).$$

Notice that for the rings of coefficients listed in Section 4.2, we have $\mathrm{Pic}(R)[2] = 0$. Therefore,

$$\mathrm{H}^1_{\mathrm{fl}}(\mathrm{Spec}(R), \mu_2) \cong R^*/(R^*)^2$$

from the Kummer sequence for μ_2.

For the last statement, notice that there is an isomorphism of R-group schemes $\mathrm{SL(V)}_R \xrightarrow{\sim} \mathrm{Spin}(\Lambda)_R$ such that the diagram

$$\begin{array}{ccc} \mathrm{SL}(V)(R) & \xrightarrow{\sim} & \mathrm{Spin}(\Lambda)(R) \\ & \searrow \quad \swarrow & \\ & \mathrm{SO}(\Lambda)(R) & \end{array}$$

commutes. The group scheme SL(V) is simply connected (its geometric fibers form a semisimple algebraic group of type A_3). Thus, the central isogeny $\mathrm{SL(V)}_R \to \mathrm{SO}(\Lambda)_R$ forms the universal covering of $\mathrm{SO}(\Lambda)_R$, which induces an isomorphism $\mathrm{SL(V)}_R \xrightarrow{\sim} \mathrm{Spin}(\Lambda)_R$ by using the isomorphism theorem over a general ring (see [17, Theorems 6.1.16 and 6.1.17], for example). □

Remark 4.4.2. When $R = \mathbb{Z}_\ell$, we have

$$\mathbb{Z}_\ell^*/(\mathbb{Z}_\ell^*)^2 \cong \begin{cases} \{\pm 1\} & \text{if } \ell \neq 2, \\ \{\pm 1\} \times \{\pm 5\} & \text{if } \ell = 2. \end{cases}$$

Thus the image of $\mathrm{SL(V)}(\mathbb{Z}_\ell)$ is a finite-index subgroup in $\mathrm{SO}(\Lambda)(\mathbb{Z}_\ell)$.

Remark 4.4.3. When $R = W(k)$, we have

$$W(k)^*/(W(k)^*)^2 \cong \begin{cases} \{1, \epsilon\} & \text{if } k = \mathbb{F}_{p^s} \text{ for } p > 2, s \geq 1, \\ \{1\} & \text{if } k = \bar{k} \text{ or } k^s = k, \operatorname{char} k > 2, \end{cases}$$

where $\epsilon \in \mathbb{Z}$ is such that $\epsilon \not\equiv y^2 \mod p^s$ for an integer y, as $W(k)$ is Henselian. Thus, the wedge map $\mathrm{SL(V)}(W) \to \mathrm{SO}(\Lambda)(W)$ is surjective when $k = \bar{k}$.

Let X and Y be abelian surfaces over k. Let $V_R = \mathrm{H}^1(X)_R$. We can see the set

$$\mathrm{Isom}^{\mathrm{ad},(d)}(\mathrm{H}^1(X)_R, \mathrm{H}^1(Y)_R)$$

is a (right) $\mathrm{SL}(V_R)$-torsor if it is nonempty. The wedge product provides a natural map

$$\textstyle\bigwedge^2 : \mathrm{Isom}^{\mathrm{ad},(d)}\big(\mathrm{H}^1(X)_R, \mathrm{H}^1(Y)_R\big) \to \mathrm{Isom}^{\mathrm{ad},(d)}\big(\mathrm{H}^2(X)_R, \mathrm{H}^2(Y)_R\big).$$

Let $\{v_i\}$ be an admissible basis of $\mathrm{H}^1(X)_R$ and let $\{v_i'\}$ be a d-admissible basis of $\mathrm{H}^1(Y)_R$. There is a d-admissible isomorphism $\psi_0 \in \mathrm{Isom}^{\mathrm{ad},(d)}(\mathrm{H}^1(X)_R, \mathrm{H}^1(Y)_R)$ such that $\psi_0(v_i) = v_i'$. For a d-admissible isometry $\varphi : \mathrm{H}^2(X, R) \to \mathrm{H}^2(Y, R)$, we can see that

$$\varphi = \textstyle\bigwedge^2(\psi_0) \circ g \quad \text{for some } g \in \mathrm{SO}(\Lambda_R).$$

In this way, any d-admissible isomorphism φ can be identified with the (unique) element $g \in \mathrm{SO}(\Lambda)(R)$ when the admissible bases are fixed. This allows us to deal with d-admissible isomorphisms group-theoretically. In particular, we have the following notion:

Definition 4.4.4. The *spinor norm* of the d-admissible isomorphism φ is the image $\mathrm{SN}(\varphi)$ of g under $\mathrm{SN} : \mathrm{SO}(\Lambda)(R) \to R^*/(R^*)^2$.

Lemma 4.4.5. *The spinor norm* $\mathrm{SN}(\varphi)$ *is independent of the choice of admissible bases.*

Proof. With a different choice of admissible bases, the new element satisfies $\tilde{g} = KgK^{-1}$, for some $K \in \mathrm{SO}(\Lambda_R)$. Therefore, $\mathrm{SN}(\tilde{g}) = \mathrm{SN}(g)$. □

Remark 4.4.6. When R is a field, the spinor norm can be computed by the Cartan–Dieudonné decomposition. That is, we can write any $g \in \mathrm{SO}(\Lambda)(R)$ as a composition of reflections

$$\boldsymbol{R}_{b_n} \circ \boldsymbol{R}_{b_{n-1}} \circ \cdots \circ \boldsymbol{R}_{b_1}$$

for some nonisotropic vectors $b_1, \dots, b_n \in \Lambda_R$, and $\mathrm{SN}(g) = \big[(b_1)^2 \cdots (b_{n-1})^2 (b_n)^2\big]$.

Lemma 4.4.7. *A d-admissible isomorphism φ is a wedge of some d-admissible isomorphism ψ : $\mathrm{H}^1(X, R) \to \mathrm{H}^1(Y, R)$ if and only if* $\mathrm{SN}(\varphi) = 1$.

Proof. The exact sequence (4.4.1) shows that if $\mathrm{SN}(\varphi) = \mathrm{SN}(g) = 1$, there is some $h \in \mathrm{SL}(V_R)$ such that $\bigwedge^2(h) = g$. Thus, we can take $\psi = \psi_0 \circ h$ when $\mathrm{SN}(\varphi) = 1$, and see that

$$\textstyle\bigwedge^2(\psi) = \bigwedge^2(\psi_0) \circ \bigwedge^2(h) = \varphi.$$

The converse is clear. □

4.4.1. *The isogeny category.* Recall that the isogeny category of abelian varieties $\mathtt{AV}_{\mathbb{Q},k}$ has as objects all abelian varieties over a field k, and as arrows the sets of homomorphisms

$$\mathrm{Hom}_{\mathtt{AV}_{\mathbb{Q},k}}(X, Y) := \mathrm{Hom}_{\mathtt{AV}_k}(X, Y) \otimes_{\mathbb{Z}} \mathbb{Q},$$

where $\mathrm{Hom}_{\mathtt{AV}_k}(X, Y)$ is the abelian group of homomorphisms from X to Y with the natural addition. We may also write $\mathrm{Hom}^0(X, Y)$ for $\mathrm{Hom}_{\mathtt{AV}_{\mathbb{Q},k}}(X, Y)$ if there is no confusion about k.

Definition 4.4.8. Let R be a commutative ring with unit. An *R-isogeny* from X to Y is an invertible element $f \in \mathrm{Hom}_{\mathtt{AV}_k}(X, Y) \otimes R$ i.e., there is $g \in \mathrm{Hom}_{\mathtt{AV}_k}(Y, X) \otimes R$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

A $\mathbb{Q}$-isogeny is called a quasi-isogeny, while a $\mathbb{Z}_{(\ell)}$-isogeny is called a *prime-to-ℓ quasi-isogeny*. For any quasi-isogeny (resp. prime-to-ℓ quasi-isogeny) f, we can find a minimal integer n (resp. one with $\ell \nmid n$) such that

$$nf : X \to Y$$

is an isogeny (resp. of degree prime-to-ℓ).

When $k = \mathbb{C}$, with the uniformization of complex abelian varieties, we have a canonical bijection

$$\mathrm{Hom}_{\mathtt{AV}_{\mathbb{Q},\mathbb{C}}}(X, Y) \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{Hdg}}\big(\mathrm{H}^1(Y, \mathbb{Q}), \mathrm{H}^1(X, \mathbb{Q})\big),$$

where the right-hand side is the set of $\mathbb{Q}$-linear morphisms that preserve Hodge structures. Then the integer n for f is also the minimal integer such that $(nf)^*(\mathrm{H}^1(Y, \mathbb{Z})) \subseteq \mathrm{H}^1(X, \mathbb{Z})$.

4.5. ***Shioda's trick for Hodge isogenies.*** Suppose $k = \mathbb{C}$. Let d be an integer. A *Hodge isogeny of degree d* is an isomorphism of $\mathbb{Q}$-Hodge structures

$$\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$$

such that

$$\langle x, y\rangle = d\langle \varphi(x), \varphi(y)\rangle.$$

When $d = 1$, we recover the classical Hodge isometry. Clearly, a d-admissible rational Hodge isomorphism is a Hodge isogeny of degree d. In terms of spinor norms, we can generalize Shioda's theorem (Theorem 4.1.1) to admissible rational Hodge isogenies.

Proposition 4.5.1 (Shioda's trick on admissible Hodge isogenies).

(1) *A d-admissible Hodge isogeny of degree d*

$$\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$$

is a wedge of some rational Hodge isomorphism $\psi : \mathrm{H}^1(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^1(Y, \mathbb{Q})$ if its spinor norm is trivial. In this case, the Hodge isogeny is induced by a quasi-isogeny of degree d^2.

(2) *When $d = 1$, any admissible Hodge isometry $\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$ is induced by an isogeny $f : Y \to X$ of degree n^2 for some integer n such that $\varphi = f^*/n$.*

Proof. Under the assumption of part (1), we can find a d-admissible isomorphism ψ by applying Lemma 4.4.7. It remains to prove that ψ preserves the Hodge structure; this is done essentially as for [63, Theorem 1].

For (2), we suppose the spinor norm $\mathrm{SN}(\varphi)$ equals $n\mathbb{Q}^{*2} \in \mathbb{Q}^*/\mathbb{Q}^{*2}$. Let $E = \mathbb{Q}(\sqrt{n})$. We can see that the base change $\mathrm{H}^2(X, E) \xrightarrow{\sim} \mathrm{H}^2(Y, E)$ is a Hodge isometry with coefficients in E such that $\mathrm{SN}(\varphi) = 1 \in E^*/(E^*)^2$. By applying Lemma 4.4.7, we obtain an admissible Hodge isomorphism $\psi : \mathrm{H}^1(X, E) \xrightarrow{\sim} \mathrm{H}^1(Y, E)$ (for some admissible bases for $\mathrm{H}^1(X, \mathbb{Q})$ and $\mathrm{H}^1(Y, \mathbb{Q})$). Let

$$\sigma : a + b\sqrt{n} \rightsquigarrow a - b\sqrt{n}$$

be the generator of $\mathrm{Gal}(E/\mathbb{Q})$. As we have fixed the $\mathbb{Q}$-linear admissible bases, the wedge map

$$\mathrm{SL}_4(E) \xrightarrow{\bigwedge^2} \mathrm{SO}(\Lambda)(E)$$

is defined over $\mathbb{Q}$, and so is σ-equivariant. Let g be the element in $\mathrm{SL}_4(E)$ that corresponds to ψ. Since $\bigwedge^2(g) \in \mathrm{SO}(\Lambda) \subset \mathrm{SO}(\Lambda_E)$, we see that

$$\textstyle\bigwedge^2(\sigma(g)) = \sigma\big(\bigwedge^2(g)\big) = \bigwedge^2(g),$$

which implies that $\sigma(g)g^{-1} = \pm\,\mathrm{id}_4$ since $\ker \bigwedge^2 = \{\pm\,\mathrm{id}_4\}$. If $\sigma(g) = g$, then $g \in \mathrm{SL}_4(\mathbb{Q})$ and the statement is trivially valid. If $\sigma(g) = -g$, then $g_0 = \sqrt{n}g$ lies in $\mathrm{GL}_4(\mathbb{Q})$. Let

$$\psi_0 : \mathrm{H}^1(X, \mathbb{Q}) \to \mathrm{H}^1(Y, \mathbb{Q})$$

be the element corresponding to g_0 in $\mathrm{Isom}^{\mathrm{ad},(n^2)}\big(\mathrm{H}^1(X,\mathbb{Q}),\mathrm{H}^1(Y,\mathbb{Q})\big)$. As $\bigwedge^2\psi_0 = n\varphi$ is a Hodge isogeny, part (1) then implies that ψ_0 is also a Hodge isomorphism. Thus, ψ_0 extends to a quasi-isogeny $f_0 : Y \to X$ and we have

$$\varphi = \bigwedge^2(\psi) = \frac{f_0^*}{n} : \mathrm{H}^2(X,\mathbb{Q}) \to \mathrm{H}^2(Y,\mathbb{Q}).$$

Replacing f_0 by the product mf_0 for some integer m, we get an isogeny of degree $(m^2n)^2$. □

Remark 4.5.2. If a Hodge isometry $\psi : \mathrm{H}^2(X,\mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y,\mathbb{Q})$ is not admissible, that is, if its determinant is -1 with respect to some admissible bases, then we can take its composition with the isometry $\psi^{\mathcal{P}}$ induced by the Poincaré bundle as in Example 4.2.3. After that, we see that $\psi^{\mathcal{P}} \circ \psi$ is admissible and is induced by an isogeny $f : \hat{Y} \to X$.

4.6. *ℓ-adic and p-adic Shioda's tricks.* For the integral ℓ-adic étale cohomology, we have the following statement similar to Shioda's trick for integral Betti cohomology.

Proposition 4.6.1 (ℓ-adic Shioda's trick). *Suppose $\ell \neq 2$. For any d-admissible $\mathbb{Z}_\ell$-linear isomorphism*

$$\varphi_\ell : \mathrm{H}^2_{\acute{e}t}(Y_{k^s},\mathbb{Z}_\ell) \xrightarrow{\sim} \mathrm{H}^2_{\acute{e}t}(X_{k^s},\mathbb{Z}_\ell),$$

there are an integer u and a (u^2d)-admissible $\mathbb{Z}_\ell$-isomorphism ψ_ℓ such that $\bigwedge^2(\psi_\ell) = u\varphi_\ell$. If φ_ℓ is $\mathrm{Gal}(k^s/k)$*-equivariant, then ψ_ℓ is also* $\mathrm{Gal}(k^s/k)$*-equivariant after replacing k with some finite extension.*

Proof. One can choose an element $u \in (\mathbb{Z}\setminus\{0\}) \cap \mathbb{Z}_\ell^*$ that is not a square in $\mathbb{Z}_\ell$, e.g., one satisfying the equation $u^{\frac{\ell-1}{2}} \equiv -1 \bmod \ell$ as $\ell \neq 2$. Since $\mathbb{Z}_\ell^*/(\mathbb{Z}_\ell^*)^2 \cong \{\pm 1\}$ for any $\ell \neq 2$, either φ_ℓ or $u\varphi_\ell$ has spinor norm one. Then the first statement follows from Lemma 4.4.7.

Suppose φ_ℓ is $\mathrm{Gal}(k^s/k)$-equivariant. We may assume $\bigwedge^2(\psi_\ell) = \varphi_\ell$ for simplicity. For any $g \in \mathrm{Gal}(k^s/k)$, we have

$$\bigwedge^2(g^{-1}\psi_\ell g) = g^{-1}\bigwedge^2(\psi_\ell)g = \bigwedge^2(\psi_\ell).$$

Therefore, $g^{-1}\psi_\ell g = \pm\psi_\ell$. By passing to a finite extension k'/k, we always have $g^{-1}\psi_\ell g = \psi_\ell$ for all $g \in \mathrm{Gal}(k^s/k')$, which proves the assertion. □

For F-crystals attached to abelian surfaces, we have another variant of Shioda's trick.

Proposition 4.6.2 (p-adic Shioda's trick). *Let k be a finite field or an algebraically closed field such that* $\mathrm{char}\, k = p > 2$. *For any d-admissible W-linear isomorphism*

$$\varphi_p : \mathrm{H}^2_{\mathrm{crys}}(Y/W) \xrightarrow{\sim} \mathrm{H}^2_{\mathrm{crys}}(X/W),$$

there exist an integer u and a (u^2d)-admissible W-linear isomorphism $\psi_p : \mathrm{H}^1_{\mathrm{crys}}(Y/W) \xrightarrow{\sim} \mathrm{H}^1_{\mathrm{crys}}(X/W)$ such that $\bigwedge^2(\psi_p) = u\varphi_p$. If k is algebraically closed, we can take $u = 1$.

Moreover, if φ_p is compatible with Frobenius and $\mathbb{F}_{p^2} \subseteq k$, then there is $\xi \in \mathbb{Z}_{p^2}^* \subseteq W(k)$ such that $\xi\psi_p$ is compatible with Frobenius and $\xi^2 \in \mathbb{Z}_p^*$.

Proof. The first statement follows from a similar reasoning as Proposition 4.6.1, since $W^*/(W^*)^2 \subseteq \{1, \epsilon\}$ (see Remark 4.4.3).

For the second statement, we assume $\bigwedge^2(\psi_p) = \varphi_p$. If φ_p commutes with the Frobenius action, then

$$\bigwedge^2(C_X^{-1} \cdot \psi_p^{(1)} \cdot C_Y) = \varphi_p.$$

as in Section 4.3. Thus $C_X^{-1} \cdot \psi_p^{(1)} \cdot C_Y = \pm\psi_p^{(1)}$, which implies

$$\psi_p \circ F_X^{(1)} = \pm F_Y^{(1)} \circ \psi_p^{(1)}$$

by Lemma 4.3.1.

If $F_X^{(1)} \circ \psi_p^{(1)} = \psi_p \circ F_Y^{(1)}$, we need do nothing. If $F_X^{(1)} \circ \psi_p^{(1)} = -\psi_p \circ F_Y^{(1)}$, we can take $\xi \in \mathbb{Z}_{p^2}^* \subseteq W(k)$ such that $\xi^{p-1} = -1$. This implies

$$F_X^{(1)} \circ (\xi\psi_p)^{(1)} = \xi^p F_X^{(1)} \circ \psi = (\xi\psi_p) \circ F_Y^{(1)}.$$

Note that $\xi^2 \in \mathbb{Z}_p^*$ as $\sigma(\xi^2) = \xi^2$ and $\xi^{2p+2} = 1$. Therefore, we can conclude. □

Combining this with Tate's isogeny theorem, we have the following consequences of Propositions 4.6.1 and 4.6.2. They include a special case of Tate's conjecture.

Corollary 4.6.3. *Suppose k is a finitely generated field over $\mathbb{F}_p$ with $p > 2$. Let $\ell \neq 2$ be a prime not equal to p.*

(1) *For any admissible isometry of* $\mathrm{Gal}(k^s/k)$*-modules*

$$\varphi_\ell : \mathrm{H}^2_{ét}(Y_{k^s}, \mathbb{Z}_\ell) \xrightarrow{\sim} \mathrm{H}^2_{ét}(X_{k^s}, \mathbb{Z}_\ell),$$

we can find a $\mathbb{Z}_\ell$-isogeny $f_\ell \in \mathrm{Hom}_{k'}(X_{k'}, Y_{k'}) \otimes \mathbb{Z}_\ell$, for some finite extension k'/k, that induces $u\varphi_\ell$ for some integer u prime-to-ℓ. In particular, φ_ℓ is algebraic.

(2) *If k is finite, then for any admissible isometry*

$$\varphi_p : \mathrm{H}^2_{\mathrm{crys}}(Y/W) \xrightarrow{\sim} \mathrm{H}^2_{\mathrm{crys}}(X/W)$$

that is compatible with Frobenius, we can find a $\mathbb{Z}_p$-isogeny $f_p \in \mathrm{Hom}_{k'}(X_{k'}, Y_{k'}) \otimes \mathbb{Z}_p$ over some finite extension k'/k, such that

$$\epsilon f_p^*|_{\mathrm{H}^2_{\mathrm{crys}}(Y/W)} = u\varphi_p$$

for some prime-to-p integer u and $\epsilon \in \mathbb{Z}_p^*$. In particular, φ_p is algebraic.

Proof. For (1), Proposition 4.6.1 implies that there is an isomorphism

$$\psi_\ell : \mathrm{H}^1_{ét}(Y_{k^s}, \mathbb{Z}_\ell) \xrightarrow{\sim} \mathrm{H}^1_{ét}(X_{k^s}, \mathbb{Z}_\ell),$$

that induces $u\varphi_\ell$ and is $\mathrm{Gal}(k^s/k)$-equivariant after a finite extension of k. Then f_ℓ exists thanks to the canonical bijection

$$\mathrm{Hom}^0(X, Y) \otimes \mathbb{Z}_\ell \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{Gal}(k^s/k)}\left(\mathrm{H}^1_{ét}(Y_{k^s}, \mathbb{Z}_\ell), \mathrm{H}^1_{ét}(X_{k^s}, \mathbb{Z}_\ell)\right)$$

(see [72] and [20, VI.3, Theorem 1]).

For (2), we may assume that $\mathbb{Z}_{p^2} \subseteq W(k)$, after taking a finite extension of k. Then Proposition 4.6.2 implies that there is an isomorphism

$$\psi_p : \mathrm{H}^1_{\mathrm{crys}}(Y/W) \xrightarrow{\sim} \mathrm{H}^1_{\mathrm{crys}}(X/W)$$

that induces $u\varphi_p$, and a $\xi \in \mathbb{Z}^*_{p^2}$ such that $\xi\psi_p$ is compatible with Frobenius.

Since k a finite field, there are canonical isomorphisms

$$\mathrm{Hom}^0(X, Y) \otimes \mathbb{Z}_p \xrightarrow{\sim} \mathrm{Hom}_k\big(X[p^\infty], Y[p^\infty]\big) \xrightarrow{\sim} \mathrm{Hom}_{F,V}\big(\mathrm{H}^1_{\mathrm{crys}}(Y/W), \mathrm{H}^1_{\mathrm{crys}}(X/W)\big). \tag{4.6.1}$$

Here the first isomorphism is from p-adic Tate's isogeny theorem [36, Theorem 2.6] and the second from the faithfulness of Dieudonné functor over W [35, Theorem]. The canonical bijection (4.6.1) implies that $\xi\psi_p$ is induced by a $\mathbb{Z}_p$-isogeny $f_p \in \mathrm{Hom}^0(X, Y) \otimes \mathbb{Z}_p$. Therefore

$$f_p^*|_{\mathrm{H}^2_{\mathrm{crys}}(Y/W)} = \xi^2 u\varphi_p.$$

The $\mathbb{Z}_p$-isogeny f_p is what we require. □

Remark 4.6.4. In [73], Zarhin introduced the notion of *almost isomorphisms*. Two abelian varieties over k are called almost isomorphic if their Tate modules T_ℓ are isomorphic as Galois modules (replaced by p-divisible groups when $\ell = p$). Propositions 4.6.1 and 4.6.2 imply that it is possible to characterize almost isomorphic abelian surfaces by their second cohomology groups.

5. Derived isogeny in characteristic zero

In this section, we follow [23] and [31] to prove the twisted Torelli theorem for abelian surfaces over algebraically closed fields of characteristic zero.

5.1. ***Over $\mathbb{C}$: Hodge isogeny versus derived isogeny.*** Let X and Y be complex abelian surfaces. Throughout this section, let $\Lambda = U^{\oplus 3}$ be the direct sum of three hyperbolic lattices.

Definition 5.1.1. A rational Hodge isometry $\varphi : \mathrm{H}^2(X, \mathbb{Q}) \to \mathrm{H}^2(Y, \mathbb{Q})$ is called *reflective* if it is a reflection on Λ along a nonisotropic vector $b \in \Lambda$:

$$\boldsymbol{R}_b : \Lambda_{\mathbb{Q}} \xrightarrow{\sim} \Lambda_{\mathbb{Q}}, \quad x \mapsto x - \frac{2(x, b)}{(b, b)} b,$$

after one chooses markings $\mathrm{H}^2(X, \mathbb{Z}) \cong \Lambda$ and $\mathrm{H}^2(Y, \mathbb{Z}) \cong \Lambda$.

Lemma 5.1.2. *Any reflective Hodge isometry φ induces a Hodge isometry on twisted Mukai lattices*

$$\tilde{\varphi} : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B) \to \widetilde{\mathrm{H}}(Y, \mathbb{Z}; B'),$$

for some $B \in \mathrm{H}^2(X, \mathbb{Q})$ and $B' = -\varphi(B)$ such that the restriction of $\tilde{\varphi}_{\mathbb{Q}} : \widetilde{\mathrm{H}}(X, \mathbb{Q}) \xrightarrow{\sim} \widetilde{\mathrm{H}}(Y, \mathbb{Q})$ on $\mathrm{H}^2(X, \mathbb{Q})$ is equal to φ.

Proof. We use [31, §1.2]. The discussion there involves a purely linear-algebraic argument for twisted Mukai lattices, so it works for abelian surfaces without changes. Let us recall the construction of $\tilde{\varphi}$. By definition, there are markings $f : \mathrm{H}^2(X, \mathbb{Z}) \cong \Lambda$ and $g : \mathrm{H}^2(Y, \mathbb{Z}) \cong \Lambda$ such that the composition

$$\Lambda_{\mathbb{Q}} \xrightarrow{f^{-1}} \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\varphi} \mathrm{H}^2(Y, \mathbb{Q}) \xrightarrow{g} \Lambda_{\mathbb{Q}}$$

is a reflection $\boldsymbol{R}_b$, with $b \in \Lambda$ a primitive vector.

Let $B = f^{-1}(b)/n \in \mathrm{H}^2(X, \mathbb{Q})$ and $B' = g^{-1}(b)/n \in \mathrm{H}^2(Y, \mathbb{Q})$, where $n = \frac{1}{2}b^2$. The map

$$\tilde{\varphi} : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B) \to \widetilde{\mathrm{H}}(Y, \mathbb{Z}; B'),$$

defined by sending a vector (r, c, s) to $\big(n(B, c)-r-ns,\ \varphi(c)-n((B, c)-s)B',\ -s\big)$ is a Hodge isometry. We have $\tilde{\varphi}((0, c, (B, c))) = (0, \varphi(c), (B', \varphi(c)))$ and $\tilde{\varphi}((0, 0, 1)) = (-n, -nB', -1)$, which gives the last assertion. □

The following result characterizes reflective Hodge isometries between abelian surfaces. The idea of the proof is based on [31, Theorem 1.1], with some necessary modifications for abelian surfaces.

Theorem 5.1.3. *Let X and Y be two complex abelian surfaces. If there is a reflective Hodge isometry*

$$\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q}),$$

then, up to sign, φ is induced (in the sense of Section 3.1) *by a derived isogeny*

$$\mathrm{D}^b(X) \sim \mathrm{D}^b(Y). \tag{5.1.1}$$

Proof. According to Lemma 5.1.2, there is a Hodge isometry

$$\tilde{\varphi} : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B) \xrightarrow{\sim} \widetilde{\mathrm{H}}(Y, \mathbb{Z}; B')$$

whose restriction on $\mathrm{H}^2(X, \mathbb{Q})$ is just φ. Let $v_{B'} = (-n, -nB', -1)$ be the image of the Mukai vector $(0, 0, 1)$ under $\tilde{\varphi}$. From our construction, the Mukai vector

$$v = \exp(-B') \cdot v_{B'} = (-n, 0, 0) \in \widetilde{\mathrm{H}}(Y, \mathbb{Z})$$

satisfies $v_{B'} = \exp(B') \cdot v$. We can assume that v is positive (see Definition 3.5.1) up to a shift of $\mathrm{D}^{(1)}(\mathscr{Y})$.

Let $\mathscr{Y} \to Y$ be a $\mathbb{G}_m$-gerbe that admits a $\boldsymbol{B}$-field lift B'. For some v-generic polarization H, the moduli stack $\mathscr{M}_H(\mathscr{Y}, v)$ of $\mathscr{Y}$-twisted sheaves on Y with Mukai vector v forms a $\mathbb{G}_m$-gerbe on its coarse moduli space $M_H(\mathscr{Y}, v)$. Let $\mathcal{E}$ be a universal $(1, 1)$-twisted sheaf on $\mathscr{Y} \times \mathscr{M}_H(\mathscr{Y}, v)$. It induces a twisted Fourier–Mukai transform

$$\Phi^{\mathcal{E}} : \mathrm{D}^{(-1)}(\mathscr{M}_H(\mathscr{Y}, v)) \to \mathrm{D}^{(1)}(\mathscr{Y})$$

(see [71, Theorem 4.3]) and a Hodge isometry

$$\varphi^{\mathcal{E}} : \widetilde{\mathrm{H}}(M_H(\mathscr{Y}, v), \mathbb{Z}; B'') \xrightarrow{\sim} \widetilde{\mathrm{H}}(Y, \mathbb{Z}; B'),$$

where B'' is a $\boldsymbol{B}$-field lift of $\mathscr{M}_H(\mathscr{Y}, v)^{(-1)} \to M_H(\mathscr{Y}, v)$. The composition

$$(\varphi^{\mathcal{E}})^{-1} \circ \tilde{\varphi} : \widetilde{\mathrm{H}}(X, \mathbb{Z}; B) \xrightarrow{\sim} \widetilde{\mathrm{H}}(M_H(\mathscr{Y}, v), \mathbb{Z}; B''), \tag{5.1.2}$$

defines a Hodge isometry that maps the Mukai vector (0, 0, 1) to (0, 0, 1) and preserves the Mukai pairing; it also sends (1, 0, 0) to $(1, b, \frac{b^2}{2})$ for some $b \in \mathrm{H}^2(Y, \mathbb{Z})$. Changing B'' to $B'' + b$, one obtains a Hodge isometry that maps (1, 0, 0) to (1, 0, 0) and (0, 0, 1) to (0, 0, 1). This restricts to a Hodge isometry

$$\mathrm{H}^2(X, \mathbb{Z}) \xrightarrow{\sim} \mathrm{H}^2(M_{H'}(\mathscr{Y}, v), \mathbb{Z}). \tag{5.1.3}$$

If the Hodge isometry (5.1.3) is admissible, we can apply Shioda's Torelli theorem for abelian surfaces (Theorem 4.1.1) to conclude that there is an isomorphism

$$f : M_{H'}(\mathscr{Y}, v) \xrightarrow{\sim} X$$

such that $(\varphi^{\mathcal{E}})^{-1} \circ \tilde{\varphi} = f^*$ up to sign. Take $\mathscr{X} \to X$ as the $\mathbb{G}_m$-gerbe $\mathscr{M}_{H'}(\mathscr{Y}, v)^{(-1)} \to M_{H'}(\mathscr{Y}, v)$. Then the Hodge realization of the derived equivalence

$$\Phi^{\mathcal{E}} \circ f^* : \mathrm{D}^{(1)}(\mathscr{X}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y}) \tag{5.1.4}$$

is $\tilde{\varphi}$ up to sign.

Otherwise, the composition

$$\mathrm{H}^2(\hat{X}, \mathbb{Z}) \xrightarrow{-\mathrm{D}} \mathrm{H}^2(X, \mathbb{Z}) \xrightarrow{\sim} \mathrm{H}^2(M_H(\mathscr{Y}, v), \mathbb{Z})$$

is admissible, as explained in Example 4.2.3, and it can be realized as the pullback under an isomorphism $f : M_H(\mathscr{Y}, v) \xrightarrow{\sim} \hat{X}$ up to sign. Thus, the Hodge realization of the derived equivalence $f^* \circ \Phi^{\mathcal{P}} : \mathrm{D}^b(X) \xrightarrow{\sim} \mathrm{D}^b(M_H(\mathscr{Y}, v))$ yields the Hodge isometry (5.1.3), where $\mathcal{P}$ is the Poincaré bundle. We can consider the derived isogeny

$$\begin{aligned} &\mathrm{D}^b(X) \xrightarrow{f^* \circ \Phi^{\mathcal{P}}} \mathrm{D}^b(M_H(\mathscr{Y}, v)), \\ &\qquad\qquad \mathrm{D}^{(-1)}(\mathscr{M}_H(\mathscr{Y}, v)) \xrightarrow{\Phi^{\mathcal{E}}} \mathrm{D}^{(1)}(\mathscr{Y}). \end{aligned} \tag{5.1.5}$$

From the construction, its rational Hodge realization on second cohomology yields φ up to sign. □

Remark 5.1.4. If φ is induced from a reflection of a vector with norm $2n$, let $\mathscr{X} \to X$ and $\mathscr{Y} \to Y$ be the equivalent twisted abelian surfaces obtained in Theorem 5.1.3. Then we have

$$[\mathscr{X}]^n = \exp(nB) = 1 \in \mathrm{Br}(X),$$

which implies $[\mathscr{X}] \in \mathrm{Br}(X)[n]$. Similarly, the order of $[\mathscr{Y}]$ divides n.

Next we show that any rational Hodge isometry can be decomposed into a chain of reflective Hodge isometries. This is a special case of the Cartan–Dieudonné theorem that says that any element $g \in \mathrm{SO}(\Lambda_{\mathbb{Q}})$

can be decomposed as a product of reflections:

$$g = \boldsymbol{R}_{b_1} \circ \boldsymbol{R}_{b_2} \circ \cdots \circ \boldsymbol{R}_{b_n}, \tag{5.1.6}$$

such that $b_i \in \Lambda$ and $(b_i)^2 \neq 0$ for each i. From the surjectivity of the period map [63, Theorem II], for any rational Hodge isometry $\mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$, we can find a sequence of abelian surfaces $\{X_i\}$ with Λ-markings and Hodge isometries

$$\varphi_i : \mathrm{H}^2(X_{i-1}, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(X_i, \mathbb{Q}),$$

where $X_0 = X$ and $X_n = Y$, such that φ_i is induced by some reflection $\boldsymbol{R}_{b_i} \in \mathrm{O}(\Lambda \otimes \mathbb{Q})$. We can arrange them as in (1.1.1):

$$\begin{aligned} &\mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\varphi_1} \mathrm{H}^2(X_1, \mathbb{Q}), \\ &\qquad \mathrm{H}^2(X_1, \mathbb{Q}) \xrightarrow{\varphi_2} \mathrm{H}^2(X_2, \mathbb{Q}), \\ &\qquad\qquad \vdots \\ &\qquad\qquad \mathrm{H}^2(X_{n-1}, \mathbb{Q}) \xrightarrow{\varphi_n} \mathrm{H}^2(Y, \mathbb{Q}). \end{aligned} \tag{5.1.7}$$

Corollary 5.1.5. *If there is a rational Hodge isometry* $\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$, *then there is a derived isogeny from X to Y that φ up to sign as in* (5.1.7).

Remark 5.1.6. One application of Corollary 5.1.5 is that any rational Hodge isometry between abelian surfaces is algebraic, which is a special case of the Hodge conjecture for products of two abelian surfaces. Unlike the case of K3 surfaces, the Hodge conjecture for products of abelian surfaces has been known for a long time; see, for example, [60, Theorem 3.15].

Corollary 5.1.7. *There is a rational Hodge isometry* $\mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$ *if and only if there is a derived isogeny from* $\mathrm{Km}(X)$ *to* $\mathrm{Km}(Y)$.

Proof. Any rational Hodge isometry induces a rational isometry of Néron–Severi lattices $\mathrm{NS}(X)_{\mathbb{Q}} \simeq \mathrm{NS}(Y)_{\mathbb{Q}}$. Let $\mathrm{T}(-)$ be the transcendental part of $\mathrm{H}^2(-)$. Applying Witt's cancellation theorem, we get

$$\mathrm{H}^2(X, \mathbb{Q}) \simeq \mathrm{H}^2(Y, \mathbb{Q}) \iff \mathrm{T}(X)_{\mathbb{Q}} \simeq \mathrm{T}(Y)_{\mathbb{Q}}$$

as Hodge isometries. According to [31, Theorem 0.1], $\mathrm{Km}(X)$ is derived isogenous to $\mathrm{Km}(Y)$ if and only if there is a Hodge isometry $\mathrm{T}(\mathrm{Km}(X))_{\mathbb{Q}} \simeq \mathrm{T}(\mathrm{Km}(Y))_{\mathbb{Q}}$. Then the statement is clear from the fact that there is a canonical integral Hodge isometry $\mathrm{T}(X)(2) \simeq \mathrm{T}(\mathrm{Km}(X))$, by [48, Proposition 4.3(i)]. □

5.2. *Prime-to-ℓ Hodge isometries.*

Definition 5.2.1. We say that a rational Hodge isometry

$$\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$$

is *prime-to-ℓ* if it descends to an isometry $\mathrm{H}^2(X, \mathbb{Z}_{(\ell)}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Z}_{(\ell)})$.

Lemma 5.2.2. *Assume $\varphi : \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(Y, \mathbb{Q})$ is a reflective Hodge isometry, induced by a primitive vector $b \in \Lambda$. Then φ is prime-to-ℓ if and only if $\ell \nmid n = \frac{(b)^2}{2}$.*

Proof. One direction is obvious. For the other, suppose φ is prime-to-ℓ. By definition, there are markings $\mathrm{H}^2(X, \mathbb{Z}) \cong \Lambda$ and $\mathrm{H}^2(X, \mathbb{Z}) \cong \Lambda$ such that the isometry

$$\Lambda \otimes \mathbb{Q} \cong \mathrm{H}^2(X, \mathbb{Q}) \xrightarrow{\varphi} \mathrm{H}^2(Y, \mathbb{Q}) \cong \Lambda \otimes \mathbb{Q}$$

is the reflection $\boldsymbol{R}_b \in \mathrm{O}(\Lambda \otimes \mathbb{Q})$. As φ is prime-to-ℓ, the reflection $\boldsymbol{R}_b$ lies in $\mathrm{O}(\Lambda \otimes \mathbb{Z}_{(\ell)})$.

If $\ell \mid n$, one must have $\ell \mid (x, b)$ for any $x \in \Lambda$. However, this is contradictory, as Λ is unimodular and any primitive vector has divisibility 1. □

Another useful tool is as follows.

Lemma 5.2.3 (prime-to-ℓ Cartan–Dieudonné decomposition). *Let Λ be an integral lattice over $\mathbb{Z}$ whose reduction mod ℓ is still nondegenerate. Any orthogonal matrix $A \in \mathrm{O}(\Lambda)(\mathbb{Z}_{(\ell)}) \subset \mathrm{O}(\Lambda)(\mathbb{Q})$, with$\ell > 2$, can be decomposed into a sequence of prime-to-ℓ reflections.*

Proof. We follow the proof of [62] to refine the Cartan–Dieudonné decomposition for any field of characteristic $\neq 2$. In general, if Λ_k is a quadratic space on a field k of characteristic $\neq 2$ with Gram matrix G, let I be the identity matrix.

The proof of the Cartan–Dieudonné decomposition in [62] relies on the following facts: for any element $A \in \mathrm{O}(\Lambda_k)$, we have:

(i) A is a reflection if $\mathrm{rank}(A - I) = 1$ (see [62, Lemma 2]).

(ii) Suppose that $\mathrm{rank}(A - I) > 1$. If $S = G(A - I)$ is not skew-symmetric, then there exists $a \in \Lambda$ satisfying $a^t Sa \neq 0$ and

$$S + S^t \neq \frac{1}{a^t Sa}(Sa \cdot a^t S + S^t a \cdot a^t S^t).$$

In this case $\mathrm{rank}(A\boldsymbol{R}_b - I) = \mathrm{rank}(A - I) - 1$ and $G(A\boldsymbol{R}_b - I)$ is not skew-symmetric with $b = (A - I)a$ satisfying $b^2 = -2a^t Sa$ (see [62, Lemmas 4 and 5]).

(iii) If $S = G(A - I)$ is skew-symmetric, then there exists $b \in \Lambda$ such that $G(A\boldsymbol{R}_b - I)$ is not skew-symmetric (see the proof of [62, Theorem 2]).

Then we can decompose A as a series of reflections using (ii) repeatedly. In our case, it suffices to show that if $k = \mathbb{Q}$ and A is coprime to ℓ, i.e., nA is integral for some n coprime to ℓ, then:

(i′) A is a prime-to-ℓ reflection if $\mathrm{rank}(A - I) = 1$.

(ii′) Suppose that $\mathrm{rank}(A - I) > 1$. If the matrix $S = G(A - I)$ modulo ℓ is not skew-symmetric, then there exists a vector $a \in \Lambda$ satisfying $\ell \nmid a^t Sa$, and

$$S + S^t \neq \frac{1}{a^t Sa}(Sa \cdot a^t S + S^t a \cdot a^t S^t).$$

In this case, $\boldsymbol{R}_b$ is prime-to-ℓ with $b=(A-I)a$, $\mathrm{rank}(A\boldsymbol{R}_b-I)=\mathrm{rank}(A-I)-1$ and $G(A\boldsymbol{R}_b-I)$ is not skew-symmetric.

(iii′) If the matrix $S=G(A-I)$ modulo ℓ is skew-symmetric, then there exists $b\in\Lambda$ such that AR_b is coprime to ℓ and the modulo ℓ reduction of $G(AR_b-I)$ is not skew-symmetric.

For (i′), this is obvious.

For (ii′), if the modulo ℓ reduction $\overline{G}(\bar{A}-\bar{I})$ of $G(A-I)$ is not skew-symmetric, we can apply (ii) to the matrix $\bar{A}\in \mathrm{O}(\Lambda_{\mathbb{F}_\ell})$ to obtain a nonzero vector $\bar{a}\in\Lambda_{\mathbb{F}_\ell}$ such that $\bar{a}^t\bar{S}\bar{a}\neq 0\in\mathbb{F}_\ell$ and

$$\bar{S}+\bar{S}^t\neq\frac{1}{\bar{a}^t\bar{S}\bar{a}}(\bar{S}\bar{a}\cdot\bar{a}^t\bar{S}+\bar{S}^t\bar{a}\cdot\bar{a}^t\bar{S}^t). \tag{5.2.1}$$

Let $a\in\Lambda$ be a lifting of $\bar{a}$. It is easy to see that this is as desired.

For (iii′), the argument is similar to (ii′). □

As a result, we get the following.

Theorem 5.2.4. *Let $\ell>2$ be a prime. If there is a prime-to-ℓ rational Hodge isometry*

$$\varphi:\mathrm{H}^2(X,\mathbb{Q})\xrightarrow{\sim}\mathrm{H}^2(Y,\mathbb{Q}),$$

then there exists a prime-to-ℓ derived isogeny from X to Y that induces φ up to sign. If X and Y are prime-to-ℓ derived isogenous, then there is a prime-to-ℓ derived isogeny in which the orders of $\mathbb{G}_m$-gerbes are all prime-to-ℓ.

Proof. By using the prime-to-ℓ Cartan–Dieudonné decomposition given in Lemma 5.2.3, one can decompose the Hodge isometry

$$\varphi:\mathrm{H}^2(X,\mathbb{Z}_{(\ell)})\xrightarrow{\sim}\mathrm{H}^2(Y,\mathbb{Z}_{(\ell)}),$$

into a chain of prime-to-ℓ reflective Hodge isometries. Then Lemma 5.2.2 implies that the lift $\tilde{\varphi}$ extends to an integral isometry

$$\widetilde{\mathrm{H}}(X,\mathbb{Z}_{(\ell)})\xrightarrow{\sim}\widetilde{\mathrm{H}}(Y,\mathbb{Z}_{(\ell)}).$$

In the first case of the proof in Theorem 5.1.3, the derived isogeny (5.1.4) induces $\tilde{\varphi}$ up to sign, and is thus prime-to-ℓ. In the second case, the derived isogeny (5.1.5) is also prime-to-ℓ, since the Poincaré dual

$$\widetilde{\mathrm{H}}(X,\mathbb{Z})\xrightarrow{\sim}\widetilde{\mathrm{H}}(\hat{X},\mathbb{Z})$$

is integral and switches $(0,0,1)$ and $(1,0,0)$.

If X and Y are prime-to-ℓ derived isogenous, there is an isometry $\mathrm{T}(X)\otimes\mathbb{Z}_{(\ell)}\cong\mathrm{T}(Y)\otimes\mathbb{Z}_{(\ell)}$. Since $\ell>2$, there is a prime-to-ℓ rational Hodge isometry $\mathrm{H}^2(X,\mathbb{Z}_{(\ell)})\xrightarrow{\sim}\mathrm{H}^2(Y,\mathbb{Z}_{(\ell)})$ by [47, Theorem 3.2]. We can use the prime-to-ℓ Cartan–Dieudonné decomposition again to obtain a derived isogeny in which all the reflexive Hodge isometries are prime-to-ℓ. Then we can conclude the assertion by Lemma 5.2.2 and Remark 5.1.4. □

5.3. ***Isogeny versus derived isogeny.*** Let us now describe derived isogenies through suitable isogenies.

The functor $\underline{\mathrm{Hom}}(X, Y)$ of group homomorphisms from X to Y (not just as scheme morphisms) is representable by an étale group scheme over k (see [19, (7.14)], for example). Therefore, via Galois descent, we have

$$\mathrm{Hom}_{\mathtt{AV}_{\bar{k}}}(X_{\bar{k}}, Y_{\bar{k}}) \xrightarrow{\sim} \mathrm{Hom}_{\mathtt{AV}_{\overline{K}}}(X_{\overline{K}}, Y_{\overline{K}}), \tag{5.3.1}$$

for any algebraically closed field $\overline{K} \supset k$. A similar statement holds for derived isogenies.

Lemma 5.3.1. *Let X and Y be abelian surfaces defined over k with* $\mathrm{char}\, k = 0$. *Let $\overline{K} \supseteq k$ be an algebraically closed field containing k. Let $\bar{k}$ be the algebraic closure of k in $\overline{K}$. Then if $X_{\overline{K}}$ and $Y_{\overline{K}}$ are twisted derived equivalent, so are $X_{\bar{k}}$ and $Y_{\bar{k}}$.*

Proof. As $X_{\overline{K}}$ is twisted derived equivalent to $Y_{\overline{K}}$, by Theorem 3.5.3, there exist finitely many abelian surfaces $X_0, X_1, \dots, X_n$ defined over $\overline{K}$ with $X_0 = X_{\overline{K}}$ and

$$X_i \cong M_{H_i}(\mathscr{X}_{i-1}, v_i) \quad Y_{\overline{K}} \cong M_{H_n}(\mathscr{X}_n, v_n)$$

for some $[\mathscr{X}_{i-1}] \in \mathrm{Br}(X_{i-1})[r]$. Let us construct abelian surfaces over $\bar{k}$ to connect $X_{\bar{k}}$ and $Y_{\bar{k}}$ as follows:

Set $X'_0 = X_{\bar{k}}$, then we take $X'_1 = M_{H'_1}(\mathscr{X}'_0, v'_1)$ where $\mathscr{X}'_0$, H'_1 and v'_1 are the descent of $\mathscr{X}_0$, H_1 and v through the isomorphisms $\mathrm{Br}(X_{\overline{K}})[r] \cong \mathrm{Br}(X_{\bar{k}})[r]$, $\mathrm{NS}(X_{\overline{K}}) \cong \mathrm{NS}(X_{\bar{k}})$ and $\widetilde{\mathrm{H}}(X_{\overline{K}}) \cong \widetilde{\mathrm{H}}(X_{\bar{k}})$. The invariance of Brauer group and (ℓ-adic) Mukai lattice under extension $\bar{k} \subseteq \overline{K}$ is from the smooth base change theorem. For Néron–Severi groups, see [44, Proposition 3.1]. Then inductively, we can define X'_i as the moduli space of twisted sheaves $M_{H'_i}(\mathscr{X}'_{i-1}, v'_i)$ (or its dual, respectively) over $\bar{k}$. Note that we have natural isomorphisms

$$(M_{H'_i}(\mathscr{X}'_{i-1}, v'_i))_{\overline{K}} \cong M_{H_i}(\mathscr{X}_{i-1}, v_i)$$

over $\overline{K}$. In particular, $(M_{H'_i}(\mathscr{X}'_n, v'_i))_{\overline{K}} \cong Y_{\overline{K}}$. It follows that $M_{H'_i}(\mathscr{X}'_n, v'_i) \cong Y_{\bar{k}}$. □

For any abelian surface $X_{\mathbb{C}}$ over $\mathbb{C}$, the spreading-out argument shows that there is a finitely generated field $k \subset \mathbb{C}$ and an abelian surface X over k such that $X \times_k \mathbb{C} \cong X_{\mathbb{C}}$. We have the Artin comparison

$$\mathrm{H}^i_{\text{ét}}(X_{\bar{k}}, \mathbb{Z}_\ell) \cong \mathrm{H}^i(X_{\mathbb{C}}, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell, \tag{5.3.2}$$

for any $i \in \mathbb{Z}$ and ℓ a prime. Suppose Y is another abelian surface defined over k. Suppose $f : Y_{\mathbb{C}} \to X_{\mathbb{C}}$ is a prime-to-ℓ quasi-isogeny. By definition, it induces an isomorphism of $\mathbb{Z}_{(\ell)}$-modules

$$f^* : \mathrm{H}^1(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)} \xrightarrow{\sim} \mathrm{H}^1(Y_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)}$$

such that there is a commutative diagram

$$\begin{array}{ccc} \mathrm{H}^i(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)} & \xrightarrow{\sim} & \mathrm{H}^i(Y_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)} \\ \downarrow & & \downarrow \\ \mathrm{H}^i_{\text{ét}}(X_{\bar{k}}, \mathbb{Z}_\ell) & \xrightarrow{\sim} & \mathrm{H}^i_{\text{ét}}(Y_{\bar{k}}, \mathbb{Z}_\ell) \end{array}$$

for any i, under the comparison (5.3.2). For the converse, we have the following simple fact given by a faithfully flat descent of modules along $\mathbb{Z}_{(\ell)} \hookrightarrow \mathbb{Z}_\ell$ and the ℓ-adic Shioda thick.

Lemma 5.3.2. *A (quasi-)isogeny $f : Y_{\mathbb{C}} \to X_{\mathbb{C}}$ is prime-to-ℓ if and only if it induces an isomorphism of integral ℓ-adic realizations*

$$f^* : \mathrm{H}^2_{\acute{e}t}(X_{\bar{k}}, \mathbb{Z}_\ell) \xrightarrow{\sim} \mathrm{H}^2_{\acute{e}t}(Y_{\bar{k}}, \mathbb{Z}_\ell).$$

Inspired by Shioda's trick for Hodge isogenies (Proposition 4.5.1), we introduce the following notions.

Definition 5.3.3. Let X and Y be g-dimensional abelian varieties over k.

- X and Y are (prime-to-ℓ) *principally isogenous* if there is a (prime-to-ℓ) isogeny f from X to Y of square degree, that is, $\deg(f) = d^2$ for some $d \in \mathbb{Z}$. This f is called a *principal isogeny*.
- An isogeny $f : X \to Y$ is *quasiliftable* if f can be written as the composition of finitely many isogenies that are liftable to characteristic zero.

Now, we can state the main result in this section, which yields in particular Theorem 1.2.1.

Theorem 5.3.4. *Suppose* $\operatorname{char} k = 0$. *Let $\ell > 2$ be a prime. The following statements are equivalent*:

(1) *X is (prime-to-ℓ) principally isogenous to Y over $\bar{k}$.*

(2) *X and Y are (prime-to-ℓ) derived isogenous over $\bar{k}$.*

Proof. (1) $\Longrightarrow$ (2): we can assume that $f : X \to Y$ is a principal isogeny defined over a finitely generated field k'. By embedding k' into $\mathbb{C}$, two complex abelian surfaces $X_{\mathbb{C}}$ and $Y_{\mathbb{C}}$ are derived isogenous since there is a rational Hodge isometry

$$(f^*/n) \otimes \mathbb{Q} : \mathrm{H}^2(Y_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Q} \cong \mathrm{H}^2(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Q},$$

where $\deg(f) = n^2$. By Lemma 5.3.1, one concludes that $X_{\bar{k}}$ and $Y_{\bar{k}}$ are derived isogenous, with rational Hodge realization $(f^*/n) \otimes \mathbb{Q}$.

If f is a prime-to-ℓ isogeny, the map f^*/n restricts to an isomorphism

$$\mathrm{H}^2(Y_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)} \xrightarrow{\sim} \mathrm{H}^2(X_{\mathbb{C}}, \mathbb{Z}) \otimes \mathbb{Z}_{(\ell)}.$$

The assertion then follows from Theorem 5.2.4.

(2) $\Longrightarrow$ (1): We may assume X and Y are derived isogenous over a finitely generated field k'. Embedding k' into $\mathbb{C}$, $X_{\mathbb{C}}$ and $Y_{\mathbb{C}}$ are derived isogenous as well by Lemma 5.3.1. According to Remark 3.1.2, there is a Hodge isometry

$$\varphi : \mathrm{H}^2(Y_{\mathbb{C}}, \mathbb{Q}) \xrightarrow{\sim} \mathrm{H}^2(X_{\mathbb{C}}, \mathbb{Q}). \tag{5.3.3}$$

According to Example 4.2.3, we can assume φ is admissible after replacing X by its dual $\hat{X}$. By Proposition 4.5.1, they are principally isogenous over $\mathbb{C}$. It follows that X and Y are principally isogenous over $\bar{k}$ by (5.3.1).

If $\mathrm{D}^b(X) \sim \mathrm{D}^b(Y)$ is prime-to-ℓ, then we can choose a motive isomorphism $\mathfrak{h}^2(X) \simeq \mathfrak{h}^2(Y)$ whose ℓ-adic realization φ_ℓ is integral by the cancellation theorem over $\mathbb{Z}_\ell$ (see [54, Theorem 92:3]). The principal isogeny that induces φ is prime-to-ℓ by Lemma 5.3.2. This proves the assertion. □

Proof of Corollary 1.2.2. Let us summarize all the results that lead to Corollary 1.2.2. Using an argument similar to the one in Theorem 5.3.4, we can reduce them to the case $k = \mathbb{C}$.

(i) $\Longleftrightarrow$ (ii): This is Theorem 5.3.4.

(i) $\Longleftrightarrow$ (vi): This is Corollary 5.1.5.

(vi) $\Longleftrightarrow$ (vii) $\Longleftrightarrow$ (viii): This follows from the Witt cancellation theorem.

(i) $\Longleftrightarrow$ (iii): This is Corollary 5.1.7.

(ii) $\Longrightarrow$ (iv) $\Longrightarrow$ (v): This is from the computation in [23, Proposition 4.6]. In fact, one may take the correspondence

$$\Gamma := \bigoplus_i \Gamma_{2i} : \mathfrak{h}^{\mathrm{even}}(X) \xrightarrow{\sim} \mathfrak{h}^{\mathrm{even}}(Y),$$

where

$$\Gamma_{2i} := (f^*/n^i) \circ \pi_X^{2i} : \mathfrak{h}^{2i}(X) \to \mathfrak{h}^{2i}(Y),$$

and $f : X \to Y$ is the given principal isogeny.

(v) $\Longrightarrow$ (ii): Let $\Gamma : \mathfrak{h}^{\mathrm{even}}(X) \xrightarrow{\sim} \mathfrak{h}^{\mathrm{even}}(Y)$ be an isomorphism of Frobenius algebra objects. The Betti realization of its second component is a Hodge isometry by the Frobenius condition [23, Theorem 3.3]. Thus, X and Y are derived isogenous by Corollary 5.1.5, and hence are principally isogenous. □

6. Derived isogeny in positive characteristic

In this section, we prove the twisted derived Torelli theorem for abelian surfaces over odd characteristic fields. The primary strategy is to lift everything to characteristic zero. Throughout this section, we let k denote an algebraically closed field with characteristic $p > 3$.

6.1. ***Lifting of derived isogenies and quasi-isogenies.*** Let us start with a lifting result for derived isogenies, which is the only place we may require $p > 3$.

Proposition 6.1.1. *Let $\mathscr{X}_0 \to X_0$ and $\mathscr{Y}_0 \to Y_0$ be twisted abelian surfaces over k, which are of finite height. If there is a derived equivalence $\Phi_0 : \mathrm{D}^{(1)}(\mathscr{X}_0) \to \mathrm{D}^{(1)}(\mathscr{Y}_0)$, then there exists a discrete valuation ring V whose residue field is k and twisted abelian surfaces*

$$\begin{array}{ccc} \mathscr{X}_V & \longrightarrow & X_V \\ & \searrow & \downarrow \\ & & \operatorname{Spec} V \end{array} \quad \textit{and} \quad \begin{array}{ccc} \mathscr{Y}_V & \longrightarrow & Y_V \\ & \searrow & \downarrow \\ & & \operatorname{Spec} V \end{array}$$

over V with the following properties:

- *The special fibers are geometrically isomorphic to $\mathscr{X}_0 \to X_0$ and $\mathscr{Y}_0 \to Y_0$ respectively.*

- *There is a Fourier–Mukai transform $\Phi_V : \mathrm{D}^{(1)}(\mathscr{X}_V) \to \mathrm{D}^{(1)}(\mathscr{Y}_V)$ whose Fourier–Mukai kernel restricted to $\mathscr{X} \times \mathscr{Y}$ induces Φ_0.*

Moreover, if Φ_0 is prime-to-p and $p > 3$, the derived equivalence $\Phi_K : \mathrm{D}^{(1)}(\mathfrak{X}_K) \to \mathrm{D}^{(1)}(\mathfrak{Y}_K)$ on the generic fiber is also prime-to-p where K is the fraction field of V.

Proof. The proof proceeds like that of [10, Theorem 5.8], which ensures the existence of liftings of derived isogenies between K3 surfaces. By Theorem 3.5.3, we know that

$$\mathscr{X}_0^{(-1)} \cong \mathscr{M}_H(\mathscr{Y}_0, v)$$

is a moduli stack of $\mathscr{Y}_0$-twisted coherent sheaves for some vector $v \in \widetilde{\mathrm{N}}(\mathscr{Y}_0)$. By Lemma 2.3.1, we can find a DVR V and a projective lift $\mathscr{Y}_V \to Y_V$ over V such that $\mathrm{NS}(Y_V) \cong \mathrm{NS}(Y_0)$. Let H_V be the element in $\mathrm{NS}(Y_V)$ that extends H. Following the description of twisted extended Néron–Severi lattices as in Proposition 3.3.2, we can see that $\widetilde{\mathrm{N}}(\mathscr{Y}_V) \cong \widetilde{\mathrm{N}}(\mathscr{Y}_0)$ and hence the twisted Mukai vector v can be extended over V, still denoted by v.

Let $\mathscr{X}_V^{(-1)} = \mathscr{M}_{H_V}(\mathscr{Y}_V, v)$ be the relative moduli stack of $\mathscr{X}_V$-twisted coherent sheaves. The universal object in $\mathrm{D}^{(-1,1)}(\mathscr{X}_V \times \mathscr{Y}_V)$ induces a derived equivalence $\Phi_V : \mathrm{D}^{(1)}(\mathscr{X}_V) \to \mathrm{D}^{(1)}(\mathscr{Y}_V)$ as desired.

For the last assertion, we need to prove that the p-adic realization of Φ_K is integral. This can be deduced from a similar argument as in the proof of Theorem 1.5 in [10], based on Cais and Liu's crystalline cohomological description for the integral p-adic Hodge theory [13]. Let us sketch the proof. As Φ is prime-to-p, its cohomological realization restricts to an isometry of F-crystals

$$\tilde{\varphi}_p : \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(X_0/W) \simeq \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(Y_0/W)$$

by our definition. The base extension $\tilde{\varphi}_p \otimes K$ can be identified with the de Rham cohomological realization of Φ_K:

$$\tilde{\varphi}_K : \mathrm{H}^{\mathrm{even}}_{\mathrm{dR}}(X_K/K) \simeq \mathrm{H}^{\mathrm{even}}_{\mathrm{dR}}(Y_K/K),$$

by the Berthelot–Ogus comparison (see [4, Corollary 2.5] or [24, Theorem B.3.1]). It also preserves Hodge filtrations. Let S be the p-completion of the divided power envelope of the pair $\big(W[\![u]\!], \ker(W[\![u]\!] \to \mathcal{O}_K)\big)$. Then the map

$$\tilde{\varphi}_p \otimes_W S : \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(X_0/S) \xrightarrow{\sim} \mathrm{H}^{\mathrm{even}}_{\mathrm{crys}}(Y_0/S) \tag{6.1.1}$$

is an isomorphism of strongly divisible S-lattices (see [13, §4]). If $p > 3$, according to [13, Theorem 5.4], one can apply Breuil's functor on (6.1.1) to see that ϕ_K restricts to an $\mathbb{Z}_p$-integral $\mathrm{Gal}(\overline{K}/K)$-equivariant isometry $\mathrm{H}^{\mathrm{even}}_{\text{ét}}(X_{\overline{K}}, \mathbb{Z}_p) \xrightarrow{\sim} \mathrm{H}^{\mathrm{even}}_{\text{ét}}(Y_{\overline{K}}, \mathbb{Z}_p)$. □

Remark 6.1.2. The technical requirement that $p > 3$ is needed in [13, Theorem 4.3(3) and (4)]. When $\mathcal{O}_K = W(k)$ is unramified, this condition can be weakened to $p > 2$ by using Fontaine's Theorem 2(iii) in [21]. In general, when $p = 3$, a possible approach is to prove Shioda's trick as in Section 4 for strongly divisible S-lattices (see [11, Definition 2.1.1]), which can reduce the statement to crystalline Galois representations of Hodge–Tate weight one.

Next, one can lift separable isogenies between abelian surfaces:

Proposition 6.1.3. *Let $f : X_0 \to Y_0$ be a separable isogeny between two abelian surfaces over k. Let $W = W(k)$ be the ring of Witt vectors. Then there exist liftings $X_W \to \operatorname{Spec} W$ and $Y_W \to \operatorname{Spec} W$ such that the isogeny f can be lifted to an isogeny $f_W : X_W \to Y_W$ such that $\deg f = \deg f_W$. Thus, every prime-to-p isogeny can be lifted to a prime-to-p isogeny.*

Proof. According to [55, Proposition 11.1], there is a projective lifting $X_W \to \operatorname{Spec} W$ of X_0. Given that f is separable, $\ker f \subset X_0$ constitutes a finite étale group scheme over k, which is liftable. Choosing a lifting $G_W \subset X_W$ of $\ker f$, we obtain an isogeny

$$f_W : X_W \to Y_W := X_W/G_W,$$

which serves as a lifting of f. If f is prime-to-p, then we have $\ker f_W \subseteq X_W[n]$ for some n that is coprime to p. Consequently, f_W is also prime-to-p. ☐

6.2. ***Specialization of prime-to-p derived isogenies.*** Next, we shall show that prime-to-p geometrically derived isogenies are preserved under reduction. The idea is to show that the specialization of a moduli space of stable twisted sheaves on an abelian surface or K3 surface remains a moduli space.

Theorem 6.2.1. *Let V be a DVR with residue field k and $K = \operatorname{Frac}(V)$. Let $X_V \to \operatorname{Spec} V$ and $Y_V \to \operatorname{Spec} V$ be projective abelian surfaces or K3 surfaces over $\operatorname{Spec} V$ satisfying*

$$\operatorname{NS}(X_{\overline{K}}) \cong \operatorname{NS}(X_k), \tag{6.2.1}$$

where X_k is the special fiber of $X_V \to \operatorname{Spec} V$. If their generic fibers X_K and Y_K are (geometrically) prime-to-p derived isogenies, so are the special fibers X_k and Y_k.

Proof. With Theorem 5.2.4, it is sufficient to consider the case where there is a derived equivalence

$$\Phi_V : \mathrm{D}^{(1)}(\mathscr{X}_{\overline{K}}) \xrightarrow{\sim} \mathrm{D}^{(1)}(\mathscr{Y}_{\overline{K}})$$

for some prime-to-p $\mathbb{G}_m$-gerbes $\mathscr{X}_K \to X_K$ and $\mathscr{Y}_K \to Y_K$. From Theorem 3.5.3, we know that there is an isomorphism

$$\mathscr{Y}_{\overline{K}} \cong \mathscr{M}_H(\mathscr{X}_{\overline{K}}, v_K)^{(-1)},$$

for some twisted Mukai vector $v_K \in \widetilde{\mathrm{N}}(\mathscr{X}_K)$ and $H_K \in \operatorname{NS}(X_{\overline{K}})$ being v-generic. Up to taking a finite extension, we may assume that everything can be defined over K.

We claim that there exists a $\mathbb{G}_m$ gerbe $\mathscr{X}_V \to X_V$ whose restriction to $\operatorname{Spec} K$ is $\mathscr{X}_K \to X_K$. It suffices to show that the class $[\mathscr{X}_K] \in \operatorname{Br}(X_K)$ can be extended to an element in $\operatorname{Br}(X_V)$. By the Chinese remainder theorem, we may assume $\operatorname{ord}([\mathscr{X}_K]) = \ell^n$ for some prime $\ell \neq p$. For each prime $\ell \neq p$, from the Kummer sequence, we have the commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{Pic}(X_V)/\ell^n & \longrightarrow & \mathrm{H}^1_{\text{ét}}(X_V, \mu_{\ell^n}) & \longrightarrow & \mathrm{Br}(X_V)[\ell^n] & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \mathrm{Pic}(X_K)/\ell^n & \longrightarrow & \mathrm{H}^1_{\text{ét}}(X_K, \mu_{\ell^n}) & \longrightarrow & \mathrm{Br}(X_K)[\ell^n] & \longrightarrow & 0 \end{array}$$

The second vertical morphism is an isomorphism by smooth and proper base change. Therefore, $\mathrm{Br}(X_V)[\ell^n] \to \mathrm{Br}(X_K)[\ell^n]$ is surjective, which proves the claim.

By our assumption (6.2.1), we can pick extensions $v_V \in \widetilde{\mathrm{N}}(\mathscr{X}_V)$ and $H_V \in \mathrm{Pic}(X_V)$ of v_K and H_K. Let $\mathscr{M}_{H_V}(X_V, v_V) \to \operatorname{Spec} V$ be the relative moduli space of H_V-stable twisted sheaves. Then we have the commutative diagram

$$\begin{array}{ccccccc} M_{H_V}(\mathscr{X}_V, v_V) & \longleftarrow & M_{H_K}(\mathscr{X}_K, v_K) & \xrightarrow{\cong} & Y_K & \longrightarrow & Y_V \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \operatorname{Spec} V & \longleftarrow & \operatorname{Spec}(K) & \longrightarrow & \operatorname{Spec}(K) & \longrightarrow & \operatorname{Spec} V \end{array}$$

According to Matsusaka and Mumford [43, Theorem 1], the isomorphism between the generic fiber can be extended to $\operatorname{Spec} V$. Thus Y_k is isomorphic to $M_{H_k}(\mathscr{X}_k, v_k)$, where $v_k = v_V|_{\operatorname{Spec} k}$ and $H_k = H_V|_{\operatorname{Spec} k}$. It follows that there is a prime-to-p derived equivalence $\mathrm{D}^{(1)}(\mathscr{X}_k) \simeq \mathrm{D}^{(-1)}(\mathscr{M}_{H_k}(\mathscr{X}_k, v_k))$. □

Remark 6.2.2. Our proof fails when the twisted derived equivalence is not prime-to-p. This is because if the associated Brauer class α has order p^n, the map

$$\mathrm{Br}(X_V)[p^n] \to \mathrm{Br}(X_K)[p^n]$$

may not be surjective (see [59, 6.8.2]).

Proof of Theorem 1.4.1. When X or Y is supersingular, the assertion follows from Theorem 3.6.6(2). So we can assume that X and Y both have finite height.

(i′) ⟹ (ii′): By Proposition 6.1.1, we can find projective liftings $X_V \to \operatorname{Spec} V$ and $Y_V \to \operatorname{Spec} V$ of X and Y over some DVR V such that there is a prime-to-p twisted derived equivalence between generic fibers X_K and Y_K.

By Theorem 5.3.4, the generic fibers X_K and Y_K are geometrically prime-to-p principally isogenous. Up to a finite extension of K, we can find a prime-to-p principal isogeny $f_K : X_K \to Y_K$. The Néron extension property of smooth models X_V, Y_V [5, §7.3, Proposition 6] ensures that f_K can be extended to an isogeny

$$f_V : X_V \to Y_V.$$

The restriction $f_k : X \to Y$ over the special fibers is still a principal isogeny and we can conclude that f_k is prime-to-p by using Tate's spreading theorem for p-divisible groups (see [67, Theorem 4]).

(i′) ⟹ (ii′): Suppose that there is an isogeny $f : X \to Y$ that is prime-to-p of degree d^2. By Proposition 6.1.3, we can lift it to a prime-to-p isogeny of degree d^2 over W:

$$f_W : X_W \to Y_W.$$

Set $K = \mathrm{Frac}(W)$. The induced isogeny f_K between the generic fibers is a prime-to-p principal isogeny, which induces a G_K-equivariant isometry

$$\frac{f_K^*}{d} : \mathrm{H}^2_{\text{ét}}(Y_{\bar{K}}, \mathbb{Z}_p) \xrightarrow{\sim} \mathrm{H}^2_{\text{ét}}(X_{\bar{K}}, \mathbb{Z}_p).$$

By Theorem 5.3.4, there exists a prime-to-p derived isogeny $\mathrm{D}^b(X_{\bar{K}}) \sim \mathrm{D}^b(Y_{\bar{K}})$ whose p-adic cohomological realization is f_K^*/d. The assertion follows from Theorem 6.2.1. □

6.3. *Further remarks.* From the proof of the implication (i′) ⇒ (ii′) of Theorem 1.4.1, we can see that the lifting-specialization argument also works for non-prime-to-p derived isogenies. Thus we have

Theorem 6.3.1. *Suppose X_0 and Y_0 are abelian surfaces over k with finite height. If X_0 and Y_0 are derived isogenous, then they are quasiliftable principally isogenous.*

Moreover, we believe that the converse of Theorem 6.3.1 also holds.

Conjecture 6.3.2. *Two abelian surfaces X_0 and Y_0 are derived isogenous over k if and only if they are quasiliftable principally isogenous.*

For this conjecture, our approach remains valid provided that there is a specialization theorem for non-prime-to-p derived isogenies. According to the proof of Theorem 6.2.1, it suffices to establish the existence of specialization of Brauer classes of order p. Adhering to the notations in Theorem 6.2.1, this requires the restriction map

$$\mathrm{Br}(X_V) \to \mathrm{Br}(X_K)$$

to be surjective. See Remark 6.2.2 for further details.

6.4. *Derived isogeny for Kummer surfaces.* We now explore the interrelations between the derived isogenies of abelian surfaces and their associated Kummer surfaces. Using the lifting argument, the following theorem is an immediate consequence of the result in characteristic 0.

Theorem 6.4.1. *Assume $p > 2$. If X_0 and Y_0 are prime-to-p derived isogenous abelian surfaces over k, then the associated Kummer surfaces $\mathrm{Km}(X_0)$ and $\mathrm{Km}(Y_0)$ are prime-to-p derived isogenous. If there is a derived equivalence*

$$\mathrm{D}^b(\mathrm{Km}(X_0), \alpha_0) \simeq \mathrm{D}^b(\mathrm{Km}(Y_0), \beta_0) \tag{6.4.1}$$

with $\mathrm{ord}(\alpha_0)$ and $\mathrm{ord}(\beta_0)$ prime-to-p, then X and Y are prime-to-p derived isogenous.

Proof. For the first assertion, as before, we can quasi-lift the prime-to-p derived isogeny between X and Y to characteristic 0. By Theorem 1.4.1 and Proposition 6.1.1, their liftings are geometrically prime-to-p derived isogenous. According to [66, Corollary 4.3], the associated Kummer surfaces are prime-to-p

derived isogenous. It follows from Theorem 6.2.1 that $\mathrm{Km}(X_0)$ and $\mathrm{Km}(Y_0)$ are prime-to-p derived isogenous.

For the last assertion, if X_0 and Y_0 are supersingular, then α_0 and β_0 are trivial under our assumptions. In this case, the result follows from [38, Theorem 1.2]. Suppose X_0 or Y_0 is of finite height (then both are of finite height). According to [10, Theorem 5.8], we can find a DVR V with residue field k and projective twisted K3 surfaces over V

$$(S_V, \alpha_V) \to \mathrm{Spec}\, V \quad \text{and} \quad (S'_V, \beta_V) \to \mathrm{Spec}\, V$$

satisfying the following conditions:

- The special fibers are $(\mathrm{Km}(X_0), \alpha_0)$ and $(\mathrm{Km}(Y_0), \beta_0)$ respectively.
- The generic fibers (S_K, α_K) and (S'_K, β_K) are geometrically derived equivalent.
- $\mathrm{NS}(S_{\overline{K}}) \cong \mathrm{NS}(\mathrm{Km}(X_0))$ and $\mathrm{NS}(S'_{\overline{K}}) \cong \mathrm{NS}(\mathrm{Km}(Y_0))$.

Note that $\mathrm{NS}(S_K)$ and $\mathrm{NS}(S'_K)$ contain Kummer lattices. As seen in the proof of Lemma 2.3.1, this implies that there exist projective liftings of X_0 and Y_0, denoted by $X_V \to \mathrm{Spec}\, V$ and $Y_V \to \mathrm{Spec}\, V$, such that

$$S_{\overline{K}} \cong \mathrm{Km}(X_{\overline{K}}) \quad \text{and} \quad S'_{\overline{K}} \cong \mathrm{Km}(Y_{\overline{K}}).$$

Choose an embedding $K \hookrightarrow \mathbb{C}$, set $X_{\mathbb{C}} = X_K \otimes_K \mathbb{C}$ and $Y_{\mathbb{C}} = Y_K \otimes_K \mathbb{C}$. Then we have a prime-to-p Hodge isometry

$$\mathrm{H}^2(\mathrm{Km}(X_{\mathbb{C}}), \mathbb{Z}_{(p)}) \to \mathrm{H}^2(\mathrm{Km}(Y_{\mathbb{C}}), \mathbb{Z}_{(p)}) \tag{6.4.2}$$

induced from the prime-to-p derived equivalence. Based on the Kummer construction, for any abelian surface $X_{\mathbb{C}}$, as $p > 2$, there is a natural Hodge isometry

$$\mathrm{H}^2(\mathrm{Km}(X_{\mathbb{C}}), \mathbb{Z}_{(p)}) \cong \mathrm{H}^2(X_{\mathbb{C}}, \mathbb{Z}_{(p)}) \oplus (\Sigma_{X_{\mathbb{C}}} \otimes \mathbb{Z}_{(p)}),$$

where $\Sigma_{X_{\mathbb{C}}} \cong \bigoplus_{i=1}^{16} \mathbb{Z} e_i$ with $(e_i, e_j) = -2\delta_{ij}$ being the Kummer lattice. Then one obtains a Hodge isometry

$$\mathrm{H}^2(X_{\mathbb{C}}, \mathbb{Z}_{(p)}) \to \mathrm{H}^2(Y_{\mathbb{C}}, \mathbb{Z}_{(p)})$$

from (6.4.2) through the Witt cancellation procedure. By Theorem 5.3.4, X_K and Y_K are geometrically prime-to-p derived isogenous. The assertion follows from Theorem 6.2.1. □

Remark 6.4.2. It is natural to consider if one can apply the lifting method to prove the converse of Theorem 6.4.1. Specifically, one may wonder if $\mathrm{Km}(X_0)$ and $\mathrm{Km}(Y_0)$ are prime-to-p derived isogenous, as are X and Y.

However, the issue is that the derived isogeny between $\mathrm{Km}(X_0)$ and $\mathrm{Km}(Y_0)$ is merely quasiliftable, not known to be liftable. In other words, although we can lift every derived equivalence between twisted abelian surfaces or K3 surface to characteristic 0, we cannot necessarily find liftings of X_0 and Y_0 respectively such that the generic fibers of their associated Kummer surfaces are prime-to-p geometrically derived isogenous.

Acknowledgement

The authors are grateful to Ziquan Yang for useful comments. The authors also thank the referees for their careful reading and valuable suggestions, which improved this article.

References

[1] M. Artin, "Algebraization of formal moduli, I", pp. 21–71 in *Global analysis*: *Papers in honor of K. Kodaira*, Univ. Tokyo Press, Tokyo, 1969. MR

[2] M. Artin, "Supersingular K3 surfaces", *Ann. Sci. École Norm. Sup.* (4) **7** (1974), 543–567. MR

[3] P. Berthelot and A. Ogus, *Notes on crystalline cohomology*, Princeton University Press, 1978. MR

[4] P. Berthelot and A. Ogus, "F-isocrystals and de Rham cohomology, I", *Invent. Math.* **72**:2 (1983), 159–199. MR

[5] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR

[6] D. Bragg, "Derived equivalences of twisted supersingular K3 surfaces", *Adv. Math.* **378** (2021), art. id. 107498, 45 pp. MR

[7] D. Bragg, "Lifts of twisted K3 surfaces to characteristic 0", *Int. Math. Res. Not.* **2023**:5 (2023), 4337–4407. MR

[8] D. Bragg and M. Lieblich, "Twistor spaces for supersingular K3 surfaces", preprint, 2018. arXiv 1804.07282

[9] D. Bragg and M. Olsson, "Representability of cohomology of finite flat abelian group schemes", preprint version 2, 2025. arXiv 2107.11492v2

[10] D. Bragg and Z. Yang, "Twisted derived equivalences and isogenies between K3 surfaces in positive characteristic", *Algebra Number Theory* **17**:5 (2023), 1069–1126. MR

[11] C. Breuil, "Représentations semi-stables et modules fortement divisibles", *Invent. Math.* **136**:1 (1999), 89–122. MR

[12] T. Bridgeland and A. Maciocia, "Complex surfaces with equivalent derived categories", *Math. Z.* **236**:4 (2001), 677–697. MR

[13] B. Cais and T. Liu, "Breuil–Kisin modules via crystalline cohomology", *Trans. Amer. Math. Soc.* **371**:2 (2019), 1199–1230. Corrigendum in **373**:3 (2020), 2251–2252. MR

[14] A. H. Caldararu, *Derived categories of twisted sheaves on Calabi–Yau manifolds*, Ph.D. thesis, Cornell University, 2000, available at https://www.proquest.com/docview/304592237. MR

[15] A. Canonaco and P. Stellari, "Twisted Fourier–Mukai functors", *Adv. Math.* **212**:2 (2007), 484–503. MR

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

[17] B. Conrad, "Reductive group schemes", pp. 93–444 in *Autour des schémas en groupes*, Panor. Synthèses **42**, Soc. Math. France, Paris, 2014. MR

[18] C. Deninger and J. Murre, "Motivic decomposition of abelian schemes and the Fourier transform", *J. Reine Angew. Math.* **422** (1991), 201–219. MR

[19] B. Edixhoven, G. van der Geer, and B. Moonen, "Abelian varieties", book in progress, available at http://van-der-geer.nl/~gerard/AV.pdf.

[20] G. Faltings, "Complements to Mordell", pp. 204–227 in *Rational points*: *Seminar Bonn/Wuppertal*, 1983/1984, 3rd ed., Aspects of Mathematics **E6**, Vieweg, Braunschweig, 1992. MR

[21] J.-M. Fontaine, "Cohomologie de de Rham, cohomologie cristalline et représentations p-adiques", pp. 86–108 in *Algebraic geometry* (Tokyo/Kyoto, 1982), Lecture Notes in Math. **1016**, Springer, 1983. MR

[22] L. Fu and Z. Li, "Supersingular irreducible symplectic varieties", pp. 147–200 in *Rationality of algebraic varieties*, Progr. Math. **342**, Birkhäuser, 2021. MR

[23] L. Fu and C. Vial, "A motivic global Torelli theorem for isogenous K3 surfaces", *Adv. Math.* **383** (2021), art. id. 107674, 44 pp. MR

[24] H. Gillet and W. Messing, "Cycle classes and Riemann–Roch for crystalline cohomology", *Duke Math. J.* **55**:3 (1987), 501–538. MR

[25] J. Giraud, *Cohomologie non abélienne*, Grundl. Math. Wissen. **179**, Springer, 1971. MR

[26] 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

[27] K. Honigs, L. Lombardi, and S. Tirabassi, “Derived equivalences of canonical covers of hyperelliptic and Enriques surfaces in positive characteristic”, *Math. Z.* **295**:1-2 (2020), 727–749. MR

[28] K. Honigs, M. Lieblich, and S. Tirabassi, “Fourier–Mukai partners of Enriques and bielliptic surfaces in positive characteristic”, *Math. Res. Lett.* **28**:1 (2021), 65–91. MR

[29] D. Huybrechts, *Fourier–Mukai transforms in algebraic geometry*, The Clarendon Press, Oxford, 2006. MR

[30] D. Huybrechts, “Motives of derived equivalent K3 surfaces”, *Abh. Math. Semin. Univ. Hambg.* **88**:1 (2018), 201–207. MR

[31] D. Huybrechts, “Motives of isogenous K3 surfaces”, *Comment. Math. Helv.* **94**:3 (2019), 445–458. MR

[32] D. Huybrechts and P. Stellari, “Equivalences of twisted K3 surfaces”, *Math. Ann.* **332**:4 (2005), 901–936. MR

[33] D. Huybrechts, E. Macrì, and P. Stellari, “Stability conditions for generic K3 categories”, *Compos. Math.* **144**:1 (2008), 134–162. MR

[34] L. Illusie, “Complexe de de Rham–Witt et cohomologie cristalline”, *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR

[35] A. J. de Jong, “Finite locally free group schemes in characteristic p and Dieudonné modules”, *Invent. Math.* **114**:1 (1993), 89–137. MR

[36] A. J. de Jong, “Homomorphisms of Barsotti–Tate groups and crystals in positive characteristic”, *Invent. Math.* **134**:2 (1998), 301–333. MR

[37] K. Künnemann, “On the Chow motive of an abelian scheme”, pp. 189–205 in *Motives* (Seattle, WA, 1991), Proc. Sympos. Pure Math. **55, Part 1**, Amer. Math. Soc., Providence, RI, 1994. MR

[38] Z. Li and H. Zou, “A note on Fourier–Mukai partners of abelian varieties over positive characteristic fields”, *Kyoto J. Math.* **63**:4 (2023), 893–913. MR

[39] M. Lieblich, “Moduli of twisted sheaves”, *Duke Math. J.* **138**:1 (2007), 23–118. MR

[40] M. Lieblich and M. Olsson, “Fourier–Mukai partners of K3 surfaces in positive characteristic”, *Ann. Sci. Éc. Norm. Supér.* (4) **48**:5 (2015), 1001–1033. MR

[41] M. Lieblich and M. Olsson, “A stronger derived Torelli theorem for K3 surfaces”, pp. 127–156 in *Geometry over nonclosed fields*, Springer, 2017. MR

[42] M. Lieblich, D. Maulik, and A. Snowden, “Finiteness of K3 surfaces and the Tate conjecture”, *Ann. Sci. Éc. Norm. Supér.* (4) **47**:2 (2014), 285–308. MR

[43] T. Matsusaka and D. Mumford, “Two fundamental theorems on deformations of polarized varieties”, *Amer. J. Math.* **86** (1964), 668–684. MR

[44] D. Maulik and B. Poonen, “Néron–Severi groups under specialization”, *Duke Math. J.* **161**:11 (2012), 2167–2206. MR

[45] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, 1980. MR

[46] H. Minamide, S. Yanagida, and K. Yoshioka, “The wall-crossing behavior for Bridgeland’s stability conditions on abelian and K3 surfaces”, *J. Reine Angew. Math.* **735** (2018), 1–107. MR

[47] K. A. Morin-Strom, “Witt’s theorem for modular lattices”, *Amer. J. Math.* **101**:6 (1979), 1181–1192. MR

[48] D. R. Morrison, “On K3 surfaces with large Picard number”, *Invent. Math.* **75**:1 (1984), 105–121. MR

[49] S. Mukai, “On the moduli space of bundles on K3 surfaces, I”, pp. 341–413 in *Vector bundles on algebraic varieties* (Bombay, 1984), Tata Inst. Fund. Res. Stud. Math. **11**, Tata Inst. Fund. Res., Bombay, 1987. MR

[50] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik (2) **34**, Springer, 1994. MR

[51] A. Ogus, “Supersingular K3 crystals”, pp. 3–86 in *Journées de Géométrie Algébrique de Rennes* (Rennes, 1978), Astérisque **64**, Soc. Math. France, Paris, 1979. MR

[52] A. Ogus, "A crystalline Torelli theorem for supersingular K3 surfaces", pp. 361–394 in *Arithmetic and geometry*, Progr. Math. **36**, Birkhäuser, Boston, 1983. MR

[53] M. Olsson, "Twisted derived categories and Rouquier functors", *Int. Math. Res. Not.* **2025**:18 (2025), art. id. rnaf290, 18 pp. MR

[54] O. T. O'Meara, *Introduction to quadratic forms*, Grundl. Math. Wissen. **117**, Springer, 1963. MR

[55] F. Oort, "Lifting algebraic curves, abelian varieties, and their endomorphisms to characteristic zero", pp. 165–195 in *Algebraic geometry* (Brunswick, ME, 1985), Proc. Sympos. Pure Math. **46:**2, Amer. Math. Soc., 1987. MR

[56] D. O. Orlov, "Derived categories of coherent sheaves on abelian varieties and equivalences between them", *Izv. Ross. Akad. Nauk Ser. Mat.* **66**:3 (2002), 131–158. In Russian; translated in Izvestiya Math. **66**:3 (2002), 569–594. MR

[57] I. I. Piatetskii-Shapiro and I. R. Shafarevich, "Torelli's theorem for algebraic surfaces of type K3", *Izv. Akad. Nauk SSSR Ser. Mat.* **35** (1971), 530–572. In Russian; translated in *Math. USSR-Izv.* **5**:3 (1971), 547–588. MR

[58] A. Polishchuk, "Symplectic biextensions and a generalization of the Fourier–Mukai transform", *Math. Res. Lett.* **3**:6 (1996), 813–828. MR

[59] B. Poonen, *Rational points on varieties*, Graduate Studies in Mathematics **186**, Amer. Math. Soc., 2017. MR

[60] J. J. Ramón Marí, "On the Hodge conjecture for products of certain surfaces", *Collect. Math.* **59**:1 (2008), 1–26. MR

[61] A. N. Rudakov and I. R. Shafarevich, "Surfaces of type K3 over fields of finite characteristic", pp. 115–207 in *Current problems in mathematics*, vol. 18, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1981. In Russian; translated in Shafarevich's *Collected mathematical papers*, Springer, Berlin, 1989, pp. 657–714. MR

[62] P. Scherk, "On the decomposition of orthogonalities into symmetries", *Proc. Amer. Math. Soc.* **1** (1950), 481–491. MR

[63] T. Shioda, "The period map of Abelian surfaces", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **25**:1 (1978), 47–59. MR

[64] T. Shioda, "Supersingular K3 surfaces", pp. 564–591 in *Algebraic geometry* (Copenhagen, 1978), Lecture Notes in Math. **732**, Springer, 1979. MR

[65] A. N. Skorobogatov and Y. G. Zarhin, "The Brauer group of Kummer surfaces and torsion of elliptic curves", *J. Reine Angew. Math.* **666** (2012), 115–140. MR

[66] P. Stellari, "Derived categories and Kummer varieties", *Math. Z.* **256**:2 (2007), 425–441. MR

[67] J. T. Tate, "p-divisible groups", pp. 158–183 in *Proc. Conf. Local Fields* (Driebergen, 1966), Springer, 1967. MR

[68] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR

[69] Z. Yang, "Isogenies between K3 surfaces over $\overline{\mathbb{F}}_p$", *Int. Math. Res. Not.* **2022**:6 (2022), 4407–4450. MR

[70] K. Yoshioka, "Moduli spaces of stable sheaves on abelian surfaces", *Math. Ann.* **321**:4 (2001), 817–884. MR

[71] K. Yoshioka, "Moduli spaces of twisted sheaves on a projective variety", pp. 1–30 in *Moduli spaces and arithmetic geometry*, Adv. Stud. Pure Math. **45**, Math. Soc. Japan, Tokyo, 2006. MR

[72] Y. G. Zarhin, "Abelian varieties in characteristic p", *Mat. Zametki* **19**:3 (1976), 393–400. In Russian; translated in *Math. Notes* **19**:3 (1976), 240–244. MR

[73] Y. G. Zarhin, "Almost isomorphic abelian varieties", *Eur. J. Math.* **3**:1 (2017), 22–33. MR

Communicated by Hélène Esnault
Received 2023-11-13 Revised 2025-05-04 Accepted 2025-06-27

zhiyuan_li@fudan.edu.cn *Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China*

hzou@math.uni-bielefeld.de *Universität Bielefeld, Bielefeld, Germany*

ALGEBRA AND NUMBER THEORY 20:6 (2026)
https://doi.org/10.2140/ant.2026.20.1235

Injectivity and vanishing for the Du Bois complexes of isolated singularities

Mihnea Popa, Wanchun Shen and Anh Duc Vo

We prove an injectivity theorem for the cohomology of the Du Bois complexes of varieties with isolated singularities. We use this to deduce vanishing statements for the cohomologies of higher Du Bois complexes of such varieties. Besides some extensions and conjectures in the nonisolated case, we also provide analogues for intersection complexes.

A. Introduction

Let X be a complex algebraic variety of dimension n, and for each $k \geq 0$ let $\underline{\Omega}_X^k$ be the k-th associated graded term of the filtered de Rham complex $\underline{\Omega}_X^\bullet$ with respect to the Hodge filtration, also called the k-th Du Bois complex of X. Given their growing importance in the study of singularities via Hodge theory, it has become essential to understand the finer homological properties of these complexes. This paper addresses the case of isolated singularities, by focusing on vanishing theorems for the cohomologies of Du Bois complexes, and injectivity theorems for the cohomologies of their duals.

Vanishing of cohomologies. By definition, the Du Bois complexes $\underline{\Omega}_X^k$ have nontrivial cohomologies only in degrees in the range $[0, n]$. Something better is in fact true: a well-known result of Steenbrink [St2, (4.1)] says that

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } i > n - k,$$

and without further assumptions on k or the singularities, this is optimal.

The vanishing of other higher cohomologies $\mathcal{H}^i \underline{\Omega}_X^k$ in the possible nonvanishing range is one way to measure how good the singularities of X are. One of our main goals is to describe concrete conditions under which this holds. Previous results in this direction were obtained in [MOPW] for hypersurfaces, and more generally in [MP1] for local complete intersections. Here we address this in the case of isolated singularities, providing some results in the nonisolated case along the way as well.

The appropriate language for studying this problem is that of *higher Du Bois singularities*, studied in [MOPW; JKSY; MP1; SVV], among others. At least in the local complete intersection case, this condition means that $\underline{\Omega}_X^p$ can be identified with the sheaf of Kähler differentials Ω_X^p for certain p. Along

The authors were partially supported by the NSF grant DMS-2040378. Popa was also supported by NSF grant DMS-2401498.
MSC2020: 14B05, 14F17, 32S35.
Keywords: Du Bois complex, higher singularity, injectivity theorem.

the way towards a general definition, the weaker notion of *pre-k-Du Bois* singularities was introduced in [SVV]; this simply means the vanishing of higher cohomologies, i.e., $\mathcal{H}^i \underline{\Omega}_X^p = 0$ for all $i > 0$ and $p \le k$. See also [Tig], using different terminology.

For fixed k, the question of whether the cohomology sheaf $\mathcal{H}^i \underline{\Omega}_X^k$ is zero for some $i > 0$ is therefore by definition nontrivial only if X is not pre-k-Du Bois. Our main vanishing result studies the first degree where this is the case. The answer is influenced by the algebraic properties of the 0-th cohomology sheaf $\mathcal{H}^0 \underline{\Omega}_X^k$, which is known to always be torsion-free.

Theorem A. *Let X be a variety with pre-$(k-1)$-Du Bois isolated singularities. Then*

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for} \quad 0 < i < \operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k - 1.$$

A reformulation of this theorem in terms of resolution of singularities and higher direct images of logarithmic forms can be found in Remark 7.3.

When $k = 0$, when the hypothesis means that there are no assumptions on X, this is due to Steenbrink [St3, Proposition 1]. Using our method of proof, however, combined with an injectivity result from [KS16], we can extend this case to arbitrary singular sets.[1]

Corollary B. *If a variety X is (pre-)Du Bois away from a closed subset of dimension s, then*

$$\mathcal{H}^i \underline{\Omega}_X^0 = 0 \quad \text{for} \quad 0 < i < \operatorname{depth} \mathcal{O}_X - s - 1.$$

A quick consequence is that a Cohen–Macaulay variety of dimension n, with isolated singularities, is Du Bois if and only if the natural morphism $H^n(X, \mathcal{O}_X) \to H^n(X, \underline{\Omega}_X^0)$ is injective (hence an isomorphism); see Corollary 7.6.

The picture of vanishing results for higher cohomology sheaves is completed by the following "sliding" rule, which is quite simple but seems to not have been noted before; it holds with no assumption on the singular locus.

Proposition C. *Let X be an n-dimensional variety. If $k < n$ and $\mathcal{H}^{n-p-1} \underline{\Omega}_X^p = 0$ for all $p \le k-1$, then*

$$\mathcal{H}^{n-k} \underline{\Omega}_X^k = 0.$$

In particular $\mathcal{H}^n \underline{\Omega}_X^0 = 0$, and more generally if X is pre-$(k-1)$-Du Bois, with $k < n$, then $\mathcal{H}^{n-k} \underline{\Omega}_X^k = 0$.

Further vanishing results and conjectures in the nonisolated case are discussed later in the Introduction. Note in particular Conjecture H for an extension of Theorem A to the general case.

Injectivity for the cohomologies of duals. The key technical result of the paper, used in the proof of Theorem A and in other applications, is the following injectivity theorem for the cohomologies of the dual of the first Du Bois complex that is not a sheaf.

[1] We will see that for $k = 0$ one can safely replace $\mathcal{H}^0 \underline{\Omega}_X^0$ by $\mathcal{O}_X$.

Theorem D. *Let X be a variety with isolated pre-$(k-1)$-Du Bois singularities. Then the morphism*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet)$$

in $\mathbf{D}^b_{\mathrm{coh}}(X)$, obtained by dualizing the canonical map $\mathcal{H}^0\underline{\Omega}_X^k \to \underline{\Omega}_X^k$, is injective on cohomology.[2]

In Conjecture G we predict that the statement should hold even without the isolated singularities hypothesis.

An injectivity theorem of this type first appeared in the inspiring paper [KS16] by Kovács and Schwede for $k = 0$. It was then reinterpreted in terms of the Hodge filtration on local cohomology, and extended to arbitrary k in the case of local complete intersections, in [MP1] and [MP2]. Fundamentally, such injectivity theorems are degeneration at E_1 phenomena for appropriate Hodge-theoretic objects, and are now understood to be one of the most essential properties of Du Bois complexes.

Note a subtlety: when $k = 0$, or when X is LCI, Kähler differentials are rather well behaved under the $(k-1)$-Du Bois hypothesis, and therefore the right-hand side in the previously known injectivity theorems is expressed in terms of Ω_X^k; see e.g., Theorem 4.6. This is not the case anymore for arbitrary singularities, where we found that the natural formulation of injectivity is in terms of $\mathcal{H}^0\underline{\Omega}_X^k$. Nevertheless, one can deduce from Theorem D statements about Kähler differentials as well. Here we only include a special case that is easier to state, while the general result is Corollary 6.2; the local cohomological defect $\mathrm{lcdef}(X)$ is defined in the next subsection, and the definition of k-Du Bois singularities is recalled in Section 2.

Corollary E. *Let X be a variety with isolated $(k-1)$-Du Bois singularities, with $\dim X \geq 2$ and $\mathrm{lcdef}(X) = 0$. Then the map*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

is injective on cohomology. Here $\Omega_{X,\mathrm{tf}}^k$ denotes the quotient of Ω_X^k by its torsion subsheaf.

As for the proof of Theorem D, it is immediate to reduce to the case of projective varieties, in which case we show a more general fact, namely that the statement holds when X is pre-$(k-1)$-Du Bois with possibly higher dimensional singular locus, but pre-k-Du Bois except at finitely many points; see Theorem 5.1. We use the degeneration at E_1 of the Du Bois version of the Hodge-to-de Rham spectral sequence, inspired by the approach in [KS16] rather than that in [MP2] (as we do not yet have a good theory of the Hodge filtration on local cohomology in the non-LCI case).

Another application of Theorem D is a very quick proof of the known fact that k-rational singularities are k-Du Bois, in the case of isolated singularities. We show this in Section 8, where we also recall what is known in this direction.

More on vanishing. Going back to vanishing statements, our approach also provides a somewhat weaker statement for nonisolated singularities, which is essentially due to formal homological algebra.

[2] Here $\omega_X^\bullet$ is the dualizing complex of X.

Proposition F. *Let X be a variety that is pre-k-Du Bois away from a closed subset of dimension s. Then*

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \textit{for } 0 < i < m_k - s - 1,$$

where $m_k := \min\{\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k,\ n - k - \operatorname{lcdef}(X) + 1\}$.

Here $\operatorname{lcdef}(X)$ denotes the *local cohomological defect* of X, defined as

$$\operatorname{lcdef}(X) = \operatorname{lcd}(X, Y) - \operatorname{codim}_Y(X),$$

where $X \subseteq Y$ is an embedding in a smooth variety, with local cohomological dimension $\operatorname{lcd}(X, Y)$. It does not depend on the embedding, and $\operatorname{lcdef}(X) = 0$ for local complete intersections, but also for other interesting classes of varieties; for more details see Section 4. One can make sense of the depth of an object in the derived category, and what is really proven in Proposition F is a consequence of a general fact about arbitrary such objects, namely the same statement but with m_k replaced by the more abstract

$$m_k' := \min\{\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k,\ \operatorname{depth} \underline{\Omega}_X^k + 1\}.$$

One then uses one of the main results of [MP1], which implies that

$$\operatorname{depth} \underline{\Omega}_X^k \geq n - k - \operatorname{lcdef}(X),$$

with equality for some k. To prove Theorem A, we need to combine this abstract formulation with the main injectivity result, Theorem D.

We explain in Example 7.4 how this formal statement can be used when X is a local complete intersection in order to recover [MP1, Corollary 13.9] (see also [MOPW] for hypersurfaces), stating that if $s = \dim X_{\text{sing}}$, then for all k we have

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < n - k - s - 1.$$

At the end of Section 7 we give examples where certain intermediate cohomologies do not vanish, showing the failure of some possible extensions of this result to the general setting.

Conjectures. The main results of this paper are likely to admit natural extensions to the case of nonisolated singularities. The most important extends Theorem D.

Conjecture G. *Let X be a variety with pre-$(k-1)$-Du Bois singularities. Then the map*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0 \underline{\Omega}_X^k, \omega_X^\bullet)$$

is injective on cohomology.

This is proven in [MP2] in the local complete intersection case, when X has $(k-1)$-Du Bois singularities. The proof makes use however of the relationship between the Hodge filtration and the Ext filtration on local cohomology, shown in [MP1], which is not available in general.

The natural extension of Theorem A is the statement below. It follows from Conjecture G with the same argument that derives Theorem A from Theorem D.

Conjecture H. *Let X be a variety that is pre-$(k-1)$-Du Bois, and pre-k-Du Bois away from a closed subset of dimension s. Then*

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \textit{for } 0 < i < \operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k - s - 1.$$

Intersection complex. In Ch.E we establish analogues of the results described above, where the Du Bois complexes are replaced by *intersection Du Bois complexes*;[3] in other words, from the point of view of Hodge modules and constructible sheaves, we are replacing the constant sheaf by the intersection complex. In this case, the higher Du Bois singularities conditions are replaced by higher rational singularities analogues; for local complete intersections, the link between these types of singularities and intersection complexes was already observed in [CDM]. Since their shape is rather similar, we refer to Sections 10 and 11 in the body of the paper for these statements. The main injectivity result is Corollary 10.4, while the main vanishing result is Corollary 11.3.

What is perhaps more fundamental here is that along the way we establish an injectivity result, Theorem 10.3, that holds unconditionally and relates the duals of Du Bois complexes and their intersection analogues. This in turn uses a more technical variant, Theorem 10.5, communicated to us by S. G. Park. Using this theorem, the main result in the intersection complex setting is a consequence of Theorem D for Du Bois complexes.

B. Preliminaries

Throughout this chapter, X is a complex variety of dimension n.

1. ***Du Bois complexes.*** We recall the notion of *filtered de Rham complex*, meant as a replacement for the standard de Rham complex on smooth varieties. Denoted $(\underline{\Omega}_X^\bullet, F)$, it is an object in the bounded derived category of filtered differential complexes on X, introduced by Du Bois in [DB] along the lines suggested by work of Deligne. For each $k \geq 0$, the shifted associated graded quotient

$$\underline{\Omega}_X^k := \operatorname{Gr}_F^k \underline{\Omega}_X^\bullet[k],$$

is an object in $\mathbf{D}^b_{\mathrm{coh}}(X)$, called the *$k$-th Du Bois complex* of X. For a hyperresolution $\epsilon_\bullet : X_\bullet \to X$ of X, it can be computed as

$$\underline{\Omega}_X^k \simeq \mathbf{R}\epsilon_{\bullet *} \Omega_{X_\bullet}^k.$$

Besides [DB], one can find a detailed treatment of hyperresolutions and the construction of Du Bois complexes in [GNPP, Chapter V] or [PS, Chapter 7.3]. We only recall here a few basic facts that will be used freely throughout the paper. Note that in these statements we jump freely between the algebraic and analytic setting without changing the notation, as we hope that the context is clear in each case; the analytic results will be used in the projective setting, when the cohomology groups are the same due to GAGA.

[3] These are studied more thoroughly in the upcoming [PP].

- For each $k \geq 0$, there is a canonical morphism $\Omega_X^k \to \underline{\Omega}_X^k$, which is an isomorphism if X is smooth; here Ω_X^k are the sheaves of Kähler differentials on X; see [DB, Section 4.1] or [PS, Page 175]. In particular, $\mathcal{H}^i \underline{\Omega}_X^k$ are supported on the singular locus of X, for all $i > 0$.
- There exists a Hodge-to-de Rham spectral sequence

$$E_1^{p,q} = \mathbb{H}^q(X, \underline{\Omega}_X^p) \Longrightarrow H^{p+q}(X, \mathbb{C}),$$

which degenerates at E_1 if X is projective; see [DB, Theorem 4.5(iii)] or [PS, Proposition 7.24].

- For each $k \geq 0$, the sheaf $\mathcal{H}^0 \underline{\Omega}_X^k$ embeds in $f_* \Omega_{\widetilde{X}}^k$, where $f : \widetilde{X} \to X$ is a resolution of X, so in particular it is torsion-free; see [HJ, Remark 3.8].

2. ***Higher singularities.*** Following [MOPW; JKSY; FL2], if X is a local complete intersection (lci) subvariety of a smooth variety Y, then it is said to have *k-Du Bois singularities* if the canonical morphisms $\Omega_X^p \to \underline{\Omega}_X^p$ are isomorphisms for all $0 \leq p \leq k$, and *k-rational singularities* if the canonical morphisms $\Omega_X^p \to \mathbf{D}_X(\underline{\Omega}_X^{n-p})$ are isomorphisms for all $0 \leq p \leq k$, where $\mathbf{D}_X(\,\cdot\,) := \mathbf{R}\mathcal{H}om(\,\cdot\,, \omega_X)$.

For non-lci varieties, however, even the condition $\Omega_X^1 \xrightarrow{\sim} \underline{\Omega}_X^1$ turns out to be quite restrictive; as explained in [SVV] the definitions above are not suitable anymore. In the general setting, new definitions of k-Du Bois and k-rational singularities are introduced in *loc. cit.*. As it is often sufficient, one can first consider weaker notions obtained by removing the conditions in cohomological degree 0.

Definition 2.1. We say that X has *pre-k-Du Bois* singularities if

$$\mathcal{H}^i \underline{\Omega}_X^p = 0 \quad \text{for all} \quad i > 0 \quad \text{and} \quad 0 \leq p \leq k.$$

We say that X has *pre-k-rational* singularities if

$$\mathcal{H}^i(\mathbf{D}_X(\underline{\Omega}_X^{n-p})) = 0 \quad \text{for all} \quad i > 0 \quad \text{and} \quad 0 \leq p \leq k.$$

Several other conditions are imposed in the full definition of general k-Du Bois and k-rational singularities. They agree with the classical notions of Du Bois and rational singularities when $k = 0$, and with the definitions mentioned above in the local complete intersection case. See [SVV, Proposition 5.5, 5.6] for more details.

Definition 2.2. We say that X has *k-Du Bois singularities* if it is seminormal, and

(1) $\operatorname{codim}_X(X_{\mathrm{sing}}) \geq 2k + 1$;

(2) X has pre-k-Du Bois singularities;

(3) $\mathcal{H}^0 \underline{\Omega}_X^p$ is reflexive, for all $p \leq k$.

Definition 2.3. We say that X has *k-rational singularities* if it is normal, and

(1) $\operatorname{codim}_X(X_{\mathrm{sing}}) > 2k + 1$;

(2) X has pre-k-rational singularities.

3. ***A new vanishing result.*** Steenbrink's vanishing theorem [St2, (4.1)] states that

$$\mathcal{H}^q \underline{\Omega}_X^p = 0 \quad \text{for } p+q > n. \tag{3.1}$$

In general, this result is the best possible. Indeed, [MOPW, Example 1.7] shows that the vanishing does not necessarily hold when $p+q=n$. However, we have:

Proposition 3.2 (Proposition C). *If $k<n$ and $\mathcal{H}^{n-p-1}\underline{\Omega}_X^p = 0$ for all $p \le k-1$, then*

$$\mathcal{H}^{n-k}\underline{\Omega}_X^k = 0.$$

Proof. Consider the spectral sequence associated to the Hodge filtration on the filtered de Rham complex

$$E_1^{p,q} := \mathcal{H}^q \underline{\Omega}_X^p \Longrightarrow \mathcal{H}^{p+q}\underline{\Omega}_X^\bullet.$$

Since $\underline{\Omega}_X^\bullet$ is quasi-isomorphic to $\mathbb{C}_X$, the spectral sequence converges to $\mathbb{C}_X$, placed in cohomological degree 0. Note that for any $\ell \ge 1$, the term $E_{\ell+1}^{k,n-k}$ is obtained as the cohomology of the complex

$$E_\ell^{k-\ell,n-k+\ell-1} \to E_\ell^{k,n-k} \to E_\ell^{k+\ell,n-k-\ell+1},$$

and the right-hand side is 0 by (3.1), while the left-hand side is 0 by assumption. Therefore

$$\mathcal{H}^{n-k}\underline{\Omega}_X^k = E_1^{k,n-k} = E_\infty^{k,n-k} = 0. \qquad \square$$

Corollary 3.3. *If X is pre-$(k-1)$-Du Bois, with $k<n$, then*

$$\mathcal{H}^{n-k}\underline{\Omega}_X^k = 0.$$

In particular, X is pre-k-Du Bois for all k if and only if it is pre-$(n-2)$-Du Bois.

Proof. The first part is an immediate consequence of Proposition C. For the second part, note that we always have $\underline{\Omega}_X^n \simeq \mathcal{H}^0\underline{\Omega}_X^n$, while in the previous row, the only term that needs checking is $\mathcal{H}^1\underline{\Omega}_X^{n-1}$, which is covered by the first part. $\square$

When X has isolated singularities, this is [FL1, Lemma 2.5]. When X is a hypersurface, a slightly weaker result is contained in [MOPW, Theorem 1.4].

Remark 3.4. The result does not hold for $k=n$, as for any variety X we have $\underline{\Omega}_X^n \simeq \mathcal{H}^0\underline{\Omega}_X^n \simeq \pi_*\omega_{\widetilde{X}}$, where $\pi : \widetilde{X} \to X$ is a resolution of singularities.

Example 3.5. If X is a pre-0-Du Bois surface, then $\mathcal{H}^i\underline{\Omega}_X^k = 0$ for all k and all $i>0$. Hence if X is a surface with rational singularities, then

$$\underline{\Omega}_X^k \simeq \mathcal{H}^0\underline{\Omega}_X^k \simeq \Omega_X^{[k]} \quad \text{for all } k.$$

The last isomorphism is a consequence of the main result of [KS21].

Example 3.6. Corollary 3.3 cannot be improved, without further assumptions, by moving to the left in the Du Bois table. For instance, any 3-fold X with an isolated rational (hence Du Bois) hypersurface singularity that is not a double point, has $\mathcal{H}^1\underline{\Omega}_X^1 \neq 0$; see [NS, Theorem 2.2].

4. ***Local cohomological dimension and depth.*** Let X be a complex variety. If Y is a smooth variety containing X (locally), the local cohomological dimension of X in Y is

$$\mathrm{lcd}(X, Y) := \max\{q \mid \mathcal{H}^q_X \mathcal{O}_Y \neq 0\},$$

where the sheaf in parenthesis is the q-th local cohomology sheaf of $\mathcal{O}_Y$ along X. It is also known that if $r = \mathrm{codim}_Y X$, then $\mathcal{H}^q_X \mathcal{O}_Y = 0$ for $q < r$ and $\mathcal{H}^r_X \mathcal{O}_Y \neq 0$. See [MP1], for example, for details and references.

As in [PSh], we consider the *local cohomological defect* $\mathrm{lcdef}(X)$ of X as

$$\mathrm{lcdef}(X) := \mathrm{lcd}(X, Y) - \mathrm{codim}_Y X.$$

A reinterpretation of the characterization of local cohomological dimension in [MP1, Theorem E] can be stated as follows:

Theorem 4.1 [MP1, Corollary 12.6]. *Let X be a subvariety of a smooth variety Y. For any integer c we have*

$$\mathrm{lcd}(X, Y) \leq c \iff \mathcal{E}xt^{j+k+1}_{\mathcal{O}_Y}(\underline{\Omega}^k_X, \omega_Y) = 0 \quad \textit{for all } \; j \geq c \textit{ and } k \geq 0.$$

or equivalently

$$\mathrm{lcdef}(X) \leq c \iff \mathcal{E}xt^{j+k+1}_{\mathcal{O}_X}(\underline{\Omega}^k_X, \omega^\bullet_X) = 0 \quad \textit{for all } \; j \geq c - \dim X \textit{ and } k \geq 0.$$

The second equivalence follows from the first thanks to Grothendieck duality for the inclusion $X \hookrightarrow Y$.

We now recall that the notion of depth of a module has a natural extension to objects in the derived category. If C is an element of the bounded derived category of finitely generated R-modules, where $(R, \mathfrak{m})$ is a noetherian local ring endowed with a dualizing complex $\omega^\bullet_R$, then one can define

$$\mathrm{depth}(C) := \min\{i \mid \mathrm{Ext}^{-i}_R(C, \omega^\bullet_R) \neq 0\}.$$

with the convention that the depth is $-\infty$ if $C = 0$. This notion is studied extensively in [FY], where it is shown to be equivalent to other natural generalizations of the usual notion of depth. When X is a variety and $\mathcal{C}$ is an element in $D^b_{\mathrm{coh}}(X)$, then we set

$$\mathrm{depth}(\mathcal{C}) := \min_{x \in \mathrm{Supp}(\mathcal{C})} \mathrm{depth}(\mathcal{C}_x),$$

where the minimum is taken over the closed points in the support of $\mathcal{C}$. The first interpretation takes the form

$$\mathrm{depth}(\mathcal{C}) = \min\{i \mid \mathcal{E}xt^{-i}_{\mathcal{O}_X}(\mathcal{C}, \omega^\bullet_X) \neq 0\}. \tag{4.2}$$

This is of course a standard interpretation of depth when $\mathcal{C}$ is a sheaf.

Using this, from Theorem 4.1 we conclude:

Corollary 4.3. *We have the identity*

$$\mathrm{lcdef}(X) = \dim X - \min_{k \geq 0}\{\mathrm{depth}\, \underline{\Omega}^k_X + k\}.$$

This shows in particular that $\mathrm{lcdef}(X)$ depends only on X, and not on the embedding, and that $\dim X \geq \mathrm{lcdef}(X) \geq 0$. These consequences can also be deduced from the topological interpretation of $\mathrm{lcdef}(X)$ as the number of nonzero perverse cohomologies of the constant sheaf $\mathbb{Q}_X$, shown in [RSW].

Example 4.4 (varieties with $\mathrm{lcdef}(X) = 0$). The condition $\mathrm{lcdef}(X) = 0$ is equivalent to $\mathrm{lcd}(X, Y) = \mathrm{codim}_Y X$ in any embedding, or equivalently to the nonvanishing of a single local cohomology sheaf $\mathcal{H}^r_X \mathcal{O}_Y$. This holds of course when X is a local complete intersection. In addition, it is known to hold when X has quotient singularities [MP1, Corollary 11.22], for affine varieties with Stanley–Reisner coordinate algebras that are Cohen–Macaulay [MP1, Corollary 11.26], for arbitrary Cohen–Macaulay surfaces [Og, Remark, pp. 338–339] and threefolds [DT, Corollary 2.8], and for Cohen–Macaulay fourfolds whose local analytic Picard groups are torsion [DT, Theorem 1.3].

Note that, according to Corollary 4.3, for such varieties we have

$$\operatorname{depth} \underline{\Omega}^k_X \geq n - k \quad \text{for all } k \geq 0.$$

We finish the chapter with a well-known vanishing result for Ext sheaves, for later use.

Lemma 4.5 (e.g., [Sta, Tag 0A7U]). *Let $\mathcal{F}$ be a coherent sheaf on X. Then*

$$\mathcal{E}xt^i_{\mathcal{O}_X}(\mathcal{F}, \omega^\bullet_X) = 0 \quad \textit{for } i < -\dim \operatorname{Supp}(\mathcal{F}).$$

C. Injectivity theorems

In this chapter we address natural injectivity theorems for the cohomologies of the duals of the graded quotients of the Du Bois complex. The first appearance of such a result was in [KS16, Theorem 3.3], where Kovács and Schwede proved that for every variety X the morphism

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}^0_X, \omega^\bullet_X) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{O}_X, \omega^\bullet_X)$$

obtained by dualizing the canonical morphism $\mathcal{O}_X \to \underline{\Omega}^0_X$ is injective on cohomology. Using the Hodge filtration on local cohomology, a slightly stronger version of this fact was obtained in [MP1, Theorem A], and then extended to higher Du Bois complexes in [MP2, Theorem A] in the case of local complete intersections:

Theorem 4.6 [MP2, Theorem A]. *If X is local complete intersection and k is a nonnegative integer such that X has $(k-1)$-Du Bois singularities, then the morphism*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}^k_X, \omega_X) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega^k_X, \omega_X)$$

in $\mathbf{D}^b_{\mathrm{coh}}(X)$, obtained by dualizing the canonical morphism $\Omega^k_X \to \underline{\Omega}^k_X$, is injective at the level of cohomology.

Here we prove an injectivity theorem for arbitrary isolated singularities, by going back to the basic Hodge-theoretic properties of Du Bois complexes, as in [KS16]. Currently we do not have a sufficiently good understanding of the Hodge filtration on local cohomology, as a mixed Hodge module, beyond the local complete intersection case treated in [MP1]. This is something highly desirable, which may clarify the picture in the general nonisolated case and lead to a proof of Conjecture G.

5. *Proof of Theorem D.* The statement of the injectivity theorem is local; hence we may assume first that X is quasiprojective. Since the singular locus S of X is a finite set, we may choose a compactification $\overline{X}$ of X such that the singular locus of $\overline{X}$ is still S, and prove the statement for $\overline{X}$. Hence it suffices to assume that X is projective to begin with. With this assumption, we prove a stronger statement:

Theorem 5.1. *Let X be a projective variety which is pre-$(k-1)$-Du Bois, and pre-k-Du Bois away from a finite set. Then the morphism*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet)$$

in $\mathbf{D}^b_{\mathrm{coh}}(X)$, obtained by dualizing the canonical map $\mathcal{H}^0\underline{\Omega}_X^k \to \underline{\Omega}_X^k$, is injective on cohomology.

The key point in the proof is the following:

Proposition 5.2. *Let X be a projective variety with pre-$(k-1)$-Du Bois singularities. Then for each i, the natural map*

$$H^i(X, \mathcal{H}^0\underline{\Omega}_X^k) \to \mathbb{H}^i(X, \underline{\Omega}_X^k),$$

obtained by applying cohomology to $\mathcal{H}^0\underline{\Omega}_X^k \to \underline{\Omega}_X^k$, is surjective.

Proof. For each $p \geq 0$, we set

$$\underline{\Omega}_X^{\leq p} := \underline{\Omega}_X^\bullet / F^{p+1}\underline{\Omega}_X^\bullet.$$

So we have an exact triangle

$$\underline{\Omega}_X^p[-p] \longrightarrow \underline{\Omega}_X^{\leq p} \longrightarrow \underline{\Omega}_X^{\leq p-1} \xrightarrow{+1} . \tag{5.3}$$

We also denote by $\Omega_{X,h}^{\leq p}$ the object in the derived category of differential complexes on X,[4] represented by the complex

$$\big[\mathcal{H}^0\underline{\Omega}_X^0 \xrightarrow{d} \mathcal{H}^0\underline{\Omega}_X^1 \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{H}^0\underline{\Omega}_X^p\big],$$

placed in cohomological degrees $0, \ldots, p$. This is not to be confused with $\mathcal{H}^0(\underline{\Omega}_X^{\leq p})$. Here we have an exact triangle

$$\mathcal{H}^0\underline{\Omega}_X^p[-p] \longrightarrow \Omega_{X,h}^{\leq p} \longrightarrow \Omega_{X,h}^{\leq p-1} \xrightarrow{+1} . \tag{5.4}$$

As in [SVV, Proposition 2.3], there exists a natural map $\Omega_{X,h}^{\leq p} \to \underline{\Omega}_X^{\leq p}$. When X is projective, the E_1-degeneration of the Hodge-to-de Rham spectral sequence for the filtered de Rham complex of X implies that the induced composition

$$H^i(X, \mathbb{C}) \to \mathbb{H}^i(X, \Omega_{X,h}^{\leq p}) \to \mathbb{H}^i(X, \underline{\Omega}_X^{\leq p})$$

is surjective for each i; hence so is the second map.

[4]The notation is motivated by the fact that $\mathcal{H}^0\underline{\Omega}_X^k$ agrees with the h-differentials $\Omega_{X,h}^k$ studied in [HJ].

Let's now consider the integer k in the statement. The map $\Omega_{X,h}^{\leq k} \to \underline{\Omega}_X^{\leq k}$ and its analogue for $k-1$, combined with the two exact triangles described above, give rise to a morphism of exact triangles

$$\begin{array}{ccccc} \mathcal{H}^0\underline{\Omega}_X^k[-k] & \longrightarrow & \Omega_{X,h}^{\leq k} & \longrightarrow & \Omega_{X,h}^{\leq k-1} \xrightarrow{+1} \\ \downarrow & & \downarrow & & \downarrow \\ \underline{\Omega}_X^k[-k] & \longrightarrow & \underline{\Omega}_X^{\leq k} & \longrightarrow & \underline{\Omega}_X^{\leq k-1} \xrightarrow{+1} \end{array}$$

Since X is pre-$(k-1)$-Du Bois, the right-most vertical map is an isomorphism. Passing to hypercohomology, we obtain a morphism of long exact sequences

$$\begin{array}{ccccccc} \mathbb{H}^{i-1}(\Omega_{X,h}^{\leq k-1}) & \longrightarrow & H^i(\mathcal{H}^0\underline{\Omega}_X^k[-k]) & \longrightarrow & \mathbb{H}^i(\Omega_{X,h}^{\leq k}) & \longrightarrow & \mathbb{H}^i(\Omega_{X,h}^{\leq k-1}) \\ \downarrow\simeq & & \downarrow & & \downarrow & & \downarrow\simeq \\ \mathbb{H}^{i-1}(\underline{\Omega}_X^{\leq k-1}) & \longrightarrow & \mathbb{H}^i(\underline{\Omega}_X^k[-k]) & \longrightarrow & \mathbb{H}^i(\underline{\Omega}_X^{\leq k}) & \longrightarrow & \mathbb{H}^i(\underline{\Omega}_X^{\leq k-1}) \end{array}$$

where the first and last vertical maps are isomorphisms. Since the third vertical map is surjective for all i, basic homological algebra shows that so is the second. □

We now consider the exact triangle

$$\mathcal{H}^0\underline{\Omega}_X^k \to \underline{\Omega}_X^k \to C \xrightarrow{+1} .$$

By definition X is pre-k-Du Bois away from a finite set of points if and only if C is supported on a finite set. After dualizing, we obtain an exact triangle

$$K \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet) \xrightarrow{+1},$$

where again K is supported on a finite set. Applying Grothendieck-Serre duality to the surjections in Proposition 5.2, we obtain that the induced morphisms

$$\mathbb{H}^i\big(X, \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet)\big) \to \mathbb{H}^i\big(X, \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet)\big)$$

are injective for all integers i.

Theorem 5.1 is then a consequence of the following general result:

Lemma 5.5. *Let X be a projective variety, and let*

$$K \to F \to G \xrightarrow{+1}$$

be an exact triangle in $\mathbf{D}^b_{\mathrm{coh}}(X)$. Suppose that K has zero-dimensional support, and that the induced maps on hypercohomology

$$\mathbb{H}^i(X, F) \to \mathbb{H}^i(X, G)$$

are injective for all i. Then the induced maps on cohomology

$$\mathcal{H}^i F \to \mathcal{H}^i G$$

are injective for all i.

Proof. The injectivity on hypercohomology implies that for each i we have short exact sequences:

$$0 \to \mathbb{H}^i(X, F) \to \mathbb{H}^i(X, G) \to \mathbb{H}^{i+1}(X, K) \to 0.$$

Now the hypercohomology of G is computed by a spectral sequence

$$E_2^{p,q} = H^p(X, \mathcal{H}^q G) \Longrightarrow \mathbb{H}^{p+q}(X, G),$$

while the similar spectral sequence for K shows that

$$\mathbb{H}^{i+1}(X, K) \simeq H^0(X, \mathcal{H}^{i+1} K),$$

because of the assumption that K is supported in dimension zero. Passing to the first associated graded term of the filtration on the total object in each of these two cases leads to a commutative diagram

$$\begin{array}{ccc} \mathbb{H}^i(X, G) & \longrightarrow & \mathbb{H}^{i+1}(X, K) \\ \downarrow & & \downarrow \\ E_\infty^{0,i} & \longrightarrow & H^0(X, \mathcal{H}^{i+1} K) \end{array}$$

and by the observations above, it follows that the bottom horizontal map is surjective. On the other hand, note that in fact this map has a factorization

$$E_\infty^{0,i} \hookrightarrow E_2^{0,i} = H^0(X, \mathcal{H}^i G) \xrightarrow{\varphi} H^0(X, \mathcal{H}^{i+1} K),$$

where φ comes from the connecting homomorphism $\mathcal{H}^i G \to \mathcal{H}^{i+1} K$ induced by the original triangle. Since the support of $\mathcal{H}^{i+1} K$ is zero-dimensional, this connecting homomorphism is surjective for each i, which is equivalent to our assertion. $\square$

6. *Injectivity results involving Kähler differentials.* Recall that for any $k \geq 0$ we have natural maps

$$\Omega_X^k \longrightarrow \mathcal{H}^0 \underline{\Omega}_X^k \longrightarrow \underline{\Omega}_X^k.$$

Since $\mathcal{H}^0 \underline{\Omega}_X^k$ is known to be torsion-free, this arises in fact from a sequence of maps

$$\Omega_{X,\mathrm{tf}}^k \xrightarrow{\alpha} \mathcal{H}^0 \underline{\Omega}_X^k \xrightarrow{\beta} \underline{\Omega}_X^k, \tag{6.1}$$

where $\Omega_{X,\mathrm{tf}}^k := \Omega_X^k / \mathrm{tors}(\Omega_X^k)$ is the canonical torsion-free quotient of the sheaf of Kähler differentials, and α is an inclusion which is an isomorphism away from X_{sing}. Our main injectivity theorem has the following consequence:

Corollary 6.2. *Let X be a variety with isolated pre-$(k-1)$-Du Bois singularities. Then the dual*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

of the canonical morphism (6.1) *is injective on the i-th cohomology for all $i \neq 0$. Moreover*, *if* $\mathrm{lcdef}(X) \leq \dim X - k - 1$, *then it is injective on all cohomologies.*

Proof. Given Theorem D, for the first statement it is enough to have the injectivity on cohomology of the map

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

obtained by dualizing α, in the range $i \neq 0$. But this is clear; if we complete α to a short exact sequence

$$0 \longrightarrow \Omega_{X,\mathrm{tf}}^k \longrightarrow \mathcal{H}^0\underline{\Omega}_X^k \longrightarrow Q \longrightarrow 0,$$

the cokernel Q is supported on X_{sing}, hence Lemma 4.5 and (4.2) imply $\mathcal{E}xt^i(Q, \omega_X^\bullet) = 0$ for $i \neq 0$. Dualizing the short exact sequence above then shows that in this range

$$\mathcal{E}xt^i_{\mathcal{O}_X}(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathcal{E}xt^i_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

is injective. For the second statement, simply note that Theorem 4.1 implies that

$$\mathcal{E}xt^i_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) = 0 \quad \text{for } i > \mathrm{lcdef}(X) - \dim X + k,$$

so that the assumption takes care of the remaining case $i = 0$. □

We compare this with the previous injectivity statements obtained in the literature.

- When $k = 0$, we recall that the injectivity on cohomology of the canonical morphism

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^0, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om(\mathcal{O}_X, \omega_X^\bullet)$$

holds in full generality thanks to [KS16].

- When $k \geq 1$, for isolated singularities we obtain the extension of a result shown in the local complete intersection case in [MP2]. Note first that by definition, when X is $(k-1)$-Du Bois, rather than just pre-$(k-1)$-Du Bois, we have $\mathrm{codim}\, X_{\mathrm{sing}} \geq 2k-1$; when X is a local complete intersection, this is not part of the definition, but holds automatically by [MP1, Theorem F and Corollary 9.26]. In our case this simply means $n-k-1 \geq k-2$; hence as a consequence of Corollary 6.2 we first obtain:

Corollary 6.3. *When X has isolated $(k-1)$-Du Bois singularities*, *and* $\mathrm{lcdef}(X) \leq k-2$, *the map in Corollary 6.2 is injective on all cohomologies.*

We deduce the promised analogue of the result in [MP2]:

Corollary 6.4. *Let X be a variety with isolated $(k-1)$-Du Bois singularities. If* $\mathrm{lcdef}(X) = 0$, *then the map*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

is injective on cohomology.

Proof. For $k = 0$ the result holds with no assumptions, so we may assume $k \geq 1$. When $k \geq 2$, the result is a consequence of Corollary 6.3. When $k = 1$ and $n \geq 2$, we go back directly to the statement of Corollary 6.2. When $k = n = 1$, since $\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^1 = \operatorname{depth} \Omega_{X,\mathrm{tf}}^1 = 1$, the map

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_{X,\mathrm{tf}}^k, \omega_X^\bullet)$$

is nontrivial on the i-th cohomology only when $i = -1$, in which case its injectivity follows from Corollary 6.2. □

This applies in particular when X is a local complete intersection. When $\dim X \geq 2$ or $k \geq 2$, under our assumptions [MV, Corollary 3.1] implies that Ω_X^k is torsion-free, hence the injectivity on cohomology holds directly for

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\Omega_X^k, \omega_X^\bullet).[5]$$

However, other interesting classes of varieties satisfy $\operatorname{lcdef}(X) = 0$ as well, and Corollary 6.4 also holds for those; see Example 4.4.

Remark 6.5 (nonisolated singularities). Assuming Conjecture G (or under the hypothesis of Theorem 5.1), one has analogues of the results in this section for projective varieties with possibly nonisolated singularities. The conclusion of Corollary 6.2 becomes the fact that injectivity holds on i-th cohomology for:

(1) $i > \operatorname{lcdef}(X) - \dim X + k$; in this case in fact $\mathcal{E}xt^i_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) = 0$.

(2) $i < -\dim X_{\mathrm{sing}}$.

D. Applications of injectivity

7. *Vanishing of higher cohomology.* In this section we prove Theorem A and explain some related points.

We first state a general homological result about the vanishing of cohomologies of objects in the derived category of coherent sheaves, in terms of their depth.

Proposition 7.1. *Let $A^\bullet$ be an object in $\mathbf{D}^b_{\mathrm{coh}}(X)$ such that*

(1) *$A^\bullet$ has nontrivial cohomology only in nonnegative degrees.*

(2) *The support of all $\mathcal{H}^i A^\bullet$ with $i > 0$ is contained in a closed subset of X of dimension s.*

Then we have

$$\mathcal{H}^i A^\bullet = 0 \quad \text{for } 0 < i < \min\{\operatorname{depth} \mathcal{H}^0 A^\bullet, \operatorname{depth} A^\bullet + 1\} - s - 1.$$

Proof. The first assumption implies that there is a natural morphism $\mathcal{H}^0 A^\bullet \to A^\bullet$, and taking its Grothendieck dual leads to

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(A^\bullet, \omega_X^\bullet) \xrightarrow{\varphi} \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathcal{H}^0 A^\bullet, \omega_X^\bullet) \to C^\bullet \xrightarrow{+1}, \tag{7.2}$$

[5] In fact this can be easily seen to hold when $\dim X = k = 1$ as well.

where $C^\bullet$ is the cone of the morphism φ. For each q, we obtain exact sequences

$$\mathcal{E}xt^q_{\mathcal{O}_X}(\mathcal{H}^0 A^\bullet, \omega_X^\bullet) \to \mathcal{H}^q C^\bullet \to \mathcal{E}xt^{q+1}_{\mathcal{O}_X}(A^\bullet, \omega_X^\bullet).$$

Using (4.2), the first term vanishes for $q > -\operatorname{depth} \mathcal{H}^0 A^\bullet$, and the third for $q > -\operatorname{depth} A^\bullet - 1$. We conclude that

$$\mathcal{H}^q C^\bullet = 0 \text{ for } q > -m,$$

where $m := \min\{\operatorname{depth} \mathcal{H}^0 A^\bullet, \operatorname{depth} A^\bullet\}$.

Next, applying $\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(-, \omega_X^\bullet)$ to (7.2), we obtain an exact triangle

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(C^\bullet, \omega_X^\bullet) \to \mathcal{H}^0 A^\bullet \to A^\bullet \xrightarrow{+1} .$$

It follows that for $i > 0$,

$$\mathcal{H}^i A^\bullet \simeq \mathcal{E}xt^{i+1}_{\mathcal{O}_X}(C^\bullet, \omega_X^\bullet).$$

Now for the spectral sequence

$$E_2^{p,q} = \mathcal{E}xt^p_{\mathcal{O}_X}(\mathcal{H}^q C^\bullet, \omega_X^\bullet) \Longrightarrow \mathcal{E}xt^{p-q}_{\mathcal{O}_X}(C^\bullet, \omega_X^\bullet),$$

we have $\mathcal{H}^q C^\bullet = 0$ for $q > -m$, as noted above. Further, $\mathcal{E}xt^p_{\mathcal{O}_X}(\mathcal{H}^q C^\bullet, \omega_X^\bullet) = 0$ for $p < -\dim \operatorname{Supp} \mathcal{H}^q C^\bullet$, by Lemma 4.5. This holds for $p < -s$, since by definition $C^\bullet$ is supported on the locus where $\mathcal{H}^0 A^\bullet$ and $A^\bullet$ are not quasi-isomorphic, which has dimension s.

Combining these facts, we see that $\mathcal{E}xt^i_{\mathcal{O}_X}(C^\bullet, \omega_X^\bullet) = 0$ for $i < m - s$. Thus

$$\mathcal{H}^i A^\bullet = 0 \quad \text{for } 0 < i < m - s - 1. \qquad \square$$

Proof of Proposition F. We simply take $A^\bullet = \underline{\Omega}_X^k$ in Proposition 7.1. Its higher cohomologies are supported on the non-pre-k-Du Bois locus of X, so we obtain

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < m_k' - s - 1,$$

where

$$m_k' := \min\{\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k, \operatorname{depth} \underline{\Omega}_X^k + 1\}.$$

The result then follows from the characterization of the local cohomological defect in Corollary 4.3. $\square$

Proof of Theorem A. To deduce the stronger vanishing statement in the case of isolated singularities, the key new ingredient is that, thanks to Theorem D, with the notation in (1) the long exact sequence on cohomology associated to the triangle (7.2) breaks into short exact sequences

$$0 \to \mathcal{E}xt^i_{\mathcal{O}_X}(A^\bullet, \omega_X^\bullet) \to \mathcal{E}xt^i_{\mathcal{O}_X}(\mathcal{H}^0 A^\bullet, \omega_X^\bullet) \to \mathcal{H}^i C^\bullet \to 0$$

for all i.

The inclusion of Ext sheaves gives

$$\operatorname{depth} \underline{\Omega}_X^k \geq \operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k$$

on any variety with pre-$(k-1)$-Du Bois singularities. Hence in the proof of Proposition F we have $m_k' = \operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k$, and $s = 0$, which gives the desired result. □

The exact same argument, using the injectivity theorem of Kovács and Schwede [KS16] in place of Theorem D, proves Corollary B.

Remark 7.3. Let X be a variety with isolated singular locus S, and let $f : \widetilde{X} \to X$ be a resolution of singularities with simple normal crossings exceptional divisor $E = f^{-1}(S)_{\mathrm{red}}$. The vanishing result for the Du Bois complexes of X in Theorem A can be reformulated as saying that

$$R^i f_* \Omega_{\widetilde{X}}^k (\log E)(-E) = 0 \quad \text{for } 0 < i < \operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k - 1,$$

assuming that X has pre-$(k-1)$-Du Bois singularities. Indeed, by [St2, Proposition 3.3] there is an exact triangle

$$\mathbf{R} f_* \Omega_{\widetilde{X}}^k (\log E)(-E) \to \underline{\Omega}_X^k \to \Omega_S^k \xrightarrow{+1} .$$

Thus, for $k > 0$, it is immediate that

$$\mathcal{H}^i \underline{\Omega}_X^k \simeq R^i f_* \Omega_{\widetilde{X}}^k (\log E)(-E)$$

for all i. When $k = 0$ and $i > 0$, this isomorphism still holds: the map $\mathcal{H}^0 \underline{\Omega}_X^0 \to \mathcal{O}_S$ is surjective, since the composition $\mathcal{O}_X \to \mathcal{H}^0 \underline{\Omega}_X^0 \to \mathcal{O}_S$ is the natural surjection.

Example 7.4 (the LCI case). If X is a local complete intersection with $\dim X_{\mathrm{sing}} = s$, it is shown in [MP1, Corollary 13.9], using the Hodge filtration on local cohomology, that

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < n - s - k - 1. \tag{7.5}$$

We now explain that this fact is a special case of our results here. The statement is vacuous when $\operatorname{codim} X_{\mathrm{sing}} = n - s \le k + 2$, hence we may assume $n - s \ge k + 3$, which in particular implies that X is normal and moreover, by [MV, Corollary 3.1], that the sheaf of Kähler differentials Ω_X^k is reflexive. This in turn implies that $\mathcal{H}^0 \underline{\Omega}_X^k = \Omega_X^{[k]} = \Omega_X^k$. By Lemma 1.8 of [Gre], within this range we then have

$$\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k = \operatorname{depth} \Omega_X^k \ge n - k.$$

Since $\operatorname{lcdef}(X) = 0$, (7.5) then follows from Proposition F.

A criterion for the Du Bois condition. A simple but intriguing consequence of vanishing in the form of Corollary B is the next statement. A result of a similar flavor appears in [Ko12, Corollary 1.8], where there is no initial assumption on the singularities of X, but all cohomology groups are considered.

Corollary 7.6. *Let X be a projective seminormal Cohen–Macaulay variety of dimension n, with isolated singularities, or more generally Du Bois away from a finite set of points. If $H^n(X, \mathcal{O}_X) = 0$, then X is Du Bois.*

More precisely, we have $h^n(X, \mathcal{O}_X) \ge h^n(X, \underline{\Omega}_X^0)$, and X is Du Bois $\iff h^n(X, \mathcal{O}_X) = h^n(X, \underline{\Omega}_X^0)$ $\iff$ the natural map $H^n(X, \mathcal{O}_X) \to \mathbb{H}^n(X, \underline{\Omega}_X^0)$ is injective (hence an isomorphism).

Proof. We consider the cone

$$\mathcal{O}_X \to \underline{\Omega}_X^0 \to C^\bullet \xrightarrow{+1} .$$

Note that $C^\bullet$ is supported on a finite set. We clearly have $\mathcal{H}^i C^\bullet = 0$ for $i \leq 0$ (since the seminormality condition is equivalent to $\mathcal{O}_X \simeq \mathcal{H}^0 \underline{\Omega}_X^0$), while $\mathcal{H}^i C^\bullet \simeq \mathcal{H}^i \underline{\Omega}_X^0$ for $i \geq 1$. In particular, by Proposition C we have $\mathcal{H}^i C^\bullet = 0$ for $i > n-1$. Moreover, since X is Cohen–Macaulay, by Corollary B we have $\mathcal{H}^i C^\bullet = 0$ for $i < n-1$.

Note now that we have a short exact sequence

$$0 \to \mathbb{H}^{n-1}(X, C^\bullet) \to H^n(X, \mathcal{O}_X) \to \mathbb{H}^n(X, \underline{\Omega}_X^0) \to 0.$$

The last map is surjective thanks to the degeneration of the Hodge-to-de Rham spectral sequence, as in Section 5, as it sits in the surjective composition

$$H^n(X, \mathbb{C}) \to H^n(X, \mathcal{O}_X) \to \mathbb{H}^n(X, \underline{\Omega}_X^0).$$

The hypercohomology group $\mathbb{H}^{n-1}(X, C^\bullet)$ is computed by a spectral sequence whose E_2-terms are

$$E_2^{p,q} = H^p(X, \mathcal{H}^q C^\bullet), \quad \text{with } p+q = n-1.$$

Since $\mathcal{H}^i C^\bullet = 0$ for $i \neq n-1$, this gives

$$H^n(X, \mathcal{O}_X) \simeq \mathbb{H}^n(X, \underline{\Omega}_X^0) \iff H^0(X, \mathcal{H}^{n-1} C^\bullet) = 0.$$

As the support of $C^\bullet$ is finite, this last condition is equivalent to $\mathcal{H}^{n-1} C^\bullet = 0$, hence to $C^\bullet = 0$, i.e., to X being Du Bois. □

For instance, this applies to any low-degree normal complete intersection with isolated singularities in $\mathbb{P}^N$.

Remark 7.7. The criterion above has a (rather technical) analogue for higher k: using Theorem A, the same proof shows that if X is pre-$(k-1)$-Du Bois, and pre-k-Du Bois away from a finite set, and if $\operatorname{depth} \mathcal{H}^0 \underline{\Omega}_X^k = n-k$, then the cohomology vanishing $H^{n-k}(X, \mathcal{H}^0 \underline{\Omega}_X^k) = 0$ implies that X is pre-k-Du Bois.

Examples of nonvanishing. In [MP1, Question 13.10] it is asked whether the vanishing result for local complete intersections in Example 7.4, namely

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < n-s-k-1$$

with $s = \operatorname{codim} X_{\text{sing}}$, continues to hold when X is arbitrary, or at least Cohen–Macaulay.

The study of Du Bois complexes of cones in [SVV] and [PSh] provides simple counterexamples.

Example 7.8. First a very simple example that is not Cohen–Macaulay. Let $X = C(Y, L)$ be the abstract affine cone over a smooth projective threefold X endowed with an ample line bundle L such that $H^1(Y, L) \neq 0$.[6] Then, according to [SVV, Proposition 7.2], we have $\mathcal{H}^1 \underline{\Omega}_X^0 \neq 0$ and $\mathcal{H}^1 \underline{\Omega}_X^1 \neq 0$.

[6] For example, take $X = C \times C \times C$ for some smooth projective curve C of genus $g \geq 2$, and $L = \mathcal{O}_C(p) \boxtimes \mathcal{O}_C(p) \boxtimes \mathcal{O}_C(p)$ for some $p \in C$.

Example 7.9. In this example X has rational, hence Cohen–Macaulay, singularities. Let Y be a smooth Fano threefold for which

$$H^1(Y, \Omega_Y^2 \otimes L) \neq 0,$$

where $L = \omega_Y^{-1}$. The existence of such Y is shown in [Tot, Section 2].

Let now $X = C(Y \times Y, L \boxtimes L)$ be the abstract cone over $Y \times Y$ associated to the ample line bundle $L \times L$. Since $Y \times Y$ is still a Fano variety, we have

$$H^i(Y \times Y, (L \boxtimes L)^m) = 0 \quad \text{for all } i > 0, m \geq 0,$$

hence X has rational singularities; see e.g., [SVV, Remark 7.8]. Furthermore, we have

$$H^1(Y \times Y, \Omega^2_{Y\times Y} \otimes (L \boxtimes L)) \neq 0,$$

as it contains $H^1(Y, \Omega_Y^2 \otimes L) \neq 0$ as a direct summand. Using [SVV, Proposition 7.2], it follows that

$$\mathcal{H}^1 \underline{\Omega}_X^2 \neq 0.$$

(Note that for a counterexample we needed $\mathcal{H}^i \underline{\Omega}_X^2 \neq 0$ for some $i \leq 3$.)

In view of these examples, and of the results of this paper, in retrospect the question on vanishing in [MP1] should have been more restrictive. Namely, is it true that for X arbitrary we have

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < n - k - 1 - \operatorname{lcdef}(X) - s? \tag{7.10}$$

It turns out that even this statement is false, again already for cones over special smooth varieties. In this case (or whenever the singularities are isolated) thanks to Corollary 4.3 the question becomes whether

$$\mathcal{H}^i \underline{\Omega}_X^k = 0 \quad \text{for } 0 < i < \min_{p \geq 0}\{\operatorname{depth} \underline{\Omega}_X^p + p\} - k - 1.$$

Example 7.11. Let Y be a smooth projective variety with $\dim Y \geq 3$ and $H^1(Y, \mathcal{O}_Y) = 0$, endowed with a very ample line bundle L such that $H^1(Y, L) \neq 0$. We will show the existence of such a variety below; for now we draw some conclusions about the abstract cone $X = C(Y, L)$.

We claim that the modified question in (7.10) has a negative answer when $k = 0$ and $i = 1$. First, by [SVV, Proposition 7.2], the hypothesis $H^1(Y, L) \neq 0$ implies that $\mathcal{H}^1 \underline{\Omega}_X^0 \neq 0$. On the other hand, the computation of the depth of Du Bois complexes of cones is addressed in [PSh, Theorem 3.1(2)]. In our example it implies that

- $\operatorname{depth} \underline{\Omega}_X^2 > 0$,
- $\operatorname{depth} \underline{\Omega}_X^1 > 1 \iff \begin{cases} H^1(Y, \Omega_Y^1 \otimes L^m) = 0 & \text{for } m \leq -1, \\ H^1(Y, \mathcal{O}_Y) = H^0(Y, \Omega_Y^1) = 0, \\ H^0(Y, \mathcal{O}_Y) \xrightarrow{\cup c_1(L)} H^1(Y, \Omega_Y^1) & \text{is injective,} \end{cases}$
- $\operatorname{depth} \underline{\Omega}_X^0 > 2 \iff \begin{cases} H^0(Y, L^m) = 0 & \text{for } m \leq -1, \\ H^1(Y, L^m) = 0 & \text{for } m \leq 0. \end{cases}$

All of these conditions, other than $H^1(Y, L) \neq 0$ and $H^1(Y, \mathcal{O}_Y) = 0$ provided by the hypothesis, are satisfied by Kodaira–Nakano vanishing and hard Lefschetz. It follows that

$$\min_{p\geq 0}\{\text{depth}\, \underline{\Omega}_X^p + p\} - 1 > 1,$$

while $\mathcal{H}^1\underline{\Omega}_X^0 \neq 0$.

Here is an example of a threefold Y satisfying the required properties: let $Y \subset \mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ be a general hypersurface in the linear system $|\mathcal{O}_{\mathbb{P}^2}(-d) \boxtimes \mathcal{O}_{\mathbb{P}^1}(-1) \boxtimes \mathcal{O}_{\mathbb{P}^1}(-1)|$, with $d \gg 0$, and let $L = \mathcal{O}_Y(1, 1, 1)$. Then chasing cohomology through the short exact sequences

$$0 \to \mathcal{O}_{\mathbb{P}^2\times\mathbb{P}^1\times\mathbb{P}^1}(-d, -1, -1) \to \mathcal{O}_{\mathbb{P}^2\times\mathbb{P}^1\times\mathbb{P}^1} \to \mathcal{O}_Y \to 0$$

and

$$0 \to \mathcal{O}_{\mathbb{P}^2\times\mathbb{P}^1\times\mathbb{P}^1}(-d+1, 0, 0) \to \mathcal{O}_{\mathbb{P}^2\times\mathbb{P}^1\times\mathbb{P}^1}(1, 1, 1) \to L \to 0$$

and using the Künneth formula shows that $H^1(Y, \mathcal{O}_Y) = 0$ and $H^1(Y, L) \neq 0$.

Examples of any dimension can be obtained as follows: take products $Y \times Z$ and line bundles $L \boxtimes M$, where Y is the variety above, and Z is such that $H^1(Z, \mathcal{O}_Z) = 0$ and has an ample line bundle M with $H^0(Z, M) \neq 0$.

As a general conclusion to this section, the answer to the question regarding which higher cohomologies $\mathcal{H}^i\underline{\Omega}_X^k$ vanish is dictated by the depth of $\mathcal{H}^0\underline{\Omega}_X^k$, which can sometimes be smaller than $n - k - \text{lcdef}(X)$.

8. *k-rational implies k-Du Bois.* Theorem D leads to a very quick alternative proof of the fact that normal, pre-k-rational isolated singularities are pre-k-Du Bois, obtained (even for nonisolated singularities) in [SVV, Theorem B]. As explained in [SVV, Corollary 5.8], it then follows easily that k-rational implies k-Du Bois, in the same setting. Recall that when $k = 0$ this implication was studied in [St1, Proposition 3.7] for isolated singularities, and in [Ko99; Sa3] in general. Later, [FL2, Theorem 1.6] and [MP2, Theorem B] proved that k-rational implies k-Du Bois for local complete intersections.

If X is normal and pre-k-rational, then it has rational singularities. Therefore, using the main result of [KS21], one has that the composition

$$\mathcal{H}^0\underline{\Omega}_X^k \to \underline{\Omega}_X^k \to \mathbf{D}(\underline{\Omega}_X^{n-k})$$

is a quasi-isomorphism; see [SVV, Remark 2.5] for details. Here we set

$$\mathbf{D}_X(\underline{\Omega}_X^{n-k}) := \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^{n-k}, \omega_X^\bullet[-n]).$$

Dualizing, this gives that the composition

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\mathbf{D}(\underline{\Omega}_X^{n-k}), \omega_X^\bullet) \longrightarrow \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet) \xrightarrow{\varphi} \mathbf{R}\mathcal{H}om(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet)$$

is a quasi-isomorphism as well, which in turn implies that φ is surjective on cohomology. By induction we may assume however that X has pre-$(k-1)$-Du Bois singularities; hence it is also injective on cohomology by Theorem D. It follows that φ is a quasi-isomorphism, and dualizing again we obtain that X is pre-k-Du Bois. Conjecture G would of course make the same proof work even in the nonisolated case.

E. Analogues for the intersection complex

9. ***On the relationship between Du Bois and intersection complexes.*** Let X be a complex variety of dimension n. Recall from [Sa2, Section 4.5] that we have an object $\mathbb{Q}_X^H[n] := a_X^*\mathbb{Q}_{\mathrm{pt}}^H$ in the derived category of mixed Hodge modules on X, with cohomologies in degrees ≤ 0; moreover, the top degree in the weight filtration on $\mathcal{H}^0\mathbb{Q}_X^H[n]$ is n. We also have the intersection complex $\mathrm{IC}_X\,\mathbb{Q}^H$, a simple pure Hodge module of weight n; moreover, there is a composition of quotient morphisms

$$\gamma_X : \mathbb{Q}_X^H[n] \to \mathcal{H}^0\mathbb{Q}_X^H[n] \to \mathrm{IC}_X\,\mathbb{Q}^H \simeq \mathrm{gr}_n^W\,\mathcal{H}^0\mathbb{Q}_X^H[n].$$

We first record a simple lemma for later use. Here $\mathbb{D}_X(-)$ denotes the duality functor on the derived category of filtered D-modules underlying mixed Hodge modules (see [Sa1, Section 2.4]), and we abuse the notation by continuing to use the Hodge module notation for the respective filtered D-modules. Moreover $M(\ell)$ denotes the Tate twist of M, which at the level of filtered D-modules shifts the filtration down by ℓ, i.e., $F_\bullet M(\ell) = F_{\bullet-\ell}M$. Note that since we are interested in the case when X is singular, in order to consider filtrations F_kM, we appeal to the standard procedure of taking X to be (locally) embedded in a smooth variety Y, and working with filtered D_Y-modules supported on X.

Lemma 9.1. *For a fixed integer k, the composition*

$$F_k\mathbb{Q}_X^H[n] \to F_k\,\mathrm{IC}_X\,\mathbb{Q}^H \to F_k\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)$$

is an isomorphism if and only if

$$F_k\mathbb{Q}_X^H[n] \to F_k\,\mathrm{IC}_X\,\mathbb{Q}^H \quad \text{and} \quad F_k\,\mathrm{IC}_X\,\mathbb{Q}^H \to F_k\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)$$

are both isomorphisms.

Proof. The "if" part is obvious, so we focus on the "only if" part. Moreover, it suffices to prove that the first map is an isomorphism.

By strictness, the Hodge filtration commutes with taking cohomology; hence the composition

$$F_k(\mathcal{H}^i\mathbb{Q}_X^H[n]) \to F_k(\mathcal{H}^i\,\mathrm{IC}_X\,\mathbb{Q}^H) \to F_k\big(\mathcal{H}^i\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)$$

is an isomorphism for all $i \in \mathbb{Z}$. Thus for $i \neq 0$ we get

$$F_k(\mathcal{H}^i\mathbb{Q}_X^H[n]) = \mathcal{H}^i(F_k\mathbb{Q}_X^H[n]) = 0.$$

When $i = 0$ we obtain an isomorphism

$$F_k(\mathcal{H}^0\mathbb{Q}_X^H[n]) \to F_k\,\mathrm{IC}_X\,\mathbb{Q}^H \to F_k\big(\mathcal{H}^0\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big),$$

which implies that the first map is an injection. On the other hand, the fact that $\mathrm{IC}_X\,\mathbb{Q}^H \simeq \mathrm{gr}_n^W\,\mathcal{H}^0\mathbb{Q}_X^H[n]$ implies that the morphism $\mathcal{H}^0\mathbb{Q}_X^H[n] \to \mathrm{IC}_X\,\mathbb{Q}^H$ is surjective at the level of filtered D-modules. Putting everything together, we obtain isomorphisms

$$F_k\mathbb{Q}_X^H[n] \simeq F_k(\mathcal{H}^0\mathbb{Q}_X^H[n]) \simeq F_k\,\mathrm{IC}_X\,\mathbb{Q}^H$$

which implies what we want. □

For what follows, recall that we use the notation

$$\mathbf{D}_X(-) := \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(-, \omega_X^\bullet[-n]).$$

We will make repeated use of the well-known commutation of the two duality functors via the graded de Rham functor, proved in [Sa1, Section 2.4]: if $M^\bullet$ is an object in $D^b\mathrm{MHM}(X)$, and p is an integer, then

$$\mathbf{D}_X\big(\mathrm{gr}_p^F \mathrm{DR}(M^\bullet)\big) \simeq \mathrm{gr}_{-p}^F \mathrm{DR}\big(\mathbb{D}_X(M^\bullet)\big)[-n]. \tag{9.2}$$

It is a consequence of [Sa3, Theorem 4.2] that for each p, we have the identification

$$\underline{\Omega}_X^p \simeq \mathrm{gr}_{-p}^F \mathrm{DR}(\mathbb{Q}_X^H[n])[p-n].$$

We introduce the following notation for simplicity:

$$I\underline{\Omega}_X^p := \mathrm{gr}_{-p}^F \mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H)[p-n].^7$$

Taking the composition of the morphism γ_X with its dual, we get the natural morphisms in the derived category of mixed Hodge modules $D^b\mathrm{MHM}(X)$:

$$\mathbb{Q}_X^H[n] \to \mathrm{IC}_X\,\mathbb{Q}^H \to (\mathbb{D}_X(\mathbb{Q}_X^H[n]))(-n)$$

due to the self-duality $\mathbb{D}_X(\mathrm{IC}_X\,\mathbb{Q}^H) \cong \mathrm{IC}_X\,\mathbb{Q}^H(n)$; see [Sa2, 4.5.13]. Applying the functor $\mathrm{gr}_{-p}^F \mathrm{DR}$ to this composition, we obtain the natural morphisms

$$\underline{\Omega}_X^p \xrightarrow{\varphi_p} I\underline{\Omega}_X^p \to \mathbf{D}_X(\underline{\Omega}_X^{n-p}). \tag{9.3}$$

in the derived category of coherent sheaves on X.

An important point is that the higher rationality conditions say something about these maps. When X is a local complete intersection, this is essentially contained in the proof of [CDM, Theorem 3.1].

Proposition 9.4. *Let X be a normal variety with pre-k-rational singularities. Then φ_p is an isomorphism for all $p \le k$.*

Proof. As discussed in Section 8, if X is normal with pre-k-rational singularities, of dimension n, then the natural morphisms

$$\underline{\Omega}_X^p \to \mathbf{D}_X(\underline{\Omega}_X^{n-p})$$

are isomorphisms for $p \le k$. Since

$$\underline{\Omega}_X^p \simeq \mathrm{gr}_{-p}^F \mathrm{DR}(\mathbb{Q}_X^H[n])[p-n],$$

[7] These objects are called *intersection Du Bois complexes* in [PP], where they are used extensively.

and moreover

$$\begin{aligned}\mathbf{D}_X(\underline{\Omega}_X^{n-p}) &\simeq \mathbf{D}_X(\mathrm{gr}^F_{p-n}\,\mathrm{DR}(\mathbb{Q}_X^H[n])[-p])\\ &\simeq \mathrm{gr}^F_{n-p}\,\mathrm{DR}(\mathbb{D}_X(\mathbb{Q}_X^H[n]))[p-n]\\ &\simeq \mathrm{gr}^F_{-p}\,\mathrm{DR}\big(\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)[p-n],\end{aligned}$$

we obtain

$$\mathrm{gr}^F_{-p}\,\mathrm{DR}(\mathbb{Q}_X^H[n]) \simeq \mathrm{gr}^F_{-p}\,\mathrm{DR}\big(\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)$$

for all $p \le k$. After dualizing via (9.2), this is equivalent to

$$\mathrm{gr}^F_{p}\,\mathrm{DR}(\mathbb{Q}_X^H[n]) \simeq \mathrm{gr}^F_{p}\,\mathrm{DR}\big(\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)$$

for all $p \le k-n$. According to the general Lemma 9.6 below, this is equivalent to

$$F_p\mathbb{Q}_X^H[d_X] \simeq F_p\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)$$

for all $p \le k-n$. Lemma 9.1 implies in turn

$$F_p\,\mathrm{IC}_X\,\mathbb{Q}^H \simeq F_p\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)$$

for all $p \le k-n$, which again by Lemma 9.6 is equivalent to

$$\mathrm{gr}^F_{p}\,\mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H) \simeq \mathrm{gr}^F_{p}\,\mathrm{DR}\big(\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)$$

for $p \le k-n$. Applying (9.2) one more time, we see that

$$\underline{\Omega}_X^p \simeq I\underline{\Omega}_X^p$$

for all $p \le k$. ◻

Remark 9.5. Consequences of the isomorphisms in Proposition 9.4 regarding the topology of X are studied in the upcoming [DOR] and [PP]. In particular, it is shown in these papers that if φ_p is an isomorphism for all $p \le \lceil (n-2)/2 \rceil$, then X is a rational homology manifold.

The following useful lemma is a rather straightforward application of the definitions and the strictness of the Hodge filtration.

Lemma 9.6. *Let $M^\bullet, N^\bullet \in \mathrm{D}^b\mathrm{MHM}(X)$ be objects in the bounded derived category of mixed Hodge modules on X. Then the following are equivalent*:

(1) $\mathrm{gr}^F_p\,\mathrm{DR}(M^\bullet)$ *and* $\mathrm{gr}^F_p\,\mathrm{DR}(N^\bullet)$ *are quasi-isomorphic for all $p \le k$.*

(2) $F_pM^\bullet$ *and* $F_pN^\bullet$ *are quasi-isomorphic for all $p \le k$.*

As mentioned earlier, since X is (locally) embedded into a smooth variety Y of dimension d, one uses filtered right D_Y-modules (supported on X) in order to define the objects in the statement; thus we have

$$\mathrm{gr}^F_p\,\mathrm{DR}(M^\bullet) = \big[\mathrm{gr}^F_{p-d}\,M^\bullet \otimes \wedge^d T_Y \to \cdots \to \mathrm{gr}^F_{p-1}\,M^\bullet \otimes T_Y \to \mathrm{gr}^F_p\,M^\bullet\big],$$

placed in degrees $-d$ to 0. Here, and in the rest of this paper, we use this notation to denote the total complex associated to the double complex where the vertical maps come from the differentials of a representative of $M^\bullet$, while the horizontal maps are the usual de Rham maps on each term of that representative. The lemma then follows by induction on k.

10. ***Injectivity results and conjecture for the intersection complex.*** We start with an injectivity conjecture which is the intersection complex analogue of the main Conjecture G.

Conjecture 10.1. *If X has normal and pre-$(k-1)$-rational singularities, the natural morphism*

$$\mathbf{R}\mathcal{H}om(I\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om(\mathcal{H}^0 I\underline{\Omega}_X^k, \omega_X^\bullet)$$

obtained by dualizing the canonical morphism $\mathcal{H}^0 I\underline{\Omega}_X^k \to I\underline{\Omega}_X^k$ is injective on cohomology.

Remark 10.2. It is shown in [KS21, Proposition 8.1] that for all k we have an isomorphism

$$\mathcal{H}^0 I\underline{\Omega}_X^k \simeq f_*\Omega_{\widetilde{X}}^k,$$

where $f : \widetilde{X} \to X$ is a resolution of singularities.

There is a commutative diagram

$$\begin{array}{ccc} \mathbf{R}\mathcal{H}om(I\underline{\Omega}_X^k, \omega_X^\bullet) & \longrightarrow & \mathbf{R}\mathcal{H}om(\underline{\Omega}_X^k, \omega_X^\bullet) \\ \downarrow & & \downarrow \\ \mathbf{R}\mathcal{H}om(\mathcal{H}^0 I\underline{\Omega}_X^k, \omega_X^\bullet) & \longrightarrow & \mathbf{R}\mathcal{H}om(\mathcal{H}^0\underline{\Omega}_X^k, \omega_X^\bullet). \end{array}$$

If X has normal pre-$(k-1)$-rational singularities, then it also has pre-$(k-1)$-Du Bois singularities by [SVV, Theorem B]. Therefore Conjecture 10.1 is in fact implied by Conjecture G, thanks to the following:

Theorem 10.3. *If X has normal pre-$(k-1)$-rational singularities, the morphism*

$$\mathbf{R}\mathcal{H}om(I\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om(\underline{\Omega}_X^k, \omega_X^\bullet)$$

obtained by dualizing φ_k is injective on cohomology.

Corollary 10.4. *Conjecture 10.1 holds in any of the following cases*:

(1) $k = 0$.

(2) *X has isolated singularities.*

(3) *X is a local complete intersection with $(k-1)$-rational singularities.*

Proof. In each of these cases we have the corresponding injectivity theorem for the Du Bois complex, answering Conjecture G: for $k = 0$ by [KS16], for isolated singularities by Theorem D here, and for local complete intersections by [MP2] (note that $(k-1)$-rational singularities are $(k-1)$-Du Bois). ☐

Theorem 10.3 is in turn a consequence of Proposition 9.4, combined with the following injectivity theorem, which is the main technical result of this section. It was first communicated to us by Sung Gi Park, whom we thank, using the technique of [Pa, Lemma 3.7]; we follow the method of the previous section.

Theorem 10.5 (Sung Gi Park). *Assume that the variety X satisfies the property that $\varphi_p : \underline{\Omega}_X^p \to I\underline{\Omega}_X^p$ is an isomorphism for $p \le k-1$. Then the morphism*

$$\mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(I\underline{\Omega}_X^k, \omega_X^\bullet) \to \mathbf{R}\mathcal{H}om_{\mathcal{O}_X}(\underline{\Omega}_X^k, \omega_X^\bullet)$$

obtained by dualizing φ_k is injective on cohomology. More precisely, it is an isomorphism on i-th cohomology for $i \le k-n-1$, injective for $i = k-n$, and $\mathcal{E}xt^i_{\mathcal{O}_X}(I\underline{\Omega}_X^k, \omega_X^\bullet) = 0$ for $i > k-n$.

Proof. By (9.2), we have $\mathbf{D}_X(I\underline{\Omega}_X^k) \simeq I\underline{\Omega}_X^{n-k}$, so

$$\mathcal{E}xt^i_{\mathcal{O}_X}(I\underline{\Omega}_X^k, \omega_X^\bullet) \cong \mathcal{H}^{i+n} I\underline{\Omega}_X^{n-k} = 0 \quad \text{for } i > k-n.$$

For the last vanishing, note that for any p we have that $\mathrm{gr}^F_{-p}\,\mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H)$ has nontrivial cohomologies only in nonpositive degrees (since it comes from the de Rham complex of a single D-module); hence

$$\mathcal{H}^q I\underline{\Omega}_X^p = 0 \quad \text{for } p+q > n.^8$$

Thus we focus on the statements for $i \le k-n$.

As in the proof of Proposition 9.4, the assumption implies that we have isomorphisms

$$F_p\,\mathrm{IC}_X\,\mathbb{Q}^H \simeq F_p\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n) \quad \text{for } p \le k-1-n. \tag{10.6}$$

Using (9.2), the conclusion is equivalent to the fact that the map

$$\mathrm{gr}^F_{k-n}\,\mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H) \to \mathrm{gr}^F_{k-n}\,\mathrm{DR}\big(\mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)\big)$$

is an isomorphism on i-th cohomology for $i < 0$, and injective for $i = 0$.

To simplify the notation, we set $M^\bullet := \mathbb{D}_X(\mathbb{Q}_X^H[n])(-n)$. We also think of X as being (locally) embedded in a smooth variety Y of dimension d, so that we have

$$\mathrm{gr}^F_{k-n}\,\mathrm{DR}(M^\bullet) = \big[\mathrm{gr}^F_{k-n-d}\,M^\bullet \otimes \wedge^d T_Y \to \cdots \to \mathrm{gr}^F_{k-n-1}\,M^\bullet \otimes T_Y \to \mathrm{gr}^F_{k-n}\,M^\bullet\big],$$

placed in degrees $-d$ to 0. In other words, we have an exact triangle

$$\mathrm{gr}^F_{k-n}\,M^\bullet \to \mathrm{gr}^F_{k-n}\,\mathrm{DR}(M^\bullet) \to A^\bullet \xrightarrow{+1},$$

where

$$A^\bullet := \big[\mathrm{gr}^F_{k-n-d}\,M^\bullet \otimes \wedge^d T_Y \to \cdots \to \mathrm{gr}^F_{k-n-1}\,M^\bullet \otimes T_Y\big],$$

placed in degrees $-d$ to -1. Note that (10.6) implies the isomorphism

$$\big[\mathrm{gr}^F_{k-n-d}\,\mathrm{IC}_X\,\mathbb{Q}^H \otimes \wedge^d T_Y \to \cdots \to \mathrm{gr}^F_{k-n-1}\,\mathrm{IC}_X\,\mathbb{Q}^H \otimes T_Y\big] \xrightarrow{\simeq} A^\bullet;$$

[8] This is the analogue of Steenbrink vanishing for $\underline{\Omega}_X^p$, but it holds for simpler reasons.

hence we also have a similar exact triangle with $M^\bullet$ replaced by $\mathrm{IC}_X\,\mathbb{Q}^H$.

Now consider the exact triangle

$$\mathrm{gr}^F_{k-n}\,\mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H) \to \mathrm{gr}^F_{k-n}\,\mathrm{DR}(M^\bullet) \to C^\bullet \xrightarrow{+1}$$

where $C^\bullet$ denotes the cone of the morphism on the left. Since $\mathrm{gr}^F_{k-n}\,\mathrm{DR}(\mathrm{IC}_X\,\mathbb{Q}^H)$ is a complex placed in nonpositive degrees, we only need to show that

$$\mathcal{H}^i C^\bullet = 0 \quad \text{for } i < 0.$$

Given the triangles described above, the octahedral axiom implies that we have an exact triangle

$$\mathrm{gr}^F_{k-n}\,\mathrm{IC}_X\,\mathbb{Q}^H \to \mathrm{gr}^F_{k-n}\,M^\bullet \to C^\bullet \xrightarrow{+1}$$

as well. Hence the needed statement about $\mathcal{H}^i C^\bullet$ follows immediately from the following facts. On the one hand,

$$\mathcal{H}^i\,\mathrm{gr}^F_{k-n}\,M^\bullet \simeq \mathrm{gr}^F_{k-n}\,\mathcal{H}^i M^\bullet = 0 \quad \text{for } i < 0,$$

since $M^\bullet$ has nontrivial cohomologies only in nonnegative degrees (being essentially the dual of $\mathbb{Q}^H_X[n]$, which has nontrivial cohomologies in nonpositive degrees). On the other hand, $\mathcal{H}^i\,\mathrm{IC}_X\,\mathbb{Q}^H = 0$ for $i \neq 0$, and $F_{k-n}\,\mathrm{IC}_X\,\mathbb{Q}^H \hookrightarrow F_{k-n}\mathcal{H}^0 M^\bullet$ (while at the level of F_{k-n-1} we have equality by (10.6)). Indeed, we have seen in the proof of Lemma 9.1 that the morphism $\mathcal{H}^0\mathbb{Q}^H_X \to \mathrm{IC}_X\,\mathbb{Q}^H$ is surjective at the level of filtered D-modules; similarly, by duality, the morphism $\mathrm{IC}_X\,\mathbb{Q}^H \to \mathcal{H}^0 M^\bullet$ is injective at the level of filtered D-modules. All of this implies that we have an injection

$$\mathrm{gr}^F_{k-n}\,\mathrm{IC}_X\,\mathbb{Q}^H \to \mathrm{gr}^F_{k-n}\,\mathcal{H}^0 M^\bullet.$$

This completes the proof. □

11. ***Vanishing of higher cohomologies for intersection complexes.*** We finish with results about the vanishing of higher cohomologies of $I\underline{\Omega}^k_X$. We start with the following proposal, analogous to Conjecture H:

Conjecture 11.1. *Let X be a normal variety with pre-$(k-1)$-rational singularities, and pre-k-rational away from a closed subset of dimension s. Then*

$$\mathcal{H}^i I\underline{\Omega}^k_X = 0 \quad \textit{for } 0 < i < \operatorname{depth} \mathcal{H}^0 I\underline{\Omega}^k_X - s - 1.$$

When X is a normal variety with pre-k-rational singularities, it follows from Proposition 9.4 that $\mathcal{H}^i I\underline{\Omega}^p_X = 0$ for all $i > 0$ and $p \le k$. If it is so only away from a closed set of dimension s, then an argument completely analogous to that of Proposition F implies that

$$\mathcal{H}^i I\underline{\Omega}^k_X = 0 \quad \text{for } 0 < i < n_k - s - 1, \tag{11.2}$$

where $n_k := \min\{\operatorname{depth} \mathcal{H}^0 I\underline{\Omega}^k_X, \operatorname{depth} I\underline{\Omega}^k_X + 1\}$.

If moreover X has normal pre-$(k-1)$-rational isolated or $(k-1)$-rational local complete intersection singularities, or if $k=0$, then Corollary 10.4 implies that Conjecture 10.1 holds, hence exactly as in the proof of Theorem A we have in addition that $\operatorname{depth} I\underline{\Omega}_X^k \geq \operatorname{depth} \mathcal{H}^0 I\underline{\Omega}_X^k$. Therefore:

Corollary 11.3. *Conjecture 11.1 holds when X has isolated, or $(k-1)$-rational local complete intersection singularities, or when $k=0$.*

We also have the analogue of the vanishing result for the Du Bois complex in [MP1, Corollary 13.9]; see Example 7.4.

Corollary 11.4. *Let X be a local complete intersection with $\dim X_{\text{sing}} = s$. Then*

$$\mathcal{H}^i I\underline{\Omega}_X^k = 0, \quad \textit{for all } 0 < i < n-k-s-1.$$

Proof. We use (11.2). For the intersection complex, it is always the case that $\operatorname{depth} I\underline{\Omega}_X^k \geq n-k$ by (4.2). Indeed, we have

$$\mathcal{E}xt^i_{\mathcal{O}_X}(I\underline{\Omega}_X^k, \omega_X^\bullet) = 0 \quad \text{for } i > k-n,$$

due to the self-duality (up to twist) of the intersection complex, as explained in Theorem 10.5.

Following precisely the steps in Example 7.4, under the current hypotheses we also have that $\operatorname{depth} \mathcal{H}^0 I\underline{\Omega}_X^k \geq n-k$; indeed, if $f: \widetilde{X} \to X$ is a resolution of singularities, we know that $\mathcal{H}^0 I\underline{\Omega}_X^k \simeq f_*\Omega_{\widetilde{X}}^k$ (see Remark 10.2), which in this case is reflexive, so that

$$f_*\Omega^k_{\widetilde{X}} \simeq \Omega_X^{[k]} \simeq \Omega_X^k. \qquad \square$$

Remark 11.5 (analogue of Proposition C). For completeness, we conclude by noting that there is also an analogue of this basic vanishing result for $\underline{\Omega}_X^p$ stated in the introduction. Namely, if $k < n$ and $\mathcal{H}^{n-p-1} I\underline{\Omega}_X^p = 0$ for all $p \leq k-1$, then

$$\mathcal{H}^{n-k} I\underline{\Omega}_X^k = 0.$$

In particular, if X has pre-$(k-1)$-rational singularities, with $k<n$, then $\mathcal{H}^{n-k} I\underline{\Omega}_X^k = 0$. The proof is very similar.

Acknowledgements

We thank Brad Dirks, Mircea Mustaţă, and especially Sung Gi Park, for valuable conversations and suggestions. We are grateful to the referee as well, for a number of useful comments.

References

[CDM] Q. Chen, B. Dirks, and M. Mustaţă, "The minimal exponent and k-rationality for local complete intersections", *J. Éc. polytech. Math.* **11** (2024), 849–873. MR

[DB] P. Du Bois, "Complexe de de Rham filtré d'une variété singulière", *Bull. Soc. Math. France* **109**:1 (1981), 41–81. MR

[DOR] B. Dirks, S. Olano, and D. Raychaudhury, "A Hodge theoretic generalization of $\mathbb{Q}$-homology manifolds", preprint, 2025. arXiv 2501.14065

[DT] H. Dao and S. Takagi, "On the relationship between depth and cohomological dimension", *Compos. Math.* **152**:4 (2016), 876–888. MR

[FL1] R. Friedman and R. Laza, "The higher Du Bois and higher rational properties for isolated singularities", *J. Algebraic Geom.* **33**:3 (2024), 493–520. MR

[FL2] R. Friedman and R. Laza, "Higher Du Bois and higher rational singularities", *Duke Math. J.* **173**:10 (2024), 1839–1881. MR

[FY] H.-B. Foxby and S. Iyengar, "Depth and amplitude for unbounded complexes", pp. 119–137 in *Commutative algebra* (Grenoble/Lyon, 2001), Contemp. Math. **331**, Amer. Math. Soc., 2003. MR

[GNPP] F. Guillén, V. Navarro Aznar, P. Pascual Gainza, and F. Puerta, *Hyperrésolutions cubiques et descente cohomologique* (Barcelona, 1982), Lecture Notes in Mathematics **1335**, Springer, 1988. MR

[Gre] G.-M. Greuel, "Der Gauss–Manin-Zusammenhang isolierter Singularitäten von vollständigen Durchschnitten", *Math. Ann.* **214** (1975), 235–266. MR

[HJ] A. Huber and C. Jörder, "Differential forms in the h-topology", *Algebr. Geom.* **1**:4 (2014), 449–478. MR

[JKSY] S.-J. Jung, I.-K. Kim, M. Saito, and Y. Yoon, "Higher Du Bois singularities of hypersurfaces", *Proc. Lond. Math. Soc.* (3) **125**:3 (2022), 543–567. MR

[Ko12] S. J. Kovács, "The intuitive definition of Du Bois singularities", pp. 257–266 in *Geometry and arithmetic*, Eur. Math. Soc., Zürich, 2012. MR

[Ko99] S. J. Kovács, "Rational, log canonical, Du Bois singularities: on the conjectures of Kollár and Steenbrink", *Compositio Math.* **118**:2 (1999), 123–133. MR

[KS16] S. J. Kovács and K. Schwede, "Du Bois singularities deform", pp. 49–65 in *Minimal models and extremal rays* (Kyoto, 2011), Adv. Stud. Pure Math. **70**, Math. Soc. Japan, 2016. MR

[KS21] S. Kebekus and C. Schnell, "Extending holomorphic forms from the regular locus of a complex space to a resolution of singularities", *J. Amer. Math. Soc.* **34**:2 (2021), 315–368. MR

[MOPW] M. Mustaţă, S. Olano, M. Popa, and J. Witaszek, "The Du Bois complex of a hypersurface and the minimal exponent", *Duke Math. J.* **172**:7 (2023), 1411–1436. MR

[MP1] M. Mustaţă and M. Popa, "Hodge filtration on local cohomology, Du Bois complex and local cohomological dimension", *Forum Math. Pi* **10** (2022), art. id. e22, 58 pp. MR

[MP2] M. Mustaţă and M. Popa, "On k-rational and k–Du Bois local complete intersections", *Algebr. Geom.* **12**:2 (2025), 237–261. MR

[MV] C. Miller and S. Vassiliadou, "(Co)torsion of exterior powers of differentials over complete intersections", *J. Singul.* **19** (2019), 131–162. MR

[NS] Y. Namikawa and J. H. M. Steenbrink, "Global smoothing of Calabi–Yau threefolds", *Invent. Math.* **122**:2 (1995), 403–419. MR

[Og] A. Ogus, "Local cohomological dimension of algebraic varieties", *Ann. of Math.* (2) **98** (1973), 327–365. MR

[Pa] S. G. Park, "Du Bois complex and extension of forms beyond rational singularities", preprint, 2023. arXiv 2311.15159

[PP] S. G. Park and M. Popa, "Hodge symmetry and Lefschetz theorems for singular varieties", preprint, 2024. arXiv 2410.15638

[PS] C. A. M. Peters and J. H. M. Steenbrink, *Mixed Hodge structures*, Ergebnisse der Math. (3) **52**, Springer, 2008. MR

[PSh] M. Popa and W. Shen, "Du Bois complexes of cones over singular varieties, local cohomological dimension, and K-groups", *Rev. Roumaine Math. Pures Appl.* **70**:1-2 (2025), 133–155. MR

[RSW] T. Reichelt, M. Saito, and U. Walther, "Topological calculation of local cohomological dimension", *J. Singul.* **26** (2023), 13–22. MR

[Sa1] M. Saito, "Modules de Hodge polarisables", *Publ. Res. Inst. Math. Sci.* **24**:6 (1988), 849–995. MR

[Sa2] M. Saito, "Mixed Hodge modules", *Publ. Res. Inst. Math. Sci.* **26**:2 (1990), 221–333. MR

[Sa3] M. Saito, "Mixed Hodge complexes on algebraic varieties", *Math. Ann.* **316**:2 (2000), 283–331. MR

[St1] J. H. M. Steenbrink, "Mixed Hodge structures associated with isolated singularities", pp. 513–536 in *Singularities* (Arcata, CA, 1981), vol. 2, Proc. Sympos. Pure Math. **40**.2, Amer. Math. Soc., 1983. MR

[St2] J. H. M. Steenbrink, “Vanishing theorems on singular spaces”, pp. 330–341 in *Differential systems and singularities* (Luminy, 1983), Astérisque **130**, Soc. math. de France, 1985. MR

[St3] J. H. M. Steenbrink, “Du Bois invariants of isolated complete intersection singularities”, *Ann. Inst. Fourier* (*Grenoble*) **47**:5 (1997), 1367–1377. MR

[Sta] “The Stacks project”, electronic resource, 2024, available at https://stacks.math.columbia.edu.

[SVV] W. Shen, S. Venkatesh, and A. D. Vo, “On k-Du Bois and k-rational singularities”, preprint, 2023. To appear in *Ann. Inst. Fourier*. arXiv 2306.03977

[Tig] B. Tighe, “The holomorphic extension property for higher Du Bois singularities”, preprint, 2023. arXiv 2312.01245

[Tot] B. Totaro, “Bott vanishing for Fano threefolds”, *Math. Z.* **307**:1 (2024), art. id. 14, 31 pp. MR

Communicated by Shunsuke Takagi
Received 2024-10-17 Revised 2025-04-18 Accepted 2025-05-22

mpopa@math.harvard.edu *Department of Mathematics, Harvard University, Cambridge, MA, United States*

wshen@math.harvard.edu *Department of Mathematics, Harvard University, Cambridge, MA, United States*

ducvo@math.harvard.edu *Department of Mathematics, Harvard University, Cambridge, MA, United States*

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 20 No. 6 2026