Algebra & Number Theory

Volume 8

2014

No. 3

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2014 is US $225/year for the electronic version, and $400/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

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

ANT peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.513

Derived invariants of irregular varieties and Hochschild homology

Luigi Lombardi

We study the behavior of cohomological support loci of the canonical bundle under derived equivalence of smooth projective varieties. This is achieved by investigating the derived invariance of a generalized version of Hochschild homology. Furthermore, using techniques coming from birational geometry, we establish the derived invariance of the Albanese dimension for varieties having nonnegative Kodaira dimension. We apply our machinery to study the derived invariance of the holomorphic Euler characteristic and of certain Hodge numbers for special classes of varieties. Further applications concern the behavior of particular types of fibrations under derived equivalence.

1. Introduction

It is now well-known that derived equivalent varieties share quite a few invariants. For instance, the dimension, the Kodaira dimension, the numerical dimension and the canonical ring are examples of derived invariants. By describing the behavior under derived equivalence of the Picard variety, Popa and Schnell [2011] establish the derived invariance of the number of linearly independent holomorphic one-forms. In this paper, we study the behavior under derived equivalence of other fundamental objects in the geometry of *irregular* varieties, i.e., those with positive *irregularity* $q(X) := h^0(X, \Omega_X^1)$, such as the cohomological support loci and the Albanese dimension. Applications of our techniques concern the derived invariance of the holomorphic Euler characteristic of varieties with large Albanese dimension and the derived invariance of some of the Hodge numbers of fourfolds again with large Albanese dimension. A further application concerns the behavior of fibrations of derived equivalent threefolds onto irregular varieties. This work is motivated by a well-known conjecture predicting the derived invariance of all Hodge numbers and by a conjecture of Popa (see Conjectures 1.2 and 1.3 and [Popa 2013]).

MSC2010: 14F05.
Keywords: equivalences of derived categories, support loci, Hochschild homology, Hodge numbers, Picard variety, Rouquier isomorphism.

The main tool we use to approach the problems described above is the comparison of the cohomology groups of twists by topologically trivial line bundles of the canonical bundles of the varieties in play. This is achieved by studying a generalized version of Hochschild homology that takes into account an important isomorphism due to Rouquier related to derived autoequivalences (see [Rouquier 2011, Théorème 4.18]). In this way, we obtain a theoretical result of independent interest in the study of derived equivalences of smooth projective varieties, which we now present. To begin with, we recall the *Hochschild cohomology and homology* of a smooth projective variety X:

$$HH^*(X) := \bigoplus_k \operatorname{Ext}^k_{X\times X}(i_*\mathcal{O}_X, i_*\mathcal{O}_X),$$
$$HH_*(X) := \bigoplus_k \operatorname{Ext}^k_{X\times X}(i_*\mathcal{O}_X, i_*\omega_X),$$

where $i : X \hookrightarrow X \times X$ is the diagonal embedding of X. The space $HH^*(X)$ has a structure of ring under composition of morphisms, and $HH_*(X)$ is a graded $HH^*(X)$-module with the same operation. Results of Căldăraru [2003, Theorem 8.1] and Orlov [2003, Theorem 2.1.8] show that the Hochschild cohomology and homology are derived invariants. More precisely, if $\Phi : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ is an equivalence of derived categories of smooth projective varieties, then it induces an isomorphism of rings $HH^*(X) \cong HH^*(Y)$ and an isomorphism of graded modules $HH_*(X) \cong HH_*(Y)$ compatible with the isomorphism $HH^*(X) \cong HH^*(Y)$. We now present the generalization of Hochschild homology mentioned above. For a triple $(\varphi, L, m) \in \operatorname{Aut}^0(X) \times \operatorname{Pic}^0(X) \times \mathbb{Z}$, we define the graded $HH^*(X)$-module

$$HH_*(X, \varphi, L, m) := \bigoplus_k \operatorname{Ext}^k_{X\times X}\big(i_*\mathcal{O}_X, (1, \varphi)_*(\omega_X^{\otimes m} \otimes L)\big)$$

with module structure given by composition of morphisms. We think of these spaces as a "twisted" version of the Hochschild homology of X. Lastly, we recall that a derived equivalence $\boldsymbol{D}(X) \cong \boldsymbol{D}(Y)$ induces an isomorphism of algebraic groups, called *Rouquier's isomorphism*

$$F : \operatorname{Aut}^0(X) \times \operatorname{Pic}^0(X) \to \operatorname{Aut}^0(Y) \times \operatorname{Pic}^0(Y). \tag{1}$$

(An explicit description of F is given in (3) (see [Rouquier 2011, Théorème 4.18; Huybrechts 2006, Proposition 9.45; Rosay 2009, Theorem 3.1]; cf. [Popa and Schnell 2011, footnote, p. 531]).) The following theorem describes the behavior of the twisted Hochschild homology under derived equivalence. Its proof follows the general strategy of the proofs of Orlov and Căldăraru, but further technicalities appear due to the possible presence of nontrivial automorphisms of X and Y; see Section 2 for its proof.

Theorem 1.1. *Let $\Phi : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ be an equivalence of derived categories of smooth projective varieties defined over an algebraically closed field, and let $m \in \mathbb{Z}$.*

If $F(\varphi, L) = (\psi, M)$ (where F is the Rouquier isomorphism), then Φ induces an isomorphism of graded modules

$$HH_*(X, \varphi, L, m) \cong HH_*(Y, \psi, M, m)$$

compatible with the isomorphism $HH^*(X) \cong HH^*(Y)$.

We now move our attention to the main application of Theorem 1.1, namely, the behavior of *cohomological support loci* under derived equivalence. These loci are defined as

$$V^k(\omega_X) := \{L \in \mathrm{Pic}^0(X) \mid h^k(X, \omega_X \otimes L) > 0\}$$

where X is a smooth projective variety and $k \geq 0$ is an integer. From here on, we work over the field of the complex numbers. The $V^k(\omega_X)$'s have been studied for instance in [Green and Lazarsfeld 1987; 1991; Ein and Lazarsfeld 1997; Arapura 1992; Hacon 2004; Pareschi and Popa 2011]. They are one of the most important tools in the birational study of irregular varieties; roughly speaking, they control the geometry of the Albanese map and the fibrations onto lower-dimensional irregular varieties. The following conjecture, and its weaker variant, predicts the behavior of cohomological support loci under derived equivalence. As a matter of notation, we denote by $V^k(\omega_X)_0$ the union of the irreducible components of $V^k(\omega_X)$ passing through the origin.

Conjecture 1.2 [Popa 2013, Conjecture 1.2]. *If X and Y are smooth projective derived equivalent varieties, then*

$$V^k(\omega_X) \cong V^k(\omega_Y) \quad \text{for all } k \geq 0.$$

Conjecture 1.3 [Popa 2013, Variant 1.3]. *Under the assumptions of Conjecture 1.2, there exist isomorphisms*

$$V^k(\omega_X)_0 \cong V^k(\omega_Y)_0 \quad \text{for all } k \geq 0.$$

It is important to emphasize that for all the applications we are interested in (e.g., invariance of the Albanese dimension, invariance of the holomorphic Euler characteristic and invariance of Hodge numbers) it is in fact enough to verify Conjecture 1.3. We also remark that Conjecture 1.2 holds for varieties of general type since the cohomological support loci are birational invariants while derived equivalent varieties of general type are birational by [Kawamata 2002, Theorem 1.4]. Moreover, in [Popa 2013, §2], it has been shown that Conjecture 1.2 holds for surfaces as well.

In Section 3, we try to attack the above conjectures for varieties of arbitrary dimension. To begin with, we show that Theorem 1.1 implies the derived invariance of $V^0(\omega_X)$ (see Proposition 3.1). On the other hand, due to the possible presence of nontrivial automorphisms, the study of the derived invariance of the higher cohomological support loci is more involved. Nonetheless, by using a version of the Hochschild–Kostant–Rosenberg isomorphism and Brion's structural results on the actions of nonaffine groups on smooth varieties, we are able to show the derived invariance of $V^1(\omega_X)_0$ (see Corollary 3.4). The next theorem summarizes the main results on the derived invariance of these loci.

Theorem 1.4. *Let X and Y be smooth projective derived equivalent varieties. Then the Rouquier isomorphism induces isomorphisms of algebraic sets*

(i) $V^0(\omega_X) \cong V^0(\omega_Y)$,

(ii) $V^0(\omega_X) \cap V^1(\omega_X) \cong V^0(\omega_Y) \cap V^1(\omega_Y)$ *and*

(iii) $V^1(\omega_X)_0 \cong V^1(\omega_Y)_0$.

We note that (i) also holds if we consider arbitrary powers of the canonical bundle (see Proposition 3.1). We point out also that cases in which the Rouquier isomorphism induces the full isomorphism $V^1(\omega_X) \cong V^1(\omega_Y)$ occur for instance when either X is of maximal Albanese dimension (see Corollary 5.2) or when the neutral component of the automorphism group, $\mathrm{Aut}^0(X)$, is affine (see Remark 3.6); Theorem 1.4 is proved in Section 3.

Next we study Conjectures 1.2 and 1.3 for varieties of dimension three. In the process, we recover Conjecture 1.2 in dimension two as well, making the isomorphisms on cohomological support loci explicit. In the following theorem, we collect all results concerning the behavior of cohomological support loci of derived equivalent threefolds. We denote by $\mathrm{alb}_X : X \to \mathrm{Alb}(X)$ the Albanese map of X, and we say that X is of *maximal Albanese dimension* if $\dim \mathrm{alb}_X(X) = \dim X$, i.e., alb_X is generically finite onto its image.

Theorem 1.5. *Let X and Y be smooth projective irregular derived equivalent threefolds. Then*:

(i) *Conjecture 1.3 holds.*

(ii) *Conjecture 1.2 holds if one of the following hypotheses is satisfied*:

- (a) *X is of maximal Albanese dimension.*
- (b) $V^k(\omega_X) = \mathrm{Pic}^0(X)$ *for some* $k \geq 0$ (*for instance*, *by* [Pareschi and Popa 2011, Theorem E], $V^0(\omega_X) = \mathrm{Pic}^0(X)$ *whenever* $\mathrm{alb}_X(X)$ *is not fibered in subtori and* $V^0(\omega_X) \neq \varnothing$).
- (c) $\mathrm{Aut}^0(X)$ *is affine* (*for instance*, *by a theorem of Nishi* [Matsumura 1963, Theorem 2], *this again happens when* $\mathrm{alb}_X(X)$ *is not fibered in subtori*).

(iii) *If* $q(X) \geq 2$, *then* $\dim V^k(\omega_X) = \dim V^k(\omega_Y)$ *for all* $k \geq 0$.

Point (iii) brings evidence to a further variant of Conjecture 1.2 predicting the invariance of the dimensions of cohomological support loci [Popa 2013, Variant 1.4]; partial results for the case $q(X) = 1$ are described in Remark 6.10. Since the proofs of Theorems 1.4 and 1.5 extend to analogous results regarding cohomological support loci of bundles of holomorphic p-forms, when possible, we will prove them in such generality. Please refer to Theorem 4.2 and Section 6 for the proof of Theorem 1.5.

Finally, we move our attention to applications of Theorems 1.4 and 1.5. The first regards the behavior of the Albanese dimension, $\dim \mathrm{alb}_X(X)$, under derived equivalence. According to Conjecture 1.3, the Albanese dimension is expected to be preserved under derived equivalence as it can be read off from the dimensions of the $V^k(\omega_X)_0$'s (see (5)), which is the case in dimension three thanks to Theorem 1.5. In higher dimension, we establish this invariance for varieties having nonnegative Kodaira dimension $\kappa(X)$ by using the derived invariance of the irregularity and an extension of a result due to Chen, Hacon and Pardini [Hacon and Pardini 2002, Proposition 2.1; Chen and Hacon 2004, Corollary 3.6] on the study of the geometry of the Albanese map via the Iitaka fibration; see Section 5.

Theorem 1.6. *Let X and Y be smooth projective derived equivalent varieties. If* $\dim X \leq 3$, *or if* $\dim X > 3$ *and* $\kappa(X) \geq 0$, *then*

$$\dim \mathrm{alb}_X(X) = \dim \mathrm{alb}_Y(Y).$$

The second application concerns the holomorphic Euler characteristic. This is expected to be the same for arbitrary derived equivalent smooth projective varieties since the Hodge numbers are expected to be preserved (which is known to hold in dimension up to three [Popa and Schnell 2011, Corollary C]). We deduce this for varieties of large Albanese dimension as a consequence of the previous results and generic vanishing.

Corollary 1.7. *Let X and Y be smooth projective derived equivalent varieties. If* $\dim \mathrm{alb}_X(X) = \dim X$, *or if* $\dim \mathrm{alb}_X(X) = \dim X - 1$ *and* $\kappa(X) \geq 0$, *then*

$$\chi(\omega_X) = \chi(\omega_Y).$$

An immediate consequence is the derived invariance of two of the Hodge numbers for fourfolds satisfying the hypotheses of Corollary 1.7.

Corollary 1.8. *Let X and Y be smooth projective derived equivalent fourfolds. If* $\dim \mathrm{alb}_X(X) = 4$, *or if* $\dim \mathrm{alb}_X(X) = 3$ *and* $\kappa(X) \geq 0$, *then*

$$h^{0,2}(X) = h^{0,2}(Y) \quad \textit{and} \quad h^{1,3}(X) = h^{1,3}(Y).$$

We remark that Popa and Schnell [2011, Corollary 3.4] establish the invariance of $h^{0,2}$ and $h^{1,3}$ under different hypotheses, namely, when $\mathrm{Aut}^0(X)$ is not affine (we recall that $h^{0,4}$, $h^{0,3}$, $h^{0,1}$ and $h^{1,2}$ are always known to be invariant; see [Popa and Schnell 2011]). Corollaries 1.7 and 1.8 are proved in Section 7.

We now present our last application in a direction that is one of the main motivations for Conjectures 1.2 and 1.3 as explained in [Popa 2013]. From the classification of Fourier–Mukai equivalences for surfaces [Kawamata 2002; Bridgeland and Maciocia 2001], it is known that, if X admits a fibration $f : X \to C$ onto a smooth curve of genus ≥ 2, then any of its Fourier–Mukai partners admits a fibration onto the same curve. Here we use our analysis, and a theorem of Green and Lazarsfeld regarding the properties of positive-dimensional irreducible components of the cohomological support loci, to investigate the behavior of fibrations of derived equivalent threefolds onto irregular varieties. Recall that a smooth variety X is called of *Albanese general type* if alb_X is nonsurjective and generically finite onto its image. The proof of the next corollary is contained in Proposition 7.3 and Remark 7.4.

Corollary 1.9. *Let X and Y be smooth projective derived equivalent threefolds. There exists a morphism $f : X \to W$ with connected fibers onto a normal variety W of dimension ≤ 2 such that any smooth model of W is of Albanese general type if and only if Y has a fibration of the same type. Moreover, there exists a morphism $f : X \to C$ with connected fibers onto a smooth curve C of genus ≥ 2 if and only if there exists a morphism $h : Y \to D$ with connected fibers onto a smooth curve D of genus ≥ 2.*

To conclude, we remark that, while the approach in this paper relies in part on techniques of [Popa and Schnell 2011], the key new ingredient is their interaction with the twisted Hochschild homology, introduced and studied here. We are hopeful that this general method will find further applications in the future.

2. Derived invariance of the twisted Hochschild homology

In this section, we aim to prove Theorem 1.1. Its proof is based on a technical lemma extending previous computations carried out by Căldăraru [2003, Proposition 8.1] and Orlov [2003, Isomorphism (10)].

Let X and Y be smooth projective varieties defined over an algebraically closed field K, and let p and q be the projections from $X \times Y$ onto the first and second factor, respectively. We denote by $\boldsymbol{D}(X) := \boldsymbol{D}^b(\mathscr{C}oh(X))$ the bounded derived category of coherent sheaves on a smooth projective variety X. When there is no possibility of ambiguity, we use the same symbol to denote a functor and its associated derived functor. An object $\mathscr{E}$ in $\boldsymbol{D}(X \times Y)$ defines *Fourier–Mukai functors* with *kernel* $\mathscr{E}$ as

$$\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y), \quad \mathscr{F} \mapsto q_*(p^*\mathscr{F} \otimes \mathscr{E}),$$
$$\Psi_{\mathscr{E}} : \boldsymbol{D}(Y) \to \boldsymbol{D}(X), \quad \mathscr{G} \mapsto p_*(q^*\mathscr{G} \otimes \mathscr{E}).$$

We say that X and Y are *derived equivalent* if there exists a K-linear exact equivalence of triangulated categories $\Phi : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$. By a fundamental result of

Orlov, any such equivalence is of Fourier–Mukai type; i.e., there exists an object $\mathscr{E}$ in $\boldsymbol{D}(X \times Y)$ such that $\Phi \cong \Phi_{\mathscr{E}}$. Furthermore, the object $\mathscr{E}$ is unique up to isomorphism.

We recall that an equivalence $\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ induces an equivalence

$$\Phi_{\mathscr{E}^* \boxtimes \mathscr{E}} : \boldsymbol{D}(X \times X) \to \boldsymbol{D}(Y \times Y)$$

with kernel

$$\mathscr{E}^* \boxtimes \mathscr{E} := p_{13}^* \mathscr{E}^* \otimes p_{24}^* \mathscr{E},$$

where $\mathscr{E}^* := \boldsymbol{R}\mathscr{H}om(\mathscr{E}, \mathscr{O}_{X \times Y}) \otimes p^* \omega_X[\dim X]$ and p_{rs} is the projection from $X \times X \times Y \times Y$ onto the (r, s)-factor [Orlov 2003, Proposition 2.1.7]. Moreover, for any automorphisms $\varphi \in \mathrm{Aut}^0(X)$ and $\psi \in \mathrm{Aut}^0(Y)$ (here the superscript 0 denotes the neutral component of the corresponding group), we define the embeddings $(1, \varphi) : X \hookrightarrow X \times X,\ x \mapsto (x, \varphi(x))$ and $(1, \psi) : Y \hookrightarrow Y \times Y,\ y \mapsto (y, \psi(y))$. Finally, we denote by i and j the diagonal embeddings of X and Y, respectively.

Lemma 2.1. *Let X and Y be smooth projective varieties defined over an algebraically closed field, and let $\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ be an equivalence. Denote by F the induced Rouquier isomorphism* (*see* (1)), *and let $m \in \mathbb{Z}$. If $F(\varphi, L) = (\psi, M)$, then*

$$\Phi_{\mathscr{E}^* \boxtimes \mathscr{E}}\big((1, \varphi)_*(\omega_X^{\otimes m} \otimes L)\big) \cong (1, \psi)_*(\omega_Y^{\otimes m} \otimes M).$$

Proof. We denote by t_r and t_{rs} the projections from $Y \times X \times Y$ onto the r-th and (r, s)-th factors, respectively. Moreover, we define the morphism $\lambda : Y \times X \times Y \to X \times X \times Y \times Y$ as $(y_1, x, y_2) \mapsto (x, \varphi(x), y_1, y_2)$, and we look at the fiber product diagram

$$\begin{tikzcd} Y \times X \times Y \arrow[r, "\lambda"] \arrow[d, "t_2"] & X \times X \times Y \times Y \arrow[d, "p_{12}"] \\ X \arrow[r, "(1,\varphi)"] & X \times X \end{tikzcd}$$

so that, by base change and the projection formula, we get

$$\begin{aligned} \Phi_{\mathscr{E}^* \boxtimes \mathscr{E}}\big((1, \varphi)_*(\omega_X^{\otimes m} \otimes L)\big) &\cong p_{34*}\big(p_{12}^*(1, \varphi)_*(\omega_X^{\otimes m} \otimes L) \otimes (\mathscr{E}^* \boxtimes \mathscr{E})\big) \\ &\cong p_{34*}\big(\lambda_* t_2^*(\omega_X^{\otimes m} \otimes L) \otimes p_{13}^* \mathscr{E}^* \otimes p_{24}^* \mathscr{E}\big) \\ &\cong p_{34*}\lambda_*\big(t_2^*(\omega_X^{\otimes m} \otimes L) \otimes \lambda^* p_{13}^* \mathscr{E}^* \otimes \lambda^* p_{24}^* \mathscr{E}\big) \\ &\cong t_{13*}\big(t_2^*(\omega_X^{\otimes m} \otimes L) \otimes t_{21}^* \mathscr{E}^* \otimes t_{23}^*(\varphi \times 1)^* \mathscr{E}\big). \end{aligned} \tag{2}$$

By [Orlov 2003, p. 535], the equivalence $\Phi_{\mathscr{E}}$ induces an isomorphism $\mathscr{E} \otimes p^* \omega_X \cong \mathscr{E} \otimes q^* \omega_Y$. Moreover, by [Popa and Schnell 2011, Lemma 3.1], the condition $F(\varphi, L) = (\psi, M)$ is equivalent to an isomorphism of objects in $\boldsymbol{D}(X \times Y)$

$$(\varphi \times 1)^* \mathscr{E} \otimes p^* L \cong (1 \times \psi)_* \mathscr{E} \otimes q^* M. \tag{3}$$

Therefore, we get an isomorphism of objects

$$p^*(\omega_X^{\otimes m} \otimes L) \otimes (\varphi \times 1)^*\mathscr{E} \cong q^*(\omega_Y^{\otimes m} \otimes M) \otimes (1 \times \psi)_*\mathscr{E},$$

and by pulling it back via $t_{23} : Y \times X \times Y \to X \times Y$, we finally obtain

$$t_2^*(\omega_X^{\otimes m} \otimes L) \otimes t_{23}^*(\varphi \times 1)^*\mathscr{E} \cong t_3^*(\omega_Y^{\otimes m} \otimes M) \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}. \tag{4}$$

At this point, we rewrite the morphism $t_3 : Y \times X \times Y \to Y$ as $t_3 = \sigma_2 \circ t_{13}$, where $\sigma_2 : Y \times Y \to Y$ is the projection onto the second factor. Moreover, we denote by $\rho : Y \times X \to X \times Y$ the inversion morphism $(y, x) \mapsto (x, y)$. Then by (2) and (4), we obtain

$$\begin{aligned}
\Phi_{\mathscr{E}^* \boxtimes \mathscr{E}}\big((1, \varphi)_*(\omega_X^{\otimes m} \otimes L)\big) &\cong t_{13*}\big(t_3^*(\omega_Y^{\otimes m} \otimes M) \otimes t_{21}^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}\big) \\
&\cong t_{13*}\big(t_{13}^*\sigma_2^*(\omega_Y^{\otimes m} \otimes M) \otimes t_{21}^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}\big) \\
&\cong \sigma_2^*(\omega_Y^{\otimes m} \otimes M) \otimes t_{13*}\big(t_{21}^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}\big) \\
&\cong \sigma_2^*(\omega_Y^{\otimes m} \otimes M) \otimes t_{13*}\big(t_{12}^*\rho^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}\big).
\end{aligned}$$

Finally, by [Orlov 2003, Proposition 2.1.2], the object $t_{13*}(t_{12}^*\rho^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E})$ in $\boldsymbol{D}(Y \times Y)$ is the kernel of the composition

$$\Phi_{(1\times\psi)_*\mathscr{E}} \circ \Phi_{\rho^*\mathscr{E}^*} \cong \psi_* \circ \Phi_{\mathscr{E}} \circ \Psi_{\mathscr{E}^*} \cong \psi_* \circ \mathrm{id}_{\boldsymbol{D}(Y)} \cong \psi_*,$$

where we used the fact that $\Psi_{\mathscr{E}^*}$ is the right adjoint to $\Phi_{\mathscr{E}}$. On the other hand, since the kernel of the derived functor $\psi_* : \boldsymbol{D}(Y) \to \boldsymbol{D}(Y)$ is the structure sheaf of the graph of ψ, i.e., $\mathscr{O}_{\Gamma_\psi} \cong (1, \psi)_*\mathscr{O}_Y$ [Huybrechts 2006, Example 5.4], we have an isomorphism

$$t_{13*}\big(t_{12}^*\rho^*\mathscr{E}^* \otimes t_{23}^*(1 \times \psi)_*\mathscr{E}\big) \cong (1, \psi)_*\mathscr{O}_Y$$

as the kernel of an equivalence is unique up to isomorphism. To recap,

$$\begin{aligned}
\Phi_{\mathscr{E}^* \boxtimes \mathscr{E}}\big((1, \varphi)_*(\omega_X^{\otimes m} \otimes L)\big) &\cong \sigma_2^*(\omega_Y^{\otimes m} \otimes M) \otimes (1, \psi)_*\mathscr{O}_Y \\
&\cong (1, \psi)_*\big((1, \psi)^*\sigma_2^*(\omega_Y^{\otimes m} \otimes M)\big) \\
&\cong (1, \psi)_*\big(\psi^*(\omega_Y^{\otimes m} \otimes M)\big) \\
&\cong (1, \psi)_*(\omega_Y^{\otimes m} \otimes M).
\end{aligned}$$

The last isomorphism follows as the action of $\mathrm{Aut}^0(X)$ on $\mathrm{Pic}^0(X)$ is trivial [Popa and Schnell 2011, Footnote, p. 531]. □

Proof of Theorem 1.1. Let $\mathscr{E}$ be the kernel of the equivalence Φ so that $\Phi \cong \Phi_{\mathscr{E}}$. By Lemma 2.1, the equivalence $\Phi_{\mathscr{E}^* \boxtimes \mathscr{E}}$ induces isomorphisms between the graded components of $HH_*(X, \varphi, L, m)$ and $HH_*(Y, \psi, M, m)$ as follows:

$$\begin{aligned}\operatorname{Ext}^k_{X\times X}\big(i_*\mathcal{O}_X, (1,\varphi)_*(\omega_X^{\otimes m}\otimes L)\big)&\\ &\cong \operatorname{Ext}^k_{Y\times Y}\big(\Phi_{\mathcal{E}^*\boxtimes\mathcal{E}}(i_*\mathcal{O}_X), \Phi_{\mathcal{E}^*\boxtimes\mathcal{E}}\big((1,\varphi)_*(\omega_X^{\otimes m}\otimes L)\big)\big)\\ &\cong \operatorname{Ext}^k_{Y\times Y}\big(j_*\mathcal{O}_Y, (1,\psi)_*(\omega_Y^{\otimes m}\otimes M)\big).\end{aligned}$$

Moreover, since $\Phi_{\mathcal{E}^*\boxtimes\mathcal{E}}$ is a functor, it follows that it induces an isomorphism of graded modules. □

Theorem 1.1 will be often used in the following weaker form:

Corollary 2.2. *Let X and Y be smooth projective derived equivalent varieties defined over an algebraically closed field of characteristic zero. If $F(1, L) = (1, M)$, then for any integers m and $k \geq 0$ there exist isomorphisms*

$$\bigoplus_{q=0}^{k} H^{k-q}(X, \Omega_X^{\dim X - q}\otimes\omega_X^{\otimes m}\otimes L)\cong\bigoplus_{q=0}^{k} H^{k-q}(Y, \Omega_Y^{\dim Y - q}\otimes\omega_Y^{\otimes m}\otimes M).$$

Proof. The corollary is a consequence of Theorem 1.1 and of the general fact that the groups $\operatorname{Ext}^k_{X\times X}(i_*\mathcal{O}_X, i_*\mathcal{F})$ decompose as $\bigoplus_{q=0}^{k} H^{k-q}(X, \Omega_X^{\dim X-q}\otimes\omega_X^{-1}\otimes\mathcal{F})$ for any coherent sheaf $\mathcal{F}$ and for all $k \geq 0$ [Yekutieli 2003, Corollary 4.7; Swan 1996, Corollary 2.6]. □

3. Behavior of cohomological support loci under derived equivalence

In this section, we study the behavior of cohomological support loci under derived equivalence. Applications of our analysis will be provided in Section 7. From now on, we work over the field of the complex numbers.

3A. *Cohomological support loci.* Let X be a complex smooth projective irregular variety. Given a coherent sheaf $\mathcal{F}$ on X, we define the *cohomological support loci of* $\mathcal{F}$ as

$$V_r^k(\mathcal{F}) := \{L \in \operatorname{Pic}^0(X) \mid h^k(X, \mathcal{F}\otimes L) \geq r\}$$

for all integers $k \geq 0$ and $r \geq 1$. By semicontinuity, these loci are algebraic closed subsets in $\operatorname{Pic}^0(X)$. We set $V^k(\mathcal{F}) := V_1^k(\mathcal{F})$, and we denote by $V_r^k(\mathcal{F})_0$ the union of all the irreducible components of $V_r^k(\mathcal{F})$ passing through the origin of $\operatorname{Pic}^0(X)$. By following the work of Pareschi and Popa [2011], we say that $\mathcal{F}$ is a *GV-sheaf* if

$$\operatorname{codim}_{\operatorname{Pic}^0(X)} V^k(\mathcal{F}) \geq k \quad \text{for all } k > 0.$$

In the following, we study the behavior of the loci $V_r^k(\mathcal{F})$ under equivalence of derived categories where $\mathcal{F} = \omega_X, \omega_X^{\otimes m}, \Omega_X^p\otimes\omega_X^{\otimes m}$ with $m, p \in \mathbb{Z}$ and $p \geq 0$. We recall that the cohomological support loci $V^k(\omega_X)$ associated to the canonical bundle are invariant under birational modifications for all $k \geq 0$. Furthermore, they detect the Albanese dimension of X, namely, the dimension of the image of the Albanese map $\mathrm{alb}_X : X \to \mathrm{Alb}(X)$, thanks to the following formula [Popa 2013, p. 7] deduced from results of [Green and Lazarsfeld 1987; Lazarsfeld and Popa 2010]:

$$\dim \mathrm{alb}_X(X) = \min_{k=0,\dots,\dim X} \{\dim X - k + \operatorname{codim} V^k(\omega_X)_0\}. \tag{5}$$

Finally, we point out that, if $\dim \mathrm{alb}_X(X) = \dim X - k$, then there are inclusions

$$V^k(\omega_X) \supset V^{k+1}(\omega_X) \supset \cdots \supset V^{\dim X}(\omega_X) = \{\mathcal{O}_X\} \tag{6}$$

(see [Pareschi and Popa 2011, Proposition 3.14; Green and Lazarsfeld 1987, Theorem 1] or [Ein and Lazarsfeld 1997, Lemma 1.8] for the case $k = 0$).

3B. ***Derived invariance of the zeroth cohomological support locus.*** The following proposition proves and extends Theorem 1.4(i):

Proposition 3.1. *Let X and Y be smooth projective varieties, and let $\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ be an equivalence. Denote by F the induced Rouquier isomorphism, and let m and r be integers such that $r \geq 1$. If $L \in V_r^0(\omega_X^{\otimes m})$ and $F(1, L) = (\psi, M)$, then $\psi = 1$ and $M \in V_r^0(\omega_Y^{\otimes m})$. Moreover, F induces an isomorphism of algebraic sets*

$$V_r^0(\omega_X^{\otimes m}) \cong V_r^0(\omega_Y^{\otimes m}).$$

Proof. Let L be a line bundle in $V_r^0(\omega_X^{\otimes m})$, and suppose that $F(1, L) = (\psi, M)$ for some $\psi \in \mathrm{Aut}^0(Y)$ and $M \in \mathrm{Pic}^0(Y)$. By Theorem 1.1 and the adjunction formula, we have

$$\begin{aligned} r \leq h^0(X, \omega_X^{\otimes m} \otimes L) &= \dim \mathrm{Hom}_{X \times X}\big(i_*\mathcal{O}_X, i_*(\omega_X^{\otimes m} \otimes L)\big) \\ &= \dim \mathrm{Hom}_{Y \times Y}\big(j_*\mathcal{O}_Y, (1, \psi)_*(\omega_Y^{\otimes m} \otimes M)\big) \\ &= \dim \mathrm{Hom}_Y\big((1, \psi)^* j_*\mathcal{O}_Y, \omega_Y^{\otimes m} \otimes M\big). \end{aligned}$$

Since $(1, \psi)^* j_*\mathcal{O}_Y$ is supported on the locus of fixed points of ψ (which is of codimension ≥ 1 if $\psi \neq 1$) and since there are no nonzero morphisms from a torsion sheaf to a locally free sheaf, we must have that ψ is the identity automorphism on Y and consequently that $M \in V_r^0(\omega_Y^{\otimes m})$. Therefore, we have an inclusion of algebraic sets $F(1, V_r^0(\omega_X^{\otimes m})) \subset (1, V_r^0(\omega_Y^{\otimes m}))$.

In order to show the reverse inclusion, we consider the right adjoint $\Psi_{\mathscr{E}^*}$ to $\Phi_{\mathscr{E}}$ so that $\Psi_{\mathscr{E}^*} \circ \Phi_{\mathscr{E}} \cong 1_{\boldsymbol{D}(X)}$ and $\Phi_{\mathscr{E}} \circ \Psi_{\mathscr{E}^*} \cong 1_{\boldsymbol{D}(Y)}$. An easy computation shows that, if F' is the Rouquier isomorphism induced by $\Psi_{\mathscr{E}^*}$, then $F' = F^{-1}$ [Lombardi 2013, Lemma 2.1.9]. Hence, by repeating the previous argument, we get an inclusion $F^{-1}(1, V_r^0(\omega_Y^{\otimes m})) \subset (1, V_r^0(\omega_X^{\otimes m}))$ inducing the wanted isomorphism. $\square$

3C. ***Behavior of higher cohomological support loci under derived equivalence.*** In this section, we establish the isomorphism $V^1(\omega_X)_0 \cong V^1(\omega_Y)_0$ of Theorem 1.4. It turns out that, by using the same techniques (i.e., invariance of twisted Hochschild homology and Brion's results on actions of nonaffine groups), one can show a more general result involving cohomological support loci associated to bundles of holomorphic p-forms, which we now present.

Theorem 3.2. *Let X and Y be smooth projective varieties of dimension d, and let $\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ be an equivalence. Denote by F be the induced Rouquier isomorphism, and let m be an integer. If $L \in \bigcup_{p,q\geq 0} V^p(\Omega_X^q \otimes \omega_X^{\otimes m})_0$ and $F(1, L) = (\psi, M)$, then $\psi = 1$ and $M \in \bigcup_{p,q\geq 0} V^p(\Omega_Y^q \otimes \omega_Y^{\otimes m})_0$. Moreover, F induces isomorphisms of algebraic sets*

$$\bigcup_{q=0}^{k} V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})_0 \cong \bigcup_{q=0}^{k} V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m})_0 \quad \textit{for any } k \geq 0.$$

Proof. To begin with, we recall some notation and facts from [Popa and Schnell 2011, Theorem A]. Let $\alpha : \mathrm{Pic}^0(Y) \to \mathrm{Aut}^0(X)$ and $\beta : \mathrm{Pic}^0(X) \to \mathrm{Aut}^0(Y)$ be morphisms defined as

$$\alpha(M) = \mathrm{pr}_1(F^{-1}(1, M)) \quad \text{and} \quad \beta(L) = \mathrm{pr}_1(F(1, L))$$

(pr_1 denotes the projection onto the first factor from the product $\mathrm{Aut}^0(\cdot) \times \mathrm{Pic}^0(\cdot)$). We denote by A and B the images of α and β, respectively. We recall that A and B are isogenous abelian varieties.

We first consider the case when A is trivial. Then $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$, and by Corollary 2.2, we get inclusions

$$F\big(1, \textstyle\bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m} r)\big) \subset \big(1, \textstyle\bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m} r)\big) \quad \text{for any } k \geq 0.$$

In order to prove the reverse inclusions, we note that B is trivial as well and that the Rouquier isomorphism induced by the right adjoint $\Psi_{\mathscr{E}^*}$ to $\Phi_{\mathscr{E}}$ is F^{-1}. Therefore, a second application of Corollary 2.2 yields inclusions

$$F^{-1}\big(1, \textstyle\bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m} r)\big) \subset \big(1, \textstyle\bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})\big) \quad \text{for any } k \geq 0,$$

concluding the proof of this case.

We suppose now that both A and B are nontrivial. We first show the following:

Claim 3.3. *There are inclusions $F\big(1, \bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})_0\big) \subset (1, \mathrm{Pic}^0(Y))$ for all integers m and $k \geq 0$.*

Proof. Brion's results on actions of nonaffine algebraic groups imply that X is an étale locally trivial fibration $\xi : X \to A/H$ where H is a finite subgroup of A (the proof of this fact is analogous to the one of [Popa and Schnell 2011, Lemma 2.4];

see also [Brion 2010]). Let Z be the smooth and connected fiber of ξ over the origin of A/H. Via base change, we get a commutative diagram

$$\begin{tikzcd} A \times Z \arrow[r, "g"] \arrow[d] & X \arrow[d, "\xi"] \\ A \arrow[r] & A/H \end{tikzcd}$$

where $g(\varphi, z) = \varphi(z)$. Let $(z_0, y_0) \in Z \times Y$ be an arbitrary point, and let

$$f = (f_1 \times f_2) : A \times B \to X \times Y$$

be the orbit map $(\varphi, \psi) \mapsto (\varphi(z_0), \psi(y_0))$. In [Popa and Schnell 2011, p. 533], it is shown that

$$L \in (\operatorname{Ker} f_1^*)_0 \implies F(1, L) = (1, M) \quad \text{for some } M \in \operatorname{Pic}^0(Y)$$

(here $(\operatorname{Ker} f_1^*)_0$ denotes the neutral component of $\operatorname{Ker} f_1^*$). So it is enough to show the inclusion

$$\bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})_0 \subset (\operatorname{Ker} f_1^*)_0 \quad \text{for any } k \geq 0. \tag{7}$$

This is achieved by computing cohomology groups on $A \times Z$ via the étale morphism g and by using the fact that these computations are straightforward on A. Let p_1 and p_2 be the projections from the product $A \times Z$ onto the first and second factors, respectively. By denoting by $\nu : A \times \{z_0\} \hookrightarrow A \times Z$ the inclusion morphism, we have $g \circ \nu = f_1$. Moreover, via the isomorphism $\operatorname{Pic}^0(A \times Z) \cong \operatorname{Pic}^0(A) \times \operatorname{Pic}^0(Z)$, we obtain $g^*L \cong p_1^*L_1 \otimes p_2^*L_2$, where $L_1 \in \operatorname{Pic}^0(A)$ and $L_2 \in \operatorname{Pic}^0(Z)$. Note also that $f_1^*L \cong \nu^*g^*L \cong L_1$. Finally, for all $L \in \bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})$, there are inclusions

$$0 \neq \bigoplus_{q=0}^{k} H^{k-q}(X, \Omega_X^{d-q} \otimes \omega_X^{\otimes m} \otimes L) \subset \bigoplus_{q=0}^{k} H^{k-q}(A \times Z, \Omega_{A\times Z}^{d-q} \otimes \omega_{A\times Z}^{\otimes m} \otimes g^*L) \tag{8}$$

[Lazarsfeld 2004, Injectivity Lemma 4.1.14]. Therefore, thanks to Künneth's formula, the sum on the right-hand side of (8) is nonzero only if $f_1^*L \cong \mathcal{O}_A$, i.e., $L \in \operatorname{Ker} f_1^*$. This shows (7). □

By Claim 3.3 and Corollary 2.2, we obtain that for any $k \geq 0$ the Rouquier isomorphism maps

$$1 \times \bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})_0 \mapsto 1 \times \bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m})_0.$$

In complete analogy, one can also show that

$$M \in \bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m})_0 \implies F^{-1}(1, M) = (1, L) \quad \text{for some } L \in \operatorname{Pic}^0(X).$$

This concludes the proof since, by Corollary 2.2, F^{-1} maps

$$1 \times \bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m})_0 \mapsto 1 \times \bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m})_0$$

for any $k \geq 0$. $\square$

The following corollaries yield the proofs of Theorem 1.4(iii) and (ii):

Corollary 3.4. *Under the assumptions of Theorem 3.2, the Rouquier isomorphism F induces isomorphisms of algebraic sets*

$$V_r^1(\omega_X)_0 \cong V_r^1(\omega_Y)_0 \quad \textit{for any } r \geq 1.$$

Proof. Let $L \in V_r^1(\omega_X)_0$. By Theorem 3.2, we have $F(1, L) = (1, M)$ for some $M \in \mathrm{Pic}^0(Y)$, and by Corollary 2.2, we get an isomorphism

$$H^1(X, \omega_X \otimes L) \oplus H^0(X, \Omega_X^{d-1} \otimes L) \cong H^1(Y, \omega_Y \otimes M) \oplus H^0(Y, \Omega_Y^{d-1} \otimes M).$$

Moreover, by Serre duality and the Hodge linear-conjugate isomorphism, we obtain equalities

$$h^0(X, \Omega_X^{d-1} \otimes L) = h^1(X, \omega_X \otimes L) \quad \text{and} \quad h^0(Y, \Omega_Y^{d-1} \otimes M) = h^1(Y, \omega_Y \otimes M).$$

Hence, $h^1(X, \omega_X \otimes L) = h^1(Y, \omega_Y \otimes M) \geq r$, and therefore, F induces the wanted isomorphisms as in the proof of Theorem 3.2. $\square$

Corollary 3.5. *Under the assumptions of Theorem 3.2, and for any integers l, m, r and s with $r, s \geq 1$, the Rouquier isomorphism F induces isomorphisms of algebraic sets*

$$V_r^0(\omega_X^{\otimes m}) \cap \big(\bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes l})\big) \cong V_r^0(\omega_Y^{\otimes m}) \cap \big(\bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes l})\big),$$
$$V_r^0(\omega_X^{\otimes m}) \cap V_s^1(\omega_X) \cong V_r^0(\omega_Y^{\otimes m}) \cap V_s^1(\omega_Y).$$

Proof. In Proposition 3.1, we have seen that, if $L \in V_r^0(\omega_X^{\otimes m})$, then $F(1, L) = (1, M)$ for some $M \in V_r^0(\omega_Y^{\otimes m})$. We argue then as in the proofs of Theorem 3.2 and Corollary 3.4. $\square$

Remark 3.6. It is important to note that, whenever $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$, the proofs of Theorem 3.2 and Corollary 3.4 yield full isomorphisms

$$\bigcup_q V^{k-q}(\Omega_X^{d-q} \otimes \omega_X^{\otimes m}) \cong \bigcup_q V^{k-q}(\Omega_Y^{d-q} \otimes \omega_Y^{\otimes m}) \quad \text{for any } k \geq 0,$$
$$V_r^1(\omega_X) \cong V_r^1(\omega_Y).$$

By Theorem 3.2, this occurs either if $V^p(\Omega_X^q \otimes \omega_X^{\otimes m}) = \mathrm{Pic}^0(X)$ for some $p, q \geq 0$ and $m \in \mathbb{Z}$ or if $\mathrm{Aut}^0(X)$ is affine (since in this case the abelian variety A in the proof of Theorem 3.2 is trivial).

4. Popa's conjectures in dimensions two and three

In this section, we aim to prove Theorem 1.5(i). In other words, we show that Conjecture 1.3, predicting the derived invariance of cohomological support loci of type $V^k(\omega_X)_0$, holds in dimension three. The proofs of (ii) and (iii) of the same theorem are postponed to Section 6 since they use the derived invariance of the Albanese dimension, which will be proved in Section 5. Before starting with the proof of Theorem 1.5(i), we make a couple of considerations regarding the case of surfaces.

4A. ***The case of surfaces.*** In dimension two, Popa [2013, Theorem 2.1] proves the derived invariance of the full cohomological support loci $V^k(\omega_X)$. His proof is based on an explicit computation of cohomological support loci according to the classification of surfaces up to Fourier–Mukai equivalences [Bridgeland and Maciocia 2001]. As an application of Proposition 3.1 and Corollary 3.4, we recover this result by making the isomorphisms between cohomological support loci explicit. More precisely, if F is the Rouquier isomorphism induced by an equivalence of derived categories, then $F(1, V_r^k(\omega_X)) = (1, V_r^k(\omega_Y))$ for all integers $k \geq 0$ and $r \geq 1$. Moreover, by using the same techniques, it is possible to show that F induces further isomorphisms $V_r^1(\Omega_X^1) \cong V_r^1(\Omega_Y^1)$ for all $r \geq 1$ (see [Lombardi 2013, Theorem 5.1.2] for a detailed analysis).

Example 4.1 (Elliptic surfaces). Let X be an elliptic surface of Kodaira dimension one and of maximal Albanese dimension (i.e., an isotrivial elliptic surface fibered onto a curve of genus ≥ 2). By following [Beauville 1992], we recall an invariant attached to this type of surfaces. First of all, we note that X admits a unique fibration $f : X \to C$ onto a curve of genus ≥ 2 (see for instance [Popa 2013, p. 5]). We then denote by G the general fiber of f and by $\mathrm{Pic}^0(X, f)$ the kernel of the pull-back of the inclusion $u : G \hookrightarrow X$

$$0 \to \mathrm{Pic}^0(X, f) \to \mathrm{Pic}^0(X) \xrightarrow{u^*} \mathrm{Pic}^0(G).$$

In [Beauville 1992, (1.6)], it is shown that there exists a finite group $\Gamma^0(f)$ and an isomorphism

$$\mathrm{Pic}^0(X, f) \cong f^* \mathrm{Pic}^0(C) \times \Gamma^0(f).$$

The group $\Gamma^0(f)$ is the invariant mentioned above; it is identified with the group of the connected components of $\mathrm{Pic}^0(X, f)$.

We now consider another smooth projective surface Y such that $\boldsymbol{D}(X) \cong \boldsymbol{D}(Y)$. Then, by [Bridgeland and Maciocia 2001, Proposition 4.4], Y is an elliptic surface fibered onto C. Moreover, Y is of maximal Albanese dimension as well. To see this, we observe that, since the cohomological support loci are derived invariant in dimension two, we have $\dim \mathrm{alb}_Y(Y) = \dim \mathrm{alb}_X(X) = 2$ thanks to (5). Hence, we denote by $g : Y \to C$ the unique fibration of Y and by $\Gamma^0(g)$ its invariant. Pham [2011, Theorem 5.2.7] proves that the invariant $\Gamma^0(\cdot)$ attached to this kind of surface is a derived invariant; in other words, he proves that

$$\Gamma^0(f) \cong \Gamma^0(g). \tag{9}$$

Here we note that (9) also follows from the derived invariance of the zeroth cohomological support locus. In fact, by results of Popa [2013, p. 5], we know that

$$V^0(\omega_X) = \mathrm{Pic}^0(X, f) \cong f^* \mathrm{Pic}^0(C) \times \Gamma^0(f)$$

and similarly for $V^0(\omega_Y)$. Therefore, Proposition 3.1 implies

$$f^* \mathrm{Pic}^0(C) \times \Gamma^0(f) \cong V^0(\omega_X) \cong V^0(\omega_Y) \cong g^* \mathrm{Pic}^0(C) \times \Gamma^0(g),$$

which in particular yields (9).

4B. ***Proof of Theorem 1.5(i).***

Theorem 4.2. *Let X and Y be smooth projective threefolds and $\Phi_{\mathscr{E}} : \boldsymbol{D}(X) \to \boldsymbol{D}(Y)$ an equivalence, and let F be the induced Rouquier isomorphism. Then F induces isomorphisms of algebraic sets*

$$V_r^p(\Omega_X^q)_0 \cong V_r^p(\Omega_Y^q)_0 \quad \textit{for any } p, q \geq 0 \textit{ and } r \geq 1.$$

Proof. The isomorphisms $V_r^0(\omega_X) \cong V_r^0(\omega_Y)$ and $V_r^1(\omega_X)_0 \cong V_r^1(\omega_Y)_0$ have been proved in Proposition 3.1 and Corollary 3.4, respectively. On the other hand, the isomorphisms $V_r^3(\omega_X) \cong V_r^3(\omega_Y)$ are trivial and follow by Serre duality. We now show the isomorphisms $V_r^2(\omega_X)_0 \cong V_r^2(\omega_Y)_0$. To begin with, we note that, by Claim 3.3, if $L \in V_r^2(\omega_X)_0$, then necessarily $F(1, L) = (1, M)$ for some line bundle $M \in \mathrm{Pic}^0(Y)$. Moreover, for $k = 0, 1$, we have equalities $h^k(X, \omega_X \otimes L) = h^k(Y, \omega_Y \otimes M)$ whenever $L \in V_r^2(\omega_X)_0$ and $F(1, L) = (1, M)$ (see Corollary 2.2). Finally, since the holomorphic Euler characteristic is both a derived invariant in dimension three [Popa and Schnell 2011, Corollary C] and invariant under deformation, we have equalities $\chi(\omega_X \otimes L) = \chi(\omega_X) = \chi(\omega_Y) = \chi(\omega_Y \otimes M)$, from which we easily deduce $h^2(X, \omega_X \otimes L) = h^2(Y, \omega_Y \otimes M)$. Thus, if $L \in V_r^2(\omega_X)_0$, then $M \in V_r^2(\omega_Y)_0$ and consequently F induces inclusions $F(1, V_r^2(\omega_X)_0) \subset (1, V_r^2(\omega_Y)_0)$. Since F^{-1} is the Rouquier isomorphism induced by the right adjoint $\Psi_{\mathscr{E}^*}$ to $\Phi_{\mathscr{E}}$, we can repeat the previous argument to obtain the reverse inclusions $F^{-1}(1, V_r^2(\omega_Y)_0) \subset (1, V_r^2(\omega_X)_0)$. This in turn yields isomorphisms $V_r^0(\Omega_X^1)_0 \cong V_r^0(\Omega_Y^1)_0$ thanks to Serre duality and the Hodge linear-conjugate isomorphism.

We now prove the isomorphisms $V_r^1(\Omega_X^q)_0 \cong V_r^1(\Omega_Y^q)_0$ for $q = 1, 2$. By Claim 3.3, we have $F(1, V_r^1(\Omega_X^q)_0) \subset (1, \mathrm{Pic}^0(Y))$. By Serre duality and the Hodge linear-conjugate isomorphism, $h^0(X, \Omega_X^1 \otimes L) = h^2(X, \omega_X \otimes L)$ and $h^0(Y, \Omega_Y^1 \otimes M) = h^2(Y, \omega_Y \otimes M)$ for all line bundles $L \in \mathrm{Pic}^0(X)$ and $M \in \mathrm{Pic}^0(Y)$. Consequently, if

$L \in V^0(\Omega^1_X)_0$ and $F(1, L) = (1, M)$, then by Corollary 2.2 with $m = 0$ and $k = 2$ we have $h^1(X, \Omega^2_X \otimes L) = h^1(Y, \Omega^2_Y \otimes M)$. At this point, in order to prove the wanted isomorphisms, it is enough to proceed as before. In complete analogy, one can also prove the isomorphisms $V^1_r(\Omega^1_X)_0 \cong V^1_r(\Omega^1_Y)_0$, this time by using Corollary 2.2 with $m = 0$ and $k = 3$. □

5. Behavior of the Albanese dimension under derived equivalence

In this section, we prove Theorem 1.6. Our main tool is a generalization of a result due to Chen, Hacon and Pardini saying that, if $f : X \to Z$ is a nonsingular representative of the Iitaka fibration of a smooth projective variety X of maximal Albanese dimension, then

$$q(X) - q(Z) = \dim X - \dim Z$$

[Hacon and Pardini 2002, Proposition 2.1; Chen and Hacon 2004, Corollary 3.6]. We generalize this fact in two ways: (i) we consider all possible values of the Albanese dimension of X, and (ii) we replace the Iitaka fibration with a more general class of morphisms.

Lemma 5.1. *Let X and Z be smooth projective varieties and $f : X \to Z$ a surjective morphism with connected fibers. If the general fiber of f is a smooth variety with surjective Albanese map, then*

$$q(X) - q(Z) = \dim \mathrm{alb}_X(X) - \dim \mathrm{alb}_Z(Z).$$

Proof. We follow [Hacon and Pardini 2002, Proposition 2.1; Chen and Hacon 2004, Corollary 3.6]. Due to the functoriality of the Albanese map, we get a commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{alb}_X} & \mathrm{Alb}(X) \\ \Big\downarrow f & & \Big\downarrow f_* \\ Z & \xrightarrow{\mathrm{alb}_Z} & \mathrm{Alb}(Z) \end{array}$$

where f_* is surjective since f is [Beauville 1996, Remark V.14]. Furthermore, f_* has connected fibers. To see this, we denote by K the connected component of $\mathrm{Ker}\, f_*$ through the origin and set $A := \mathrm{Alb}(X)/K$. Then the natural map $\nu : A \to \mathrm{Alb}(Y)$ is étale and f factors through the induced map $Y \times_{\mathrm{Alb}(Y)} A \to Y$, which is étale of the same degree as ν. Since f has connected fibers, we see that ν is an isomorphism and $K = \mathrm{Ker}\, f_*$.

We now show that the image of a general fiber P of f via alb_X is a translate of $\mathrm{Ker}\, f_*$. Since alb_P is surjective, the image of P via alb_X is a translate of a subtorus of $\mathrm{Ker}\, f_*$. Furthermore, since P moves in a continuous family, such images are all translates of a fixed subtorus $T \subset \mathrm{Ker}\, f_*$. Our next step is to show $T = \mathrm{Ker}\, f_*$. By setting $B := \mathrm{Alb}(X)/T$, we see that the induced morphism $X \to B$ maps a general fiber of f to a point. Therefore, it induces a rational map $h : Z \dashrightarrow B$, which is a morphism since B is an abelian variety. Furthermore, $h(Z)$ generates the abelian variety B since the image of the Albanese map generates the Albanese variety. This leads to the inequality

$$\dim B \le q(Z) = q(X) - \dim \operatorname{Ker} f_*,$$

which in turn yields $\dim T \ge \dim \operatorname{Ker} f_*$ as $\dim B = q(X) - \dim T$. For dimension reasons, we get then $T = \operatorname{Ker} f_*$. In particular, this says that $\mathrm{alb}_X(X)$ is fibered in tori of dimension $q(X) - q(Z)$ over $\mathrm{alb}_Z(Z)$, and by the theorem on the dimension of the fibers of a morphism, we get the stated equality. □

Proof of Theorem 1.6. We begin with the case $\dim X \le 3$. In Sections 4A and 4B, we have seen that in dimension up to three the cohomological support loci associated to the canonical bundle around the origin are derived invariant, i.e., $V^k(\omega_X)_0 \cong V^k(\omega_Y)_0$ for all $k \ge 0$. Therefore, (5), in combination with the fact that derived equivalent varieties have the same dimension, immediately leads to $\dim \mathrm{alb}_X(X) = \dim \mathrm{alb}_Y(Y)$.

We now assume $\dim X > 3$ and $\kappa(X) \ge 0$. If $\kappa(X) = \kappa(Y) = 0$, then the Albanese maps of X and Y are surjective by [Kawamata 1981, Theorem 1]. Thus, the Albanese dimensions of X and Y are $q(X)$ and $q(Y)$, respectively, which are equal by work of Popa and Schnell [2011, Corollary B].

We now suppose $\kappa(X) = \kappa(Y) > 0$. Since the problem is invariant under birational modification, with a little abuse of notation, we consider nonsingular representatives $f : X \to Z$ and $g : Y \to W$ of the Iitaka fibrations of X and Y, respectively [Mori 1987, (1.10)]. As the canonical rings of X and Y are isomorphic [Orlov 2003, Corollary 2.1.9], it turns out that Z and W are birational varieties (see [Mori 1987, Proposition 1.4] or [Toda 2006, p. 13]). By [Kawamata 1981, Theorem 1], the morphisms f and g satisfy the hypotheses of Lemma 5.1, which yields

$$\begin{aligned} q(X) - \dim \mathrm{alb}_X(X) &= q(Z) - \dim \mathrm{alb}_Z(Z) \\ &= q(W) - \dim \mathrm{alb}_W(W) = q(Y) - \dim \mathrm{alb}_Y(Y). \end{aligned}$$

We conclude as $q(X) = q(Y)$. □

As an application of Theorem 1.6, we have the following:

Corollary 5.2. *Let X and Y be smooth projective derived equivalent varieties with X of maximal Albanese dimension. If F denotes the induced Rouquier isomorphism and $F(1, L) = (\psi, M)$ with $L \in V^1_r(\omega_X)$, then $\psi = 1$ and $M \in V^1_r(\omega_Y)$. Moreover, F induces isomorphisms of algebraic sets*

$$V^1_r(\omega_X) \cong V^1_r(\omega_Y) \quad \textit{for any } r \ge 1.$$

Proof. We have $\kappa(X) \geq 0$ since X is of maximal Albanese dimension. Hence, Theorem 1.6 ensures that Y is of maximal Albanese dimension as well. We apply then Corollary 3.5 after having noted the inclusions $V_r^1(\omega_X) \subset V^0(\omega_X)$ and $V_r^1(\omega_Y) \subset V^0(\omega_Y)$ (see (6)). □

6. End of the proof of Theorem 1.5

6A. ***Proof of Theorem 1.5(ii).*** The following two propositions prove and extend Theorem 1.5(ii):

Proposition 6.1. *Let X and Y be smooth projective derived equivalent threefolds, and let F be the induced Rouquier isomorphism. Assume that either* $\mathrm{Aut}^0(X)$ *is affine or that* $V^p(\Omega_X^q \otimes \omega_X^{\otimes m}) = \mathrm{Pic}^0(X)$ *for some $m, p, q \in \mathbb{Z}$ with $p, q \geq 0$. Then F induces isomorphisms of algebraic sets*

$$V_r^p(\Omega_X^q) \cong V_r^p(\Omega_Y^q) \quad \textit{for all } p, q \geq 0 \textit{ and } r \geq 1.$$

Proof. By Remark 3.6, we have $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$. The isomorphisms $V_r^0(\omega_X) \cong V_r^0(\omega_Y)$ and $V_r^1(\omega_X) \cong V_r^1(\omega_Y)$ hold by Proposition 3.1 and Remark 3.6, respectively. The isomorphisms $V_r^2(\omega_X) \cong V_r^2(\omega_Y)$ follow since in dimension three $\chi(\omega_X) = \chi(\omega_Y)$ [Popa and Schnell 2011, Corollary C].

We now establish the isomorphisms $V_r^1(\Omega_X^2) \cong V_r^1(\Omega_Y^2)$. Let $L \in V_r^1(\Omega_X^2)$ so that $F(1, L) = (1, M)$ for some $M \in \mathrm{Pic}^0(Y)$. By Corollary 2.2 with $m = 0$ and $k = 2$, Serre duality and the Hodge linear-conjugate isomorphism, we get $h^1(X, \Omega_X^2 \otimes L) = h^1(Y, \Omega_Y^2 \otimes M)$. This shows that F maps $1 \times V_r^1(\Omega_X^2) \mapsto 1 \times V_r^1(\Omega_Y^2)$, inducing the wanted isomorphisms as in Proposition 3.1. Finally, the isomorphisms $V_r^1(\Omega_X^1) \cong V_r^1(\Omega_Y^1)$ are deduced in the same way by using Corollary 2.2 with $m = 0$ and $k = 3$. □

Proposition 6.2. *Let X and Y be smooth projective derived equivalent threefolds, and let F be the induced Rouquier isomorphism. If X is of maximal Albanese dimension, then F induces isomorphisms of algebraic sets*

$$V_r^k(\omega_X) \cong V_r^k(\omega_Y) \quad \textit{for all } k \geq 0 \textit{ and } r \geq 1.$$

Proof. Proposition 3.1 and Corollary 5.2 yield the isomorphisms $V_r^k(\omega_X) \cong V_r^k(\omega_Y)$ for any $k \neq 2$, so we only focus on the remaining case. Since X is of maximal Albanese dimension, we obtain an inclusion $V_r^2(\omega_X) \subset V^0(\omega_X)$ (see (6)) leading to a further inclusion $F(1, V_r^2(\omega_X)) \subset (1, \mathrm{Pic}^0(Y))$ thanks to Proposition 3.1. Hence, by Corollary 2.2, $h^k(X, \omega_X \otimes L) = h^k(Y, \omega_Y \otimes M)$ whenever $F(1, L) = (1, M)$ with $L \in V_r^2(\omega_X)$ and $k = 0, 1$. Moreover, we get $h^2(X, \omega_X \otimes L) = h^2(Y, \omega_Y \otimes M)$ since $\chi(\omega_X) = \chi(\omega_Y)$ [Popa and Schnell 2011, Corollary C]. Therefore, F maps $1 \times V_r^2(\omega_X) \mapsto 1 \times V_r^2(\omega_Y)$, and by arguing as in Proposition 3.1, F^{-1} maps $1 \times V_r^2(\omega_Y) \mapsto 1 \times V_r^2(\omega_X)$, finishing the proof. □

6B. ***Proof of Theorem 1.5(iii).*** We show now the proof of Theorem 1.5(iii). Before jumping into technicalities, we first present the plan of its proof.

Thanks to Propositions 6.1 and 6.2, we can assume that X is a threefold with $\dim \mathrm{alb}_X(X) \leq 2$ and $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ and with nonaffine automorphism group $\mathrm{Aut}^0(X)$. In particular, we can suppose that X is not of general type and that $\chi(\omega_X) = 0$ [Popa and Schnell 2011, Corollary 2.6]. Thanks to Proposition 3.1, Theorem 1.6 and [Popa and Schnell 2011, Theorem A(1)], the Fourier–Mukai partner Y of X satisfies the same hypotheses as X. Hence, Theorem 1.5(iii) follows as soon as we classify $\dim V^i(\omega_X)$ in terms of derived invariants. This classification is carried out in the following Propositions 6.5–6.9 where $\dim V^1(\omega_X)$ and $\dim V^2(\omega_X)$ are computed in terms of $\kappa(X)$, $q(X)$, $\dim \mathrm{alb}_X(X)$ and $\dim V^0(\omega_X)$.

The main tools we use towards the proofs of Propositions 6.5–6.9 are generic vanishing theorems [Green and Lazarsfeld 1987, Theorem 1; Pareschi and Popa 2011, Theorem 5.8], Kollár's result on higher direct images of the canonical bundle [Kollár 1986b, Theorem 3.1; 1986a, Theorem 2.1 and Proposition 7.6] and the classification of smooth projective surfaces (see for instance [Beauville 1996]). The following two lemmas will be useful to our analysis:

Lemma 6.3. *Let X and Y be smooth projective varieties and $f : X \to Y$ be a surjective morphism with connected fibers. If h denotes the dimension of the general fiber of f, then*

$$f^* V^k(\omega_Y) \subset V^{k+h}(\omega_X) \quad \text{for any } k = 0, \dots, \dim Y.$$

Proof. By [Kollár 1986a, Theorem 2.1 and Proposition 7.6], we have $R^h f_* \omega_X \cong \omega_Y$ and $R^k f_* \omega_X = 0$ for $k > h$. Moreover, by [Kollár 1986b, Theorem 3.1], we obtain decompositions

$$H^{k+h}(X, \omega_X \otimes f^* L) \cong H^k(Y, \omega_Y \otimes L) \oplus \bigoplus_{l \neq k} H^l(Y, R^{h+k-l} f_* \omega_X \otimes L)$$

for any $L \in \mathrm{Pic}^0(Y)$. At this point, it is enough to note that the pull-back homomorphism $f^* : \mathrm{Pic}^0(Y) \to \mathrm{Pic}^0(X)$ is injective as the fibers of f are connected. □

Lemma 6.4. *Let X be a smooth projective variety with $\kappa(X) = -\infty$. Then $V^0(\omega_X^{\otimes m}) = \varnothing$ for any $m > 0$.*

Proof. Suppose that $L \in V^0(\omega_X^{\otimes m})$ for some $m > 0$. By [Chen and Hacon 2004, Theorem 3.2], we can assume that L is a line bundle of finite order, say, of order e. If $\mathcal{O}_X \to \omega_X^{\otimes m} \otimes L$ is a nonzero section of $\omega_X^{\otimes m} \otimes L$, then it induces a nonzero section $\mathcal{O}_X \to \omega_X^{\otimes me}$; this yields a contradiction as $\kappa(\omega_X) = -\infty$. □

Proposition 6.5. *Let X be a smooth projective threefold such that $\kappa(X) = 2$, $\dim \mathrm{alb}_X(X) = 2$, $\chi(\omega_X) = 0$ and $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$. If $q(X) = 2$, then we have* (i) $\dim V^2(\omega_X) = 0$, (ii) $\dim V^1(\omega_X) = 1$ if and only if $\dim V^0(\omega_X) = 1$ and

(iii) $\dim V^1(\omega_X) = 0$ if and only if $\dim V^0(\omega_X) \leq 0$. If $q(X) > 2$, then we have $\dim V^1(\omega_X) = \dim V^2(\omega_X) = q(X) - 1$.

Proof. Since the problem is invariant under birational modification, with a little abuse of notation, we consider a nonsingular representative $f : X \to S$ of the Iitaka fibration of X [Mori 1987, (1.10)] so that X and S are smooth varieties and f is an algebraic fiber space. We divide the proof into three cases according to the values of the Albanese dimension of S.

Case I: $\dim \mathrm{alb}_S(S) = 2$. By the classification theory of smooth projective surfaces, S is either a surface of general type, or birational to an abelian surface, or birational to an elliptic surface fibered onto a curve of genus ≥ 2. Moreover, by Lemma 5.1, we have $q(X) = q(S)$.

If S is of general type, then by Castelnuovo's theorem [Beauville 1996, Theorem X.4] we have $\chi(\omega_S) > 0$ and hence $V^0(\omega_S) = \mathrm{Pic}^0(S)$. Therefore, by Lemma 6.3, we get $V^1(\omega_X) = \mathrm{Pic}^0(X)$, and consequently, $V^0(\omega_X) = \mathrm{Pic}^0(X)$ since $\chi(\omega_X) = 0$ and $V^2(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ (see (5)). This contradicts our hypotheses, and hence, this case does not occur.

If S is birational to an abelian surface, then we have $q(X) = q(S) = 2$ and $f^* \mathrm{Pic}^0(S) = \mathrm{Pic}^0(X)$. By using [Kollár 1986b, Theorem 3.1], we obtain decompositions

$$H^2(X, \omega_X \otimes f^* L) \cong H^2(S, f_*\omega_X \otimes L) \oplus H^1(S, R^1 f_*\omega_X \otimes L)$$

for any $L \in \mathrm{Pic}^0(S)$. Moreover, we note that $R^1 f_*\omega_X \cong \omega_S$ and $R^2 f_*\omega_X = 0$ [Kollár 1986a, Proposition 7.6 and Theorem 2.1]. Therefore, since by [Pareschi and Popa 2011, Theorem 5.8] $f_*\omega_X$ is a GV-sheaf on S (i.e., $\mathrm{codim}_{\mathrm{Pic}^0(S)} V^k(f_*\omega_X) \geq k$ for $k > 0$), we get $\dim V^2(\omega_X) = 0$. At this point, the statements (ii) and (iii) of the proposition follow as $\chi(\omega_X) = 0$ and $\dim V^1(\omega_X) \geq 0$ (note that $\mathcal{O}_X \in V^1(\omega_X)$ since $q(X) = 2$).

If S is birational to an elliptic surface $h : S \to C$ fibered onto a curve C of genus $g(C) = q(S) - 1 = q(X) - 1 \geq 2$, then X is fibered onto C as well. Therefore, we have $V^0(\omega_C) = \mathrm{Pic}^0(C)$, and consequently, $V^2(\omega_X)$ is of codimension one in $\mathrm{Pic}^0(X)$ by Lemma 6.3 and (5). Since $\chi(\omega_X) = 0$, $V^1(\omega_X)$ is of codimension one as well.

Case II: $\dim \mathrm{alb}_S(S) = 1$. We have $q(X) = q(S) + 1$ by Lemma 5.1. Moreover, alb_S has connected fibers, and by [Beauville 1996, Proposition V.15], $\mathrm{alb}_S(S)$ is a smooth curve of genus $q(S)$. We distinguish two subcases: $q(S) = 1$ and $q(S) \geq 2$.

If $q(S) = 1$, then $q(X) = 2$ and alb_X is surjective. Let $X \xrightarrow{b} Z \to \mathrm{Alb}(X)$ be the Stein factorization of alb_X, and let $b' : X' \to Z'$ be a nonsingular representative of b. We note that Z' is a smooth surface with $q(Z') = 2$ and of maximal Albanese dimension. Therefore, either Z' is of general type or it is birational to an abelian surface. However, we have just seen that Z' cannot possibly be of general type;

therefore, Z' is birational to an abelian surface, and the same calculations of the previous case apply.

If $q(S) \geq 2$, then the Albanese map of S induces a fibration of S onto a smooth curve C of genus $g(C) = q(S)$. Therefore, X is fibered onto C as well, and we conclude as in the previous case.

Case III: $\dim \mathrm{alb}_S(S) = 0$. As we have seen in the proof of Lemma 5.1, the image of a general fiber of f is mapped via alb_X onto a fiber of the induced morphism $f_* : \mathrm{Alb}(X) \to \mathrm{Alb}(S)$. On the other hand, if $\dim \mathrm{alb}_S(S) = 0$, then $\mathrm{Alb}(S)$ is trivial. This yields a contradiction, and therefore, this case does not occur. □

Proposition 6.6. *Let X be a smooth projective threefold such that $\kappa(X) = 2$, $\dim \mathrm{alb}_X(X) = 1$, $\chi(\omega_X) = 0$ and $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$. If $q(X) = 1$, then we have $\dim V^1(\omega_X) \leq 0$ and $\dim V^2(\omega_X) = 0$. On the other hand, if $q(X) > 1$, then we have $V^1(\omega_X) = V^2(\omega_X) = \mathrm{Pic}^0(X)$.*

Proof. As in the previous proof, we denote by $f : X \to S$ a nonsingular representative of the Iitaka fibration of X. We distinguish two cases: $\dim \mathrm{alb}_S(S) = 0$ and $\dim \mathrm{alb}_S(S) = 1$.

If $\dim \mathrm{alb}_S(S) = 0$, then we have $q(S) = 0$ and therefore $q(X) = 1$ by Lemma 5.1. Moreover, by [Ueno 1973, Lemma 2.11], alb_X is surjective and has connected fibers. We set $E := \mathrm{Alb}(X)$ and $a := \mathrm{alb}_X$ and note that by [Kollár 1986a, Proposition 7.6] there is an isomorphism $R^2 a_* \omega_X \cong \mathcal{O}_E$. Finally, by [Kollár 1986b, Theorem 3.1], we get isomorphisms

$$H^2(X, \omega_X \otimes a^* L) \cong H^1(E, R^1 a_* \omega_X \otimes L) \oplus H^0(E, L)$$

for any $L \in \mathrm{Pic}^0(E) \cong \mathrm{Pic}^0(X)$. By [Hacon 2004, Corollary 4.2], $R^1 a_* \omega_X$ is a GV-sheaf on E. Hence, $\dim V^2(\omega_X) = 0$, and consequently, $V^1(\omega_X)$ is either empty or zero-dimensional as $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ and $\chi(\omega_X) = 0$.

We now suppose $\dim \mathrm{alb}_S(S) = 1$. In this case, alb_S has connected fibers and its image is a smooth curve B of genus $g(B) = q(S) > 1$. Moreover, we have $q(X) = q(S)$ by Lemma 5.1. We distinguish two subcases: $q(S) = 1$ and $q(S) > 1$. If $q(S) = 1$, then the image of alb_X is an elliptic curve and the same argument of the previous case applies. If $q(S) = g(B) > 1$, then we get $V^0(\omega_B) = \mathrm{Pic}^0(B)$ and $\mathrm{Pic}^0(X) \cong \mathrm{Pic}^0(S) \cong \mathrm{Pic}^0(B)$. Hence, by Lemma 6.3, there are inclusions

$$\mathrm{alb}_S^* \mathrm{Pic}^0(B) = \mathrm{alb}_S^* V^0(\omega_B) \subset V^1(\omega_S) \subset \mathrm{Pic}^0(S)$$

leading to $V^1(\omega_S) = \mathrm{Pic}^0(S)$. Moreover, a second application of Lemma 6.3 gives

$$f^* V^1(\omega_S) \subset V^2(\omega_X) \subset \mathrm{Pic}^0(X),$$

showing that $V^2(\omega_X) = \mathrm{Pic}^0(X)$. Finally, we also have $V^1(\omega_X) = \mathrm{Pic}^0(X)$ as $\chi(\omega_X) = 0$. □

Proposition 6.7. *Let X be a smooth projective threefold such that $\kappa(X) = 1$, $\chi(\omega_X) = 0$ and $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$.*

(i) *Assume* $\dim \mathrm{alb}_X(X) = 2$. *If* $q(X) = 2$, *then we have* (i) $\dim V^2(\omega_X) = 0$, (ii) $\dim V^1(\omega_X) = 1$ if and only if $\dim V^0(\omega_X) = 1$ and (iii) $\dim V^1(\omega_X) = 0$ if and only if $\dim V^0(\omega_X) \leq 0$. If $q(X) \geq 3$, then we have $\dim V^1(\omega_X) = \dim V^2(\omega_X) = q(X) - 1$.

(ii) Assume $\dim \mathrm{alb}_X(X) = 1$. If $q(X) = 1$, then we have $\dim V^1(\omega_X) \leq 0$ and $\dim V^2(\omega_X) = 0$. If $q(X) \geq 2$, then we obtain $V^1(\omega_X) = V^2(\omega_X) = \mathrm{Pic}^0(X)$.

Proof. We start with the case $\dim \mathrm{alb}_X(X) = 2$. Let $f : X \to C$ be a nonsingular representative of the Iitaka fibration of X where C is a smooth curve.

If $g(C) \geq 2$, then by Lemma 5.1 we have $q(X) = g(C)+1 \geq 3$, and by Lemma 6.3, we obtain a series of inclusions $f^* \mathrm{Pic}^0(C) = f^* V^0(\omega_C) \subset V^2(\omega_X) \subset \mathrm{Pic}^0(X)$. We conclude that

$$\dim V^2(\omega_X) = q(X) - 1$$

since $V^2(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ by (5). Therefore, we see that $V^1(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ as $\chi(\omega_X) = 0$ and $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$. Finally, thanks to the inclusion $V^1(\omega_X) \supset V^2(\omega_X)$ of (6), we obtain $\dim V^1(\omega_X) = q(X) - 1$.

If $g(C) \leq 1$, then $q(X) = 2$ and $a := \mathrm{alb}_X$ is surjective. Let $b : X' \to Z'$ be a nonsingular representative of the Stein factorization of a. Then, as we have seen in the proof of Proposition 6.5, Z' is birational to an abelian surface, and therefore, $\dim V^2(\omega_X) = 0$. Since $\mathcal{O}_X \in V^1(\omega_X)$, we obtain the statements (ii) and (iii) of part (i).

We now study the case $\dim \mathrm{alb}_X(X) = 1$. If $g(C) \geq 2$, then $q(X) = g(C)$ and $f^* \mathrm{Pic}^0(C) = \mathrm{Pic}^0(X)$. Therefore, by Lemma 6.3, we get $V^2(\omega_X) = \mathrm{Pic}^0(X)$, and hence, we have $V^1(\omega_X) = \mathrm{Pic}^0(X)$. On the other hand, if $g(C) \leq 1$, then $q(X) = 1$ and $\mathrm{alb}_X : X \to \mathrm{Alb}(X)$ is an algebraic fiber space onto an elliptic curve. We conclude then as in the proof of Proposition 6.6. □

Proposition 6.8. *Let X be a smooth projective threefold such that $\kappa(X) = 0$ and $\chi(\omega_X) = 0$. If* $\dim \mathrm{alb}_X(X) = 2$, *then we have* $\dim V^1(\omega_X) = \dim V^2(\omega_X) = 0$. *On the other hand, if* $\dim \mathrm{alb}_X(X) = 1$, *then we have* $\dim V^1(\omega_X) \leq 0$ *and* $\dim V^2(\omega_X) = 0$.

Proof. We recall that, by [Chen and Hacon 2002, Lemma 3.1], $V^0(\omega_X)$ consists of at most one point. We start with the case $\dim \mathrm{alb}_X(X) = 2$. By [Kawamata 1981, Theorem 1], alb_X is surjective and has connected fibers. Therefore, we have $q(X) = h^2(X, \omega_X) = 2$ and hence $\mathcal{O}_X \in V^1(\omega_X)$ since $\chi(\omega_X) = 0$. We set $a := \mathrm{alb}_X$, and we note that, by [Hacon 2004, Corollary 4.2], $a_*\omega_X$ is a GV-sheaf, i.e.,

$$\operatorname{codim} V^1(a_*\omega_X) \geq 1 \quad \text{and} \quad \operatorname{codim} V^2(a_*\omega_X) \geq 2.$$

By using that $R^1a_*\omega_X \cong \mathcal{O}_{\mathrm{Alb}(X)}$ and $R^2a_*\omega_X = 0$ [Kollár 1986a, Proposition 7.6 and Theorem 2.1] and by using [Kollár 1986b, Theorem 3.1], we get isomorphisms

$$H^1(X, \omega_X \otimes a^*L) \cong H^1(\mathrm{Alb}(X), a_*\omega_X \otimes L) \oplus H^0(\mathrm{Alb}(X), L)$$

for any $L \in \mathrm{Pic}^0(\mathrm{Alb}(X)) \cong \mathrm{Pic}^0(X)$. Therefore, we have

$$\operatorname{codim} V^1(\omega_X) \geq 1 \quad \text{and} \quad \operatorname{codim} V^2(\omega_X) \geq 2,$$

and consequently, the hypothesis $\chi(\omega_X) = 0$ implies $\dim V^1(\omega_X) = 0$.

If $\dim \mathrm{alb}_X(X) = 1$, then as in the previous case we have $\dim V^2(\omega_X) = 0$. Therefore, $V^1(\omega_X)$ is either empty or of dimension zero since $\chi(\omega_X) = 0$. □

Proposition 6.9. *Let X be a smooth projective threefold such that $\kappa(X) = -\infty$ and $\chi(\omega_X) = 0$.*

(i) *Suppose* $\dim \mathrm{alb}_X(X) = 2$. *If* $q(X) = 2$, *then* $V^1(\omega_X) = V^2(\omega_X) = \{\mathcal{O}_X\}$. *If* $q(X) > 2$, *then we obtain* $\dim V^1(\omega_X) = \dim V^2(\omega_X) = q(X) - 1$.

(ii) *Suppose* $\dim \mathrm{alb}_X(X) = 1$. *If* $q(X) = 1$, *then we have* $\dim V^1(\omega_X) \leq 0$ *and* $\dim V^2(\omega_X) = 0$. *If* $q(X) > 1$, *then we obtain* $V^1(\omega_X) = V^2(\omega_X) = \mathrm{Pic}^0(X)$.

Proof. We start with the case $\dim \mathrm{alb}_X(X) = 2$. Let $a : X \to S \subset \mathrm{Alb}(X)$ be the Albanese map of X and $b : X \to S'$ be the Stein factorization of a, and let $c : X' \to S''$ be a nonsingular representative of b. We can easily check that $q(X') = q(S'')$ and $\dim \mathrm{alb}_S(S) = 2$ and hence that $\kappa(S'') \geq 0$. Furthermore, we have $c_*\omega_{X'} = 0$. To see this, we point out that by [Pareschi and Popa 2011, Theorem 5.8] $c_*\omega_{X'}$ is a GV-sheaf on S'' and moreover that, by Lemma 6.4, $V^0(c_*\omega_{X'}) = V^0(\omega_{X'}) = V^0(\omega_X) = \varnothing$. This immediately implies $c_*\omega_{X'} = 0$ as a GV-sheaf $\mathcal{F}$ is nonzero if and only if $V^0(\mathcal{F}) \neq \varnothing$. We distinguish now three cases according to the values of $\kappa(S'')$.

If $\kappa(S'') = 0$, then S'' is birational to an abelian surface. This forces $q(X) = q(X') = q(S'') = 2$ and $c^* \mathrm{Pic}^0(S'') = \mathrm{Pic}^0(X')$. By [Kollár 1986b, Theorem 3.1; 1986a, Theorem 2.1 and Proposition 7.6], we obtain isomorphisms

$$H^2(X', \omega_{X'} \otimes c^*L) \cong H^1(S'', \omega_{S''} \otimes L) \quad \text{and} \quad H^1(X', \omega_{X'} \otimes c^*L) \cong H^0(S'', \omega_{S''} \otimes L)$$

for any $L \in \mathrm{Pic}^0(S'')$. Therefore, we have $V^2(\omega_X) \cong V^2(\omega_{X'}) = c^*V^1(\omega_{S''}) = \{\mathcal{O}_{X'}\}$ and $V^1(\omega_X) \cong V^1(\omega_{X'}) = c^*V^0(\omega_{S''}) = \{\mathcal{O}_{X'}\}$.

If $\kappa(S'') = 1$, then S'' is birational to an elliptic surface of maximal Albanese dimension fibered onto a curve of genus $g(C) \geq 2$. Thus, X is fibered onto C as well and $q(X') = q(S'') = g(C) + 1$. By Lemma 6.3 and (5), we deduce $\dim V^2(\omega_{X'}) = g(C) = q(X') - 1$, and therefore, we get $\dim V^1(\omega_{X'}) = q(X') - 1$ as $\chi(\omega_{X'}) = 0$ and $V^0(\omega_{X'}) = \varnothing$.

If $\kappa(S'') = 2$, then by Castelnuovo's theorem we have $\chi(\omega_{S''}) > 0$, which immediately yields $V^0(\omega_{S''}) = \mathrm{Pic}^0(S'')$. By using Lemma 6.3, we see that $\dim V^0(\omega_{X'}) > 0$. This contradicts Lemma 6.4, and hence, this case does not occur.

We now suppose $\dim \mathrm{alb}_X(X) = 1$. Let $a : X \to C \subset \mathrm{Alb}(X)$ be the Albanese map of X where $C := \mathrm{Im}\, a$. Then a has connected fibers and $q(X) = g(C)$ by [Ueno 1973, Lemma 2.11]. As in the previous case, we note that $a_*\omega_X = 0$. Moreover, by [Kollár 1986b, Theorem 3.1; 1986a, Proposition 7.6], we obtain isomorphisms

$$H^1(X, \omega_X \otimes a^*L) \cong H^0(C, R^1a_*\omega_X \otimes L),$$
$$H^2(X, \omega_X \otimes a^*L) \cong H^1(C, R^1a_*\omega_X \otimes L) \oplus H^0(C, \omega_C \otimes L)$$

for any $L \in \mathrm{Pic}^0(C)$. At this point, we distinguish two cases: $g(C) = 1$ and $g(C) > 1$. If $g(C) = q(X) > 1$, then we have $V^0(\omega_C) = \mathrm{Pic}^0(C)$, and by Lemma 6.3, we get $V^2(\omega_X) = V^1(\omega_X) = \mathrm{Pic}^0(X)$. On the other hand, if $g(C) = q(X) = 1$, then by [Hacon 2004, Corollary 4.2] $R^1a_*\omega_X$ is a GV-sheaf on $C = \mathrm{Alb}(X)$. Hence, we obtain $\dim V^2(\omega_X) = 0$, and consequently, we see that $\dim V^1(\omega_X) \leq 0$ since $\chi(\omega_X) = 0$ and $V^0(\omega_X) = \varnothing$. $\square$

Remark 6.10. In the case $q(X) = 1$, the previous propositions yield the following statement: for each k, $\dim V^k(\omega_X) = 1$ if and only if $\dim V^k(\omega_Y) = 1$. In general, we have not been able to show that, if a locus $V^k(\omega_X)$ is empty or of dimension zero, then the corresponding locus $V^k(\omega_Y)$ is empty or of dimension zero, respectively. This ambiguity is mainly caused by the possible presence of nontrivial automorphisms.

An application of a sheafified version of the derivative complex [Ein and Lazarsfeld 1997, Theorem 3; Lazarsfeld and Popa 2010] can be shown to yield Conjecture 1.2 for threefolds having $q(X) = 2$ [Lombardi 2013, Proposition 5.2.15].

7. Applications

In this final section, we prove Corollaries 1.7, 1.8 and 1.9. Moreover, we present a further result regarding the invariance of the Euler characteristic of powers of the canonical bundle for derived equivalent smooth minimal varieties of maximal Albanese dimension.

7A. *Holomorphic Euler characteristic and Hodge numbers.*

Proof of Corollary 1.7. Let $d := \dim X = \dim Y$. We begin with the case $\dim \mathrm{alb}_X(X) = d$. By Theorem 1.6, Y is of maximal Albanese dimension, and by (5), we get inequalities

$$\mathrm{codim}\, V^1(\omega_X) \geq 1 \quad \text{and} \quad \mathrm{codim}\, V^1(\omega_Y) \geq 1.$$

We distinguish two cases: $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ and $V^0(\omega_X) = \mathrm{Pic}^0(X)$. If $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$, then we also have $V^0(\omega_Y) \subsetneq \mathrm{Pic}^0(Y)$ by Proposition 3.1. Moreover, there are inclusions $\mathrm{Pic}^0(X) \supsetneq V^0(\omega_X) \supset V^1(\omega_X) \supset \cdots \supset V^d(\omega_X) = \{\mathcal{O}_X\}$ and similarly for the loci $V^k(\omega_Y)$ (see (6)). Therefore, if $L \notin V^0(\omega_X)$ and $M \notin V^0(\omega_Y)$, then

$h^k(X, \omega_X \otimes L) = h^k(Y, \omega_Y \otimes M) = 0$ for all $k \geq 0$. Since the holomorphic Euler characteristic is invariant under deformation, we finally obtain

$$\chi(\omega_X) = \chi(\omega_X \otimes L) = 0 = \chi(\omega_Y \otimes M) = \chi(\omega_Y).$$

On the other hand, if $V^0(\omega_X) = \mathrm{Pic}^0(X)$, then, by Proposition 3.1, $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$, and thus,

$$\text{there exists } L_0 \in V^0(\omega_X) \setminus \left(\bigcup_{k=1}^{d} V^k(\omega_X)\right) \\ \text{such that } F(1, L_0) = (1, M_0) \text{ with } M_0 \in V^0(\omega_Y) \setminus \left(\bigcup_{k=1}^{d} V^k(\omega_Y)\right).$$

Hence, by using Corollary 2.2 with $m = k = 0$, we have

$$\chi(\omega_X) = \chi(\omega_X \otimes L_0) = h^0(X, \omega_X \otimes L_0) = h^0(Y, \omega_Y \otimes M_0) = \chi(\omega_Y \otimes M_0) = \chi(\omega_Y).$$

We suppose now $\dim \mathrm{alb}_X(X) = d - 1$ and $\kappa(X) \geq 0$. By Theorem 1.6, we have $\dim \mathrm{alb}_Y(Y) = d - 1$, and therefore, there are inclusions $V^1(\omega_X) \supset V^2(\omega_X) \supset \cdots \supset V^d(\omega_X)$ and $V^1(\omega_Y) \supset V^2(\omega_Y) \supset \cdots \supset V^d(\omega_Y)$. We distinguish four cases.

The first case is when $V^0(\omega_X) = V^1(\omega_X) = \mathrm{Pic}^0(X)$. By Proposition 3.1 and Corollary 3.4, it turns out that $V^0(\omega_Y) = V^1(\omega_Y) = \mathrm{Pic}^0(Y)$ as well. We claim that

$$\text{there exists } \mathcal{O}_X \neq L_1 \in V^0(\omega_X) \backslash V^2(\omega_X) \\ \text{such that } F(1, L_1) = (1, M_1) \text{ with } \mathcal{O}_Y \neq M_1 \in V^0(\omega_Y) \backslash V^2(\omega_Y).$$

In fact, the Rouquier isomorphism maps $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$ by Remark 3.6, and therefore, it is enough to choose the image under F^{-1} of a generic element $(1, M)$ with $M \notin V^2(\omega_Y)$. By using Corollary 2.2 twice, first with $k = 0$ and then with $k = 1$, we obtain

$$\begin{aligned}\chi(\omega_X) &= \chi(\omega_X \otimes L_1) = h^0(X, \omega_X \otimes L_1) - h^1(X, \omega_X \otimes L_1) \\ &= h^0(Y, \omega_Y \otimes M_1) - h^1(Y, \omega_Y \otimes M_1) = \chi(\omega_Y \otimes M_1) = \chi(\omega_Y).\end{aligned}$$

The second case is when $V^0(\omega_X) = \mathrm{Pic}^0(X)$ and $V^1(\omega_X) \subsetneq \mathrm{Pic}^0(X)$. By Proposition 3.1 and Corollary 3.4, $V^0(\omega_Y) = \mathrm{Pic}^0(Y)$ and $V^1(\omega_Y) \subsetneq \mathrm{Pic}^0(Y)$. As before, $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$, and hence, we can pick an element

$$\mathcal{O}_X \neq L_2 \in V^0(\omega_X) \backslash V^1(\omega_X) \\ \text{such that } F(1, L_2) = (1, M_2) \text{ with } \mathcal{O}_Y \neq M_2 \in V^0(\omega_Y) \backslash V^1(\omega_Y).$$

Hence, equalities $\chi(\omega_X) = \chi(\omega_X \otimes L_2) = h^0(X, \omega_X \otimes M_2) = h^0(Y, \omega_Y \otimes M_2) = \chi(\omega_Y \otimes M_2) = \chi(\omega_Y)$ hold.

The third case is when $V^0(\omega_X) \subsetneq \mathrm{Pic}^0(X)$ and $V^1(\omega_X) = \mathrm{Pic}^0(X)$. By using Proposition 3.1 and Corollary 3.4, it is easy to see that $V^0(\omega_Y) \subsetneq \mathrm{Pic}^0(Y)$ and

$V^1(\omega_Y) = \mathrm{Pic}^0(Y)$. Moreover, Remark 3.6 yields $F(1, \mathrm{Pic}^0(X)) = (1, \mathrm{Pic}^0(Y))$. Therefore, similarly to the previous cases, there exists a pair $(L_3, M_3) \neq (\mathcal{O}_X, \mathcal{O}_Y)$ such that

$$F(1, L_3) = (1, M_3) \text{ with } L_3 \notin V^0(\omega_X) \cup V^2(\omega_X) \text{ and } M_3 \notin V^0(\omega_Y) \cup V^2(\omega_Y),$$

and by Corollary 2.2, we have $\chi(\omega_X) = \chi(\omega_X \otimes L_3) = -h^1(X, \omega_X \otimes L_3) = -h^1(Y, \omega_Y \otimes M_3) = \chi(\omega_Y \otimes M_3) = \chi(\omega_Y)$.

The last case is when both $V^0(\omega_X)$ and $V^1(\omega_X)$ are proper subvarieties of $\mathrm{Pic}^0(X)$. Then $V^0(\omega_Y)$ and $V^1(\omega_Y)$ are proper subvarieties as well, and hence, $\chi(\omega_X) = \chi(\omega_Y) = 0$. □

Proof of Corollary 1.8. By the derived invariance of Hochschild homologies $HH_0(X) \cong HH_0(Y)$ and $HH_1(X) \cong HH_1(Y)$, we have $h^0(X, \omega_X) = h^0(Y, \omega_Y)$ and $h^1(X, \omega_X) = h^1(Y, \omega_Y)$. Therefore, Corollary 1.7 implies $h^{0,2}(X) = h^{0,2}(Y)$ since $h^3(X, \omega_X) = q(X) = q(Y) = h^3(Y, \omega_Y)$ and $h^4(X, \omega_X) = 1 = h^4(Y \omega_Y)$.

For the second equality, we apply Corollary 2.2 with $(L, M) = (\mathcal{O}_X, \mathcal{O}_Y)$ and $k = 2$ so that $h^2(X, \omega_X) + h^1(X, \Omega^3_X) + h^0(X, \Omega^2_X) = h^2(Y, \omega_Y) + h^1(Y, \Omega^3_Y) + h^0(Y, \Omega^2_Y)$. Therefore, we obtain $h^{1,3}(X) = h^{1,3}(Y)$ since Serre duality and the Hodge linear-conjugate isomorphism yield equalities $h^2(X, \omega_X) = h^0(X, \Omega^2_X)$ and $h^2(Y, \omega_Y) = h^0(Y, \Omega^2_Y)$. □

By using a result in [Pareschi and Popa 2011], we can also derive a consequence about pluricanonical bundles.

Corollary 7.1. *Let X and Y be smooth projective derived equivalent varieties with X of maximal Albanese dimension and minimal. Then*

$$\chi(\omega_X^{\otimes m}) = \chi(\omega_Y^{\otimes m}) \quad \textit{for all } m \geq 2.$$

Proof. By [Pareschi and Popa 2011, Corollary 5.5], $\omega_X^{\otimes m}$ and $\omega_Y^{\otimes m}$ are GV-sheaves on X and Y, respectively, for any $m \geq 2$.[1] In particular, this implies that $\operatorname{codim} V^1(\omega_X^{\otimes m}) \geq 1$ and $\operatorname{codim} V^1(\omega_Y^{\otimes m}) \geq 1$. At this point, we argue as in the first part of the proof of Corollary 1.7 after having noted the inclusions $V^0(\omega_X^{\otimes m}) \supset V^1(\omega_X^{\otimes m})$ and $V^0(\omega_Y^{\otimes m}) \supset V^1(\omega_Y^{\otimes m})$ [Pareschi and Popa 2011, Proposition 3.14]. □

7B. ***Fibrations.*** In this subsection, we study the behavior of particular types of fibrations under derived equivalence. We begin by recalling some terminology from [Catanese 1991; Lazarsfeld and Popa 2010].

A smooth projective variety X is of *Albanese general type* if it is of maximal Albanese dimension and has nonsurjective Albanese map. An *irregular fibration* or a *higher irrational pencil* is a surjective morphisms with connected fibers $f : X \to Z$

[1] The minimality condition is necessary; see [Pareschi and Popa 2011, Example 5.6].

onto a normal variety Z with $0 < \dim Z < \dim X$ and such that any smooth model of Z is of maximal Albanese dimension or Albanese general type, respectively.

Popa [2013, Corollary 3.4] observes that a consequence of Conjecture 1.3 is that, if X admits a fibration onto a variety having nonsurjective Albanese map, then any Fourier–Mukai partner of X admits an irregular fibration. With Theorem 1.4 at hand, we can verify this statement under an additional hypothesis on X.

Proposition 7.2. *Let X and Y be smooth projective derived equivalent varieties with* $\dim \mathrm{alb}_X(X) \geq \dim X - 1$. *If X admits a surjective morphism $f : X \to Z$ with connected fibers onto a normal variety Z having nonsurjective Albanese map and such that* $\dim X > \dim Z$, *then Y admits an irregular fibration.*

Proof. Let $Z \xrightarrow{f'} Z' \to \mathrm{alb}_Z(Z)$ be the Stein factorization of alb_Z. By taking a nonsingular representative of f', we can assume Z' smooth. We can easily check that Z' is of maximal Albanese dimension (so that $\mathcal{O}_{Z'} \in V^0(\omega_{Z'})$) and that $\mathrm{alb}_{Z'}$ is not surjective. Hence, by [Ein and Lazarsfeld 1997, Proposition 2.2], there exists a positive-dimensional irreducible component V of $V^0(\omega_{Z'})$ passing through the origin. Moreover, by Lemma 6.3, we have $(f \circ f')^* V \subset V^k(\omega_X)_0$ where $k = \dim X - \dim Z'$, and by (5), we get $(f \circ f')^* V \subset V^k(\omega_X)_0 \subset V^1(\omega_X)_0$. Finally, by Theorem 1.4(iii), there exists a positive-dimensional irreducible component $V' \subset V^1(\omega_Y)_0$ We conclude then by applying [Green and Lazarsfeld 1991, Theorem 0.1]. □

We point out that, thanks to Theorem 1.5, we can remove the hypothesis "$\dim \mathrm{alb}_X(X) \geq \dim X - 1$" from the above proposition in the case of threefolds. The following proposition, together with the subsequent remark, provides the proof of Corollary 1.9:

Proposition 7.3. *Let X and Y be smooth projective derived equivalent threefolds. Fix k to be either* 1 *or* 2. *Then X admits a higher irrational pencil $f : X \to Z$ with* $0 < \dim Z \leq k$ *if and only if Y admits a higher irrational pencil $g : Y \to W$ with* $0 < \dim W \leq k$.

Proof. We start with the case $k = 1$, and therefore, we consider a higher irrational pencil $f : X \to Z$ onto a smooth curve Z of genus $g(Z) \geq 2$. By Lemma 6.3, we have $f^* V^0(\omega_Z) = f^* \mathrm{Pic}^0(Z) \subset V^2(\omega_X)_0$, and by Theorem 1.5(i), there exists a component $T \subset V^2(\omega_Y)_0$ such that

$$\dim T \geq q(Z) \geq 2. \tag{10}$$

Moreover, by [Green and Lazarsfeld 1991, Theorem 0.1] or by [Beauville 1992, Corollaire 2.3], there exists an irrational fibration $g : Y \to W$ onto a smooth curve W such that $T \subset g^* \mathrm{Pic}^0(W) + \gamma$ for some $\gamma \in \mathrm{Pic}^0(Y)$. Therefore, we obtain the inequality

$$q(W) = g(W) \geq \dim T \geq 2 \tag{11}$$

ensuring that g is a higher irrational pencil.

We suppose now $k=2$, and we consider a higher irrational pencil $f : X \to Z$ onto a surface. It is a general fact that, by possibly replacing Z with a lower-dimensional variety, one can furthermore assume $\chi(\omega_{Z'}) > 0$ for any smooth model Z' of Z (see [Pareschi and Popa 2009, p. 271]). If $\dim Z = 1$, then we apply the argument of the previous case. On the other hand, if $\dim Z = 2$ then by Lemma 6.3 we get

$$f^* V^0(\omega_Z) = f^* \operatorname{Pic}^0(Z) \subset V^1(\omega_X)_0.$$

Moreover, by Theorem 1.5, there exists a component $T \subset V^1(\omega_Y)_0$ such that $\dim T \geq q(Z') \geq 3$, and by [Green and Lazarsfeld 1991, Theorem 0.1], there exists an irregular fibration $g : Y \to W$ such that $T \subset g^* \operatorname{Pic}^0(W) + \gamma$ for some $\gamma \in \operatorname{Pic}^0(Y)$. Therefore, $q(W) \geq \dim T \geq 3$ and g is a higher irrational pencil. $\square$

Remark 7.4. We can slightly improve the statement of Proposition 7.3 in the case of fibrations onto curves. In fact, by going back to the proof of Proposition 7.3 in the case $k=1$, we see that from the inequalities (10) and (11) we obtain the inequality $q(W) \geq q(Z)$. Then the following holds. Fix an integer $g \geq 2$. The variety X admits a higher irrational pencil $f : X \to C$ onto a curve of genus $g(C) \geq g$ if and only if Y admits a higher irrational pencil $h : Y \to D$ onto a curve of genus $g(D) \geq g$.

Acknowledgements

I am deeply grateful to Mihnea Popa for his insights, hints and encouragement and to Christian Schnell for suggestions regarding the proof of Theorem 3.2. I also thank Chih-Chi Chou, Lawrence Ein, Víctor González-Alonso, Emanuele Macrì, Wenbo Niu and Tuan Pham for helpful conversations. This work got started and completed while I was a graduate student at the University of Illinois at Chicago.

References

[Arapura 1992] D. Arapura, "Higgs line bundles, Green–Lazarsfeld sets, and maps of Kähler manifolds to curves", *Bull. Amer. Math. Soc.* **26**:2 (1992), 310–314. MR 92m:14010 Zbl 0759.14016

[Beauville 1992] A. Beauville, "Annulation du H^1 pour les fibrés en droites plats", pp. 1–15 in *Complex algebraic varieties* (Bayreuth, 1990), edited by K. Hulek et al., Lecture Notes in Math. **1507**, Springer, Berlin, 1992. MR 94a:14048 Zbl 0792.14006

[Beauville 1996] A. Beauville, *Complex algebraic surfaces*, 2nd ed., London Mathematical Society Student Texts **34**, Cambridge University Press, 1996. MR 97e:14045 Zbl 0849.14014

[Bridgeland and Maciocia 2001] T. Bridgeland and A. Maciocia, "Complex surfaces with equivalent derived categories", *Math. Z.* **236**:4 (2001), 677–697. MR 2002a:14017 Zbl 1081.14023

[Brion 2010] M. Brion, "Some basic results on actions of nonaffine algebraic groups", pp. 1–20 in *Symmetry and spaces*, edited by H. E. A. Campbell et al., Progr. Math. **278**, Birkhäuser, Boston, MA, 2010. MR 2011c:14128 Zbl 1217.14029

[Catanese 1991] F. Catanese, "Moduli and classification of irregular Kaehler manifolds (and algebraic varieties) with Albanese general type fibrations", *Invent. Math.* **104**:2 (1991), 263–289. MR 92f:32049 Zbl 0743.32025

[Chen and Hacon 2002] J. A. Chen and C. D. Hacon, "On algebraic fiber spaces over varieties of maximal Albanese dimension", *Duke Math. J.* **111**:1 (2002), 159–175. MR 2002m:14007 Zbl 1055.14010

[Chen and Hacon 2004] J. A. Chen and C. D. Hacon, "On the irregularity of the image of the Iitaka fibration", *Comm. Algebra* **32**:1 (2004), 203–215. MR 2005c:14005 Zbl 1137.14008

[Căldăraru 2003] A. Căldăraru, "The Mukai pairing, I: The Hochschild structure", preprint, 2003. arXiv math/0308079v2

[Ein and Lazarsfeld 1997] L. Ein and R. Lazarsfeld, "Singularities of theta divisors and the birational geometry of irregular varieties", *J. Amer. Math. Soc.* **10**:1 (1997), 243–258. MR 97d:14063 Zbl 0901.14028

[Green and Lazarsfeld 1987] M. Green and R. Lazarsfeld, "Deformation theory, generic vanishing theorems, and some conjectures of Enriques, Catanese and Beauville", *Invent. Math.* **90**:2 (1987), 389–407. MR 89b:32025 Zbl 0659.14007

[Green and Lazarsfeld 1991] M. Green and R. Lazarsfeld, "Higher obstructions to deforming cohomology groups of line bundles", *J. Amer. Math. Soc.* **4**:1 (1991), 87–103. MR 92i:32021 Zbl 0735.14004

[Hacon 2004] C. D. Hacon, "A derived category approach to generic vanishing", *J. Reine Angew. Math.* **575** (2004), 173–187. MR 2005m:14026 Zbl 1137.14012

[Hacon and Pardini 2002] C. D. Hacon and R. Pardini, "On the birational geometry of varieties of maximal Albanese dimension", *J. Reine Angew. Math.* **546** (2002), 177–199. MR 2003d:14016 Zbl 0993.14005

[Huybrechts 2006] D. Huybrechts, *Fourier–Mukai transforms in algebraic geometry*, Clarendon Press, Oxford, 2006. MR 2007f:14013 Zbl 1095.14002

[Kawamata 1981] Y. Kawamata, "Characterization of abelian varieties", *Compositio Math.* **43**:2 (1981), 253–276. MR 83j:14029 Zbl 0471.14022

[Kawamata 2002] Y. Kawamata, "D-equivalence and K-equivalence", *J. Differential Geom.* **61**:1 (2002), 147–171. MR 2004m:14025 Zbl 1056.14021

[Kollár 1986a] J. Kollár, "Higher direct images of dualizing sheaves, I", *Ann. of Math.* (2) **123**:1 (1986), 11–42. MR 87c:14038 Zbl 0598.14015

[Kollár 1986b] J. Kollár, "Higher direct images of dualizing sheaves, II", *Ann. of Math.* (2) **124**:1 (1986), 171–202. MR 87k:14014 Zbl 0605.14014

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **48**, Springer, Berlin, 2004. MR 2005k:14001a Zbl 1093.14501

[Lazarsfeld and Popa 2010] R. Lazarsfeld and M. Popa, "Derivative complex, BGG correspondence, and numerical inequalities for compact Kähler manifolds", *Invent. Math.* **182**:3 (2010), 605–633. MR 2011m:32024 Zbl 1209.32011

[Lombardi 2013] L. Lombardi, *Derived equivalences of irregular varieties and constraints on Hodge numbers*, Ph.D. thesis, University of Illinois at Chicago, 2013, http://www.math.uni-bonn.de/people/lombardi/thesis.pdf.

[Matsumura 1963] H. Matsumura, "On algebraic groups of birational transformations", *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) **34** (1963), 151–155. MR 28 #3041 Zbl 0134.16601

[Mori 1987] S. Mori, "Classification of higher-dimensional varieties", pp. 269–331 in *Algebraic geometry, Bowdoin* (Brunswick, ME, 1985), part 1, edited by S. J. Bloch, Proc. Sympos. Pure Math. **46**, Amer. Math. Soc., Providence, RI, 1987. MR 89a:14040 Zbl 0656.14022

[Orlov 2003] D. O. Orlov, "Derived categories of coherent sheaves and equivalences between them", *Uspekhi Mat. Nauk* **58**:3 (2003), 89–172. In Russian; translated in *Russian Math. Surveys* **58**:3 (2003), 511–591. MR 2004g:14021 Zbl 1118.14021

[Pareschi and Popa 2009] G. Pareschi and M. Popa, "Strong generic vanishing and a higher-dimensional Castelnuovo–de Franchis inequality", *Duke Math. J.* **150**:2 (2009), 269–285. MR 2011b:14044 Zbl 1206.14067

[Pareschi and Popa 2011] G. Pareschi and M. Popa, "GV-sheaves, Fourier–Mukai transform, and generic vanishing", *Amer. J. Math.* **133**:1 (2011), 235–271. MR 2012e:14043 Zbl 1208.14015

[Pham 2011] T. D. Pham, *On Picard varieties of surfaces with equivalent derived categories*, Ph.D. thesis, University of Illinois at Chicago, 2011, http://search.proquest.com/docview/1286775745.

[Popa 2013] M. Popa, "Derived equivalence and non-vanishing loci", pp. 567–575 in *A celebration of algebraic geometry* (Cambridge, MA, 2011), edited by B. Hassett et al., Clay Math. Proc. **18**, Amer. Math. Soc., Providence, RI, 2013. MR 3114957

[Popa and Schnell 2011] M. Popa and C. Schnell, "Derived invariance of the number of holomorphic 1-forms and vector fields", *Ann. Sci. Éc. Norm. Supér.* (4) **44**:3 (2011), 527–536. MR 2839458 Zbl 1221.14020

[Rosay 2009] F. Rosay, "Some remarks on the group of derived autoequivalences", preprint, 2009. arXiv 0907.3880

[Rouquier 2011] R. Rouquier, "Automorphismes, graduations et catégories triangulées", *J. Inst. Math. Jussieu* **10**:3 (2011), 713–751. MR 2012h:16019 Zbl 1244.16009

[Swan 1996] R. G. Swan, "Hochschild cohomology of quasiprojective schemes", *J. Pure Appl. Algebra* **110**:1 (1996), 57–80. MR 97j:19003 Zbl 0865.18010

[Toda 2006] Y. Toda, "Fourier–Mukai transforms and canonical divisors", *Compos. Math.* **142**:4 (2006), 962–982. MR 2007d:14039 Zbl 1098.18005

[Ueno 1973] K. Ueno, "Classification of algebraic varieties, I", *Compositio Math.* **27** (1973), 277–342. MR 50 #13030 Zbl 0284.14015

[Yekutieli 2003] A. Yekutieli, "The continuous Hochschild cochain complex of a scheme (survey)", pp. 227–233 in *Topics in algebraic and noncommutative geometry* (Luminy, France and Annapolis, MD, 2001), edited by C. G. Melles et al., Contemp. Math. **324**, Amer. Math. Soc., Providence, RI, 2003. MR 1986127 Zbl 1064.16008

Communicated by Yujiro Kawamata
Received 2012-09-20 Revised 2013-06-03 Accepted 2013-09-28

lombardi@math.uni-bonn.de *Mathematisches Institut, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.543

Sato–Tate distributions of twists of $y^2 = x^5 - x$ and $y^2 = x^6 + 1$

Francesc Fité and Andrew V. Sutherland

We determine the limiting distribution of the normalized Euler factors of an abelian surface A defined over a number field k when A is $\overline{\mathbb{Q}}$-isogenous to the square of an elliptic curve defined over k with complex multiplication. As an application, we prove the Sato–Tate conjecture for Jacobians of $\mathbb{Q}$-twists of the curves $y^2 = x^5 - x$ and $y^2 = x^6 + 1$, which give rise to 18 of the 34 possibilities for the Sato–Tate group of an abelian surface defined over $\mathbb{Q}$. With twists of these two curves, one encounters, in fact, all of the 18 possibilities for the Sato–Tate group of an abelian surface that is $\overline{\mathbb{Q}}$-isogenous to the square of an elliptic curve with complex multiplication. Key to these results is the *twisting Sato–Tate group* of a curve, which we introduce in order to study the effect of twisting on the Sato–Tate group of its Jacobian.

1. Introduction

Let A be an abelian variety of dimension g, defined over a number field k. The generalized *Sato–Tate conjecture* predicts that the Haar measure of a certain compact subgroup G of the unitary symplectic group $\mathrm{USp}(2g)$ governs the distribution of the normalized Euler factors $\bar{L}_{\mathfrak{p}}(A, T)$ as $\mathfrak{p}$ varies over the primes of k where A has good reduction. The normalized Euler factor at a prime $\mathfrak{p}$ is the polynomial $\bar{L}_{\mathfrak{p}}(A, T) = L_{\mathfrak{p}}(A, T/q^{1/2})$, where $q = \|\mathfrak{p}\|$ is the norm of $\mathfrak{p}$ and

$$L_{\mathfrak{p}}(A, T) = \prod_{i=1}^{2g} (1 - \alpha_i T)$$

MSC2010: primary 11M50; secondary 14K15, 14G10, 11G20, 11G10.
Keywords: Sato–Tate, hyperelliptic curves, abelian surfaces, twists.

is the *L-polynomial* of A at $\mathfrak{p}$. The polynomial $L_\mathfrak{p}(A, T)$ has the defining property that for each positive integer n,

$$\#A(\mathbb{F}_{q^n}) = \prod_{i=1}^{2g}(1-\alpha_i^n).$$

In order to give a precise statement of the Sato–Tate conjecture, we need to specify the group G and to define what it means for G to "govern" the distribution of the polynomials $L_\mathfrak{p}(A, T)$. Serre [2012] defined, in terms of ℓ-adic monodromy groups, a compact real Lie subgroup of $\mathrm{USp}(2g)$ associated to the abelian variety A, denoted by $\mathrm{ST}(A)$ and called the *Sato–Tate group* of A, satisfying the following property: for each prime $\mathfrak{p}$ at which A has good reduction, there exists a conjugacy class $s(\mathfrak{p})$ of $\mathrm{ST}(A)$ whose characteristic polynomial equals $\bar{L}_\mathfrak{p}(A, T) := \sum_{i=0}^{2g} a_i(A)(\mathfrak{p})T^i$.[1] For $i = 0, 1, \ldots, 2g$, let I_i denote the interval

$$I_i = \left[-\binom{2g}{i}, \binom{2g}{i}\right],$$

and consider the map

$$\Phi_i : \mathrm{ST}(A) \subseteq \mathrm{USp}(2g) \to I_i \subseteq \mathbb{R} \tag{1-1}$$

that sends an element of $\mathrm{ST}(A)$ to the i-th coefficient of its characteristic polynomial. Let $\mu(\mathrm{ST}(A))$ denote the Haar measure of $\mathrm{ST}(A)$ and let $\Phi_{i,*}(\mu(\mathrm{ST}(A)))$ denote its image on I_i by Φ_i. We can now state the generalized Sato–Tate conjecture.

Conjecture 1.1. *For $i = 0, 1, \ldots, 2g$, the $a_i(A)(\mathfrak{p})$ are equidistributed[2] on I_i with respect to $\Phi_{i,*}(\mu(\mathrm{ST}(A)))$.*[3]

The original Sato–Tate conjecture addresses the case where A is an elliptic curve $E/\mathbb{Q}$ without complex multiplication (CM), in which case $g = 1$ and $\mathrm{ST}(A) = \mathrm{USp}(2) = \mathrm{SU}(2)$. This case of the conjecture has recently been proved; see [Serre 2012, p. 105] for a complete list of references. For elliptic curves E/k with complex multiplication, there are two cases, depending on whether the CM field M is contained in k or not. In the former case, $\mathrm{ST}(E)$ is isomorphic to the unitary group $\mathrm{U}(1)$ (embedded in $\mathrm{SU}(2)$), and in the latter case, $\mathrm{ST}(E)$ is isomorphic to the normalizer of $\mathrm{U}(1)$ in $\mathrm{SU}(2)$. Both cases follow from classical results that we recall in Section 3B.

In all three cases arising for $g = 1$, it is easy to see that the Sato–Tate group of E is invariant under twisting: if E' is isomorphic to E over $\overline{\mathbb{Q}}$, then $\mathrm{ST}(E')$ is

[1] See also [Fité et al. 2012, §2] for a brief summary of this construction; there the Sato–Tate group of A is denoted by ST_A, rather than $\mathrm{ST}(A)$.

[2] When we make equidistribution statements, we sort primes in increasing order by norm.

[3] There is a slightly stronger form of Conjecture 1.1 which asserts that in fact the conjugacy classes $s(\mathfrak{p})$ are equidistributed with respect to the projection of $\mu(\mathrm{ST}(A))$ on the set of conjugacy classes of $\mathrm{ST}(A)$; see [Fité et al. 2012, Conjecture 1.1].

isomorphic to $\mathrm{ST}(E)$. However, when $g > 1$, this is no longer true.

In this article we study the possibilities for the Sato–Tate group of the Jacobians of twists of genus-2 curves defined over $\mathbb{Q}$ with many automorphisms (these arise for curves whose Jacobians are $\overline{\mathbb{Q}}$-isogenous to the square of an elliptic curve with complex multiplication), and to prove that in these cases Conjecture 1.1 is true.[4]

The curves we consider give rise to 18 of the 34 Sato–Tate groups that can occur for an abelian surface defined over $\mathbb{Q}$, yet they all lie in one of the two $\overline{\mathbb{Q}}$-isomorphism classes corresponding to the curves listed in the title of this article. This makes apparent the importance of understanding the effect of twisting on the Sato–Tate group.

In the remainder of this section, we describe the two points in the moduli space of genus-2 curves that are the object of our study, and state our main result (Theorem 1.4). We also describe the numerical computations used to obtain explicit examples that realize all the possibilities permitted by our main theorem.

Let us first fix some notation. Throughout this paper, $\overline{\mathbb{Q}}$ denotes a fixed algebraic closure of $\mathbb{Q}$ that is assumed to include the number field k and all of its algebraic extensions. Let $G_k = \mathrm{Gal}(\overline{\mathbb{Q}}/k)$ denote the absolute Galois group of k. For any algebraic variety X defined over k and any extension L/k, we use X_L to denote the algebraic variety defined over L obtained from X by the base change $k \hookrightarrow L$. For abelian varieties A and B defined over k, we write $A \sim B$ to indicate that there is an isogeny between A and B that is defined over k. We may write $A \sim_k B$ to emphasize the field of definition, but this is redundant (to indicate an isogeny defined over an extension L/k, we write $A_L \sim B_L$).

1A. ***Genus-2 curves with many automorphisms.*** Let C be a curve of genus $g \leq 3$ defined over k. In Section 2, we define the *twisting Sato–Tate group* $\mathrm{ST}_{\mathrm{Tw}}(C)$ of C, a compact Lie group with the property that the Sato–Tate group of the Jacobian of any twist of C is isomorphic to a subgroup of $\mathrm{ST}_{\mathrm{Tw}}(C)$. There is a well-known bijection between the set of twists of C up to k-isomorphism and the cohomology group $H^1(G_k, \mathrm{Aut}(C_{\overline{\mathbb{Q}}}))$, given by associating to a twist C' of C the class of the cocycle $\xi(\tau) := \phi({}^{\tau}\phi)^{-1}$, where ϕ is an isomorphism from $C'_{\overline{\mathbb{Q}}}$ to $C_{\overline{\mathbb{Q}}}$. Thus the group $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ is a good measure of how complicated the twists of C can be.

For the rest of Section 1, we let $k = \mathbb{Q}$ and $g = 2$. The automorphism group $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ is then one of the following seven groups:

$$\mathrm{C}_2,\ \mathrm{D}_2,\ \mathrm{D}_4,\ \mathrm{D}_6,\ \mathrm{C}_{10},\ 2\mathrm{D}_6,\ \tilde{\mathrm{S}}_4.$$

Here C_n denotes the cyclic group of n elements, D_n the dihedral group of order $2n$, and S_n the symmetric group on n letters. The groups $2\mathrm{D}_6$ and $\tilde{\mathrm{S}}_4$ are 2-coverings of

[4]Using the techniques of this article, one can obtain analogous results for genus-3 curves with many automorphisms, such as the Fermat and Klein quartics; see [Fité et al. $\geq$ 2014].

D_6 and S_4, isomorphic to $C_3 \rtimes D_4$ (with action kernel V_4) and $GL_2(\mathbb{F}_3)$, respectively. In the generic case, $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ is isomorphic to C_2. This implies that every twist C' of C is quadratic, and we have $\mathrm{ST}(\mathrm{Jac}(C')) = \mathrm{ST}(\mathrm{Jac}(C)) = \mathrm{ST}_{\mathrm{Tw}}(C)$.

We are interested in the opposite situation: the two exotic cases where $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ is as large as possible: $\tilde{S}_4$ and $2D_6$. All genus-2 curves C with $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ isomorphic to $\tilde{S}_4$ (resp. $2D_6$) are isomorphic to

$$y^2 = x^5 - x \quad (\text{resp. } y^2 = x^6 + 1), \tag{1-2}$$

and thus they constitute a single $\overline{\mathbb{Q}}$-isomorphism class $\mathscr{C}_2$ (resp. $\mathscr{C}_3$) of curves.

We shall choose representative curves C_2^0 and C_3^0 for $\mathscr{C}_2$ and $\mathscr{C}_3$ that are defined over $\mathbb{Q}$ and have particularly nice arithmetic properties. We write C^0 (resp. $\mathscr{C}$) to denote either C_2^0 or C_3^0 (resp. either $\mathscr{C}_2$ or $\mathscr{C}_3$). The key arithmetic property we require of C^0 is that its Jacobian be $\mathbb{Q}$-isogenous to E^2, where E is an elliptic curve defined over $\mathbb{Q}$ (with CM). This applies only to the curve $y^2 = x^6 + 1$ listed in (1-2), which we take as our representative C_3^0 for the class $\mathscr{C}_3$, but it also applies to the curve

$$y^2 = x^6 - 5x^4 - 5x^2 + 1, \tag{1-3}$$

which we take as a *better* representative C_2^0 for the class $\mathscr{C}_2$ of $y^2 = x^5 - x$.

The classification in [Fité et al. 2012] gives an explicit description of each of the 52 Sato–Tate groups that can and do arise in genus 2, as subgroups of USp(4), of which 32 have identity component (isomorphic to) U(1). The two curves listed in (1-2) both appear in [Fité et al. 2012], where they are shown to have Sato–Tate groups with identity component U(1). It follows that if C is a twist of either of these curves, then $\mathrm{ST}(\mathrm{Jac}(C))$ also has identity component U(1). In fact, the representative curves for all 32 of the U(1) cases listed in [Fité et al. 2012] are actually twists of one of the two curves in (1-2) (possibly using an extended field of definition).

Among the 32 genus-2 Sato–Tate groups with identity component U(1), two are maximal. The first has component group $S_4 \times C_2$ and is denoted by $J(O)$, while the second has component group $D_6 \times C_2$ and is denoted by $J(D_6)$. We will prove that $\mathrm{ST}_{\mathrm{Tw}}(C_2^0) = J(O)$ and $\mathrm{ST}_{\mathrm{Tw}}(C_3^0) = J(D_6)$, and, as a consequence, that the Sato–Tate group of any twist of C_2^0 (resp. C_3^0) is isomorphic to a subgroup of $J(O)$ (resp. $J(D_6)$). Conversely, we will show that every Sato–Tate group that can occur over $\mathbb{Q}$ and is isomorphic to a subgroup of $J(O)$ (resp. $J(D_6)$) arises for some $\mathbb{Q}$-twist C of C_2^0 (resp. C_3^0), by giving explicit examples in each case.[5] Most of the Sato–Tate groups G with identity component U(1) are actually subgroups of *both* $J(O)$ and $J(D_6)$. In such cases we exhibit $\mathbb{Q}$-twists of both C_2^0 and C_3^0 that have Sato–Tate group G.

[5]We call C a $\mathbb{Q}$-*twist* of C^0 if C is defined over $\mathbb{Q}$ and $C_{\overline{\mathbb{Q}}} \simeq C^0_{\overline{\mathbb{Q}}}$.

1B. *Main result.* Recall that C^0 denotes either C_2^0 or C_3^0. These are both genus-2 curves defined over $\mathbb{Q}$ whose Jacobians are $\mathbb{Q}$-isogenous to the square of an elliptic curve $E/\mathbb{Q}$ with CM by an imaginary quadratic field M equal to $\mathbb{Q}(\sqrt{-2})$ or $\mathbb{Q}(\sqrt{-3})$, respectively. Our main result is that Conjecture 1.1 holds for the Jacobians of the $\mathbb{Q}$-twists C of C^0.

In order to state the theorem more precisely, we introduce some notation.

Definition 1.2. For any $\mathbb{Q}$-twist C of C^0, let $K/\mathbb{Q}$ (resp. $L/\mathbb{Q}$) denote the minimal extension over which all endomorphisms of $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ (resp. homomorphisms from $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ to $E_{\overline{\mathbb{Q}}}$) are defined. Then we write $T(C)$ for the isomorphism class

$$\bigl[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)\bigr].$$

We say that two triples of groups (H_1, H_2, H_3) and (H_1', H_2', H_3') are isomorphic if $H_i \simeq H_i'$ for $i = 1, 2, 3$. We write $[H_1, H_2, H_3]$ for the isomorphism class of (H_1, H_2, H_3), which we regard as a triple of abstract groups.

Definition 1.3. For any finite group H with a subgroup H_0 and a normal subgroup N, and any positive integers r and s with $r \mid s$, let $o(s, r)$ (resp. $\bar{o}(s, r)$) count the elements in H_0 (resp. $H \setminus H_0$) of order s whose projection in H/N has order r. Let $z(H, N, H_0)$ denote the vector $[z_1, z_2]$, where

$$z_1 = \bigl[o(1,1), o(2,1), o(2,2), o(3,3), o(4,2), o(6,3), o(6,6), o(8,4), o(12,6)\bigr],$$
$$z_2 = \bigl[\bar{o}(2,2), \bar{o}(4,2), \bar{o}(6,6), \bar{o}(8,4), \bar{o}(12,6)\bigr].$$

For any $\mathbb{Q}$-twist C of C^0, write

$$z(C) := [z_1(C), z_2(C)] := z\bigl(\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(L/K), \mathrm{Gal}(L/M)\bigr).$$

We also define $o(r) = \sum_s o(s, r)$ and $\bar{o}(s) = \sum_r \bar{o}(s, r)$. We note that in the cases of interest, $z(H, N, H_0)$ is $z(C)$ for some $\mathbb{Q}$-twist C of C^0. In this situation, $o(r)$ is the number of elements in $\mathrm{Gal}(L/M)$ whose projection to $\mathrm{Gal}(K/M)$ has order r, and $\bar{o}(s)$ is the number of elements of order s in $\mathrm{Gal}(L/\mathbb{Q})$ that are not in $\mathrm{Gal}(L/M)$. Clearly

$$\sum_{r,s} o(s, r) = \sum_{r,s} \bar{o}(s, r) = \frac{|\mathrm{Gal}(L/\mathbb{Q})|}{2}.$$

Moreover, we prove in Proposition 4.9 that the only pairs (s, r) for which $o(s, r)$ or $\bar{o}(s, r)$ can be nonzero are those that appear in the vectors z_1 and z_2.

Finally, let $L_p(C, T)$ denote the Euler factor of C at a prime p of good reduction. We may write the normalized Euler factor $\bar{L}_p(C, T) = L_p(C, T/p^{1/2})$ as

$$\bar{L}_p(C, T) = T^4 + a_1(C)(p)T^3 + a_2(C)(p)T^2 + a_1(C)(p)T + 1.$$

We are now ready to state our main theorem.

Theorem 1.4. *Let C be a $\mathbb{Q}$-twist of C^0.*

(i) *There are exactly* 20 *possibilities for $T(C)$ if $C^0 = C^0_2$, and* 21 *if $C^0 = C^0_3$.*

(ii) *The triple $T(C)$ and the vector $z(C)$ uniquely determine each other.*

(iii) *The triple $T(C)$ (or $z(C)$) determines the Sato–Tate group* $\mathrm{ST}(\mathrm{Jac}(C))$.

(iv) *For $i = 1, 2$, the $a_i(C)(p)$ are equidistributed on $I_i = \bigl[-\binom{4}{i}, \binom{4}{i}\bigr]$ with respect to a measure $\mu(a_i(C))$ that is uniquely determined by the vector $z(C)$. More precisely, the density function of $\mu(a_i(C))$ is continuous up to a finite number of points, and it is therefore uniquely determined by its moments*:

$$\mathrm{M}_n[\mu(a_1(C))] = \frac{1}{[L:\mathbb{Q}]}\bigl(o(1)2^n + o(3) + o(4)2^{n/2} + o(6)3^{n/2}\bigr)b_{0,n},$$

$$\begin{aligned}\mathrm{M}_n[\mu(a_2(C))] = \frac{1}{[L:\mathbb{Q}]}\bigl(&o(1)b_{4,n} + o(2)b_{0,n} + o(3)b_{1,n} + o(4)b_{2,n} + o(6)b_{3,n}\\ &+ \bar{o}(2)2^n + \bar{o}(4)(-2)^n + \bar{o}(6)(-1)^n + \bar{o}(12)\bigr).\end{aligned}$$

Here $b_{m,n}$ denotes the coefficient[6] *of X^n in $(X^2 + mX + 1)^n$.*

(v) *Conjecture 1.1 holds for C.*

We actually prove statement (iv) in greater generality, for an abelian surface A defined over a number field k with $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$, where E is an elliptic curve defined over k with CM by a quadratic imaginary field M. This is accomplished in Section 3 via Corollary 3.12, whose proof relies on a study of the structure of $\mathrm{Hom}(E_L, A_L) \otimes_M \overline{\mathbb{Q}}$ as a Galois $\overline{\mathbb{Q}}[\mathrm{Gal}(L/M)]$-module and a refined equidistribution statement of Frobenius elements of a CM elliptic curve when restricted to certain Galois conjugacy classes (see Corollary 3.8). We compute the moments

$$\mathrm{M}_n[a_i(C)] := \lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p\le x} a_i(C)(p)^n,$$

where p varies over primes of good reduction, and prove equidistribution of the $a_i(C)(p)$ with respect to a measure $\mu(a_i(C))$. It follows that $\mathrm{M}_n[\mu(a_1(C))] = \mathrm{M}_n[a_i(C)]$. We devote Section 4 to the proofs of assertions (i), (ii), and (iii), which follow from Corollary 4.18, Proposition 4.16, and Proposition 4.17, respectively. The final assertion (v) follows from (iii) and (iv): it is enough to check that for each of the 41 possibilities of $T(C)$, the formulas obtained for $\mu[a_i(C)]$ coincide with the ones obtained for $\Phi_{i,*}(\mu(\mathrm{ST}(\mathrm{Jac}(C))))$ in [Fité et al. 2012]. In fact, it has very recently been shown that the generalized Sato–Tate conjecture (in its strong form) holds in general for abelian surfaces with potential complex multiplication; see [Johansson 2013].

[6]For $m = 0, 1, 2, 3, 4$, the $b_{m,n}$ form the sequences A126869, A0002426, A000984, A026375, A081671, respectively, in the Online Encyclopedia of Integer Sequences [OEIS 2011].

1C. ***Numerical computations.*** In Section 5, we show that all 41 of the possible triples $T(C)$ determined in section Section 4 actually arise for some $\mathbb{Q}$-twist C of C^0 by exhibiting a provable example of each case. The example curves C were obtained by an extensive search that was made feasible by part (ii) of Theorem 1.4; it is computationally much easier to approximate $z(C)$ than it is to explicitly compute $T(C)$, which requires computing the Galois groups of number fields of fairly large degree (48 or 96 in the most typical cases).

For an elliptic curve E with CM, the values $a_1(E)(p)$ can be computed very quickly, and we show how to compute $a_1(C)(p)$ and $a_2(C)(p)$ from $a_1(E)(p)$ using the fact that $\mathrm{Jac}(C)$ is $\overline{\mathbb{Q}}$-isogenous to E^2 (see Proposition 4.9). This allows us to efficiently compute an approximation of $z(C)$ (using again Proposition 4.9) of precision sufficient to provisionally identify $T(C)$ (via part (ii) of Theorem 1.4). Many curves were analyzed (tens of thousands) in order to obtain 41 candidate examples, one for each possible triple $T(C)$. For each of these 41 candidates, we then proved that the provisional identification of $T(C)$ is correct by explicitly computing the Galois groups $\mathrm{Gal}(L/\mathbb{Q})$, $\mathrm{Gal}(K/\mathbb{Q})$, and $\mathrm{Gal}(L/M)$.

2. The twisting Sato–Tate group of a curve

In this section we define the *twisting Sato–Tate group*, which is our main object of study. We do so in terms of the *algebraic Sato–Tate group* defined by Banaszak and Kedlaya [2011]. Let A be an abelian variety of dimension $g \le 3$ defined over a number field k, and fix an embedding of k into $\mathbb{C}$. Fix a polarization on A and a symplectic basis for the singular homology group $H_1(A_{\mathbb{C}}^{\mathrm{top}}, \mathbb{Q})$. Use it to equip this space with an action of $\mathrm{GSp}_{2g}(\mathbb{Q})$. For each $\tau \in G_k$, define

$$\mathrm{L}(A, \tau) := \left\{\gamma \in \mathrm{Sp}_{2g} : \gamma^{-1}\alpha\gamma = {}^{\tau}\alpha \text{ for all } \alpha \in \mathrm{End}(A_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}\right\}. \tag{2-1}$$

Here we view α as an endomorphism of $H_1(A_{\mathbb{C}}^{\mathrm{top}}, \mathbb{Q})$. The *algebraic Sato–Tate group* of A is defined by

$$\mathrm{AST}(A) := \bigcup_{\tau \in G_k} \mathrm{L}(A, \tau).$$

The Sato–Tate group $\mathrm{ST}(A)$ is a maximal compact subgroup of $\mathrm{AST}(A) \otimes_{\mathbb{Q}} \mathbb{C}$; see [Banaszak and Kedlaya 2011, Theorems 6.1 and 6.10].

Remark 2.1. As noted in the introduction, $\mathrm{ST}(A)$ is invariant under twisting when $g = 1$. This does not hold for $g > 1$; however, $\mathrm{ST}(A)$ *is* invariant under quadratic twisting. For $g \le 3$, this follows easily from the definitions above. Indeed, let $\chi : G_k \to \mathbb{C}$ be a quadratic character. For every $\tau \in G_k$, one has $\mathrm{L}(A \otimes \chi, \tau) = \mathrm{L}(A, \tau) \otimes \chi(\tau)$ (see (2-2) for a more general relation). Invariance under quadratic twisting follows from the fact that $\mathrm{L}(A, \tau) \otimes \chi(\tau) = \mathrm{L}(A, \tau)$. For A of arbitrary dimension, the invariance of $\mathrm{ST}(A)$ under quadratic twisting follows easily from the definition of $\mathrm{ST}(A)$ given in [Serre 2012] (see also [Fité et al. 2012]), in terms of the image of the ℓ-adic representation attached to A.

We now assume that A is the Jacobian $\mathrm{Jac}(C)$ of a curve C defined over k, and view $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ as a subgroup of $\mathrm{GL}(H_1(\mathrm{Jac}(C)^{\mathrm{top}}_{\mathbb{C}}, \mathbb{Q}))$.

Definition 2.2. The *twisting algebraic Sato–Tate group of* C is the algebraic subgroup of $\mathrm{Sp}_{2g}/\mathbb{Q}$ defined by

$$\mathrm{AST}_{\mathrm{Tw}}(C) := \mathrm{AST}(\mathrm{Jac}(C)) \cdot \mathrm{Aut}(C_{\overline{\mathbb{Q}}}).$$

Observe that $\mathrm{AST}_{\mathrm{Tw}}(C)$ is indeed a group: for any $\gamma_1, \gamma_2 \in \mathrm{AST}(\mathrm{Jac}(C))$ and $\alpha_1, \alpha_2 \in \mathrm{Aut}(C_{\overline{\mathbb{Q}}})$, we have

$$\gamma_1\alpha_1(\gamma_2\alpha_2)^{-1} = \gamma_1\gamma_2^{-1}\gamma_2[\alpha_1\alpha_2^{-1}]\gamma_2^{-1} = \gamma_1(\gamma_2^{-1})^{\tau_2^{-1}}(\alpha_1\alpha_2) \in \mathrm{AST}_{\mathrm{Tw}}(C).$$

We will make the notational convention that the τ_i are such that $\gamma_i \in \mathrm{L}(A, \tau_i)$ until the end of the section. Now let C' be a *twist* of C, a curve defined over k for which $C'_L \simeq C_L$ for some finite Galois extension L/k. Let $\phi : C'_L \to C_L$ be a fixed isomorphism. It is easy to check that

$$\mathrm{L}(\mathrm{Jac}(C'), \tau) = \phi^{-1}\,\mathrm{L}(\mathrm{Jac}(C), \tau)({}^{\tau}\phi). \tag{2-2}$$

Here ϕ is seen as a homomorphism from $H_1\big(\mathrm{Jac}(C')^{\mathrm{top}}_{\mathbb{C}}, \mathbb{Q}\big)$ to $H_1\big(\mathrm{Jac}(C)^{\mathrm{top}}_{\mathbb{C}}, \mathbb{Q}\big)$.

Lemma 2.3. *Let* $\gamma' \in \mathrm{L}(\mathrm{Jac}(C'), \tau) \subseteq \mathrm{AST}(\mathrm{Jac}(C'))$. *Write* γ' *as* $\phi^{-1}\gamma({}^{\tau}\phi)$ *with* γ *in* $\mathrm{L}(\mathrm{Jac}(C), \tau)$ *as in* (2-2). *The map*

$$\Lambda_\phi : \mathrm{AST}(\mathrm{Jac}(C')) \to \mathrm{AST}_{\mathrm{Tw}}(C), \quad \Lambda_\phi(\gamma') = \gamma({}^{\tau}\phi)\phi^{-1}$$

is a (*well-defined*) *monomorphism of groups.*

Proof. Let $\gamma'_1 = \phi^{-1}\gamma_1({}^{\tau_1}\phi)$ and $\gamma'_2 = \phi^{-1}\gamma_2({}^{\tau_2}\phi)$ be elements of $\mathrm{L}(\mathrm{Jac}(C'), \tau_1)$ and $\mathrm{L}(\mathrm{Jac}(C'), \tau_2)$, respectively. Then

$$\begin{aligned}
\Lambda_\phi(\gamma'_1\gamma'_2) &= \Lambda_\phi\big(\phi^{-1}\gamma_1\gamma_2\gamma_2^{-1}[({}^{\tau_1}\phi)\phi^{-1}]\gamma_2({}^{\tau_2}\phi)\big)\\
&= \Lambda_\phi\big(\phi^{-1}\gamma_1\gamma_2({}^{\tau_2\tau_1}\phi)({}^{\tau_2}\phi)^{-1}({}^{\tau_2}\phi)\big)\\
&= \Lambda_\phi\big(\phi^{-1}\gamma_1\gamma_2({}^{\tau_2\tau_1}\phi)\big) = \gamma_1\gamma_2({}^{\tau_2\tau_1}\phi)\phi^{-1}\\
&= \gamma_1\gamma_2\big[({}^{\tau_2\tau_1}\phi)({}^{\tau_2}\phi)^{-1}\big]\gamma_2^{-1}\gamma_2({}^{\tau_2}\phi)\phi^{-1}\\
&= \gamma_1({}^{\tau_1}\phi)\phi^{-1}\gamma_2({}^{\tau_2}\phi)\phi^{-1} = \Lambda_\phi(\gamma'_1)\Lambda_\phi(\gamma'_2).
\end{aligned}$$

It is clear that Λ_ϕ is both well-defined and injective: $\Lambda_\phi(\gamma'_1) = \Lambda_\phi(\gamma'_2)$ if and only if $\gamma'_1 = \gamma'_2$. □

We now define the *twisting Sato–Tate group* $\mathrm{ST}_{\mathrm{Tw}}(C)$ of C.

Definition 2.4. The *twisting Sato–Tate group* $\mathrm{ST}_{\mathrm{Tw}}(C)$ of C is a maximal compact subgroup of $\mathrm{AST}_{\mathrm{Tw}}(C) \otimes \mathbb{C}$.

Remark 2.5. It follows from the previous lemma that for any twist C' of C, the Sato–Tate group $\mathrm{ST}(\mathrm{Jac}(C))$ is isomorphic to a subgroup of $\mathrm{ST}_{\mathrm{Tw}}(C)$. We also note that the component groups of $\mathrm{ST}_{\mathrm{Tw}}(C)$ and $\mathrm{AST}_{\mathrm{Tw}}(C) \otimes \mathbb{C}$ must be isomorphic, and the identity components of $\mathrm{ST}_{\mathrm{Tw}}(C)$ and $\mathrm{ST}(\mathrm{Jac}(C))$ are equal.

Our next goal is to study the component group of $\mathrm{ST}_{\mathrm{Tw}}(C)$ when C is a hyperelliptic curve (of genus $g \leq 3$). Consider the group[7]

$$\big(\mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \rtimes \mathrm{AST}(\mathrm{Jac}(C))\big)/Z,$$

where Z is the normal subgroup of $\mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \rtimes \mathrm{AST}(\mathrm{Jac}(C))$ consisting of the pairs (α, γ) with $\alpha = \gamma$, where $\alpha \in \mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ and $\gamma \in \mathrm{AST}(\mathrm{Jac}(C))$.

Lemma 2.6. *The map*

$$\Phi : \mathrm{AST}_{\mathrm{Tw}}(C) \to \big(\mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \rtimes \mathrm{AST}(\mathrm{Jac}(C))\big)/Z, \quad \Phi(\gamma\alpha) = (\alpha^{-1}, \gamma)$$

is a (well-defined) isomorphism.

Proof. For any $\gamma_1, \gamma_2 \in \mathrm{AST}(\mathrm{Jac}(C))$ and $\alpha_1, \alpha_2 \in \mathrm{Aut}(C_{\overline{\mathbb{Q}}})$, we have

$$\begin{aligned}\Phi(\gamma_1\alpha_1\gamma_2\alpha_2) &= \Phi\big(\gamma_1\gamma_2({}^{\tau_2}\alpha_1)\alpha_2\big) = \big(\alpha_2^{-1}({}^{\tau_2}\alpha_1)^{-1}, \gamma_1\gamma_2\big)\\ &= (\alpha_1^{-1}, \gamma_1)(\alpha_2^{-1}, \gamma_2) = \Phi(\gamma_1\alpha_1)\Phi(\alpha_2\gamma_2).\end{aligned}$$

The surjectivity of Φ is clear. It remains to prove that $\Phi(\gamma_1\alpha_1) = \Phi(\gamma_2\alpha_2)$ if and only if $\gamma_1\alpha_1 = \gamma_2\alpha_2$. On the one hand, $\Phi(\gamma_1\alpha_1) = \Phi(\gamma_2\alpha_2)$ if and only if

$$Z \ni (\alpha_2^{-1}, \gamma_2)(\alpha_1^{-1}, \gamma_1)^{-1} = \big({}^{\tau_1^{-1}}(\alpha_1\alpha_2^{-1}), \gamma_2\gamma_1^{-1}\big).$$

On the other hand, $\gamma_1\alpha_1 = \gamma_2\alpha_2$ if and only if $\alpha_1\alpha_2^{-1} = \gamma_1^{-1}\gamma_2$, or equivalently, ${}^{\tau_1^{-1}}(\alpha_1\alpha_2^{-1}) = \gamma_2\gamma_1^{-1}$. But then $\big({}^{\tau_1^{-1}}(\alpha_1\alpha_2^{-1}), \gamma_2\gamma_1^{-1}\big) \in Z$. □

We now assume C is a hyperelliptic curve (of genus $g \leq 3$). As an endomorphism of $H_1\big(\mathrm{Jac}(C)^{\mathrm{top}}_{\mathbb{C}}, \mathbb{Q}\big)$, the hyperelliptic involution w of C corresponds to the matrix $-1 \in \mathrm{Sp}_{2g}(\mathbb{Q})$. Recall that $\mathrm{AST}(\mathrm{Jac}(C))$ contains the matrix -1. Thus $(-1, -1) \in Z$, and Lemma 2.6 implies that $\mathrm{AST}_{\mathrm{Tw}}(C)$ is isomorphic to a subgroup of

$$\big(\mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \rtimes \mathrm{AST}_{\mathrm{Jac}(C)}\big)/\langle(-1, -1)\rangle.$$

Let K/k denote the minimal field extension over which all the endomorphisms of $\mathrm{Jac}(C)$ are defined. Then, since the component group of $\mathrm{ST}(\mathrm{Jac}(C))$ is isomorphic to $\mathrm{Gal}(K/k)$ (see [Banaszak and Kedlaya 2011, Remark 6.4, Theorem 6.10]), and

[7]The product of elements (α_1, γ_1) and (α_2, γ_2) in $\mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \rtimes \mathrm{AST}(\mathrm{Jac}(C))$ is defined to be $(\alpha_2\gamma_2^{-1}\alpha_1\gamma_2, \gamma_1\gamma_2) = (\alpha_2 \cdot {}^{\tau_2}\alpha_1, \gamma_1\gamma_2)$, where $\gamma_2 \in \mathrm{L}(A, \tau_2) \subset \mathrm{AST}(\mathrm{Jac}(C))$.

the identity component of $\mathrm{ST}(\mathrm{Jac}(C))$ contains the matrix -1, the component group of $\mathrm{ST}_{\mathrm{Tw}}(C)$ is isomorphic to a subgroup of

$$\mathrm{Aut}(C_{\overline{\mathbb{Q}}})/\langle w\rangle \rtimes \mathrm{Gal}(K/k).$$

By Lemma 2.3, for any twist C' of C, there exists a monomorphism of groups

$$\overline{\lambda}_{\phi} : \mathrm{Gal}(K/k) \to \mathrm{Aut}(C_{\overline{\mathbb{Q}}})/\langle w\rangle \rtimes \mathrm{Gal}(K/k). \tag{2-3}$$

It follows that if there exists a twist $\tilde{C}$ of C such that

$$|\mathrm{Gal}(\tilde{K}/k)| = |\mathrm{Aut}(C_{\overline{\mathbb{Q}}})| \cdot |\mathrm{Gal}(K/k)|/2, \tag{2-4}$$

where $\tilde{K}/k$ is the minimal extension over which all the endomorphisms of $\mathrm{Jac}(\tilde{C})$ are defined, then $\mathrm{ST}_{\mathrm{Tw}}(C) = \mathrm{ST}(\mathrm{Jac}(\tilde{C}))$, and for every twist C' of C, the Sato–Tate group $\mathrm{ST}(\mathrm{Jac}(C'))$ is a subgroup of $\mathrm{ST}(\mathrm{Jac}(\tilde{C}))$.

Remark 2.7. Let C_2^0 and C_3^0 be the two curves defined in Section 1A. If $\tilde{C}$ is a twist of C_2^0 (resp. C_3^0) such that $\mathrm{ST}(\tilde{C}) = J(O)$ (resp. $J(D_6)$), then (2-4) is satisfied. It follows that $\mathrm{ST}_{\mathrm{Tw}}(C_2^0) = J(O)$ and $\mathrm{ST}_{\mathrm{Tw}}(C_3^0) = J(D_6)$.

3. Squares of CM elliptic curves

We shall work in the category of abelian varieties up to isogeny, so we call the elements of $\mathrm{Hom}(A, B) \otimes \mathbb{Q}$ homomorphisms, the elements of $\mathrm{End}(A) \otimes \mathbb{Q}$ endomorphisms, and the surjective elements in $\mathrm{Hom}(A, B) \otimes \mathbb{Q}$ isogenies.

We henceforth assume that A is an abelian variety over k such that $A_{\overline{\mathbb{Q}}} \sim E_{\overline{\mathbb{Q}}}^2$, where E is an elliptic curve defined over k with CM by an imaginary quadratic field M (except in Section 3D, where we do not assume E has CM). Let L/k be the minimal extension over which all the homomorphisms from $E_{\overline{\mathbb{Q}}}$ to $A_{\overline{\mathbb{Q}}}$ are defined, and let K/k be the minimal extension over which all the endomorphisms of $A_{\overline{\mathbb{Q}}}$ are defined. We note that $kM \subseteq K \subseteq L$, and we have $\mathrm{Hom}(E_{\overline{\mathbb{Q}}}, A_{\overline{\mathbb{Q}}}) \simeq \mathrm{Hom}(E_L, A_L)$ and $A_L \sim E_L^2$.

3A. *The Galois modules* $\mathbf{Hom}(E_L, A_L)$ *and* $\mathbf{End}(A_L)$.

Let σ and $\overline{\sigma}$ denote the two embeddings of M into $\overline{\mathbb{Q}}$. Consider

$$\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}} \quad \text{and} \quad \mathrm{End}(A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}},$$

where the tensor products are taken via the embedding $\sigma : M \hookrightarrow \overline{\mathbb{Q}}$. If we let $\mathrm{Gal}(L/kM)$ act trivially on $\overline{\mathbb{Q}}$ and naturally on $\mathrm{Hom}(E_L, A_L)$, these products become $\overline{\mathbb{Q}}[\mathrm{Gal}(L/kM)]$-modules of dimensions 2 and 4, respectively, over $\overline{\mathbb{Q}}$, and similarly for $\overline{\sigma}$.

Definition 3.1. Let $\theta := \theta_{M,\sigma}(E, A)$ (resp. $\theta_{M,\sigma}(A)$) denote the representation afforded by the module $\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$ (resp. $\mathrm{End}(A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$), and similarly define $\overline{\theta} := \theta_{M,\overline{\sigma}}(E, A)$ and $\theta_{M,\overline{\sigma}}(A)$. Let $\theta_{\mathbb{Q}} := \theta_{\mathbb{Q}}(E, A)$ (resp. $\theta_{\mathbb{Q}}(A)$) denote the representation afforded by the $\mathbb{Q}[\mathrm{Gal}(L/k)]$-module $\mathrm{Hom}(E_L, A_L) \otimes \mathbb{Q}$ (resp. $\mathrm{End}(A_L) \otimes \mathbb{Q}$).

For each $\tau \in \mathrm{Gal}(L/kM)$, we write

$$\det(1 - \theta(\tau)T) = 1 + a_1(\theta)(\tau)T + a_2(\theta)(\tau)T^2,$$

where $a_1(\theta) = \mathrm{Tr}\,\theta$ and $a_2(\theta) = \det(\theta)$ are elements of M. Observe that

$$\mathrm{Tr}\,\theta_{\mathbb{Q}}(\tau) = \mathrm{Tr}_{M/\mathbb{Q}}\,\mathrm{Tr}\,\theta(\tau) \quad \text{if } \tau \in \mathrm{Gal}(L/kM). \tag{3-1}$$

For $z \in M$, let $|z| := \sqrt{\sigma(z)\overline{\sigma}(z)}$.

Proposition 3.2. *There is an isomorphism of* $\overline{\mathbb{Q}}[\mathrm{Gal}(L/kM)]$-*modules*

$$\mathrm{End}(A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}} \simeq \big(\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}\big)^* \otimes \mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}.$$

Thus $\mathrm{Tr}\,\theta_{M,\sigma}(A) = \mathrm{Tr}\,\theta_{M,\overline{\sigma}}(E, A) \cdot \mathrm{Tr}\,\theta_{M,\sigma}(E, A) = |\,\mathrm{Tr}(\theta)|^2 \in \mathbb{Q}$, *and therefore* $\theta_{M,\sigma}(A) \simeq \theta_{M,\overline{\sigma}}(A)$.

Proof. Consider the natural inclusion of $\overline{\mathbb{Q}}[\mathrm{Gal}(L/kM)]$-modules

$$\mathrm{End}(A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}} \hookrightarrow \mathrm{Hom}_{\overline{\mathbb{Q}}}\big(\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}, \mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}\big),$$

which sends an element ψ in $\mathrm{End}(A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$ to the linear map of $\overline{\mathbb{Q}}$-vector spaces that sends f in $\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$ to $\psi \circ f$ in $\mathrm{Hom}(E_L, A_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$. Both spaces have dimension 4 over $\overline{\mathbb{Q}}$, and thus must be isomorphic as $\overline{\mathbb{Q}}[\mathrm{Gal}(L/kM)]$-modules. □

Let $\pi : \mathrm{Gal}(L/kM) \to \mathrm{Gal}(K/kM)$ be the natural projection. For each τ in $\mathrm{Gal}(L/kM)$, let $s = s(\tau)$ denote the order of τ and let $r = r(\tau)$ denote the order of $\pi(\tau)$ in $\mathrm{Gal}(K/kM)$. The possible values of r are 1, 2, 3, 4, and 6; see [Fité et al. 2012, §4.5].

Proposition 3.3. *Suppose* $\tau \in \mathrm{Gal}(L/k)$ *does not lie in* $\mathrm{Gal}(L/kM)$. *Then the eigenvalues of* $\theta_{\mathbb{Q}}(E, A)(\tau)$ *are as follows*:

$$\begin{array}{ll} s = 2\colon\ -1, -1, 1, 1 & s = 8\colon\ \zeta_8, \zeta_8^3, \zeta_8^5, \zeta_8^7 \\ s = 4\colon\ i, i, -i, -i & s = 12\colon\ \zeta_{12}, \zeta_{12}^5, \zeta_{12}^7, \zeta_{12}^{11} \\ s = 6\colon\ \zeta_3, \zeta_3^2, \zeta_6, \zeta_6^5 & \end{array}$$

Here, ζ_r *stands for an* r-*th root of unity.*

Proof. We can assume that kM/k is quadratic; otherwise there is nothing to prove. We first show the following properties of $\theta_{\mathbb{Q}}(E, A)$:

(i) The least common multiple of the orders of the eigenvalues of $\theta_{\mathbb{Q}}(E, A)(\tau)$ is equal to s.

(ii) If $\tau \in \mathrm{Gal}(L/k) \setminus \mathrm{Gal}(L/kM)$, then $\mathrm{Tr}\,\theta_{\mathbb{Q}}(E, A)(\tau) = 0$.

It follows from the definition of L/k that the representation $\theta_{\mathbb{Q}}(E, A)$ is faithful, which implies (i). Let χ be the quadratic character of $\mathrm{Gal}(L/k)$ associated to the quadratic extension kM/k. Then $E \otimes \chi \sim_k E$ (and, in fact, $A \otimes \chi \sim_k A$), which implies that $\mathrm{Hom}(E_L, A_L) = \mathrm{Hom}(E_L, A_L) \otimes \chi$ (by [Mazur et al. 2007, Proposition 1.6], for example). This proves (ii).

For $s = 2, 6, 8, 12$, the proposition follows from (i) and (ii). For $s = 4$, (i) implies that i is an eigenvalue of $\theta_{\mathbb{Q}}(E, A)(\tau)$, and (ii) leaves just two possibilities for the four eigenvalues: $i, -i, 1, -1$, or $i, -i, i, -i$. We now show that only the latter can arise. The eigenvalues of $\theta_{\mathbb{Q}}(E, A)(\tau)$ are quotients of roots of $\bar{L}_{\mathfrak{p}}(E, T)$ and roots of $\bar{L}_{\mathfrak{p}}(A, T)$, where $\mathfrak{p}$ is a prime of k, inert in kM, of good reduction for A and E. We can further assume that $\mathfrak{p}$ has absolute degree 1. Then $\bar{L}_{\mathfrak{p}}(E, T) = 1 + T^2$, and the polynomial $\bar{L}_{\mathfrak{p}}(A, T)$ is one of the following:

$$(1 - T^2)^2, \quad 1 - T^2 + T^4, \quad 1 + T^4, \quad 1 + T^2 + T^4, \quad (1 + T^2)^2. \tag{3-2}$$

In no case can both 1 and i arise as quotients of a root of $\bar{L}_{\mathfrak{p}}(E, T) = 1 + T^2$ and roots of $\bar{L}_{\mathfrak{p}}(A, T)$. □

In view of Proposition 3.2, we write $\theta_M(A)$ for $\theta_{M,\sigma}(A) \simeq \theta_{M,\bar{\sigma}}(A)$.

Proposition 3.4. *For each $\tau \in \mathrm{Gal}(L/kM)$, we have*

$$\mathrm{Tr}\,\theta_M(A)(\tau) = 2 + \zeta_r + \bar{\zeta}_r.$$

Proof. It follows from [Fité et al. 2012, Proposition 9] that the eigenvalues of $\theta_{\mathbb{Q}}(A)(\tau)$ are $1, 1, 1, 1, \zeta_r, \zeta_r, \bar{\zeta}_r, \bar{\zeta}_r$. Equation (3-1) leaves three possibilities for the eigenvalues of $\theta_M(A)$: they must be either $1, 1, \zeta_r, \bar{\zeta}_r$, or $1, 1, \zeta_r, \zeta_r$, or $1, 1, \bar{\zeta}_r, \bar{\zeta}_r$. By Proposition 3.2, $\mathrm{Tr}\,\theta_M(A)$ is rational, so only the first possibility can occur. □

3B. ***Equidistribution for Frobenius conjugacy classes.*** We first recall the well-known notion of equidistribution on a compact topological space X (see [Serre 1998, Chapter 1]). Let $\mathscr{C}(X)$ denote the Banach space of continuous, complex valued functions f on X, with norm $\|f\| = \sup_{x \in X} |f(x)|$. Let μ be a Radon measure on X, a continuous linear form on $\mathscr{C}(X)$. Let $\{x_i\}_{i \geq 1}$ be a sequence of points of X. The sequence $\{x_i\}_{i \geq 1}$ is said to be equidistributed with respect to μ if for every $f \in \mathscr{C}(X)$, we have

$$\mu(f) = \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} f(x_i).$$

Note that if $\{x_i\}_{i \geq 1}$ is equidistributed with respect to μ, then μ is positive and has total mass 1. We are particularly interested in the case where X is an interval I of $\mathbb{R}$. In this case, the n-th moment $\mathrm{M}_n[\mu]$ of μ is the value $\mu(\varphi_n)$, where φ_n is the function of $\mathscr{C}(I)$ defined by $\varphi_n(z) = z^n$. Analogously, the n-th moment of a sequence $\{x_i\}_{i\geq 1}$ on I, if it exists, is defined by

$$\mathrm{M}_n[\{x_i\}_{i\geq 1}] = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} x_i^n.$$

Thus if the sequence $\{x_i\}_{i\geq 1}$ is equidistributed with respect to μ on I, then its n-th moment exists and is equal to $\mathrm{M}_n[\mu]$.

Let F/k be a field extension, and let P_{E_F} denote the set of primes of F at which the elliptic curve E_F has good reduction. We write the normalized L-polynomial for E_F at a prime $\mathfrak{p}$ of P_{E_F} as

$$\bar{L}_{\mathfrak{p}}(E_F, T) = 1 + a_1(E_F)(\mathfrak{p})T + T^2.$$

Choose an *ordering by norm* $\{\mathfrak{p}_i\}_{i\geq 1}$ of P_{E_F}, that is, an ordering for which $\|\mathfrak{p}\|_i \leq \|\mathfrak{p}\|_j$ for all $1 \leq i \leq j$, and let $a_1(E_F)$ denote the sequence

$$\{a_1(E_F)(\mathfrak{p}_i)\}_{i\geq 1}$$

of real numbers in the interval $[-2, 2]$. Equidistribution statements about $a_1(E_F)$ do not depend on the particular ordering by norm we have chosen.

Until the end of this section, we assume that F contains kM. We begin by recalling classical results of Hecke and Deuring that yield equidistribution for $a_1(E_F)$ with respect to the measure

$$\mu_{\mathrm{cm}} = \frac{1}{\pi} \frac{dz}{\sqrt{4 - z^2}},$$

supported on $[2, -2]$. Here dz denotes the restriction of the Lebesgue measure on $\mathbb{R}$ to the interval $[-2, 2]$. The measure μ_{cm} is uniquely characterized by the fact that it is continuous and its n-th moment is $b_n := b_{0,n}$ (as in Theorem 1.4).

We actually require a slightly stronger equidistribution statement than the one above. Let c be a Frobenius conjugacy class of an arbitrary finite Galois extension F'/F, and let P_c denote the set of primes in P_{E_F} that are unramified in F' and whose Frobenius conjugacy class is c. We will show that the subsequence $a_{1,c}(E_F)$ of $a_1(E_F)$ obtained by restricting to the primes in P_c is also equidistributed with respect to μ_{cm}.

Remark 3.5. Henceforth, for a compact group G, let $\mu(G)$ denote its Haar measure. In terms of the (generalized) Sato–Tate conjecture, the measure μ_{cm} is seen as $\Phi_{1,*}(\mu(\mathrm{ST}(E_F)))$, where Φ_1 is the trace map defined in (1-1) and $\mathrm{ST}(E_F) = \mathrm{U}(1)$. Recall that the Sato–Tate group $\mathrm{ST}(E)$ of an elliptic curve E defined over k with CM by M is $\mathrm{U}(1)$ (embedded in $\mathrm{SU}(2)$) if M is contained in k, and the normalizer of $\mathrm{U}(1)$ in $\mathrm{SU}(2)$ if M is not contained in k.

We follow the presentation in [Gross 1980, Chapter 1]. Let $\mathfrak{p}$ be a prime of F of good reduction for E_F. Let $\overline{\mathbb{F}}_{\mathfrak{p}}$ denote the algebraic closure of the residue field of F at $\mathfrak{p}$. The image of the injection

$$\operatorname{End}(E_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q} = M \hookrightarrow \operatorname{End}(E_{\overline{\mathbb{F}}_{\mathfrak{p}}}) \otimes \mathbb{Q}$$

contains the Frobenius endomorphism $\mathrm{Fr}_{\mathfrak{p}} : E_{\overline{\mathbb{F}}_{\mathfrak{p}}} \to E_{\overline{\mathbb{F}}_{\mathfrak{p}}}$, which acts on a point by raising its coordinates to the q-th power, where $q = \|\mathfrak{p}\|$. Let $\alpha(\mathfrak{p}) := \alpha(E_F)(\mathfrak{p}) \in M^*$ denote the preimage of $\mathrm{Fr}_{\mathfrak{p}}$ under this injection. Since the characteristic polynomial of $\mathrm{Fr}_{\mathfrak{p}}$ is reciprocal to the L-polynomial of E_F at $\mathfrak{p}$, we have

$$a_1(E_F)(\mathfrak{p}) = -\frac{1}{\|\mathfrak{p}\|^{1/2}}\big(\sigma(\alpha(\mathfrak{p})) + \overline{\sigma}(\alpha(\mathfrak{p}))\big). \tag{3-3}$$

For any place v of F, let F_v denote the completion of F at v and let $\mathbb{O}_v$ denote the ring of integers of F_v. Let $I_F = \prod_v' F_v$ denote the group of idèles of F. Here the product runs over all places v of F, and the prime means that if $\mathfrak{s} = (\mathfrak{s}_v)$ belongs to I_F, then $\mathfrak{s}_v$ is in $\mathbb{O}_v^*$ for all but finitely many v. We write $v_{\mathfrak{p}}$ for the valuation associated to a finite prime $\mathfrak{p}$ of F. We then attach to E_F the group homomorphism

$$\chi_{E_F} : I_F \to M^*$$

uniquely characterized by the following three properties:

(i) $\operatorname{Ker}(\chi_{E_F})$ is an open subgroup of I_F.

(ii) If $\mathfrak{s} = (a)$ is a principal idèle ($a \in F^*$), then $\chi_{E_F}(\mathfrak{s}) = N_{F/M}(a)$.

(iii) If $\mathfrak{s} = (\mathfrak{s}_v)$ is an idèle with $\mathfrak{s}_v = 1$ at all infinite places of F and at those finite places where E_F has bad reduction, then

$$\chi_{E_F}(\mathfrak{s}) = \prod_{v_{\mathfrak{p}}} \alpha(\mathfrak{p})^{v_{\mathfrak{p}}(\mathfrak{s}_{\mathfrak{p}})}.$$

3B1. *The* 1*-dimensional* ℓ*-adic representation attached to* E_F. Fix a prime ℓ different from the characteristic of $\overline{\mathbb{F}}_{\mathfrak{p}}$ and an embedding of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_\ell$, and let $V_\ell(E_F)$ denote the (rational) ℓ-adic Tate module of E_F. Define

$$V_\sigma(E) := V_\ell(E_F) \otimes_{M,\sigma} \overline{\mathbb{Q}}_\ell, \tag{3-4}$$

where the tensor product is taken via the embedding $M \hookrightarrow \overline{\mathbb{Q}}_\ell$ induced by σ. Similarly define $V_{\overline{\sigma}}(E)$. We then have an isomorphism of $\overline{\mathbb{Q}}_\ell[G_F]$-modules:

$$V_\ell(E_F) \otimes \overline{\mathbb{Q}}_\ell \simeq V_\sigma(E) \oplus V_{\overline{\sigma}}(E). \tag{3-5}$$

Let $\varrho_{\ell,\sigma} : G_F \to \operatorname{Aut}(V_\sigma(E))$ denote the ℓ-adic character corresponding to the action of G_F on $V_\sigma(E)$. If $\mathrm{Frob}_{\mathfrak{p}}$ is an arithmetic Frobenius at $\mathfrak{p}$ in G_F, then the value of $\varrho_{\ell,\sigma}(\mathrm{Frob}_\mathfrak{p})$ is $\sigma(\alpha(\mathfrak{p}))$. Define

$$\psi_{\ell,\sigma} : I_F \to \big(M \otimes_{M,\sigma} \overline{\mathbb{Q}}_\ell\big)^*, \quad \psi_{\ell,\sigma}(\mathfrak{s}) = \chi_{E_F}(\mathfrak{s}) \otimes \big(N_{F/M}(\mathfrak{s}^{-1})\big)_\ell,$$

where for an idèle $\mathfrak{s}$ in I_F, the component of the idèle $N_{F/M}(\mathfrak{s})$ in I_M corresponding to the place w is $\prod_{v|w} N_{F_v/M_w}(\mathfrak{s}_v)$, where the product runs over all places v of F lying over w. We then have $\psi_{\ell,\sigma}(F^*) = 1$, by property (ii). Thus $\psi_{\ell,\sigma}$ is a continuous character on the group $C_F = I_F/F^*$ of classes of idèles. Since its image is totally disconnected, it is a character of C_F/C_F^0, where C_F^0 is the identity component of C_F. Artin reciprocity yields an isomorphism $\mathrm{Rec} : G_F^{\mathrm{ab}} \to C_F/C_F^0$. Property (iii) then implies that $\psi_{\ell,\sigma} \circ \mathrm{Rec}(\mathrm{Frob}_\mathfrak{p}) = \sigma(\alpha(\mathfrak{p}))$, and thus

$$\psi_{\ell,\sigma} \circ \mathrm{Rec}(\mathrm{Frob}_\mathfrak{p}) = \varrho_{\ell,\sigma} \tag{3-6}$$

as ℓ-adic characters of G_F.

3B2. *The Hecke character attached to E_F.* A *Hecke character* of F is a continuous homomorphism $\psi : I_F \to \mathbb{C}^*$ such that $\psi(F^*) = 1$. For primes $\mathfrak{p}$ where ψ is unramified, let $\psi(\mathfrak{p})$ denote $\psi(\mathfrak{s})$, where $\mathfrak{s}_\mathfrak{p}$ is a uniformizer of $\mathcal{O}_\mathfrak{p}$ and $\mathfrak{s}_v = 1$ for $v \neq \mathfrak{p}$, and let $\psi(\mathfrak{p}) = 0$ when ψ is ramified at $\mathfrak{p}$. The L-function of ψ is defined as

$$L(\psi, s) := \prod_\mathfrak{p} \big(1 - \psi(\mathfrak{p}) \|\mathfrak{p}\|^{-s}\big)^{-1}.$$

Hecke [1920] showed that if ψ is nontrivial and takes values in $\mathrm{U}(1)$, then $L(\psi, s)$ is a nonzero holomorphic function for $\Re(s) \geq 1$. Let us fix an embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}$, so that we may view σ and $\overline{\sigma}$ as embeddings of M into $\mathbb{C}$. Define

$$\psi_{\infty,\sigma} : I_F \to \big(M \otimes_{M,\sigma} \mathbb{C}\big)^*, \quad \psi_{\infty,\sigma}(\mathfrak{s}) = \chi_{E_F}(\mathfrak{s}) \otimes \big(N_{F/M}(\mathfrak{s}^{-1})\big)_\infty,$$

where ∞ denotes the only infinite place of M. Property (ii) of χ_{E_F} implies that $\psi_{\infty,\sigma}$ is a Hecke character. It is unramified at the primes of good reduction for E_F, and we note that $\overline{\psi}_{\infty,\sigma} = \psi_{\infty,\overline{\sigma}}$. Let $|z|$ denote the absolute value of a complex number z and define

$$\psi^1_{\infty,\sigma} : I_F \to \mathrm{U}(1), \quad \psi^1_{\infty,\sigma}(\mathfrak{s}) = \psi_{\infty,\sigma}(\mathfrak{s})/|\psi_{\infty,\sigma}(\mathfrak{s})|.$$

For every prime $\mathfrak{p}$ of good reduction for E_F, let

$$\alpha_1(\mathfrak{p}) := \alpha_1(E_F)(\mathfrak{p}) := \psi^1_{\infty,\sigma} \circ \mathrm{Rec}(\mathrm{Frob}_\mathfrak{p}) = \sigma(\alpha(\mathfrak{p}))/\|\mathfrak{p}\|^{1/2}. \tag{3-7}$$

Let α_1 denote the sequence $\{\alpha_1(\mathfrak{p}_i)\}_{i\geq 1}$.

3B3. *Equidistribution statements.* For a finite Galois extension F'/F and a conjugacy class c of $\mathrm{Gal}(F'/F)$, let P_c be as above. Let $\alpha_{1,c} := \alpha_{1,c}(E_F)$ denote the subsequence of α_1 obtained by restricting to the primes of P_c. Our goal is to prove the following proposition.

Proposition 3.6. *Let c be any conjugacy class of* $\mathrm{Gal}(F'/F)$. *Then $\alpha_{1,c}$ is equidistributed with respect to* $\mu(\mathrm{U}(1))$.

We first recall a theorem of Serre. Let G be a compact group and X the set of its conjugacy classes. Let P be an infinite subset of the primes of F, and let $\{\mathfrak{p}_i\}_{i\geq 1}$ be an ordering by norm of P. Assume that each prime $\mathfrak{p}$ in P has been assigned a corresponding element $x_{\mathfrak{p}}$ in X.

Theorem 3.7 [Serre 1998, p. I-23]. *The sequence $\{x_{\mathfrak{p}_i}\}_{i\geq 1}$ is equidistributed over X with respect to the image on X of the Haar measure of G if and only if $L(\varrho, s)$ is holomorphic and nonzero for $\Re(s) \geq 1$ for every irreducible and nontrivial representation ϱ of G. Here $L(\varrho, s)$ stands for the infinite product*

$$\prod_{\mathfrak{p}\in P} \det\bigl(1 - \varrho(x_{\mathfrak{p}})\|\mathfrak{p}\|^{-s}\bigr)^{-1}.$$

We now use Theorem 3.7 to prove Proposition 3.6.

Proof. We first reduce to the case that F'/F is abelian (in fact, cyclic). Let τ be an element of c, and let f denote its order. Define

$$I(\tau) = \bigl\{i \in \{0, 1, \dots, f-1\} \mid [\tau^i] = c\bigr\}.$$

Let H be the subfield of F' fixed by $\langle\tau\rangle$. The residue degree over F of a prime $\mathfrak{P}$ of H lying over $\mathfrak{p} \in P_c$ is 1, and thus $\alpha_1(E_H)(\mathfrak{P}) = \alpha_1(E_F)(\mathfrak{p})$. Then $\alpha_{1,c}(E_F)$ is the disjoint union

$$\bigsqcup_{i\in I(\tau)} \alpha_{1,\tau^i}(E_H),$$

where we identify τ^i with its conjugacy class in the cyclic group $\mathrm{Gal}(F'/H)$. To show that $\alpha_{1,c} = \alpha_{1,c}(E_F)$ is $\mu(\mathrm{U}(1))$-equidistributed, it suffices to show that all its subsequences $\alpha_{1,\tau^i}(E_H)$ are (any sequence that can be partitioned into a finite set of subsequences that are all equidistributed with respect to a fixed common measure is clearly equidistributed with respect to the same measure), and if we assume the proposition holds for abelian extensions, then this is true.

So suppose that F'/F is abelian, and define $G := \mathrm{U}(1) \times \mathrm{Gal}(F'/F)$ and $x_{\mathfrak{p}} := \alpha_1(\mathfrak{p}) \times \mathrm{Frob}_{\mathfrak{p}}$ for each prime in P_{E_F} unramified in F'/F. Since for such a prime, $x_{\mathfrak{p}} \in \mathrm{U}(1) \times \{c\}$ if and only if $\mathfrak{p} \in P_c$, proving the proposition is equivalent to showing that $\{x_{\mathfrak{p}_i}\}_{i\geq 1}$ is equidistributed over the set X of conjugacy classes of G with respect to the measure induced by the Haar measure of G. The irreducible characters of G are of the form $\phi_a \otimes \chi$, where $\phi_a : \mathrm{U}(1) \to \mathbb{C}^*$ is a character of $\mathrm{U}(1)$, which is of the form $\phi_a(z) = z^a$ for some integer a, and χ is an irreducible character of $\mathrm{Gal}(F'/F)$, which is 1-dimensional since $\mathrm{Gal}(F'/F)$ is abelian. By

Theorem 3.7, it is enough to show that if $\phi_a \otimes \chi$ is nontrivial, then

$$L(\phi_a \otimes \chi, s) = \prod_{\mathfrak{p}} \bigl(1 - \psi^1_{\infty,\sigma}(\mathfrak{p})^a \chi(\mathfrak{p}) \|\mathfrak{p}\|^{-s}\bigr)^{-1}$$

is holomorphic and nonzero for $\Re(s) \geq 1$. Via Artin reciprocity, we may view $(\psi^1_{\infty,\sigma})^a \otimes \chi$ as a Hecke character (with values in U(1)), and then $L(\phi_a \otimes \chi, s)$ is equal, up to a finite number of factors, to the Hecke L-function $L((\psi^1_{\infty,\sigma})^a \otimes \chi, s)$, which is holomorphic and nonzero for $\Re(s) \geq 1$. □

Recalling that

$$\mu_{\mathrm{cm}} = \frac{1}{\pi} \frac{dz}{\sqrt{4 - z^2}}$$

supported on $[-2, 2]$ is the image by Φ_1 of the Haar measure of U(1), we obtain the following.

Corollary 3.8. *Let E be an elliptic curve defined over k with CM by an imaginary quadratic field M. Let F be any field containing kM, let F'/F be a finite Galois extension, and let c be a conjugacy class of* $\mathrm{Gal}(F'/F)$. *Then*:

(i) *The sequence $a_{1,c}(E_F)$ is equidistributed with respect to the measure μ_{cm}.*

(ii) $\mathrm{M}_n[a_{1,c}(E_F)] = \mathrm{M}_n[a_1(E_F)]$.

3C. ***Equidistribution of $a_1(A)$ and $a_2(A)$.*** As in Section 3A, A is an abelian surface defined over k with $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$, where E is an elliptic curve defined over k with CM by M, and we have the tower of fields $kM \subseteq K \subseteq L$, where L/k is the minimal extension over which all the homomorphisms from $A_{\overline{\mathbb{Q}}}$ to $E_{\overline{\mathbb{Q}}}$ are defined, and K/k is the minimal extension over which all the endomorphisms of $A_{\overline{\mathbb{Q}}}$ are defined.

For any field extension F/k, let P_{A_F} denote the set of primes of F at which A_F has good reduction. For $\mathfrak{p}$ in P_{A_F}, we write the normalized L-polynomial for A_F at $\mathfrak{p}$ as

$$\bar{L}_{\mathfrak{p}}(A_F, T) = 1 + a_1(A_F)(\mathfrak{p})T + a_2(A_F)(\mathfrak{p})T^2 + a_1(A_F)(\mathfrak{p})T^3 + T^4.$$

Let P be the set of primes lying in P_{A_F} and P_{E_F} that are unramified in FL. Choose an ordering by norm $\{\mathfrak{p}_i\}_{i\geq 1}$ of P, and let $a_1(A_F)$ and $a_2(A_F)$ denote the sequences

$$\{a_1(A_F)(\mathfrak{p}_i)\}_{i\geq 1}, \quad \{a_2(A_F)(\mathfrak{p}_i)\}_{i\geq 1},$$

respectively. In this section, we use the results in Sections 3A and 3B to prove equidistribution for $a_1(A)$ and $a_2(A)$.

Lemma 3.9. *Let $\mathfrak{p}$ be a prime of good reduction for A and E that splits in kM and is unramified in L.*

(i) *With* $u_1 = \operatorname{Re} a_1(\theta)(\mathrm{Frob}_{\mathfrak{p}})$ *and* $u_2 = \operatorname{Im} a_1(\theta)(\mathrm{Frob}_{\mathfrak{p}})$, *we have*

$$a_1(A)(\mathfrak{p}) = u_1 a_1(E)(\mathfrak{p}) \pm u_2\sqrt{4 - a_1(E)(\mathfrak{p})^2}.$$

(ii) *With* $v_1 = \operatorname{Re} a_2(\theta)(\mathrm{Frob}_{\mathfrak{p}})$ *and* $v_2 = \operatorname{Im} a_2(\theta)(\mathrm{Frob}_{\mathfrak{p}})$, *we have*

$$a_2(A)(\mathfrak{p}) = v_1 a_1(E)(\mathfrak{p})^2 - 2v_1 + |a_1(\theta)(\mathrm{Frob}_{\mathfrak{p}})|^2 \mp v_2 a_1(E)(\mathfrak{p})\sqrt{4 - a_1(E)(\mathfrak{p})^2}.$$

Proof. Define $V_\sigma(A)$ and $V_{\overline{\sigma}}(A)$ as in (3-4). We then have the following isomorphism of $\overline{\mathbb{Q}}_\ell[G_{kM}]$-modules:

$$V_\ell(A_{kM}) \otimes \overline{\mathbb{Q}}_\ell \simeq V_\sigma(A) \oplus V_{\overline{\sigma}}(A).$$

By arguments analogous to those in [Fité 2013, Theorem 3.1], we have

$$V_\sigma(A) \simeq \theta_{M,\sigma}(E, A) \otimes V_\sigma(E), \quad V_{\overline{\sigma}}(A) \simeq \theta_{M,\overline{\sigma}}(E, A) \otimes V_{\overline{\sigma}}(E).$$

Thus there is an isomorphism of $\overline{\mathbb{Q}}_\ell[G_{kM}]$-modules:

$$V_\ell(A) \otimes \overline{\mathbb{Q}}_\ell \simeq \theta_{M,\sigma}(E, A) \otimes V_\sigma(E) \oplus \theta_{M,\overline{\sigma}}(E, A) \otimes V_{\overline{\sigma}}(E). \tag{3-8}$$

To shorten notation, we write $\alpha_1(\mathfrak{p})$ for $\alpha_1(E_{kM})(\mathfrak{p}) = \sigma(\alpha(E_{kM})(\mathfrak{p}))/\|\mathfrak{p}\|^{1/2}$, as defined in (3-7). Then $\overline{\alpha_1(\mathfrak{p})} = \overline{\sigma}(\alpha(E_{kM})(\mathfrak{p}))/\|\mathfrak{p}\|^{1/2}$, and (3-8) implies that

$$\begin{aligned} a_1(A_{kM})(\mathfrak{p}) &= -a_1(\mathfrak{p})\alpha_1(\mathfrak{p}) - \overline{a_1(\mathfrak{p})\alpha_1(\mathfrak{p})}, \\ a_2(A_{kM})(\mathfrak{p}) &= a_2(\mathfrak{p})\alpha_1(\mathfrak{p})^2 + \overline{a_2(\mathfrak{p})\alpha_1(\mathfrak{p})}^2 + a_1(\mathfrak{p})\overline{a_1(\mathfrak{p})}, \end{aligned} \tag{3-9}$$

where $a_i(\mathfrak{p})$ denotes $a_i(\theta)(\mathrm{Frob}_{\mathfrak{p}})$. The proposition then follows from the fact that $a_1(E_{kM})(\mathfrak{p}) = -\alpha_1(\mathfrak{p}) - \overline{\alpha_1(\mathfrak{p})}$. □

Proposition 3.10. *For* $\tau \in \mathrm{Gal}(L/kM)$, *let* $u(\tau) = |a_1(\theta)(\tau)|$. *Then* $a_1(A_{kM})$ *and* $a_2(A_{kM})$ *are equidistributed with respect to the measures*

(i) $\displaystyle \mu(a_1(A_{kM})) := \frac{1}{[L:kM]}\frac{1}{\pi}\sum_\tau \frac{dz}{\sqrt{4u(\tau)^2 - z^2}}\mathbf{1}_{[-2u(\tau),2u(\tau)]}$,

(ii) $\displaystyle \mu(a_2(A_{kM})) := \frac{1}{[L:kM]}\frac{1}{\pi}\sum_\tau \frac{dz}{\sqrt{4 - (u(\tau)^2 - z)^2}}\mathbf{1}_{[u(\tau)^2-2,u(\tau)^2+2]}$,

whose support lies in the intervals $I_1 = [-4, 4]$ *and* $I_2 = [-6, 6]$, *respectively. In each sum,* τ *ranges over* $\mathrm{Gal}(L/kM)$ *and* $\mathbf{1}_{[a,b]}$ *is the characteristic function of the interval* $[a, b] \subseteq \mathbb{R}$. *Moreover, we have*

(i) $\displaystyle \mathrm{M}_n[a_1(A_{kM})] = \frac{1}{[L:kM]}\sum_\tau b_{0,n}u(\tau)^n$,

(ii) $\displaystyle \mathrm{M}_n[a_2(A_{kM})] = \frac{1}{[L:kM]}\sum_\tau b_{u(\tau)^2,n}$,

where the integer $b_{m,n}$ *is the coefficient of* X^n *in* $(X^2 + mX + 1)^n$.

Proof. We can rewrite the equations in (3-9) as follows:

$$a_1(A_{kM})(\mathfrak{p}) = |a_1(\mathfrak{p})| \left(\frac{-a_1(\mathfrak{p})}{|a_1(\mathfrak{p})|} \alpha_1(\mathfrak{p}) + \frac{-\overline{a_1(\mathfrak{p})}}{|a_1(\mathfrak{p})|} \overline{\alpha_1(\mathfrak{p})} \right),$$

$$\begin{aligned} a_2(A_{kM})(\mathfrak{p}) &= |a_2(\mathfrak{p})| \left(\frac{a_2(\mathfrak{p})^{1/2}}{|a_2(\mathfrak{p})|^{1/2}} \alpha_1(\mathfrak{p}) + \frac{\overline{a_2(\theta)}^{1/2}}{|a_2(\mathfrak{p})|^{1/2}} \overline{\alpha_1(\mathfrak{p})} \right)^2 - 2|a_2(\mathfrak{p})| + |a_1(\mathfrak{p})|^2 \\ &= \left(a_2(\mathfrak{p})^{1/2} \alpha_1(\mathfrak{p}) + \overline{a_2(\mathfrak{p})}^{1/2} \overline{\alpha_1(\mathfrak{p})} \right)^2 - 2 + |a_1(\mathfrak{p})|^2, \end{aligned}$$

where $a_i(\mathfrak{p})$ denotes $a_i(\theta)(\mathrm{Frob}_{\mathfrak{p}})$, and we have used $|a_2(\mathfrak{p})| = 1$. The equidistribution statements now follow from the Chebotarev density theorem and two facts below:

(1) For any $z \in \mathrm{U}(1)$ and any conjugacy class c of $\mathrm{Gal}(L/kM)$, the sequence $z\alpha_{1,c}$ is $\mu(\mathrm{U}(1))$-equidistributed on $\mathrm{U}(1)$. Indeed, Proposition 3.6 ensures equidistribution of $\alpha_{1,c}$, and invariance under translations is in fact the defining property of the Haar measure. Thus the sequence $z\alpha_{1,c} + \overline{z\alpha}_{1,c}$ is μ_{cm}-equidistributed on $I_1(E_{kM}) = [-2, 2]$.

(2) If a sequence $\beta = \{\beta_i\}_{i \geq 1}$ is μ_{cm}-equidistributed on $[-2, 2]$, then for $u \in \mathbb{R}_{>0}$:

- The sequence $u\beta$ is equidistributed on $[-2u, 2u]$ with respect to the measure
$$\frac{1}{\pi} \frac{dz}{\sqrt{4u^2 - z^2}}.$$
- The sequence $\{\beta_i^2 - 2 + u^2\}_{i \geq 1}$ is equidistributed on $[u^2 - 2, u^2 + 2]$ with respect to the measure
$$\frac{1}{\pi} \frac{dz}{\sqrt{4 - (u^2 - z)^2}}.$$

Regarding the moments, the Chebotarev density theorem implies that

$$\mathrm{M}_n[a_1(A_{kM})] = \frac{1}{[L : kM]} \sum_{\tau} |a_1(\theta)(\tau)|^n \cdot \mathrm{M}_n\big[z([\tau])\alpha_1 + \overline{z([\tau])}\overline{\alpha}_1 \mid P_{[\tau]}\big],$$

where $z([\tau]) = -a_1(\theta)(\tau)/|a_1(\theta)(\tau)|$. But now (1) implies that

$$\mathrm{M}_n\big[z([\tau])\alpha_1 + \overline{z([\tau])}\overline{\alpha}_1 \mid P_{[\tau]}\big] = b_{0,n}.$$

The same argument is used to compute

$$\mathrm{M}_n[a_2(A_{kM})] = \frac{1}{[L : kM]} \sum_{\tau} \sum_{i=0}^{n} \binom{n}{i} \binom{2i}{i} \big(|a_1(\theta)(\tau)|^2 - 2\big)^{n-i}.$$

One then applies

$$\sum_{i=0}^{n}\binom{n}{i}\binom{2i}{i}(m-2)^{n-i} = [X^n]\big((X+1)^2+(m-2)X\big)^n$$
$$= [X^n](X^2+mX+1)^n = b_{m,n},$$

where $[X^n]f(X)$ denotes the coefficient of X^n in the polynomial $f(X)$. □

We now generalize the definitions of $o(r)$ and $\overline{o}(s)$ given in Section 1B for $k=\mathbb{Q}$.

Definition 3.11. Let $o(r)$ count the elements in $\mathrm{Gal}(L/kM)$ whose projection in $\mathrm{Gal}(K/kM)$ has order r. Let $\overline{o}(s)$ count the elements in $\mathrm{Gal}(L/k)\setminus\mathrm{Gal}(L/kM)$ of order s.

If $k=kM$, the sequence $a_i(A)$ is equidistributed with respect to $\mu(a_i(A_{kM}))$ and $\mathrm{M}_n[a_i(A)]=\mathrm{M}_n[a_i(A_{kM})]$, for $i=1,2$.

Corollary 3.12. *Suppose $k\neq kM$. Then $a_1(A)$ and $a_2(A)$ are equidistributed with respect to the measures*

(i) $\mu(a_1(A)) := \frac{1}{2}\mu(a_1(A_{kM}))+\frac{1}{2}\delta_0$,

(ii) $\mu(a_2(A)) := \frac{1}{2}\mu(a_1(A_{kM}))$
$$+\frac{1}{2[L:kM]}\big(\bar{o}(2)\delta_2+\bar{o}(4)\delta_{-2}+\bar{o}(6)\delta_{-1}+\bar{o}(12)\delta_1\big),$$

whose support lies in the intervals $I_1=[-4,4]$ and $I_2=[-6,6]$, respectively. Here δ_z denotes the Dirac measure at z. We also have

(i) $\mathrm{M}_n[a_1(A)] = \dfrac{1}{[L:k]}\big(o(1)2^n+o(3)+o(4)2^{n/2}+o(6)3^{n/2}\big)b_{0,n}$,

(ii) $\mathrm{M}_n[a_2(A)] = \dfrac{1}{[L:k]}\big(o(1)b_{4,n}+o(2)b_{0,n}+o(3)b_{1,n}+o(4)b_{2,n}+o(6)b_{3,n}$
$$+\bar{o}(2)2^n+\bar{o}(4)(-2)^n+\bar{o}(6)(-1)^n+\bar{o}(12)\big).$$

Proof. We focus on the proof of the statements about the moments, since the arguments involved suffice to deduce the statements about the measures. Statement (i) follows from Propositions 3.2, 3.4, and 3.10, and the equality

$$\mathrm{M}_{2n}[a_1(A_{kM})] = 2\cdot\mathrm{M}_{2n}[a_1(A)],$$

which follows from the fact that if $\mathfrak{p}$ is a prime of k, where A has good reduction and $\mathfrak{p}$ is inert in kM, then A is supersingular at $\mathfrak{p}$ and $a_1(A)(\mathfrak{p})=0$.

For (ii), let ν denote the nontrivial conjugacy class of $\mathrm{Gal}(kM/k)$. Note that

$$\mathrm{M}_n[a_2(A)] = \tfrac{1}{2}\mathrm{M}_n[a_2(A)\mid P_1]+\tfrac{1}{2}\mathrm{M}_n[a_2(A)\mid P_\nu].$$

To compute $\mathrm{M}_n[a_2(A) \mid P_1] = \mathrm{M}_n[a_2(A_{kM})]$, we apply Proposition 3.10. We then claim that

$$\mathrm{M}_n[a_2(A_k) \mid P_\nu] = \frac{1}{[L:kM]}\big(\overline{o}(2)2^n + \overline{o}(4)(-2)^n + \overline{o}(6)(-1)^n + \overline{o}(12)\big).$$

We may restrict to primes $\mathfrak{p}$ of k that are inert in kM, of absolute residue degree 1, and of good reduction for both A and E. The polynomial $\bar{L}_{\mathfrak{p}}(A, T)$ must then be one of the five listed in (3-2).

We now consider the Rankin–Selberg polynomial $\bar{L}_{\mathfrak{p}}\big(E, \theta_{\mathbb{Q}}(E, A), T\big)$, whose roots are all products of roots of $\bar{L}_{\mathfrak{p}}(E, T) = 1 + T^2$, and all roots of the polynomial $\det\big(1 - \theta_{\mathbb{Q}}(E, A)(\mathrm{Frob}_{\mathfrak{p}})T\big)$. More explicitly, if s is the order of $\mathrm{Frob}_{\mathfrak{p}}$ in $\mathrm{Gal}(L/k)$, one may apply Proposition 3.3 to compute $\bar{L}_{\mathfrak{p}}\big(E, \theta_{\mathbb{Q}}(E, A), T\big)$. This yields:

$$\begin{array}{lll} s=2:\ (1+T^2)^4 & s=6:\ (1-T^2+T^4)^2 & s=12:\ (1+T^2+T^4)^2 \\ s=4:\ (1-T^2)^4 & s=8:\ (1+T^4)^2 & \end{array}$$

By arguments analogous to those of [Fité 2013, Theorem 3.1], there is an inclusion of $\mathbb{Q}_\ell[G_k]$-modules

$$V_\ell(A) \subseteq V_\ell(E) \otimes \theta_{\mathbb{Q}}(E, A).$$

This implies that $\bar{L}_{\mathfrak{p}}(A, T)$ divides $\bar{L}_{\mathfrak{p}}(E, \theta_{\mathbb{Q}}(E, A), T)$. It immediately follows that $\bar{L}_{\mathfrak{p}}(A, T)$ is

$$\begin{array}{lll} s=2:\ (1+T^2)^2 & s=6:\ 1-T^2+T^4 & s=12:\ 1+T^2+T^4 \\ s=4:\ (1-T^2)^2 & s=8:\ 1+T^4 & \end{array}$$

Finally, we observe that the condition $\bar{L}_{\mathfrak{p}}(A, T)$ divides $\bar{L}_{\mathfrak{p}}(E, \theta_{\mathbb{Q}}(E, A), T)$ implies that s can not attain any value other than the ones considered. □

3D. ***Additional remarks.*** As noted in the introduction, all 32 of the genus-2 Sato–Tate groups with identity component isomorphic to U(1) can arise as the Sato–Tate group of an abelian variety A defined over k with $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$, where E is an elliptic curve defined over k (with CM).

However, not all 10 of the genus-2 Sato–Tate groups with identity component isomorphic to SU(2) can arise as the Sato–Tate group of an abelian variety A defined over k such that $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$, where E is an elliptic curve defined over k (without CM).[8] The Sato–Tate groups for which this is not true are the four whose component group contains an element of order 4 or 6. Indeed, recall that $\theta_{\mathbb{Q}}(E, A)$ and $\theta_{\mathbb{Q}}(A)$ are the representations afforded by $\mathrm{Hom}(E_L, A_L) \otimes \mathbb{Q}$ and $\mathrm{End}(A_L) \otimes \mathbb{Q}$. As in the proof of Proposition 3.2, one can then show that $\theta_{\mathbb{Q}}(A) = \theta_{\mathbb{Q}}(E, A)^{\otimes 2}$,

[8]All Sato–Tate groups with identity component SU(2) can occur for an A over k such that $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$ for *some* elliptic curve E, but this curve need not be defined over k.

that is, $a_1(\theta_{\mathbb{Q}}(A)) = a_1(\theta_{\mathbb{Q}}(E, A))^2$. But if $\tau \in \mathrm{Gal}(K/k)$ has order 4 or 6, then $a_1(\theta_{\mathbb{Q}}(A))(\tau) = 2$ or 3, which are not squares in $\mathbb{Q}$.

We end this section by computing the density $z_1(A_k)$ of zero traces of an abelian variety A defined over k such that $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$ for some elliptic curve E defined over $\overline{\mathbb{Q}}$.

Lemma 3.13. *Let A be an abelian variety defined over k such that $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$, where E is an elliptic curve defined over $\overline{\mathbb{Q}}$ (not necessarily over k). Let M denote the CM field if E has CM, and let $M = \mathbb{Q}$ otherwise. Then*

$$z_1(A_k) = \begin{cases} \dfrac{o(2)}{|\mathrm{Gal}(L/kM)|} & \text{if } [kM : k] = 1, \\[2ex] \dfrac{1}{2} + \dfrac{1}{2}\dfrac{o(2)}{|\mathrm{Gal}(L/kM)|} & \text{if } [kM : k] = 2. \end{cases}$$

Proof. Except for a set of density zero, any prime $\mathfrak{p}$ of k that does not split in kM is supersingular, in which case $a_1(A)(\mathfrak{p}) = 0$. This gives density 0 in the first case and density $\frac{1}{2}$ in the second case. Among the primes that split in kM, we wish to show that exactly the proportion $o(2)/|\mathrm{Gal}(L/kM)|$ have trace 0. Among these primes, the density of the supersingular primes is zero. Let $\mathfrak{p}$ be a nonsupersingular prime of good reduction for A that splits in kM. From Remark 4.8 in [Fité et al. 2012] in the non-CM case, and from Proposition 3.4 in the CM case, the roots of $L_{\mathfrak{p}}(A, T)$ are α, $\overline{\alpha}$, $\zeta_r\alpha$, $\overline{\zeta}_r\overline{\alpha}$, where r is the order of $\mathrm{Frob}_{\mathfrak{p}}$ in $\mathrm{Gal}(K/k)$ and where $\alpha/\overline{\alpha}$ is not a root of unity. It follows that $\alpha + \overline{\alpha} + \zeta_r\alpha + \overline{\zeta}_r\overline{\alpha} = 0$ if and only if $r = 2$. One then applies the Chebotarev density theorem. □

4. Twists of $y^2 = x^5 - x$ and $y^2 = x^6 + 1$

In this section, we strengthen the results of Section 3 in the particular case that $k = \mathbb{Q}$ and $A \sim_{\mathbb{Q}} \mathrm{Jac}(C)$, where C is a twist of the curve $y^2 = x^5 - x$ or $y^2 = x^6 + 1$. We first introduce some convenient notation. Let C_2^0 and C_3^0 denote the curves defined over $\mathbb{Q}$ by the equations

$$C_2^0\colon\ y^2 = x^6 - 5x^4 - 5x^2 + 1, \quad C_3^0\colon\ y^2 = x^6 + 1.$$

The curve C_2^0 is a twist of $y^2 = x^5 - x$, as one may verify by computing their respective Igusa invariants, as defined in [Igusa 1960]. As shown below, the Jacobian of C_2^0 is $\mathbb{Q}$-isogenous to the square of an elliptic curve defined over $\mathbb{Q}$, a property that the curve $y^2 = x^5 - x$ does not enjoy. We also note that the minimal field of definition of the endomorphisms of the Jacobian of C_2^0 is $\mathbb{Q}(\sqrt{-2})$, but for $y^2 = x^5 - x$ it is $\mathbb{Q}(i, \sqrt{-2})$.

Let E_2^0 and E_3^0 denote the elliptic curves defined over $\mathbb{Q}$ by the equations

$$E_2^0\colon\ Y^2 = X^3 - 5X^2 - 5X + 1, \quad E_3^0\colon\ Y^2 = X^3 + 1.$$

We note that $j(E_2^0) = 2^6 5^3$ and $j(E_3^0) = 0$, and thus E_2^0 has CM by $\mathbb{Q}(\sqrt{-2})$ and E_3^0 has CM by $\mathbb{Q}(\sqrt{-3})$.

To simplify notation, throughout this section d denotes either 2 or 3, and we write C^0 for C_d^0, E^0 for E_d^0, and M for $\mathbb{Q}(\sqrt{-d})$. We use C to denote a twist of C^0 defined over $\mathbb{Q}$. In the context of Section 3, we are specializing $A_{\overline{\mathbb{Q}}} \sim E^2_{\overline{\mathbb{Q}}}$ to the case where $A = \mathrm{Jac}(C)$ and $E = E^0$, as we now show.

4A. ***Fields of definition of isomorphisms.***

Lemma 4.1. $\mathrm{Jac}(C_d^0)$ *is $\mathbb{Q}$-isogenous to* $(E_d^0)^2$.

Proof. We proceed as in the proof of Lemma 4.1 in [Fité and Lario 2013]. The quotient of C_d^0 by the nonhyperelliptic involution $\alpha(x, y) = (-x, y)$ is precisely the elliptic curve E_d^0, and thus $\mathrm{Jac}(C_d^0) \sim_{\mathbb{Q}} E_d^0 \times E$, where E is also an elliptic curve defined over $\mathbb{Q}$. The automorphism $\gamma(x, y) = (1/x, y/x^3)$ does not commute with α, which implies that $\mathrm{End}(\mathrm{Jac}(C^0))$ is nonabelian, and therefore $\mathrm{Jac}(C_d^0) \sim_{\mathbb{Q}} (E_d^0)^2$. □

Lemma 4.2. *The minimal number field over which all the automorphisms of $C_{\overline{\mathbb{Q}}}$ are defined coincides with the minimal number field over which all the endomorphisms of* $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ *are defined.*

Proof. Let K_a (resp. K_e) denote the minimal number field over which all the automorphisms of $C_{\overline{\mathbb{Q}}}$ (resp. all the endomorphisms of $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$) are defined. The fact that $\mathrm{Aut}(C_{K_a})$ is nonabelian and contains a nonhyperelliptic involution implies that $\mathrm{Jac}(C)_{K_a} \sim E^2$, where E is an elliptic curve defined over K_a. Since E has CM by M, it follows that $K_e = K_a M$. But [Cardona 2001, Proposition 7.3.1] asserts that $M = \mathbb{Q}(\sqrt{-3})$ is already contained in K_a if C is a twist of C_3^0, whereas [Cardona 2006, Proposition 8] states that $M = \mathbb{Q}(\sqrt{-2})$ is already contained in K_a if C is a twist of C_2^0. □

We use K to denote the field given by Lemma 4.2. We note that K is a Galois extension of $\mathbb{Q}$, and we have $M \subseteq K$, with equality in the case $C = C^0$.

Lemma 4.3. *Let ϕ be an isomorphism from $C^0_{\overline{\mathbb{Q}}}$ to $C_{\overline{\mathbb{Q}}}$. The following number fields coincide*:

(i) *the minimal field over which all isomorphisms from $C^0_{\overline{\mathbb{Q}}}$ to $C_{\overline{\mathbb{Q}}}$ are defined*;

(ii) *the compositum of K (or even just M) and the minimal field L_ϕ over which ϕ is defined*;

(iii) *the minimal field over which all homomorphisms from* $\mathrm{Jac}(C^0)_{\overline{\mathbb{Q}}}$ *to* $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ *are defined*;

(iv) *the minimal field over which all homomorphisms from $E^0_{\overline{\mathbb{Q}}}$ to* $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ *are defined.*

Proof. Let L_1, L_2, L_3, and L_4 denote the fields defined by (i), (ii), (iii), and (iv), respectively. Any isomorphism ψ from $C^0_{\overline{\mathbb{Q}}}$ to $C_{\overline{\mathbb{Q}}}$ can be written as $\psi = \alpha \circ \phi$ and $\phi \circ \alpha^0$ for some $\alpha \in \mathrm{Aut}(C_K)$ and some $\alpha^0 \in \mathrm{Aut}(C^0_M)$. This implies that $L_1 \subseteq ML_\phi \subseteq KL_\phi = L_2$. Conversely, for any $\alpha^0 \in \mathrm{Aut}(C^0_M)$ and $\alpha \in \mathrm{Aut}(C_K)$, the compositions $\alpha \circ \phi$ and $\phi \circ \alpha^0$ are isomorphisms from $C^0_{\overline{\mathbb{Q}}}$ to $C_{\overline{\mathbb{Q}}}$. It follows that $L_2 \subseteq L_1$. Thus we have shown $L_1 = ML_\phi = KL_\phi = L_2$.

The isomorphism from $C^0_{L_\phi}$ to C_{L_ϕ} induces an isogeny $\mathrm{Jac}(C^0)_{L_\phi} \sim \mathrm{Jac}(C)_{L_\phi}$, which we also denote by ϕ. Any homomorphism from $\mathrm{Jac}(C^0)_{\overline{\mathbb{Q}}}$ to $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ can be written as $\psi \circ \phi$ for some $\psi \in \mathrm{End}(\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}) \otimes \mathbb{Q}$. This implies that $L_3 \subseteq L_\phi K, L_\phi M = L_2$. Conversely, it is clear that L_1 is contained in L_3.

Any endomorphism ϕ from $\mathrm{Jac}(C^0)_{\overline{\mathbb{Q}}}$ to $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ can be written as $\phi_2 \circ \phi_1$, where $\phi_1 \in \mathrm{Hom}(\mathrm{Jac}(C^0), (E^0)^2) \otimes \mathbb{Q}$ and $\phi_2 \in (\mathrm{Hom}(E^0_{L_4}, \mathrm{Jac}(C)_{L_4}) \otimes \mathbb{Q})^2$. Thus $L_3 \subseteq L_4$. Conversely, any homomorphism from $E^0_{\overline{\mathbb{Q}}}$ to $\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}$ can be written as $\phi_2 \circ \phi_1$, where $\phi_1 \in \mathrm{Hom}(E^0, \mathrm{Jac}(C^0)) \otimes \mathbb{Q}$ and ϕ_2 is an element of $\mathrm{Hom}(\mathrm{Jac}(C^0)_{L_3}, \mathrm{Jac}(C)_{L_3}) \otimes \mathbb{Q}$. Thus $L_4 \subseteq L_3$. □

We use L to denote the field given by Lemma 4.3, and we note that L is a Galois extension of $\mathbb{Q}$ that contains K.

Remark 4.4. If A is the abelian three-fold $E^0 \times \mathrm{Jac}(C)$, we observe that L coincides with the minimal field over which all the endomorphisms of $A_{\overline{\mathbb{Q}}}$ are defined. It follows that the component group of $\mathrm{ST}(A)$ is isomorphic to $\mathrm{Gal}(L/\mathbb{Q})$.

4B. ***The Galois module*** $\mathbf{Hom}(\boldsymbol{E^0_L}, \mathbf{Jac}(\boldsymbol{C})_{\boldsymbol{L}})$**.** We now compute $\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))$, strengthening Lemma 3.9 in the case where $A \sim_{\mathbb{Q}} \mathrm{Jac}(C)$. We take advantage of the following fact: the group $\mathrm{Gal}(L/\mathbb{Q})$ is isomorphic to a subgroup of $G_{C^0} := \mathrm{Aut}(C^0_M) \rtimes \mathrm{Gal}(M/\mathbb{Q})$. Here the action of $\mathrm{Gal}(M/\mathbb{Q})$ on $\mathrm{Aut}(C^0_M)$ is the natural one (see [Fité and Lario 2013, §2]).

More precisely, let $\phi : C_L \to C^0_L$ be an isomorphism. Then

$$\lambda_\phi : \mathrm{Gal}(L/\mathbb{Q}) \hookrightarrow G_{C^0}, \quad \lambda_\phi(\sigma) = \big(\phi ({}^\sigma\phi)^{-1}, \pi_{L/M}(\sigma)\big)$$

is a monomorphism of groups, where $\pi_{L/M} : \mathrm{Gal}(L/\mathbb{Q}) \to \mathrm{Gal}(M/\mathbb{Q})$ is the natural projection, as in [Fité and Lario 2013, Lemma 2.1]. Now let

$$\mathrm{Res}^{\mathbb{Q}}_M \lambda_\phi : \mathrm{Gal}(L/M) \hookrightarrow \mathrm{Aut}(C^0_M)$$

be the restriction of λ_ϕ at $\mathrm{Gal}(L/M)$. Consider the 2-dimensional M-rational representation

$$\theta_{E^0,C^0} : \mathrm{Aut}(C^0_M) \to \mathrm{Aut}_{\overline{\mathbb{Q}}}\big(\mathrm{Hom}(E^0_M, \mathrm{Jac}(C^0)_M) \otimes_{M,\sigma} \overline{\mathbb{Q}}\big)$$

defined by $\theta_{E^0,C^0}(\alpha)(\psi) = \alpha \circ \psi$. As in [Fité and Lario 2013, Theorem 2.1], one then has

$$\theta_{E^0,C^0} \circ \mathrm{Res}^{\mathbb{Q}}_M \lambda_\phi \simeq \theta_{M,\sigma}(E^0, \mathrm{Jac}(C)), \tag{4-1}$$

Class	1a	2a	2b	3a	4a	6a	8a	8b
Size	1	1	12	8	6	8	6	6
χ_1	1	1	1	1	1	1	1	1
χ_2	1	1	−1	1	1	1	−1	−1
χ_3	2	2	0	−1	2	−1	0	0
χ_4	2	−2	0	−1	0	1	$\sqrt{-2}$	$-\sqrt{-2}$
χ_5	2	−2	0	−1	0	1	$-\sqrt{-2}$	$\sqrt{-2}$
χ_6	3	3	1	0	−1	0	−1	−1
χ_7	3	3	−1	0	−1	0	1	1
χ_8	4	−4	0	1	0	−1	0	0

Table 1. Character table of $\mathrm{Aut}((C_2^0)_M) \simeq \langle 48, 29\rangle$.

Class	1a	2a	2b	2c	3a	4a	6a	6b	6c
Size	1	1	2	6	2	6	2	2	2
χ_1	1	1	1	1	1	1	1	1	1
χ_2	1	1	1	−1	1	−1	1	1	1
χ_3	1	1	−1	−1	1	1	−1	−1	1
χ_4	1	1	−1	1	1	−1	−1	−1	1
χ_5	2	2	−2	0	−1	0	1	1	−1
χ_6	2	−2	0	0	2	0	0	0	−2
χ_7	2	2	2	0	−1	0	−1	−1	−1
χ_8	2	−2	0	0	−1	0	$-\sqrt{-3}$	$\sqrt{-3}$	1
χ_9	2	−2	0	0	−1	0	$\sqrt{-3}$	$-\sqrt{-3}$	1

Table 2. Character table of $\mathrm{Aut}((C_3^0)_M) \simeq \langle 24, 8\rangle$.

where $\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))$ is the representation of $\mathrm{Gal}(L/M)$ in Definition 3.1.

Lemma 4.5. *Let C be a twist of C^0. Then*

$$\mathrm{Tr}\,\theta_{E^0,C^0} = \begin{cases} \chi_4 \text{ or } \chi_5 & \text{if } C^0 = C_2^0 \text{ (see Table 1),} \\ \chi_8 \text{ or } \chi_9 & \text{if } C^0 = C_3^0 \text{ (see Table 2).} \end{cases}$$

Proof. A glance at Tables 1 and 2 tells us that any M-rational faithful representation of degree 2 must have trace χ_4 or χ_5 when $C^0 = C_2^0$, or trace χ_8 or χ_9 when $C^0 = C_3^0$. The two possibilities in each case correspond to the two different embeddings of M into $\overline{\mathbb{Q}}$. □

Proposition 4.6. *The index of K in L is at most* 2.

Proof. As in Lemma 4.1, let α be the nonhyperelliptic involution $\alpha(x, y) = (-x, y)$ of C^0. Let E be the elliptic curve $C_K/\langle\phi^{-1}\alpha\phi\rangle$ defined over K (note that $\phi^{-1}\alpha\phi$ is an automorphism of C, all of which are defined over K). The isomorphism $\phi : C_L \to C_L^0$ induces an isomorphism $\tilde{\phi} : E_L \to E_L^0$. Thus E is a K-twist of E^0.

From characterization (iii) of L in Lemma 4.3, it is clear that L is the compositum of K and the minimal field $L_{\tilde{\phi}}$ over which $\tilde{\phi}$ is defined.

When $C^0 = C_2^0$, we have $j(E) \neq 0, 1728$, and by [Silverman 2009, p. 304], it follows that $\tilde{\phi}$ is then defined over a quadratic extension of K and $[L : K] \leq 2$. When $C^0 = C_3^0$, we have $j(E) = 0$, and in this case $L = K(\sqrt[6]{\gamma})$, for some $\gamma \in K$. Let $L_0 = K(\sqrt{\gamma})$. It suffices to show that $\sqrt[3]{\gamma} \in L_0$.

Suppose for the sake of contradiction that $\sqrt[3]{\gamma} \notin L_0$. Then $\mathrm{Gal}(L/L_0) \simeq \mathrm{C}_3$. Lemma 4.5 then implies that if τ is a nontrivial element of $\mathrm{Gal}(L/L_0)$, then $\mathrm{Tr}\,\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))(\tau) = -1$. Therefore, the restriction of the representation afforded by the $\mathrm{Gal}(L/M)$-module $\mathrm{Hom}(E_L^0, \mathrm{Jac}(C)_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$ to $\mathrm{Gal}(L/L_0)$ is

$$\mathrm{Res}^M_{L_0}\, \theta_{M,\sigma}(E^0, \mathrm{Jac}(C)) \simeq \chi \oplus \overline{\chi},$$

where χ is any of the two nontrivial characters of $\mathrm{Gal}(L/L_0)$. As in [Fité 2013, Theorem 3.1], we have

$$\mathrm{Res}^M_{L_0}\, \theta_{M,\sigma}(E^0, \mathrm{Jac}(C)) \otimes V_\sigma(E^0) \simeq V_\sigma(\mathrm{Jac}(C)),$$

as $\overline{\mathbb{Q}}_\ell[G_{L_0}]$-modules. This implies that

$$V_\sigma(\mathrm{Jac}(C)) \simeq \big(\chi \otimes V_\sigma(E^0)\big) \oplus \big(\overline{\chi} \otimes V_\sigma(E^0)\big), \tag{4-2}$$

as $\overline{\mathbb{Q}}_\ell[G_{L_0}]$-modules. However, as seen in Lemma 4.2, $\mathrm{Jac}(C)_{L_0} \sim E^2_{L_0}$, which implies the following isomorphism of $\overline{\mathbb{Q}}_\ell[G_{L_0}]$-modules:

$$V_\sigma(\mathrm{Jac}(C)) \simeq V_\sigma(E)^{2\oplus}. \tag{4-3}$$

But now (4-2) and (4-3) together imply $V_\sigma(E^0) \simeq \chi \otimes V_\sigma(E^0)$, which is impossible. (We remark that if $\mathrm{Res}^M_{L_0}\, \theta_{M,\sigma}(E^0, \mathrm{Jac}(C)) \simeq \chi^{2\oplus}$, one does not reach a contradiction; see Example 4.12). $\square$

Proposition 4.7. *Let w be the hyperelliptic involution of C^0. Then $[L : K] = 2$ if and only if $(w, 1) \in G_{C^0}$ lies in the image of λ_ϕ. If $[L : K] = 2$, then the preimage of $(w, 1)$ by λ_ϕ is the nontrivial element ω of* $\mathrm{Gal}(L/K)$.

Proof. We first suppose that $[L : K] = 2$. Observe that for both C_2^0 and C_3^0, if $\alpha \in \mathrm{Aut}(C_M^0)$ and $\mathrm{Tr}\,\theta_{E^0,C^0}(\alpha) = -2$, then $\alpha = w$. In view of the isomorphism in (4-1), it thus suffices to prove that $\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))(\omega) = -2$. From the proof of Proposition 4.6, we know that $\mathrm{Jac}(C)_K \sim E^2$, where E is an elliptic curve defined over K with CM by M. Fix an isomorphism $\psi_1 : E_L^0 \to E_L$. Fix an isogeny $\psi_2 : E_K \times E_K \to \mathrm{Jac}(C)_K$. For $i = 1, 2$, let $\iota_i : E_K \to E_K \times E_K$ denote the natural injection to the i-th factor. Then $\psi_2 \circ \iota_1 \circ \psi_1$ and $\psi_2 \circ \iota_2 \circ \psi_1$ constitute a basis of the $\overline{\mathbb{Q}}[\mathrm{Gal}(L/M)]$-module $\mathrm{Hom}(E_L^0, \mathrm{Jac}(C)_L) \otimes_{M,\sigma} \overline{\mathbb{Q}}$. The claim follows from the fact that ${}^\omega\psi_1 = -\psi_1$, ${}^\omega\psi_2 = \psi_2$, and ${}^\omega\iota_i = \iota_i$.

Now suppose that $[L:K]=1$. Recall the monomorphism

$$\overline{\lambda}_\phi : \mathrm{Gal}(K/\mathbb{Q}) \hookrightarrow \mathrm{Aut}(C_M^0)/\langle w\rangle \rtimes \mathrm{Gal}(M/\mathbb{Q})$$

of (2-3). The commutativity of the diagram

$$\begin{tikzcd}
 & & G_{C^0} \arrow[dr] & \\
\mathrm{Gal}(L/\mathbb{Q}) \arrow[urr, hook, "\lambda_\phi"] \arrow[r, equal] & \mathrm{Gal}(K/\mathbb{Q}) \arrow[rr, hook, "\overline{\lambda}_\phi"] & & G_{C^0}/\langle (w,1)\rangle
\end{tikzcd}$$

implies that $(w,1)$ does not lie in the image of λ_ϕ. □

Remark 4.8. Let $H_0 := \lambda_\phi(\mathrm{Gal}(L/M))$. If $(1,\tau)$ lies in the image of λ_ϕ, then $\lambda_\phi(\mathrm{Gal}(L/\mathbb{Q})) = H_0 \rtimes \langle (1,\tau)\rangle$; indeed, H_0 is normal in $\lambda_\phi(\mathrm{Gal}(L/\mathbb{Q}))$, since its index is 2, and $H_0 \cap \langle (1,\tau)\rangle$ is trivial. In this case, H_0 is stable under the action of $\mathrm{Gal}(M/\mathbb{Q})$. However, it is not true in general that $\mathrm{Gal}(L/\mathbb{Q}) \simeq H_0 \rtimes \langle (1,\tau)\rangle$ or that H_0 is stable under the action of $\mathrm{Gal}(M/\mathbb{Q})$.

Proposition 4.9. *For τ in $\mathrm{Gal}(L/\mathbb{Q})$, let $s=s(\tau)$, $r=r(\tau)$, and $t=t(\tau)$ denote the orders of τ, the projection of τ on $\mathrm{Gal}(K/\mathbb{Q})$, and the projection of τ on $\mathrm{Gal}(M/\mathbb{Q})$, respectively. The following hold*:

(i) *The triple (s,r,t) is one of the 13 triples listed in Table 3.*

(ii) *If τ fixes M, then the triple $(s,r,1)$ determines, up to sign, the quantities*

$$a_1(\theta)(\tau) = \mathrm{Tr}\,\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))(\tau), \tag{4-4}$$

$$a_2(\theta)(\tau) = \det\theta_{M,\sigma}(E^0, \mathrm{Jac}(C))(\tau), \tag{4-5}$$

as specified in Table 3.

(iii) *For each triple (s,r,t), let $F_{(s,r,t)} : [-2,2] \to [-4,4]\times[-2,6]$ be the map defined in Table 3. For every prime $p>3$ unramified in L of good reduction for both $\mathrm{Jac}(C)$ and E^0, there exists a unique triple (s,r,t) such that*

$$F_{(s,r,t)}\big(a_1(E^0)(p)\big) = \big(u\cdot a_1(\mathrm{Jac}(C))(p), a_2(\mathrm{Jac}(C))(p)\big), \tag{4-6}$$

with $u=\pm1$ (in fact, $u=1$ for $(s,r,t)\neq(6,6,1)$ and $(8,4,1)$). Moreover, the unique triple (s,r,t) for which (4-6) holds is $(f_L(p), f_K(p), f_M(p))$, where $f_F(p)$ is the residue degree of p in F.

Remark 4.10. For a prime p unramified in L such that $f_M(p)=1$, we have

$$a_2(\mathrm{Jac}(C))(p) = a_2(\theta)(\mathrm{Frob}_p)\cdot a_1(E^0)(p)^2 + |a_1(\theta)(\mathrm{Frob}_p)|^2 - 2a_2(\theta)(\mathrm{Frob}_p),$$

where $a_2(\theta)(\mathrm{Frob}_p)=\pm1$. It follows that for any two twists C and C' of C^0, we have

$$\hat{a}_2(\mathrm{Jac}(C))(p) \equiv \pm\hat{a}_2(\mathrm{Jac}(C'))(p) \pmod{p},$$

(s,r,t)	$F_{(s,r,t)}(x)$	$a_1(\theta)(\tau)$	$a_2(\theta)(\tau)$
$(1,1,1)$	$(2x, x^2+2)$	2	1
$(2,1,1)$	$(-2x, x^2+2)$	-2	1
$(2,2,1)$	$(0, -x^2+2)$	0	-1
$(3,3,1)$	$(-x, x^2-1)$	-1	1
$(4,2,1)$	$(0, x^2-2)$	0	1
$(6,3,1)$	(x, x^2-1)	1	1
$(6,6,1)$	$\left(\sqrt{3(4-x^2)}, -x^2+5\right)$	$\pm\sqrt{-3}$	-1
$(8,4,1)$	$\left(\sqrt{2(4-x^2)}, -x^2+4\right)$	$\pm\sqrt{-2}$	-1
$(2,2,2)$	$(0,2)$	$-$	$-$
$(4,2,2)$	$(0,-2)$	$-$	$-$
$(6,6,2)$	$(0,-1)$	$-$	$-$
$(8,4,2)$	$(0,0)$	$-$	$-$
$(12,6,2)$	$(0,1)$	$-$	$-$

Table 3. The triples for (s,r,t) associated to $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ (as defined in Proposition 4.9), and corresponding values of $F_{(s,r,t)}(x)$, $a_1(\theta)(\tau)$, and $a_2(\theta)(\tau)$.

where $\hat{a}_i(A)(p) = p^{i/2}a_i(A)(p)$ is the integer that appears as the coefficient of T^i in the (unnormalized) L-polynomial $L_p(A,T)$.

Proof. For assertion (i), assume first that $t=1$. Observe that s is the order of $\lambda_\phi(\tau)$ in $\mathrm{Aut}(C_M^0)$, and r is the order of the projection of $\lambda_\phi(\tau)$ in $\mathrm{Aut}(C_M^0)/\langle w\rangle$. Let c denote the conjugacy class of $\lambda_\phi(\tau)$ in $\mathrm{Aut}(C_M^0)$. One finds that the pairs (s,r) are determined by the conjugacy class of τ as follows:

$$\text{if } C^0 = C_3^0,\quad \begin{cases} c: & 1a & 2a & 2b, 2c & 3a & 4a & 6a, 6b & 6c \\ (r,s): & (1,1) & (2,1) & (2,2) & (3,3) & (4,2) & (6,6) & (6,3) \end{cases}$$

$$\text{if } C^0 = C_2^0,\quad \begin{cases} c: & 1a & 2a & 2b & 3a & 4a & 6a & 8a \\ (r,s): & (1,1) & (2,1) & (2,2) & (3,3) & (4,2) & (6,3) & (8,4) \end{cases}$$

(see Tables 2 and 1 for the names of the conjugacy classes). Assertion (ii) now follows immediately by applying the isomorphism in (4-1) and Lemma 4.5. If $t=2$, then r must be 2, 4, or 6, and the fact that $[L:K]\le 2$ limits (s,r,t) to either one of the last 5 triples in Table 3, or $(4,4,2)$. The latter possibility is ruled out by Proposition 3.3: if $s=4$, then for every prime p for which Frob_p lies in the same conjugacy class of τ in $\mathrm{Gal}(L/\mathbb{Q})$, we have $\bar{L}_p(\mathrm{Jac}(C),T) = (1-T^2)^2$, and the only quotients of roots of this polynomial are 1 and -1. This implies that $\theta_M(\mathrm{Jac}(C))(\tau)$ has order 2, and since $\theta_M(\mathrm{Jac}(C))$ is faithful, we must have $r = r(\tau) = 2$, not 4.

For $t=1$, the existence statement in (iii) follows from combining Lemma 3.9 with statement (ii), and for $t=2$, it follows from the proof of Corollary 3.12.

The uniqueness of the map $F_{(s,r,t)}$ satisfying (4-6) at a prime $p>3$ may be verified by noting that the graphs of the 13 functions $F_{(s,r,t)}$ intersect in only finitely many points in $\mathbb{R}^3$, none of which corresponds to a possible value of $\big(a_1(E^0)(p), a_1(\mathrm{Jac}(C))(p), a_2(\mathrm{Jac}(C))(p)\big)$ for any prime $p>3$. Finally, we note that if $\tau=\mathrm{Frob}_p$, then $\big(s(\tau), r(\tau), t(\tau)\big)=\big(f_L(p), f_K(p), f_M(p)\big)$. $\square$

We now give two examples of abelian varieties A such that $A_{\overline{\mathbb{Q}}}\sim(E^0_{\overline{\mathbb{Q}}})^2$ for which the conclusions of Propositions 4.6 and 4.9 do not hold because A is not $\mathbb{Q}$-isogenous to the Jacobian of a twist of C^0. In the two examples below, we use the elliptic curve

$$\tilde{E}^0_3\colon\ y^2=x^3+2$$

defined over $\mathbb{Q}$, which is a twist of E^0_3.

Example 4.11. Let $A=E^0_3\times\tilde{E}^0_3$. Then $K=L=\mathbb{Q}(\sqrt[6]{2},\zeta_3)$. If $\tau\in\mathrm{Gal}(L/M)$ and $s(\tau)=3$, then $a_1(\theta)(\tau)=1+\zeta_3$ or $1+\overline{\zeta}_3$ and $a_2(\theta)(\tau)=\zeta_3$ or $\overline{\zeta}_3$, which do not lie in $\mathbb{Q}$. Thus by Proposition 4.9, A is not $\mathbb{Q}$-isogenous to the Jacobian of any $\mathbb{Q}$-twist of C^0_3. Moreover, for $p=7$, one can compute that $\hat{a}_2(A)(7)=30$, while $\hat{a}_2((E^0_3)^2)(7)=\hat{a}_2(\mathrm{Jac}(C^0_3))(7)=18$. Thus

$$\hat{a}_2(A)(7)\not\equiv\pm\hat{a}_2(\mathrm{Jac}(C^0))(7) \pmod 7.$$

Example 4.12. Let $A=(\tilde{E}^0_3)^2$; then $L=\mathbb{Q}(\sqrt[6]{2},\zeta_3)$ and $K=\mathbb{Q}(\zeta_3)$; we have $[L:K]=6$, and, by Proposition 4.6, A is not $\mathbb{Q}$-isogenous to the Jacobian of any $\mathbb{Q}$-twist of C^0_3. In the context of the proof of Proposition 4.6, $L_0=\mathbb{Q}(\sqrt{2},\zeta_3)$ and $\mathrm{Res}^{\mathbb{Q}}_{L_0}\theta(E^0,A)\simeq\chi^{2\oplus}$, rather than $\mathrm{Res}^{\mathbb{Q}}_{L_0}\theta(E^0,A)\simeq\chi\oplus\overline{\chi}$, which avoids the contradiction used in the proof. Moreover, for A we may have $s(\tau)=3$ and $r(\tau)=1$, which gives a pair (s,r) that cannot occur for the Jacobian of any $\mathbb{Q}$-twist of C^0_3, by part (ii) of Proposition 4.9.

4C. *The triples $T(C)$.* We determine the possible values of the triple $T(C)$, which denotes the isomorphism class $\big[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)\big]$. To specify triples explicitly, we use identifiers from the Small Groups Library [Besche et al. 2002] found in computer algebra systems such as GAP and Magma. These identifiers consist of a pair of positive integers $\langle n,m\rangle$, where n is the order of the group and m distinguishes the group from other groups of order n but otherwise has no meaning. We also recall from Section 4B the embeddings

$$\begin{aligned}
\lambda_\phi&:\mathrm{Gal}(L/\mathbb{Q})\hookrightarrow G_{C^0}, & \mathrm{Res}\,\lambda_\phi&:\mathrm{Gal}(L/M)\hookrightarrow\mathrm{Aut}(C^0_M),\\
\overline{\lambda}_\phi&:\mathrm{Gal}(K/\mathbb{Q})\hookrightarrow G_{C^0}/\langle(w,1)\rangle, & \mathrm{Res}\,\overline{\lambda}_\phi&:\mathrm{Gal}(K/M)\hookrightarrow\mathrm{Aut}(C^0_M)/\langle w\rangle,
\end{aligned}$$

where w denotes the hyperelliptic involution of C^0.

Lemma 4.13. *The groups G_{C^0}, $G_{C^0}/\langle(w,1)\rangle$, $\mathrm{Aut}(C^0_M)$, and $\mathrm{Aut}(C^0_M)/\langle w\rangle$ are as follows*:

C^0	G_{C^0}	$G_{C^0}/\langle(w,1)\rangle$	$\mathrm{Aut}(C^0_M)$	$\mathrm{Aut}(C^0_M)/\langle w\rangle$
C^0_2	$\langle 96, 193\rangle$	$\langle 48, 48\rangle$	$\langle 48, 29\rangle$	$\langle 24, 12\rangle$
C^0_3	$\langle 48, 38\rangle$	$\langle 24, 14\rangle$	$\langle 24, 8\rangle$	$\langle 12, 4\rangle$

Proof. We show how to compute G_{C^0} and $\mathrm{Aut}(C^0_M)$; the respective quotients are then easily obtained. Recall that if $C/\mathbb{Q}$ is a genus-2 curve, given by a hyperelliptic equation $y^2 = f(x)$, where $f(x) \in \mathbb{Q}[x]$, then for any $\alpha \in \mathrm{Aut}(C_{\overline{\mathbb{Q}}})$, there exist $m, n, p, q \in \overline{\mathbb{Q}}$ such that

$$\alpha(x, y) = \left(\frac{mx+n}{px+q}, \frac{mq-np}{(px+q)^3}y\right); \tag{4-7}$$

see, for example, [Cardona 2006]. Let

$$\iota(\alpha) := \begin{pmatrix} m & n \\ p & q \end{pmatrix}.$$

The map $\iota : \mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \to \mathrm{GL}_2(\overline{\mathbb{Q}})$ that sends α to $\iota(\alpha)$ is a $G_{\mathbb{Q}}$-equivariant monomorphism. For $d = 2, 3$, we have $\mathrm{Aut}((C^0_d)_M) = \langle U_d, V_d\rangle$, where

$$U_2 = \frac{1}{2}\begin{pmatrix} \sqrt{-2}-1 & 1 \\ 1 & 1+\sqrt{-2} \end{pmatrix}, \quad V_2 = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{-2}+1 \\ -1-\sqrt{-2} & 1 \end{pmatrix}.$$

$$U_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad V_3 = \frac{1}{2}\begin{pmatrix} 0 & -1+\sqrt{-3} \\ 1+\sqrt{-3} & 0 \end{pmatrix},$$

One can readily check that U_d and V_d represent automorphisms of $(C^0_d)_M$, and that they generate a group of order 48 if $d = 2$ and of order 24 if $d = 3$, which are known to be the orders of $\mathrm{Aut}((C^0_d)_M)$. With this explicit representation, the isomorphism type of $\langle U_d, V_d\rangle$ is then easily determined by a computer algebra system. The group $G_{C^0_d}$ is then determined by explicitly computing the semidirect product $\langle U_d, V_d\rangle \rtimes \mathrm{Gal}(\mathbb{Q}(\sqrt{-d})/\mathbb{Q})$. $\square$

Remark 4.14. We note that $\langle 48, 29\rangle \simeq \tilde{\mathrm{S}}_4 \simeq \mathrm{GL}_2(\mathbb{F}_3)$ and $\langle 24, 8\rangle \simeq 2\mathrm{D}_6 \simeq \mathrm{C}_3 \rtimes \mathrm{D}_4$ are the two automorphism groups mentioned in the introduction, with quotients $\langle 24, 12\rangle \simeq \mathrm{S}_4$ and $\langle 12, 4\rangle \simeq \mathrm{D}_6$, respectively. The group $\langle 24, 14\rangle$ is isomorphic to $\mathrm{D}_6 \times \mathrm{C}_2$, while the groups $\langle 48, 38\rangle$ and $\langle 48, 48\rangle$ are both degree-3 extensions of $\mathrm{D}_4 \times \mathrm{C}_2$ and $\langle 96, 193\rangle$ is a degree-3 extension of $\mathrm{C}_8 \rtimes \mathrm{Aut}(\mathrm{C}_8)$.

Let $\tilde{T}(C)$ denote the triple

$$\big(\lambda_\phi(\mathrm{Gal}(L/\mathbb{Q})), \lambda_\phi(\mathrm{Gal}(L/M)), \lambda_\phi(\mathrm{Gal}(L/M))\big)$$

in $G_{C^0} \times G_{C^0} \times G_{C^0}$. Since λ_ϕ is injective, the conjugacy class of $\tilde{T}(C)$ determines $T(C)$ and $z(C)$, where $z(C)$ is the vector in Definition 1.3. In order to bound the number of possibilities for $T(C)$ and $z(C)$, we first bound the number of possible triples $\tilde{T}(C)$, up to conjugation.

Lemma 4.15. *Let H, N, and H_0 be subgroups of G_{C^0}. If $\tilde{T}(C) = (H, N, H_0)$, then the following conditions must be satisfied*:

(i) *H_0 and $H \cap \mathrm{Aut}(C^0_M) \times \langle 1 \rangle$ coincide and have order $|H|/2$.*

(ii) *N and $\langle (w, 1) \rangle \cap H_0$ coincide.*

Proof. Let $\mathrm{Gal}(M/\mathbb{Q}) = \{1, \tau\}$. Then

$$H_0 = \lambda_\phi(\mathrm{Gal}(L/M)) \subseteq (\mathrm{Aut}(C^0_M) \times \{1\}) \cap H,$$
$$H_1 := \lambda_\phi\big(\mathrm{Gal}(L/\mathbb{Q}) \setminus \mathrm{Gal}(L/M)\big) \subseteq (\mathrm{Aut}(C^0_M) \times \{\tau\}) \cap H.$$

The injectivity of λ_ϕ implies that $|H_0| = |H_1| = |H|/2$, and (i) follows from the fact that $H = H_0 \sqcup H_1$.

Proving (ii) is equivalent to showing that $(w, 1)$ lies in the image of λ_ϕ if and only if $[L : K] = 2$. But this has already been proved; see Proposition 4.7. □

Proposition 4.16. *Let H, N, and H_0 be subgroups of G_{C^0} that satisfy conditions* (i) *and* (ii) *of Lemma 4.15.*

(i) *For $C^0 = C^0_2$ (resp. $C^0 = C^0_3$), there are* 27 (*resp.* 38) *possibilities for the conjugacy class of (H, N, H_0) in $G_{C^0} \times G_{C^0} \times G_{C^0}$.*

(ii) *For $C^0 = C^0_2$ (resp. $C^0 = C^0_3$), the* 27 (*resp.* 38) *possibilities for the conjugacy class of (H, N, H_0) give rise to the* 23 (*resp.* 23) *isomorphism classes $[H, H/N, H_0]$ and vectors $z(H, N, H_0)$ listed in the top* (*resp. bottom*) *half of Table 4. Moreover*, *$[H, H/N, H_0]$ and $z(H, N, H_0)$ determine each other uniquely.*

(iii) *For $C^0 = C^0_2$ (resp. $C^0 = C^0_3$), the triple $T(C)$ and the vector $z(C)$ must be among those listed in the corresponding half of Table 4, and $T(C)$ and $z(C)$ determine each other uniquely.*

Proof. For (i), recall that $G_{C^0_3} \simeq \langle 48, 38 \rangle$ and $\mathrm{Aut}((C^0_3)_M) \simeq \langle 24, 8 \rangle$. The following three facts permit us to work with $G_{C^0_3}$ and $\mathrm{Aut}((C^0_3)_M)$ as abstract groups. First, there are exactly two subgroups A_1 and A_2 of $\langle 48, 38 \rangle$ isomorphic to $\langle 24, 8 \rangle$. Second, there is a unique nontrivial central involution $\hat{w}$ in $\langle 48, 38 \rangle$, and it lies in both A_1 and A_2. Third, consider the two lists of triples of groups, up to conjugation,

$$\mathscr{L}_i = \big\{\big(H, \langle \hat{w} \rangle \cap H, H \cap A_i\big) \mid H \subseteq \langle 48, 38 \rangle, |H \cap A_i| = |H|/2\big\}/\sim, \quad i = 1, 2,$$

where $(H, \langle \hat{w} \rangle \cap H, H \cap A_i) \sim (H', \langle \hat{w} \rangle \cap H', H' \cap A_i)$ if H and H' are conjugated in $\langle 48, 38 \rangle$. Then the lists $\mathscr{L}_1$ and $\mathscr{L}_2$ coincide; write $\mathscr{L}$ for this list. For $C^0 = C^0_2$, the three previous facts can be checked to hold *verbatim* when replacing $\langle 48, 38 \rangle$ and $\langle 24, 8 \rangle$ by $\langle 96, 193 \rangle$ and $\langle 48, 29 \rangle$, respectively. For $C^0 = C^0_3$, $\mathscr{L}$ has 38 elements and, for $C^0 = C^0_2$, it has 27 elements.

For (ii), for each of C_2^0 and C_3^0, we enumerate the triples (H, N, H_0) in $\mathscr{L}$ and explicitly compute $[H, H/N, H_0]$ and $z(H, N, H_0)$ in each case using a computer algebra system (we used Magma), obtaining the values listed in Table 4. One then checks that $[H, H/N, H_0] = [H', H'/N', H_0']$ if and only if $z(H, H/N, H_0) = z(H', H'/N', H_0')$.

Statement (iii) follows immediately from (ii) and Lemma 4.15. □

Proposition 4.17. *The vector $z(C)$ and the triple $T(C)$ both uniquely determine the Sato–Tate group* $\mathrm{ST}(\mathrm{Jac}(C))$.

Proof. The 18 Sato–Tate groups G that can occur over $\mathbb{Q}$ with $G^0 \simeq \mathrm{U}(1)$ (see [Fité et al. 2012, Theorem 4.3]) are uniquely determined by the combination of:

(a) the isomorphism classes of the groups G/G^0 and $G^{ns}/G^{ns,0}$,

(b) the vector $z_2(G) = \big(z_{2,2}(G), z_{2,-2}(G), z_{2,-1}(G), z_{2,0}(G), z_{2,1}(G)\big)$,[9]

where G^{ns} is the index-2 subgroup of G obtained by removing from G those components all of whose elements have a constant characteristic polynomial.

On the one hand, the isomorphism classes of the groups G/G^0 and $G^{ns}/G^{ns,0}$ are determined by $T(C)$, since $G/G^0 \simeq \mathrm{Gal}(K/\mathbb{Q})$ and $G^{ns}/G^{ns,0} \simeq \mathrm{Gal}(K/M)$. On the other hand, $z_2(G)$ is determined by $z(C)$; indeed, it follows from the construction of the Sato–Tate group in terms of the image of the ℓ-adic representation attached to $\mathrm{Jac}(C)$ and from assertion (iii) of Proposition 4.9, that $z_2(G) \cdot [L : K] = z_2(C)$. □

Corollary 4.18. *For each triple $[H, H/N, H_0]$ in Table 4, there exists a twist C of C^0 such that $T(C) = [H, H/N, H_0]$ if and only if the corresponding row in the table is not marked with an asterisk. Thus, for $C^0 = C_2^0$ (resp. for $C^0 = C_3^0$) there are exactly* 20 (*resp.* 21) *possibilities for $T(C)$.*

Proof. Observe that the triples marked with an asterisk in Table 4 correspond to Sato–Tate groups (equivalently, Galois types) that cannot arise for abelian surfaces defined over $\mathbb{Q}$ (see [Fité et al. 2012, Proposition 4.11]). For each of the triples $[H, H/N, H_0]$ that is not marked with an asterisk, a curve C with $T(C) = [H, H/N, H_0]$ is exhibited in Tables 5 and 6 (for details on how the curves have been found, see Section 5A; for details on how $T(C)$ is computed for each of the curves, see Section 5C.) □

Remark 4.19. If the triple $[H, H, H_0]$ appears in either half of Table 4, then so does the triple $[H \times \mathrm{C}_2, H, H_0 \times \mathrm{C}_2]$. In other words, if there exists a twist C of C^0 such that $\mathrm{Gal}(L/\mathbb{Q}) = \mathrm{Gal}(K/\mathbb{Q})$, then there exists a twist C' of C^0 such that $\mathrm{Gal}(K'/\mathbb{Q}) = \mathrm{Gal}(K/\mathbb{Q})$ and $\mathrm{Gal}(L'/\mathbb{Q}) \simeq \mathrm{Gal}(K/\mathbb{Q}) \times \mathrm{C}_2$. Here K' (resp. L') is

[9]Following the notation of [Fité et al. 2012], recall that $z_{2,i}(G)$ denotes the number of connected components of G all of whose elements have a constant characteristic polynomial, for which the coefficient of the quadratic term is equal to i. Note that the components of the vector $z_2(G)$ have been permuted with respect to the definition of $z_2(G)$ given in [Fité et al. 2012].

the minimal field over which all the automorphisms of C' (resp. all the isomorphisms between C' and C^0) are defined. Indeed, if C is given by the hyperelliptic equation $y^2 = f(x)$, let C' be the curve given by $dy^2 = f(x)$, where $d \in \mathbb{Q}^*$ is not a square in K. We will use this remark in Section 5 for the computation of some of the curves.

Remark 4.20. Among the 18 Sato–Tate groups with identity component U(1) that can occur over $\mathbb{Q}$, there are 13 that are subgroups of $J(O)$ and 11 that are subgroups of $J(D_6)$ (6 are subgroups of both). From Table 6, we see that the 13 that are subgroups of $J(O)$ can all occur as $\mathbb{Q}$-twists of C_2^0, and the 11 that are subgroups of $J(D_6)$ can all occur as $\mathbb{Q}$-twists of C_3^0.

5. Numerical computations

We now describe the methods used to obtain the example curves C listed in Tables 5 and 6. As in Section 4, each curve C is a $\mathbb{Q}$-twist of $C^0 = C_d^0$, for $d = 2, 3$, where $\mathrm{Jac}(C_d^0) \sim (E_d^0)^2$ and E_d^0 is an elliptic curve with CM by $M = \mathbb{Q}(\sqrt{-d})$. For $d = 2$, we list 20 curves C that are $\mathbb{Q}$-twists of the curve C_2^0 defined by $y^2 = x^6 - 5x^4 - 5x^2 + 1$, realizing every possible triple

$$T(C) = \big[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)\big]$$

that can occur when C is a $\mathbb{Q}$-twist of C_2^0. Recall that the fields K and L are the minimal fields of definition $\mathrm{End}(\mathrm{Jac}(C)_{\overline{\mathbb{Q}}})$ and $\mathrm{Hom}(\mathrm{Jac}(C)_{\overline{\mathbb{Q}}}, E_{\overline{\mathbb{Q}}})$, respectively, as in Definition 1.2. Similarly, for $d = 3$, we list 21 curves C that are twists of the curve C_3^0 defined by $y^2 = x^6 + 1$, realizing every possible triple $T(C)$ that can occur when C is a $\mathbb{Q}$-twist of C_3^0.

For each of the two curves C^0, we followed the procedure outlined below:

(1) Generate a large set S of $\mathbb{Q}$-twists of C^0.

(2) For each $C \in S$, compute a provisional value of the triple $T(C)$.

(3) Select a single representative C for each distinct triple $T(C)$ and then verify the provisional value of $T(C)$ by explicitly computing the fields K and L and the triple $T(C) = \big[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)\big]$.

The purpose of the "provisional" computation of $T(C)$ in step (2) is to avoid computing the fields K and L for all of the curves in S, which would have been infeasible. Explicit computation of the fields K and L (and their Galois groups) for even a single curve C can be quite time-consuming, taking hours or even days of computer time, and the sets S that we used contained tens of thousands of curves.

In the rest of this section we fill in some of the details of the three steps listed above.

5A. ***Generating twists of C^0.*** Explicit parametrizations of the families of twists of C_2^0 and C_3^0 are given in [Cardona 2001; 2006]. One can easily obtain a large set S using these parametrizations. However, the resulting curves tend to have large coefficients, making the computation of K and L more difficult, and the vast majority of curves in S are likely to represent the generic case, where $\mathrm{Gal}(K/\mathbb{Q})$ and $\mathrm{Gal}(L/\mathbb{Q})$ are as large as possible. In principle, one can control the isomorphism type of $\mathrm{Gal}(K/\mathbb{Q})$ by placing appropriate constraints on the input parameters, but this is not enough to determine the Sato–Tate group, and it gives no control over $\mathrm{Gal}(L/\mathbb{Q})$.

We instead adapted the search method used in [Fité et al. 2012], generating S by enumerating all curves of the form $y^2 = \sum_{i=0}^{6} c_i x^i$ satisfying coefficient bounds $|c_i| \le B_i$. To quickly identify curves C that are twists of C^0, we first compute $a_1(C)(p)$ for a handful of small primes p that are inert in M, and immediately discard C if $a_1(C)(p) \neq 0$ for any such p. We then compute the absolute Igusa invariants of C, and compare them to the corresponding values for C^0. With the bounds B_i chosen to encompass some 2^{50} curves with small coefficients, we obtain a set S containing tens of thousands of twists of C^0 in each case.

After applying the method in Section 5B below to all of the curves in S, we had several candidate curves C for every possible triple $T(C)$ that can arise when C is defined over $\mathbb{Q}$ (the triples listed in Table 4 that are not marked with an asterisk). We then selected a single representative C for each triple and computed K and L for each of these C, as described in Section 5C, and then computed the Galois groups $\mathrm{Gal}(L/\mathbb{Q})$, $\mathrm{Gal}(K/\mathbb{Q})$, and $\mathrm{Gal}(L/M)$, using the Magma computer algebra system, to obtain the true value of the triple $T(C)$. As expected, this computation confirmed the provisional value in every case. Indeed, in all but the most time-consuming cases we were able to repeat the computations using several different candidate curves C and always obtained the expected value of $T(C)$.

Remark 5.1. The computation of the triple $T(C)$ in Magma is completely independent of the calculations used to obtain a provisional value for $T(C)$, which were performed using the `smalljac` software library [Sutherland 2011]; the purpose of the provisional computations was simply to obtain a set of candidate curves that is much smaller than the initial set S. The fact that in every case we obtained the same value for $T(C)$ using two completely different methods gives us a high degree of confidence in our numerical computations.

5B. ***Provisional computation of $T(C)$.*** To provisionally identify the triple $T(C)$, we compute an approximation of the vector $z(C)$ (see Definition 1.3), which, by Theorem 1.4, uniquely determines $T(C)$. To do this, it suffices to determine the triples (s, r, t) of residue degrees $(f_L(p), f_K(p), f_M(p))$ for a sample set of primes p (say, primes $p \le 2^{16}$ of good reduction for C), and then count how often each triple appears. The components $o(s, r)$ and $\bar{o}(s, r)$ of the vector $z(C)$ may be approximated by computing the relative frequencies of the triples $(s, r, 1)$ and $(s, r, 2)$, respectively, and normalizing so that $o(1, 1) = 1$.

We can easily compute $t = f_M(p) \in \{1, 2\}$ by checking whether p splits in M, but we also need to compute $r = f_K(p)$ and $s = f_L(p)$, and we would like to do so *without knowing K or L*. This can be achieved as follows: we first compute $a_1(E^0)(p)$ and the values $a_1(C)(p)$ and $a_2(C)(p)$, as described in Section 5B1, and then determine the unique map $F_{(s,r,t)}$ from Proposition 4.9 for which

$$F_{(s,r,t)}(a_1(E^0)(p)) = \big(\pm a_1(C)(p), a_2(C)(p)\big). \tag{5-1}$$

5B1. *Computation of $a_1(C)(p)$ and $a_2(C)(p)$.* For an arbitrary genus-2 curve, efficient computation of $a_1(C)(p)$ and $a_2(C)(p)$ is addressed in [Kedlaya and Sutherland 2008], but in the special case of interest here, where C is a $\mathbb{Q}$-twist of C^0, we use a faster approach. The Jacobian of C^0 is $\mathbb{Q}$-isogenous to the square of E^0, an elliptic curve defined over $\mathbb{Q}$. Because E^0 has complex multiplication, we can very efficiently determine $a_1(E)(p)$. Taking C_2^0 as an example, E_2^0 is defined by the Weierstrass equation $y^2 = x^3 - 5x^2 - 5x + 1$. This curve has CM by $M = \mathbb{Q}(\sqrt{-2})$, and for any prime $p > 2$ we may compute $a = a_1(E_2^0)(p)$ as follows: $a = 0$ if p is inert in M and otherwise $a = 4x/\sqrt{p}$, where the integer x satisfies $p = x^2 + 2y^2$ for some integer y. The positive integer $z = |x|$ may be determined via Cornacchia's algorithm, and then $x = (-1)^\epsilon z$, where $\epsilon = (z-1)/2 + (p-1)(p+5)/16$; see [Rubin and Silverberg 2010] for details. The computation for C_3^0 is similar: in this case E_3^0 is defined by $y^2 = x^3 + 1$, with CM by $M = \mathbb{Q}(\sqrt{-3})$.

With $a_1(E^0)(p)$ computed, there are only a handful of pairs (a_1, a_2) that are compatible with (5-1), that is, for which there exists a triple (s, r, t) such that $F_{(s,r,t)}(a_1(E^0)(p)) = (\pm a_1, a_2)$. Taking into account whether C is a twist of C_2^0 or C_3^0, whether p splits in M or not, and that the sign of a_1 is actually ambiguous in only 2 cases, there are at most 8 possibilities. Each compatible pair (a_1, a_2) determines an integer

$$n = p^2 + p^{3/2}a_1 + pa_2 + p^{1/2}a_1 + 1,$$

one of which is equal to $\#\operatorname{Jac}(C)(\mathbb{F}_p)$. In most cases, if we pick a random point $P \in \operatorname{Jac}(C)(\mathbb{F}_p)$, the equation $nP = 0$ will hold for exactly one n and uniquely determine a_1 and a_2. Even when this is not the case, after factoring the integers n, we can determine the order of any point P in $\operatorname{Jac}(C)(\mathbb{F}_p)$, using just $\tilde{O}(\log p)$ operations in $\mathbb{F}_p$; see [Sutherland 2007, Chapter 7]. This allows us to compute the order of $\operatorname{Jac}(C)(\mathbb{F}_p)$ using a probabilistic generic group algorithm (of Las Vegas type) that runs in $O(p^{1/4})$ expected time; see [Sutherland 2007; Kedlaya and Sutherland 2008, Proposition 1].[10] This compares to an $O(p^{3/4})$ expected running

[10] The $O(p^{1/4})$ bound is a worst-case estimate; it is faster than this for most p.

time for an arbitrary genus-2 curve using a generic group algorithm.[11]

Having computed $L_p(C, 1) = \#\,\mathrm{Jac}(C)(\mathbb{F}_p)$, we use the same method to determine $L_p(C, -1) = \#\,\mathrm{Jac}(\tilde{C})(\mathbb{F}_p)$, where $\tilde{C}$ is any nontrivial quadratic twist of C over $\mathbb{F}_p$, and these two values uniquely determine a_1 and a_2.

The algorithm described above is included in the most recent version of the `smalljac` software library, whose source code is available at [Sutherland 2011].

5C. ***Computation of K and L.*** In this section, we describe the procedure used to compute the fields K and L for the curves C listed in Tables 5 and 6.

For the field K, its characterization in Lemma 4.2 as the minimal field over which all the automorphisms of C are defined turns out to be the most computationally effective. For all 41 curves $C\colon y^2 = f(x)$ listed in Tables 5 and 6, one readily checks that $\mathrm{Aut}(C^0_{\overline{\mathbb{Q}}}) \simeq \mathrm{Aut}(C_{\overline{\mathbb{Q}}}) = \mathrm{Aut}(C_{F(\zeta_{24})})$, where F is the splitting field of $f(x)$ (see Remark 5.2 below). It is then a finite problem to identify the minimal subfield K of $F(\zeta_{24})$ for which $\mathrm{Aut}(C_K) = \mathrm{Aut}(C_{F(\zeta_{24})})$.

Having computed K, we determine L as follows. For any nonhyperelliptic involution $\beta \in \mathrm{Aut}(C^0_M)$, the elliptic quotient $C^0/\langle\beta\rangle$ is defined over M. If β_1 and β_2 are conjugate in $\mathrm{Aut}(C^0_M)$, then $C^0/\langle\beta_1\rangle \simeq C^0/\langle\beta_2\rangle$. For C^0_2 there is just one conjugacy class of nonhyperelliptic involutions; hence in this case every elliptic quotient $C^0/\langle\beta\rangle$ is isomorphic to E^0_M. For C^0_3 there are two conjugacy classes of nonhyperelliptic involutions, of size 2 and 6 (see Table 2). The first corresponds to the M-isomorphism class of E^0_3, and the second corresponds to the M-isomorphism class of the elliptic curve $y^2 = x^3 - 15x + 22$.

Since we know K explicitly, we can compute $\mathrm{Aut}(C_K)$ and enumerate all the nonhyperelliptic involutions α (there are 12 when $d = 2$ and 8 when $d = 3$). For $d = 2$ we pick any α, and for $d = 3$ we pick α from the conjugacy class of size 2. Define $\tilde{E} := C_K/\langle\alpha\rangle$ and $\tilde{E}^0 := C^0_M/\langle\phi\alpha\phi^{-1}\rangle$. The isomorphism ϕ induces an isomorphism $\tilde{\phi} : \tilde{E}_L \to \tilde{E}^0_L$. As in the proof of Proposition 4.6, L is the compositum of K and the minimal field over which $\tilde{\phi}$ is defined. Our choice of α ensures that $\tilde{E}^0 \simeq E^0_M$; thus $\tilde{E}_L \simeq E^0_L$.

By applying [Cardona et al. 1999, Lemma 2.2], we can compute an explicit Weierstrass equation for $\tilde{E}$ of the form

$$\tilde{E}\colon\ Y^2 = X^3 + AX + B, \quad \text{with } A, B \in K.$$

Writing E^0 in the form $Y^2 = X^3 + UX + V$, there then exists $\gamma \in L$ such that $U = \gamma^4 A$ and $V = \gamma^6 B$, and γ generates L as an (at most quadratic) extension of K. We can easily derive γ from the coefficients A, B, U, and V.

[11] As noted in [Kedlaya and Sutherland 2008], the asymptotically faster polynomial-time algorithm of Pila [1990] is not practically useful in the range of p relevant to the computations considered here.

Remark 5.2. In fact, it is true in general that for any twist C of C_2^0 (resp. C_3^0), the field K is contained in $F(\sqrt{-2})$ (resp. $F(\sqrt{-3}, i)$). We thank J. Quer for kindly providing the following argument.

Let $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})^*$ denote the subgroup of $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ generated by those elements α such that $\mathrm{Trace}(\iota(\alpha))$ is nonzero. We claim that

$$\mathrm{Aut}(C_{\overline{\mathbb{Q}}})^* = \mathrm{Aut}(C_{FM})^*.$$

Let $\mathrm{WP}(C)$ denote the set of Weierstrass points of C and let σ be an element of G_{FM}. It suffices to show that ${}^\sigma\alpha = \alpha$ for every α in $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})^*$ such that $\mathrm{Trace}(\iota(\alpha))$ is nonzero. Observe that for every P in $\mathrm{WP}(C)$, one has ${}^\sigma P = P$. Then, writing $Q = \alpha^{-1}(P)$, we have

$${}^\sigma\alpha \circ \alpha^{-1}(P) = ({}^\sigma\alpha)(Q) = {}^\sigma(\alpha(Q)) = {}^\sigma P = P,$$

which implies that ${}^\sigma\alpha$ is either α or $w\alpha$, since the action of $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})/\langle w\rangle$ on $\mathrm{WP}(C)$ is faithful. Provided that $\mathrm{Trace}(\iota(\alpha))$ is in M, the latter option is not possible, since otherwise we would have

$$\mathrm{Trace}(\iota(\alpha)) = {}^\sigma\,\mathrm{Trace}(\iota(\alpha)) = \mathrm{Trace}(\iota(w\alpha)) = -\,\mathrm{Trace}(\iota(\alpha)),$$

contradicting the fact that $\mathrm{Trace}(\iota(\alpha))$ is nonzero. Since $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})$ and $\mathrm{Aut}(C^0_{\overline{\mathbb{Q}}})$ are conjugated, the groups $\mathrm{Aut}(C_{\overline{\mathbb{Q}}})^*$ and $\mathrm{Aut}(C^0_{\overline{\mathbb{Q}}})^*$ are isomorphic. It is straightforward to check that

$$\mathrm{Aut}\big((C_2^0)_{\overline{\mathbb{Q}}}\big)^* \simeq \tilde{\mathrm{S}}_4 \quad \text{and} \quad \mathrm{Aut}\big((C_3^0)_{\overline{\mathbb{Q}}}\big)^* \simeq \mathrm{C}_2 \times \mathrm{C}_6.$$

Thus, for every twist C of C_2^0, the field K is contained in $F(\sqrt{-2})$; but for a twist C of C_3^0, the order of $\mathrm{Aut}\big((C_3^0)_{F(\sqrt{-3})}\big)$ can be 12 or 24. By considering the parametrizations given in [Cardona 2001, Proposition 7.4.1] of all the twists C of C_3^0 as well as of the corresponding embeddings $\iota(\mathrm{Aut}(C_{\overline{\mathbb{Q}}}))$ in $\mathrm{GL}_2(\overline{\mathbb{Q}})$, one may explicitly verify that K is always contained in $F(\sqrt{-3}, i)$.

5C1. *An example.* Consider the twist C of C_3^0 defined by the hyperelliptic equation

$$y^2 = f(x) = x^6 + 15x^4 + 20x^3 + 30x^2 + 18x + 5$$

over $\mathbb{Q}$. This curve is listed in Table 6 for the triple $[\langle 24, 5\rangle, \langle 12, 4\rangle, \langle 12, 1\rangle]$. Let us prove that this is in fact the triple $T(C) = \big[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)\big]$.

We first compute K. Let F denote the splitting field of $f(x)$. One checks (via Magma) that $|\mathrm{Aut}(C_{MF})| = 24$, where $M = \mathbb{Q}(\sqrt{-3})$, and therefore $K \subseteq MF$ (since we know *a priori* that $|\mathrm{Aut}(C_K)| = |\mathrm{Aut}((C_3^0)_{\overline{\mathbb{Q}}})| = 24$). By enumerating the various subfields of MF, we find that the minimal subfield K of MF for which $|\mathrm{Aut}(C_K)| = 24$ is $K = M(\sqrt{5}, a)$, where $a^3 + 3a - 1 = 0$.

To compute L, we choose the nonhyperelliptic involution α of $\mathrm{Aut}(C_K)$ whose image under the map $\iota : \mathrm{Aut}(C_{\overline{\mathbb{Q}}}) \to \mathrm{GL}_2(\overline{\mathbb{Q}})$ defined in (4-7) is

$$\iota(\alpha) = \frac{1}{5}\begin{pmatrix} \sqrt{5} & -2\sqrt{5} \\ -2\sqrt{5} & -\sqrt{5} \end{pmatrix}.$$

Applying [Cardona et al. 1999, Lemma 2.2] yields a Weierstrass equation for $\tilde{E} = C/\langle\alpha\rangle$:

$$\tilde{E}\colon\ Y^2 = X^3 + B, \quad \text{with } B = -\frac{11}{97656250}\sqrt{5} + \frac{1}{3906250}.$$

Since E^0 is the curve $y^2 = x^3 + 1$, we have $U = 0$ and $V = 1$, so $\gamma^6 = 1/B$. This implies that

$$\gamma^2 - \left(\frac{125}{2}\sqrt{5} + \frac{375}{2}\right)a^2 + \left(\frac{125}{2}\sqrt{5} + \frac{125}{2}\right)a - 125\sqrt{5} - 375 = 0,$$

and one finds that $L = K\bigl(\sqrt{2\sqrt{5}+10}\bigr)$.

Having explicitly computed the fields K and L, it is then straightforward to verify that $\mathrm{Gal}(L/\mathbb{Q}) \simeq \langle 24, 5\rangle$, $\mathrm{Gal}(K/\mathbb{Q}) \simeq \langle 12, 4\rangle$, and $\mathrm{Gal}(L/M) \simeq \langle 12, 1\rangle$ using Magma.

6. Tables

This section contains the remaining tables described earlier, whose definitions we briefly recall. Remember that C^0 is one of the two curves C_2^0: $y^2 = x^6 - 5x^4 - 5x^2 + 1$ (in which case $M = \mathbb{Q}(\sqrt{-2})$) or C_3^0: $y^2 = x^6 + 1$ (in which case $M = \mathbb{Q}(\sqrt{-3})$). Table 4 lists (up to isomorphism) the possible values of the triples $T(C)$ that can arise when C is a $\mathbb{Q}$-twist of the curve C^0.

Section 4C describes the computation of these tables. Each triple $[H, H/N, H_0]$ is a possible value for $T(C) = \bigl[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}, \mathrm{Gal}(L/M)\bigr]$, and is determined by a subgroup $H \subset G_{C^0}$ whose intersection with $\mathrm{Aut}(C^0_M)$ is an index-2 subgroup H_0 of H, where $N = H \cap Z(G_{C^0})$.

For each triple $T(C)$ we list the corresponding Sato–Tate group G and its matching Galois type, as defined in [Fité et al. 2012], as well as the vector $z(C)$ given by Definition 1.3, all of which are uniquely determined by $T(C)$, by Theorem 1.4. As proven in [Fité et al. 2012], the Sato–Tate groups $J(C_1)$, $J(C_3)$, and $C_{4,1}$ cannot arise for a genus-2 curve defined over $\mathbb{Q}$, and the corresponding rows in Table 4 are marked with an asterisk.

In Tables 5 and 6, we list representative curves that realize every triple $T(C)$ that can occur when C is defined over $\mathbb{Q}$. For each curve, we also give an explicit description of the fields K and L, where K is the minimal field for which $\mathrm{Aut}(C_K) = \mathrm{Aut}(C_{\overline{\mathbb{Q}}})$, and L is the minimal extension of K over which C is isomorphic to C^0.

G	H	H/N	H_0	Galois type	$z(H, N, H_0)$
$*J(C_1)$	$\langle 4,1\rangle$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathbb{H}]$	$[1,1,0,0,0,0,0,0,0,0,2,0,0,0]$
$J(C_2)$	$\langle 8,2\rangle$	$\langle 4,2\rangle$	$\langle 4,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathbb{H}]$	$[1,1,0,0,2,0,0,0,0,2,2,0,0,0]$
$J(C_2)$	$\langle 8,3\rangle$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathbb{H}]$	$[1,1,2,0,0,0,0,0,0,2,2,0,0,0]$
$*J(C_3)$	$\langle 12,2\rangle$	$\langle 6,2\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{C}_6,\mathrm{C}_3,\mathbb{H}]$	$[1,1,0,2,0,2,0,0,0,0,2,0,0,4]$
$J(C_4)$	$\langle 16,6\rangle$	$\langle 8,2\rangle$	$\langle 8,1\rangle$	$\mathbf{F}[\mathrm{C}_4\times\mathrm{C}_2,\mathrm{C}_4]$	$[1,1,0,0,2,0,0,4,0,2,2,0,4,0]$
$J(D_2)$	$\langle 16,11\rangle$	$\langle 8,5\rangle$	$\langle 8,3\rangle$	$\mathbf{F}[\mathrm{D}_2\times\mathrm{C}_2,\mathrm{D}_2]$	$[1,1,4,0,2,0,0,0,0,6,2,0,0,0]$
$J(D_2)$	$\langle 16,13\rangle$	$\langle 8,5\rangle$	$\langle 8,4\rangle$	$\mathbf{F}[\mathrm{D}_2\times\mathrm{C}_2,\mathrm{D}_2]$	$[1,1,0,0,6,0,0,0,0,6,2,0,0,0]$
$J(D_3)$	$\langle 24,6\rangle$	$\langle 12,4\rangle$	$\langle 12,4\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathbb{H}]$	$[1,1,6,2,0,2,0,0,0,6,2,0,0,4]$
$J(D_4)$	$\langle 32,43\rangle$	$\langle 16,11\rangle$	$\langle 16,8\rangle$	$\mathbf{F}[\mathrm{D}_4\times\mathrm{C}_2,\mathrm{D}_4]$	$[1,1,4,0,6,0,0,4,0,10,2,0,4,0]$
$J(T)$	$\langle 48,33\rangle$	$\langle 24,13\rangle$	$\langle 24,3\rangle$	$\mathbf{F}[\mathrm{A}_4\times\mathrm{C}_2,\mathrm{A}_4]$	$[1,1,0,8,6,8,0,0,0,6,2,0,0,16]$
$J(O)$	$\langle 96,193\rangle$	$\langle 48,48\rangle$	$\langle 48,29\rangle$	$\mathbf{F}[\mathrm{S}_4\times\mathrm{C}_2,\mathrm{S}_4]$	$[1,1,12,8,6,8,0,12,0,18,2,0,12,16]$
$C_{2,1}$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\langle 1,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathrm{M}_2(\mathbb{R})]$	$[1,0,0,0,0,0,0,0,0,1,0,0,0,0]$
$C_{2,1}$	$\langle 4,2\rangle$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,0,0,0,0,0,0,2,0,0,0,0]$
$*C_{4,1}$	$\langle 8,1\rangle$	$\langle 4,1\rangle$	$\langle 4,1\rangle$	$\mathbf{F}[\mathrm{C}_4,\mathrm{C}_2]$	$[1,1,0,0,2,0,0,0,0,0,0,0,4,0]$
$D_{2,1}$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,0,1,0,0,0,0,0,0,2,0,0,0,0]$
$D_{2,1}$	$\langle 8,3\rangle$	$\langle 4,2\rangle$	$\langle 4,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,0,2,0,0,0,0,4,0,0,0,0]$
$D_{2,1}$	$\langle 8,5\rangle$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,1,2,0,0,0,0,0,0,4,0,0,0,0]$
$D_{3,2}$	$\langle 6,1\rangle$	$\langle 6,1\rangle$	$\langle 3,1\rangle$	$\mathbf{F}[\mathrm{D}_3,\mathrm{C}_3]$	$[1,0,0,2,0,0,0,0,0,3,0,0,0,0]$
$D_{3,2}$	$\langle 12,4\rangle$	$\langle 6,1\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{D}_3,\mathrm{C}_3]$	$[1,1,0,2,0,2,0,0,0,6,0,0,0,0]$
$D_{4,1}$	$\langle 16,7\rangle$	$\langle 8,3\rangle$	$\langle 8,3\rangle$	$\mathbf{F}[\mathrm{D}_4,\mathrm{D}_2]$	$[1,1,4,0,2,0,0,0,0,4,0,0,4,0]$
$D_{4,1}$	$\langle 16,8\rangle$	$\langle 8,3\rangle$	$\langle 8,4\rangle$	$\mathbf{F}[\mathrm{D}_4,\mathrm{D}_2]$	$[1,1,0,0,6,0,0,0,0,4,0,0,4,0]$
$D_{4,2}$	$\langle 16,7\rangle$	$\langle 8,3\rangle$	$\langle 8,1\rangle$	$\mathbf{F}[\mathrm{D}_4,\mathrm{C}_4]$	$[1,1,0,0,2,0,0,4,0,8,0,0,0,0]$
O_1	$\langle 48,29\rangle$	$\langle 24,12\rangle$	$\langle 24,3\rangle$	$\mathbf{F}[\mathrm{S}_4,\mathrm{A}_4]$	$[1,1,0,8,6,8,0,0,0,12,0,0,12,0]$
$*J(C_1)$	$\langle 4,1\rangle$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathbb{H}]$	$[1,1,0,0,0,0,0,0,0,0,2,0,0,0]$
$J(C_2)$	$\langle 8,2\rangle$	$\langle 4,2\rangle$	$\langle 4,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathbb{H}]$	$[1,1,0,0,2,0,0,0,0,2,2,0,0,0]$
$J(C_2)$	$\langle 8,3\rangle$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathbb{H}]$	$[1,1,2,0,0,0,0,0,0,2,2,0,0,0]$
$*J(C_3)$	$\langle 12,2\rangle$	$\langle 6,2\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{C}_6,\mathrm{C}_3,\mathbb{H}]$	$[1,1,0,2,0,2,0,0,0,0,2,0,0,4]$
$J(C_6)$	$\langle 24,10\rangle$	$\langle 12,5\rangle$	$\langle 12,5\rangle$	$\mathbf{F}[\mathrm{C}_6\times\mathrm{C}_2,\mathrm{C}_6]$	$[1,1,2,2,0,2,4,0,0,2,2,4,0,4]$
$J(D_2)$	$\langle 16,11\rangle$	$\langle 8,5\rangle$	$\langle 8,3\rangle$	$\mathbf{F}[\mathrm{D}_2\times\mathrm{C}_2,\mathrm{D}_2]$	$[1,1,4,0,2,0,0,0,0,6,2,0,0,0]$
$J(D_3)$	$\langle 24,5\rangle$	$\langle 12,4\rangle$	$\langle 12,1\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathbb{H}]$	$[1,1,0,2,6,2,0,0,0,6,2,0,0,4]$
$J(D_3)$	$\langle 24,6\rangle$	$\langle 12,4\rangle$	$\langle 12,4\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathbb{H}]$	$[1,1,6,2,0,2,0,0,0,6,2,0,0,4]$
$J(D_6)$	$\langle 48,38\rangle$	$\langle 24,14\rangle$	$\langle 24,8\rangle$	$\mathbf{F}[\mathrm{D}_6\times\mathrm{C}_2,\mathrm{D}_6]$	$[1,1,8,2,6,2,4,0,0,14,2,4,0,4]$
$C_{2,1}$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\langle 1,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathrm{M}_2(\mathbb{R})]$	$[1,0,0,0,0,0,0,0,0,1,0,0,0,0]$
$C_{2,1}$	$\langle 4,2\rangle$	$\langle 2,1\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{C}_2,\mathrm{C}_1,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,0,0,0,0,0,0,2,0,0,0,0]$
$C_{6,1}$	$\langle 6,2\rangle$	$\langle 6,2\rangle$	$\langle 3,1\rangle$	$\mathbf{F}[\mathrm{C}_6,\mathrm{C}_3,\mathrm{M}_2(\mathbb{R})]$	$[1,0,0,2,0,0,0,0,0,1,0,2,0,0]$
$C_{6,1}$	$\langle 12,5\rangle$	$\langle 6,2\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{C}_6,\mathrm{C}_3,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,2,0,2,0,0,0,2,0,4,0,0]$
$D_{2,1}$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\langle 2,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,0,1,0,0,0,0,0,0,2,0,0,0,0]$
$D_{2,1}$	$\langle 8,3\rangle$	$\langle 4,2\rangle$	$\langle 4,1\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,0,2,0,0,0,0,4,0,0,0,0]$
$D_{2,1}$	$\langle 8,5\rangle$	$\langle 4,2\rangle$	$\langle 4,2\rangle$	$\mathbf{F}[\mathrm{D}_2,\mathrm{C}_2,\mathrm{M}_2(\mathbb{R})]$	$[1,1,2,0,0,0,0,0,0,4,0,0,0,0]$
$D_{3,2}$	$\langle 6,1\rangle$	$\langle 6,1\rangle$	$\langle 3,1\rangle$	$\mathbf{F}[\mathrm{D}_3,\mathrm{C}_3]$	$[1,0,0,2,0,0,0,0,0,3,0,0,0,0]$
$D_{3,2}$	$\langle 12,4\rangle$	$\langle 6,1\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{D}_3,\mathrm{C}_3]$	$[1,1,0,2,0,2,0,0,0,6,0,0,0,0]$
$D_{6,1}$	$\langle 12,4\rangle$	$\langle 12,4\rangle$	$\langle 6,1\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathrm{M}_2(\mathbb{R})]$	$[1,0,3,2,0,0,0,0,0,4,0,2,0,0]$
$D_{6,1}$	$\langle 24,8\rangle$	$\langle 12,4\rangle$	$\langle 12,1\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathrm{M}_2(\mathbb{R})]$	$[1,1,0,2,6,2,0,0,0,8,0,4,0,0]$
$D_{6,1}$	$\langle 24,14\rangle$	$\langle 12,4\rangle$	$\langle 12,4\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{D}_3,\mathrm{M}_2(\mathbb{R})]$	$[1,1,6,2,0,2,0,0,0,8,0,4,0,0]$
$D_{6,2}$	$\langle 12,4\rangle$	$\langle 12,4\rangle$	$\langle 6,2\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{C}_6]$	$[1,0,1,2,0,0,2,0,0,6,0,0,0,0]$
$D_{6,2}$	$\langle 24,14\rangle$	$\langle 12,4\rangle$	$\langle 12,5\rangle$	$\mathbf{F}[\mathrm{D}_6,\mathrm{C}_6]$	$[1,1,2,2,0,2,4,0,0,12,0,0,0,0]$

Table 4. Triples for twists of C_2^0 (top) and C_3^0 (bottom).

G	$[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)]$	K	L
$\boldsymbol{y^2 = x^5 - x}$			
$J(C_2)$	$\langle 8, 2\rangle, \langle 4, 2\rangle, \langle 4, 1\rangle$	$M(i)$	$K\big(\sqrt{\sqrt{2}+2}\big)$
$\boldsymbol{y^2 = x^5 + 4x}$			
$J(C_2)$	$\langle 8, 3\rangle, \langle 4, 2\rangle, \langle 4, 2\rangle$	$M(i)$	$K(\sqrt[4]{2})$
$\boldsymbol{y^2 = x^6 + x^5 - 5x^4 - 5x^2 - x + 1}$			
$J(C_4)$	$\langle 16, 6\rangle, \langle 8, 2\rangle, \langle 8, 1\rangle$	$M(\sqrt{\sqrt{17}+17})$	$K\big(\sqrt{(\sqrt{17}+3)\sqrt{\sqrt{17}+17}-8\sqrt{17}}\big)$
$\boldsymbol{y^2 = x^5 + 9x}$			
$J(D_2)$	$\langle 16, 11\rangle, \langle 8, 5\rangle, \langle 8, 3\rangle$	$M(i, \sqrt{3})$	$K(\sqrt[4]{3})$
$\boldsymbol{y^2 = x^5 - 9x}$			
$J(D_2)$	$\langle 16, 13\rangle, \langle 8, 5\rangle, \langle 8, 4\rangle$	$M(i, \sqrt{3})$	$K(\sqrt[4]{3i})$
$\boldsymbol{y^2 = x^6 + 10x^3 - 2}$			
$J(D_3)$	$\langle 24, 6\rangle, \langle 12, 4\rangle, \langle 12, 4\rangle$	$M(\sqrt{-3}, \sqrt[3]{-2})$	$K\big(\sqrt{\sqrt{6}-2}\big)$
$\boldsymbol{y^2 = x^5 + 3x}$			
$J(D_4)$	$\langle 32, 43\rangle, \langle 16, 11\rangle, \langle 16, 8\rangle$	$M(i, \sqrt[4]{3})$	$K(\sqrt[8]{3})$
$\boldsymbol{y^2 = x^6 + 6x^5 - 20x^4 + 20x^3 - 20x^2 - 8x + 8}$			
$J(T)$	$\langle 48, 33\rangle, \langle 24, 13\rangle, \langle 24, 3\rangle$	$M(u_1, u_2)$	$K(\sqrt{v_1})$
$\boldsymbol{y^2 = x^6 - 5x^4 + 10x^3 - 5x^2 + 2x - 1}$			
$J(O)$	$\langle 96, 193\rangle, \langle 48, 48\rangle, \langle 48, 29\rangle$	$M(\sqrt{-11}, u_3, u_4)$,	$K(\sqrt{v_2})$
$\boldsymbol{y^2 = x^6 - 5x^4 - 5x^2 + 1}$			
$C_{2,1}$	$\langle 2, 1\rangle, \langle 2, 1\rangle, \langle 1, 1\rangle$	M	K
$\boldsymbol{y^2 = -x^6 + 5x^4 + 5x^2 - 1}$			
$C_{2,1}$	$\langle 4, 2\rangle, \langle 2, 1\rangle, \langle 2, 1\rangle$	M	$K(i)$
$\boldsymbol{y^2 = x^5 + x}$			
$D_{2,1}$	$\langle 4, 2\rangle, \langle 4, 2\rangle, \langle 2, 1\rangle$	$M(i)$	K
$\boldsymbol{y^2 = x^6 + 3x^5 - 20x^4 + 30x^3 - 35x^2 + 3x + 10}$			
$D_{2,1}$	$\langle 8, 3\rangle, \langle 4, 2\rangle, \langle 4, 1\rangle$	$M(\sqrt{7})$	$K\big(\sqrt{3\sqrt{7}+7}\big)$
$\boldsymbol{y^2 = x^5 + 81x}$			
$D_{2,1}$	$\langle 8, 5\rangle, \langle 4, 2\rangle, \langle 4, 2\rangle$	$M(i)$	$K(\sqrt{3})$
$\boldsymbol{y^2 = x^6 - 18x^5 - 15x^4 - 20x^3 + 135x^2 - 498x - 89}$			
$D_{3,2}$	$\langle 6, 1\rangle, \langle 6, 1\rangle, \langle 3, 1\rangle$	$M(u_5)$	K
$\boldsymbol{y^2 = x^6 + 4x^5 - 10x^4 + 80x^3 + 140x^2 + 144x - 184}$			
$D_{3,2}$	$\langle 12, 4\rangle, \langle 6, 1\rangle, \langle 6, 2\rangle$	$M(u_6)$	$K(i)$
$\boldsymbol{y^2 = x^5 - 2x}$			
$D_{4,1}$	$\langle 16, 7\rangle, \langle 8, 3\rangle, \langle 8, 3\rangle$	$M(i\sqrt[4]{-2})$	$K(\sqrt[8]{-2})$
$\boldsymbol{y^2 = x^5 + 2x}$			
$D_{4,1}$	$\langle 16, 8\rangle, \langle 8, 3\rangle, \langle 8, 4\rangle$	$M(i\sqrt[4]{2})$	$K(\sqrt[8]{2})$
$\boldsymbol{y^2 = x^6 + x^5 + 10x^3 + 5x^2 + x - 2}$			
$D_{4,2}$	$\langle 16, 7\rangle, \langle 8, 3\rangle, \langle 8, 1\rangle$	$M\big(\sqrt{\sqrt{-7}+7}\big)$	$K\big(\sqrt{-\sqrt{-2}\sqrt{\sqrt{-7}-7}+2\sqrt{-7}}\big)$
$\boldsymbol{y^2 = x^6 + 7x^5 + 10x^4 + 10x^3 + 15x^2 + 17x + 4}$			
O_1	$\langle 48, 29\rangle, \langle 24, 12\rangle, \langle 24, 3\rangle$	$M(u_7, u_8)$	$K\big(\sqrt{-u_8^3 + u_8^2 + 5u_8 + 4}\big)$

$$u_1^3 - 7u_1 + 7 = 0 \qquad u_2^4 + 4u_2^2 + 8u_2 + 8 = 0 \qquad u_3^3 - 4u_3 + 4 = 0 \qquad u_4^4 + 22u_4 + 22 = 0$$

$$u_5^3 + 6u_5 - 8 = 0 \qquad u_6^3 + 5u_6 - 10 = 0 \qquad u_7^3 + 5u_7 + 10 = 0 \qquad u_9^4 + 4u_9^2 + 8u_9 + 2 = 0$$

$$v_1^{12} - 12v_1^{11} + 70v_1^{10} - 236v_1^9 + 337v_1^8 - 40v_1^7 - 420v_1^6 + 452v_1^5 - 150v_1^4 + 16v_1^3 - 28v_1^2 + 8v_1 + 1 = 0$$

$$v_2^{12} + 44v_2^{11} + 682v_2^{10} + 4048v_2^9 + 3135v_2^8 - 19844v_2^7 + 306614v_2^6 + 1783540v_2^5 - 5571929v_2^4 + 85184v_2^3 + 1269774v_2^2 - 1293732v_2 - 970299 = 0$$

Table 5. Twists of C_2^0: $y^2 = x^6 - 5x^4 - 5x^2 + 1$ realizing each triple.

G	$[\mathrm{Gal}(L/\mathbb{Q}), \mathrm{Gal}(K/\mathbb{Q}), \mathrm{Gal}(L/M)]$	K	L
$\boldsymbol{y^2 = x^6 + 6x^5 + 30x^4 + 120x^2 - 96x + 64}$			
$J(C_2)$	$\langle 8, 2\rangle, \langle 4, 2\rangle, \langle 4, 1\rangle$	$M(\sqrt{5})$	$K\big(\sqrt{\sqrt{5}+5}\big)$
$\boldsymbol{y^2 = x^5 + 10x^3 + 9x}$			
$J(C_2)$	$\langle 8, 3\rangle, \langle 4, 2\rangle, \langle 4, 2\rangle$	$M(i)$	$K(\sqrt[4]{3})$
$\boldsymbol{y^2 = x^6 - 15x^4 - 20x^3 + 6x + 1}$			
$J(C_6)$	$\langle 24, 10\rangle, \langle 12, 5\rangle, \langle 12, 5\rangle$	$M(i, u_1)$	$K(\sqrt[4]{3})$
$\boldsymbol{y^2 = x^5 + 20x^3 + 36x}$			
$J(D_2)$	$\langle 16, 11\rangle, \langle 8, 5\rangle, \langle 8, 3\rangle$	$M(i, \sqrt{2})$	$K(\sqrt[4]{6})$
$\boldsymbol{y^2 = x^6 + 15x^4 + 20x^3 + 30x^2 + 18x + 5}$			
$J(D_3)$	$\langle 24, 5\rangle, \langle 12, 4\rangle, \langle 12, 1\rangle$	$M(\sqrt{5}, u_1)$	$K\big(\sqrt{2\sqrt{5}+10}\big)$
$\boldsymbol{y^2 = x^6 + 6x^5 + 40x^3 - 60x^2 + 72x - 32}$			
$J(D_3)$	$\langle 24, 6\rangle, \langle 12, 4\rangle, \langle 12, 4\rangle$	$M(i, u_2)$	$K(\sqrt[4]{3})$
$\boldsymbol{y^2 = x^6 + 3x^5 + 10x^3 - 15x^2 + 15x - 6}$			
$J(D_6)$	$\langle 48, 38\rangle, \langle 24, 14\rangle, \langle 24, 8\rangle$	$M(i, \sqrt{2}, u_3)$	$K(\sqrt[4]{2})$
$\boldsymbol{y^2 = x^6 + 1}$			
$C_{2,1}$	$\langle 2, 1\rangle, \langle 2, 1\rangle, \langle 1, 1\rangle$	M	K
$\boldsymbol{y^2 = x^6 + 15x^4 + 15x^2 + 1}$			
$C_{2,1}$	$\langle 4, 2\rangle, \langle 2, 1\rangle, \langle 2, 1\rangle$	M	$K(\sqrt{2})$
$\boldsymbol{y^2 = -x^6 - 6x^5 + 30x^4 - 20x^3 - 15x^2 + 12x - 1}$			
$C_{6,1}$	$\langle 6, 2\rangle, \langle 6, 2\rangle, \langle 3, 1\rangle$	$M(u_1)$	K
$\boldsymbol{y^2 = x^6 + 6x^5 - 30x^4 + 20x^3 + 15x^2 - 12x + 1}$			
$C_{6,1}$	$\langle 12, 5\rangle, \langle 6, 2\rangle, \langle 6, 2\rangle$	$M(u_1)$	$K(i)$
$\boldsymbol{y^2 = x^6 - 1}$			
$D_{2,1}$	$\langle 4, 2\rangle, \langle 4, 2\rangle, \langle 2, 1\rangle$	$M(i)$	K
$\boldsymbol{y^2 = 11x^6 + 30x^5 + 30x^4 + 40x^3 - 60x^2 + 120x - 88}$			
$D_{2,1}$	$\langle 8, 3\rangle, \langle 4, 2\rangle, \langle 4, 1\rangle$	$M(\sqrt{-2})$	$K\big(\sqrt{\sqrt{6}-2}\big)$
$\boldsymbol{y^2 = x^6 - 15x^4 + 15x^2 - 1}$			
$D_{2,1}$	$\langle 8, 5\rangle, \langle 4, 2\rangle, \langle 4, 2\rangle$	$M(i)$	$K(\sqrt{2})$
$\boldsymbol{y^2 = x^6 + 4}$			
$D_{3,2}$	$\langle 6, 1\rangle, \langle 6, 1\rangle, \langle 3, 1\rangle$	$M(\sqrt[3]{2})$	K
$\boldsymbol{y^2 = x^6 + 12x^5 + 15x^4 + 40x^3 + 15x^2 + 12x + 1}$			
$D_{3,2}$	$\langle 12, 4\rangle, \langle 6, 1\rangle, \langle 6, 2\rangle$	$M(\sqrt[3]{3})$	$K(\sqrt{-2})$
$\boldsymbol{y^2 = x^6 + 9x^5 - 60x^4 - 120x^3 + 240x^2 + 144x - 64}$			
$D_{6,1}$	$\langle 12, 4\rangle, \langle 12, 4\rangle, \langle 6, 1\rangle$	$M(i, u_4)$	K
$\boldsymbol{y^2 = x^6 + 6x^5 - 30x^4 - 40x^3 + 60x^2 + 24x - 8}$			
$D_{6,1}$	$\langle 24, 8\rangle, \langle 12, 4\rangle, \langle 12, 1\rangle$	$M(\sqrt{-2}, u_5)$	$K\big(\sqrt{\sqrt{6}-2}\big)$
$\boldsymbol{y^2 = x^6 + 3x^5 + 15x^4 - 20x^3 + 60x^2 - 60x + 28}$			
$D_{6,1}$	$\langle 24, 14\rangle, \langle 12, 4\rangle, \langle 12, 4\rangle$	$M(\sqrt{-2}, u_2)$	$K(\sqrt{2})$
$\boldsymbol{y^2 = x^6 + 2}$			
$D_{6,2}$	$\langle 12, 4\rangle, \langle 12, 4\rangle, \langle 6, 2\rangle$	$M(\sqrt[6]{2})$	K
$\boldsymbol{y^2 = x^6 + 6x^5 - 15x^4 + 20x^3 - 15x^2 + 6x - 1}$			
$D_{6,2}$	$\langle 24, 14\rangle, \langle 12, 4\rangle, \langle 12, 5\rangle$	$M(\sqrt{-2}, u_6)$	$K(i)$

$$u_1^3 - 3u_1 + 1 = 0 \qquad u_2^3 - 3u_2 + 4 = 0 \qquad u_3^3 + 3u_3 - 2 = 0$$
$$u_4^3 - 15u_4 - 10 = 0 \qquad u_5^3 - 9u_5 - 6 = 0 \qquad u_7^3 - 6u_7 - 6$$

Table 6. Twists of $C_3^0\colon y^2 = x^6 + 1$ realizing each triple.

The methods used to obtain these curves and the computation of K and L are described in Section 5.

Acknowledgements

We thank Joan-C. Lario and Jordi Quer for helpful comments, and we are grateful to Kiran S. Kedlaya for indicating the way to prove Proposition 3.6. Fité thanks the Massachusetts Institute of Technology for its warm hospitality from September to December 2011, the period in which this project was realized. Fité is also grateful to the University of Cambridge for its welcome in the period May-July 2011, when some of the ideas of this article originated. Fité received financial support from DGICYT grant MTM2009-13060-C02-01 and from NSF grant DMS-0545904. Sutherland received financial support from NSF grant DMS-1115455.

References

[Banaszak and Kedlaya 2011] G. Banaszak and K. S. Kedlaya, "An algebraic Sato–Tate group and Sato–Tate conjecture", preprint, 2011. To appear in *Indiana Univ. Math. J.* arXiv 1109.4449v1

[Besche et al. 2002] H. U. Besche, B. Eick, and E. A. O'Brien, "A millennium project: Constructing small groups", *Internat. J. Algebra Comput.* **12**:5 (2002), 623–644. MR 2003h:20042 Zbl 1020.20013

[Cardona 2001] G. Cardona, *Models racionals de corbes de gènere* 2, Tesi Doctoral, Universitat Politècnica de Catalunya, 2001, http://dmi.uib.es/~gcardona/tesi.pdf.

[Cardona 2006] G. Cardona, "Representations of G_k-groups and twists of the genus two curve $y^2 = x^5 - x$", *J. Algebra* **303**:2 (2006), 707–721. MR 2007e:14049 Zbl 1149.14023

[Cardona et al. 1999] G. Cardona, J. González, J. C. Lario, and A. Rio, "On curves of genus 2 with Jacobian of GL_2-type", *Manuscripta Math.* **98**:1 (1999), 37–54. MR 99j:11068 Zbl 0998.11032

[Fité 2013] F. Fité, "Artin representations attached to pairs of isogenous abelian varieties", *J. Number Theory* **133**:4 (2013), 1331–1345. MR 3004003 Zbl 1275.11100

[Fité and Lario 2013] F. Fité and J.-C. Lario, "The twisting representation of the L-function of a curve", *Rev. Mat. Iberoam.* **29**:3 (2013), 749–764. MR 3089761 Zbl 06216017

[Fité et al. 2012] F. Fité, K. S. Kedlaya, V. Rotger, and A. V. Sutherland, "Sato–Tate distributions and Galois endomorphism modules in genus 2", *Compos. Math.* **148**:5 (2012), 1390–1442. MR 2982436 Zbl 1269.11094

[Fité et al. $\geq$ 2014] F. Fité, E. Lorenzo, and A. Sutherland, "Sato–Tate distributions of twists of the Fermat and Klein quartics". In preparation.

[Gross 1980] B. H. Gross, *Arithmetic on elliptic curves with complex multiplication*, Lecture Notes in Mathematics **776**, Springer, Berlin, 1980. MR 81f:10041 Zbl 0433.14032

[Hecke 1920] E. Hecke, "Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen", *Math. Z.* **6**:1-2 (1920), 11–51. MR 1544392 JFM 47.0152.01

[Igusa 1960] J.-i. Igusa, "Arithmetic variety of moduli for genus two", *Ann. of Math.* (2) **72** (1960), 612–649. MR 22 #5637 Zbl 0122.39002

[Johansson 2013] C. Johansson, "On the Sato–Tate conjecture for non-generic abelian surfaces", preprint, 2013. arXiv 1307.6478

[Kedlaya and Sutherland 2008] K. S. Kedlaya and A. V. Sutherland, “Computing L-series of hyperelliptic curves”, pp. 312–326 in *Algorithmic number theory* (Banff, AB, 2008), edited by A. J. van der Poorten and A. Stein, Lecture Notes in Comput. Sci. **5011**, Springer, Berlin, 2008. MR 2010d:11070 Zbl 1232.11078

[Mazur et al. 2007] B. Mazur, K. Rubin, and A. Silverberg, “Twisting commutative algebraic groups”, *J. Algebra* **314**:1 (2007), 419–438. MR 2008j:14089 Zbl 1128.14034

[OEIS 2011] OEIS, “The On-Line Encyclopedia of Integer Sequences”, 2011, http://oeis.org.

[Pila 1990] J. Pila, “Frobenius maps of abelian varieties and finding roots of unity in finite fields”, *Math. Comp.* **55**:192 (1990), 745–763. MR 91a:11071 Zbl 0724.11070

[Rubin and Silverberg 2010] K. Rubin and A. Silverberg, “Choosing the correct elliptic curve in the CM method”, *Math. Comp.* **79**:269 (2010), 545–561. MR 2010i:11091 Zbl 1213.11127

[Serre 1998] J.-P. Serre, *Abelian l-adic representations and elliptic curves*, Research Notes in Mathematics **7**, A K Peters, Wellesley, MA, 1998. MR 98g:11066 Zbl 0902.14016

[Serre 2012] J.-P. Serre, *Lectures on $N_X(p)$*, Research Notes in Mathematics **11**, CRC Press, Boca Raton, FL, 2012. MR 2920749 Zbl 1238.11001

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics **106**, Springer, Dordrecht, 2009. MR 2010i:11005 Zbl 1194.11005

[Sutherland 2007] A. V. Sutherland, *Order computations in generic groups*, Ph.D. thesis, Massachusetts Institute of Technology, 2007, http://groups.csail.mit.edu/cis/theses/sutherland-phd.pdf.

[Sutherland 2011] A. V. Sutherland, “`smalljac` software library, v. 4.0”, 2011, http://math.mit.edu/~drew.

Communicated by Peter Sarnak
Received 2012-11-20 Revised 2013-08-22 Accepted 2013-09-23

francesc.fite@gmail.com *Fakultät für Mathematik, Bielefeld Universität, P.O. Box 100131, D-33501 Bielefeld, Germany*

drew@math.mit.edu *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, United States*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.587

The algebraic dynamics of generic endomorphisms of $\mathbb{P}^n$

Najmuddin Fakhruddin

We investigate some general questions in algebraic dynamics in the case of generic endomorphisms of projective spaces over a field of characteristic zero. The main results that we prove are that a generic endomorphism has no nontrivial preperiodic subvarieties, any infinite set of preperiodic points is Zariski-dense and any infinite subset of a single orbit is also Zariski-dense, thereby verifying the dynamical "Manin–Mumford" conjecture of Zhang and the dynamical "Mordell–Lang" conjecture of Denis and Ghioca and Tucker in this case.

1. Introduction

The goal of this article is to study some aspects of the algebraic dynamics of generic endomorphisms[1] of $\mathbb{P}^n$ of degree $d > 1$ over a field K of characteristic zero. Properties of algebraic varieties, for example smooth projective curves or abelian varieties, are often easier to derive for generic varieties than for arbitrary varieties, the main reason being that one has a great deal of freedom in choosing specialisations. It is natural to expect that the same holds for algebraic dynamical systems; we show that this is indeed the case for generic endomorphisms of $\mathbb{P}^n$. We prove three results for such endomorphisms: two of them have analogues expected to hold much more generally, though at present this is far from being known.

Our main result is:

Theorem 1.1. *Let $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ be a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero. For each $x \in \mathbb{P}^n(K)$, every infinite subset of $O_f(x)$, the f-orbit of x, is Zariski-dense in $\mathbb{P}^n_K$.*

This implies the dynamical "Mordell–Lang" conjecture of [Denis 1994; Ghioca and Tucker 2009] for generic endomorphisms. This conjecture has been proved for étale endomorphisms of arbitrary varieties by Bell, Ghioca and Tucker [Bell et al.

MSC2010: primary 37P55; secondary 37F10.
Keywords: generic endomorphisms, projective space.

[1] The precise meaning of generic endomorphism is given in Definition 3.1, but we note here that when $K = \mathbb{C}$ this means we consider endomorphisms in the complement of a countable union of proper subvarieties of the natural parameter variety of endomorphisms of degree d.

2010], but there are only a few other cases where it is known. The proof of this theorem is based on two other results. The first is:

Theorem 1.2. *Let $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ be a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero. If $X \subset \mathbb{P}^n_K$ is an irreducible subvariety such that $f^r(X) = X$ for some $r > 0$, then X is a point or $X = \mathbb{P}^n_K$.*

This is a rather straightforward consequence of the transitivity of the monodromy action on the set of periodic points of a fixed period of a generic endomorphism, which we prove (Proposition 3.3) using a result of Bousch [1992], Lau and Schleicher [1994] and Morton [1998] for polynomials in one variable of the form $z \mapsto z^d + c$.

We then extend Proposition 3.3 to prove transitivity of the monodromy action on the set of *preperiodic* points of fixed period and preperiod of a generic endomorphism (Proposition 4.6); this does not hold for the 1-parameter family of polynomials mentioned above, and we use a 2-parameter family containing this. This allows us to prove Zhang's "Manin–Mumford" conjecture [Ghioca et al. 2011] in the case of generic endomorphisms in the following strong form:

Theorem 1.3. *For $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero, any infinite subset of $\mathbb{P}^n(K)$ consisting of f-preperiodic points is Zariski-dense in $\mathbb{P}^n_K$.*

We prove Theorem 1.1 by combining Theorems 1.2 and 1.3 with some p-adic as well as mod p arguments. Note that the statement does not involve (pre)periodic points in any way. However, using a lifting argument for periodic points, we show that any subvariety Y containing an infinite subset of $O_f(x)$ must contain infinitely many periodic points, or x can be specialised in such a way that one may apply the p-adic interpolation argument used in [Bell et al. 2010] to prove the conjecture for étale endomorphisms. Theorems 1.3 and 1.2 then force Y to be equal to $\mathbb{P}^n_K$ in either of these cases.

2. Preliminaries

Let X be a set and $f : X \to X$ any map. By f^n, we shall mean the n-fold composite of f with itself. For $x \in X$, we denote by $O_f(x)$ its orbit under f, i.e., the set $\{f^n(x)\}_{n\geq 0}$. A point $x \in X$ is said to be *f-periodic* if $f^n(x) = x$ for some $n > 0$. The smallest such integer is called the *period* of x. We denote the set of all periodic points of period b by $\mathrm{Per}_f(b)$. A point $x \in X$ is said to *f-preperiodic* if $O_f(x)$ is finite. The *preperiod* of f is the smallest nonnegative integer a such that $f^a(x)$ is periodic, and the *period* of x is the period of any periodic point in its orbit. We denote by $\mathrm{Prep}_f(a, b)$ the set of all such points. Let $\mathrm{Orb}_f(b)$ denote the set of orbits of f-periodic points of period b. If this is finite, then $|\mathrm{Per}_f(b)| = b \cdot |\mathrm{Orb}_f(b)|$. We drop f from any of the notation introduced above if there is no scope for confusion.

If X is an algebraic variety over a field K and $f : X \to X$ is a morphism over K, we use the same notation as above for the induced map on the set of L-rational points of X for any extension field L of K.

Let S be a smooth irreducible variety over a field k, and let $g : Z \to S$ be a finite flat morphism. By the *monodromy* or *Galois action* of g, we shall mean the action of $\mathrm{Gal}(\overline{k(S)}/k(S))$ on (the points of) a geometric generic fibre of g. If g is generically smooth — this is always true if $\mathrm{char}(k) = 0$ and Z is reduced — there is a Zariski open subset U of S such that g induces a finite étale morphism $g^{-1}(U) \to U$, and then the monodromy may be interpreted as an action of $\pi_1^{\text{ét}}(U, *)$, where $*$ is a geometric point of U. If $k = \mathbb{C}$, it may be interpreted as an action of $\pi_1^{\text{top}}(U, *)$.

The monodromy action is transitive if and only if Z is irreducible. If Z is generically smooth, this is equivalent to Z^0, the smooth locus of Z, being connected or, if $k = \mathbb{C}$, path-connected.

Definition 2.1. Let $\pi : \mathscr{X} \to S$ be a projective morphism and $f : \mathscr{X} \to \mathscr{X}$ a surjective morphism over S. We say that f is *quasipolarised* if there exists a line bundle $\mathscr{L}$ on $\mathscr{X}$ such that $f^*(\mathscr{L}) \otimes \mathscr{L}^{-1}$ is S-ample.

For any morphism $\pi : \mathscr{X} \to S$ and $f : \mathscr{X} \to \mathscr{X}$ a morphism over S, we denote by Γ_f the graph of f in $\mathscr{X} \times_S \mathscr{X}$. Let $P_f(n)$ be the closed subscheme of $\mathscr{X}$ defined by the intersection of Γ_{f^n} with the diagonal. A geometric point of the fibre of $P_f(n)$ over any point $s \in S$ is a periodic point of period dividing n of the map f_s of $\mathscr{X}_s$ induced by f. Similarly, let $P_f(m, n)$ be the intersection of Γ_{f^m} and Γ_{f^n}, which we view as a subscheme of $\mathscr{X}$ via the first projection.

Lemma 2.2. *Let $\pi : \mathscr{X} \to S$ be a smooth projective morphism with S a regular irreducible finite-dimensional scheme, and let $f : \mathscr{X} \to \mathscr{X}$ be a finite quasipolarised morphism. Then*:

(1) *For any $m, n \geq 0$, $m \neq n$, $P_f(m, n)$ is finite and flat over S.*

(2) *For any s_1 and s_2 in S with s_2 a specialisation of s_1, any element of $\mathrm{Prep}_{f_{s_2}}(a, b)$ can be lifted to an element of $\mathrm{Prep}_{f_{s_1}}(a, b)$.*

Proof. Since f is proper, $P_f(m, n)$, being a closed subscheme of $\mathscr{X}$, is also proper over S. The dimension of each irreducible component of $P_f(m, n)$ is at least equal to $\dim(S)$ since the codimension of Γ_{f^i} in $\mathscr{X} \times_S \mathscr{X}$ is the relative dimension of $\mathscr{X}$ over S. To prove $P_f(m, n)$ is finite over S, it suffices to show that the fibres of this map are finite since proper quasifinite morphisms are finite. Furthermore, the finiteness of the fibres implies that the dimension of each component of $P_f(m, n)$ is exactly $\dim(S)$; hence, $P_f(m, n)$ is a local complete intersection in $\mathscr{X} \times_S \mathscr{X}$ that is regular, which implies that $P_f(m, n)$ is Cohen–Macaulay. Since the dimension of each irreducible component of $P_f(m, n)$ is at least $\dim(S)$, each such component dominates S if all the fibres are finite. It then follows from the fibrewise flatness criterion that all of (1) is a consequence of the finiteness of the fibres.

Let $\mathscr{L}$ be a line bundle on $\mathscr{X}$ so that $M = f^*(\mathscr{L}) \otimes \mathscr{L}^{-1}$ is ample. Then $(f^m)^*(\mathscr{L}) = \mathscr{L} \otimes M \otimes f^*(M) \otimes \cdots \otimes (f^{m-1})^*(M)$ and similarly for n. By the construction of $P_f(m,n)$, $(f^m)^*(\mathscr{L})$ and $(f^m)^*(\mathscr{L})$ restrict to the same line bundle on it, so assuming $m > n$ without loss of generality, we get that $(f^{m-1})^*(M) \otimes \cdots \otimes (f^n)^*(M)$ is trivial on $P_f(m,n)$. But M is ample; hence, so is $(f^i)^*(M)$ for all $i \geq 0$ since f is finite. There is at least one factor in the above tensor product of line bundles, so this is only possible if all the fibres are finite.

If $x \in \mathrm{Prep}_{f_{s_2}}(a,b)$, then x occurs in the fibre of $P_f(a+b,a)$ over s_1 and does not occur in the fibre of any $P_f(m,n)$ for $m < a+b$ or $n < a$. By (1), there is a point $\tilde{x}$ in the fibre of $P_f(a+b,a)$ specialising to x. By the definition of $P_f(m,n)$, it follows that $f_{s_1}^{a+b}(\tilde{x}) = f_{s_1}^{a}(\tilde{x})$, so $\tilde{x}$ is preperiodic with preperiod $\leq a$ and period $\leq b$. Since neither the preperiod nor the period can increase under specialisation and $\tilde{x}$ specialises to x, (2) follows. □

3. Periodic points and periodic subvarieties

Let $\mathrm{Mor}_{n,d}$ be the scheme over $\mathbb{Z}$ representing morphisms of $\mathbb{P}^n_{\mathbb{Z}}$ to itself of algebraic degree d. Its k-valued points, for any field k, consist of $(n+1)$-tuples of homogeneous polynomials of degree d over k without common zeros in $\mathbb{P}^n_k$ up to a scalar. It is smooth and of finite type over $\mathbb{Z}$ and has geometrically irreducible fibres. For any field L, we denote $\mathrm{Mor}_{n,d} \times_{\mathrm{Spec}(\mathbb{Z})} \mathrm{Spec}(L)$ by $\mathrm{Mor}_{n,d/L}$.

Definition 3.1. If K is an algebraically closed field, we say that an endomorphism f of $\mathbb{P}^n_K$ is *generic* if the image of the induced map $\mathrm{Spec}(K) \to \mathrm{Mor}_{n,d}$ corresponding to some conjugate of f by an element of $\mathrm{PGL}_{n+1}(K) = \mathrm{Aut}(\mathbb{P}^n_K)$ is the generic point of a fibre of the structure morphism $\mathrm{Mor}_{n,d} \to \mathrm{Spec}(\mathbb{Z})$.

If $K = \mathbb{C}$, the set of points in $\mathrm{Mor}_{n,d}(\mathbb{C})$ corresponding to generic morphisms is the complement of a countable union of proper subvarieties.

We recall the theorem of Bousch [1992, Chapitre 3, Théorème 4], Lau and Schleicher [1994, Theorem 4.1] and Morton [1998, Theorem D] alluded to earlier; the statement below is [Morton 1998, Theorem 10] except that we have replaced the field $\overline{\mathbb{Q}}$ there by k.

Theorem 3.2. *Let k be a field of characteristic zero, and let $f(z) = z^d + t$ with t transcendental over k and $d \geq 2$. For any $b \geq 1$, the Galois group of the polynomial $f^b(z) - z$ over $k(t)$ is the direct product $\prod_{e|b}(\mathbb{Z}/e\mathbb{Z} \,\mathrm{wr}\, S_{r_e})$, where* wr *denotes the wreath product and $e \cdot r_e$ is the number of periodic points of period e over $\overline{k(t)}$.*

The theorem can be interpreted as saying that the Galois action is as large as possible given that it must commute with the action of f. One may expect that a similar result holds for generic endomorphisms of projective spaces of arbitrary dimension; however, the following proposition, for the proof of which we will use only the transitivity of the Galois action in the above theorem, suffices for our applications:

Proposition 3.3. *Let k be a field of characteristic zero, and let $k_{n,d}$ be the function field of $\mathrm{Mor}_{n,d/k}$. Let $f_{n,d}$ be the endomorphism of $\mathbb{P}^n_{k_{n,d}}$ corresponding to the generic point of $\mathrm{Mor}_{n,d/k}$, and let b be any positive integer. Then $\mathrm{Gal}(\overline{k_{n,d}}/k_{n,d})$ acts transitively on $\mathrm{Per}_{f_{n,d}}(b)$.*

For a field k and any element $\lambda \in k$, let $\phi_\lambda : \mathbb{A}^1_k \to \mathbb{A}^1_k$ be the map given by $z \mapsto z^d + \lambda$; the integer d will be assumed to be fixed whenever we use this notation. The periodic points of ϕ_0 are 0 and the roots of unity of order prime to d: if ζ is a primitive n-th root of unity with $(n, d) = 1$, then the period of ζ is the order of d in $(\mathbb{Z}/n\mathbb{Z})^\times$.

We shall need the following simple lemma for the proof of Proposition 3.3:

Lemma 3.4. *Let $d > 1$ and $m, m' \geq 1$ be integers such that $(m, d) = (m', d) = 1$. Assume that the highest powers of 2 dividing m and m' are unequal or are both equal to 1. Let a and a' be the orders of d in $(\mathbb{Z}/m\mathbb{Z})^\times$ and $(\mathbb{Z}/m'\mathbb{Z})^\times$, respectively.*

(1) *The order of d in $(\mathbb{Z}/\operatorname{lcm}(m, m')\mathbb{Z})^\times$ is divisible by $\operatorname{lcm}(a, a')$.*

(2) *There exist roots of unity ζ and ζ' of orders m and m', respectively, so that $\zeta\zeta'^{-1}$ is of order $\operatorname{lcm}(m, m')$.*

(3) *For ζ and ζ' as above, there exists a primitive $\operatorname{lcm}(m, m')$-th root of unity η so that $\eta\zeta$ is a primitive $\operatorname{lcm}(m, m')$-th root of unity.*

Proof. The natural quotient maps from $\mathbb{Z}/\operatorname{lcm}(m, m')\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/m'\mathbb{Z}$ induce group homomorphisms from $(\mathbb{Z}/\operatorname{lcm}(m, m')\mathbb{Z})^\times$ to $(\mathbb{Z}/m\mathbb{Z})^\times$ and $(\mathbb{Z}/m'\mathbb{Z})^\times$, respectively. This implies that the order of d in $(\mathbb{Z}/\operatorname{lcm}(m, m')\mathbb{Z})^\times$ is divisible by $\operatorname{lcm}(a, a')$.

To prove (2) and (3), we may reduce to the case that m and m' are powers of the same prime p. Let $P \subset \mathbb{Z}/p^r\mathbb{Z}$ be the set of generators, so $|P| = p^r - p^{r-1}$. If $p > 2$, then $|P| > p^{r-1}$, so the translate of P by any element of $\mathbb{Z}/p^r\mathbb{Z}$ has a nonempty intersection with P. If $p = 2$, the claim follows from the extra condition since the translate of P by an element not in P always intersects P nontrivially. $\square$

Proof of Proposition 3.3. If $n = 1$ and $b > 1$, the proposition follows immediately from Theorem 3.2. If $b = 1$, a much simpler version of the argument below shows transitivity; since we do not use this later, we leave the details to the reader.

We now assume $n > 1$. Consider the morphism $g_{n,d} : \mathbb{P}^n_k \to \mathbb{P}^n_k$ given by $[x_0, \ldots, x_n] \mapsto [x_0^d, \ldots, x_n^d]$. The set $\mathrm{Per}_{g_{n,d}}(b)$ consists of points that have a representative $[\xi_0, \xi_1, \ldots, \xi_n]$ with each ξ_i equal to 0 or a $(d^b - 1)$-th root of unity.

The standard affine charts of $\mathbb{P}^n_k$ given by the locus where a fixed coordinate is nonzero are preserved by $g_{n,d}$. A simple computation on each such chart shows that the eigenvalues of the differential of $g^b_{n,d}$ at a fixed point are equal to $d^b\xi$, where ξ is a root of unity or 0. This is never equal to 1 since $d > 1$, so $\Gamma_{g^b_{n,d}}$ and the diagonal intersect transversely in $\mathbb{P}^n_k \times \mathbb{P}^n_k$ for all $b > 0$. Consequently, all periodic points of $g_{n,d}$ have multiplicity 1, so we may use $g_{n,d}$ as a basepoint in $\mathrm{Mor}_{n,d}$ in order to compute the Galois action on $\mathrm{Per}_{f_{n,d}}$.

For $0 \le i \le n$, consider the family of endomorphisms $f_i : \mathbb{P}^n_k \times \mathbb{A}^n_k \to \mathbb{P}^n_k \times \mathbb{A}^n_k$ given by

$$\begin{aligned} &f_i(([x_0, \dots, x_i, \dots, x_n], (c_1, c_2, \dots, c_n))) \\ &= \big([x_0^d + c_1 x_i^d, \dots, x_{i-1}^d + c_i x_i^d, x_i^d, x_{i+1}^d + c_{i+1} x_i^d, \dots, x_n^d + c_n x_i^d], (c_1, c_2, \dots, c_n)\big). \end{aligned}$$

On the open affine U_i given by $x_i \neq 0$, f_i is the product of the n polynomials ϕ_{c_j}. On the complement of this affine, i.e., on the subvariety given by $x_i = 0$ (which is also preserved by f_i), the maps do not depend on c_j, so the monodromy action of this family on the periodic points in this locus is trivial. Let G_i be the subgroup of the monodromy group acting on $\mathrm{Per}_{g_{n,d}}$ corresponding to this family; by applying Theorem 3.2, one gets a complete description of this group. We let G be the subgroup of the monodromy group generated by all the G_i.

Let $P = [\zeta, 1, \dots, 1, 1]$ where ζ is in $\mathrm{Per}_{\phi_0}(b)$, and let $Q = [\xi_0, \xi_1, \dots, \xi_n]$ be any other element of $\mathrm{Per}_{g_{n,d}}(b)$. We may assume that some $\xi_i = 1$, so each $\xi_j \in \mathrm{Per}_{\phi_0}(b')$ for some $b' \mid b$, and also the lcm of the periods of all the ξ_j is b. We prove the transitivity of the monodromy action by showing that there exists an element in the monodromy that sends P to Q.

From the transitivity of the Galois action in Theorem 3.2, it follows that for $\xi_j \in \mathrm{Per}_{\phi_0}(b')$, as long as some $\xi_i = 1$ with $i \neq j$, we can find an element of G that fixes all coordinates of Q except that it replaces ξ_j with any other $\xi'_j \in \mathrm{Per}_{\phi_0}(b')$. Since $0, 1 \in \mathrm{Per}_{\phi_0}(0)$, we may use this to assume that all $\xi_i \neq 0$ and then also, by dividing through by ξ_n, that $\xi_n = 1$.

We now show that we may also assume that $\xi_0 \in \mathrm{Per}_{\phi_0}(b)$. Suppose ξ_0 is a primitive m-th root of unity, ξ_1 is a primitive m'-th root of unity and $\xi_0 \in \mathrm{Per}_{\phi_0}(a)$ and $\xi_1 \in \mathrm{Per}_{\phi_0}(a')$. By using the action of G, we may change ξ_0 and ξ_1 so that $m = d^a - 1$ and $m' = d^{a'} - 1$. If the highest powers of 2 dividing m and m' are equal and greater than 1, we may change ξ_0 to a primitive $(d^a - 1)/2$-th root of unity; the period a remains unchanged. By Lemma 3.4, we may then assume that $\xi_0 {\xi_1}^{-1}$ is a primitive $\mathrm{lcm}(m, m')$-th root of unity. We multiply all coordinates of Q by ξ_1^{-1}, so the zeroth coordinate becomes $\xi_0 {\xi_1}^{-1}$, and the second becomes 1. Using the action of G, we then replace the zeroth coordinate by η as in Lemma 3.4 while keeping all other coordinates fixed. We then multiply all coordinates by ξ_1. The resulting point has all coordinates except for the zeroth equal to the corresponding coordinates of Q while the zeroth coordinate is now in $\mathrm{Per}_{\phi_0}(\mathrm{lcm}(a, a'))$. Repeating this procedure with ξ_1 replaced by ξ_2, then ξ_3 and so on, since the lcm of the periods of all the ξ_i is b, it follows that eventually we have that $\xi_0 \in \mathrm{Per}_{\phi_0}(b)$.

We now inductively transform P into Q using the action of G. If ξ_i has period a_i as an element of Per_{ϕ_0}, we may use the action of G to replace it by a primitive $(d^{a_i} - 1)$-th root of unity. If the highest power of 2 dividing $d^{a_i} - 1$ is equal to the highest power of 2 dividing $d^b - 1$ and d is odd, then we use a primitive $(d^{a_i} - 1/2)$-th root of unity instead if $i > 0$.

Let $P_0 = P$, and suppose we have constructed $P_i = [\xi_{0,i}, \xi_{1,i}, \dots, \xi_{n-1,i}, 1]$ in $\mathrm{Per}_{\phi_0}(b)$ by induction, with the following properties:

(1) $\xi_{0,i} \in \mathrm{Per}_{\phi_0}(b)$.

(2) $\xi_{j,i} = \xi_j$ for $0 < j \leq i$.

(3) $\xi_{j,1} = 1$ for $j > i$.

Clearly P_0 satisfies these properties; we will show that given P_i with $i < n$ we can find an element of G that transforms it into a point P_{i+1} with the required properties.

So suppose P_i has been constructed. By Lemma 3.4, there exists $\eta \in \mathrm{Per}_{\phi_0}(b)$ so that $\eta^{-1}\xi_{i+1} \in \mathrm{Per}_{\phi_0}(b)$. Since $\xi_{0,i} \in \mathrm{Per}_{\phi_0}(b)$, we may use the action of G to replace $\xi_{0,i}$ with η while keeping the other coordinates fixed. We then multiply all coordinates by η^{-1}, so the zeroth coordinate becomes 1, and the $(i+1)$-th coordinate becomes η^{-1}. Since $\eta^{-1}\xi_{i+1} \in \mathrm{Per}_{\phi_0}(b)$, we may use the action of G to replace the $(i+1)$-th coordinate by $\eta^{-1}\xi_{i+1}$ while keeping all the other coordinates fixed. If we now multiply all coordinates by η, we obtain a point P_{i+1} with the property that all coordinates except the zeroth and $(i+1)$-th of P_i and P_{i+1} are equal, the zeroth coordinate is $\eta \in \mathrm{Per}_{\phi_0}(b)$ and the $(i+1)$-th is ξ_{i+1}, so fulfilling the requirements.

We thus obtain a point P_n with the property that all coordinates of P_n and Q are equal except possibly for the first. Since the zeroth coordinate of both is in $\mathrm{Per}_{\phi_0}(b)$ and the n-th is 1, we may use an element of G to transform P_n into Q. It follows that the action of G, hence of the full monodromy group, is transitive on $\mathrm{Per}_{\phi_0}(b)$. □

Corollary 3.5. *No preperiodic point of $f_{n,d}$ lies in the ramification locus.*

Proof. The ramification locus of $f_{n,d}$ is defined over $k_{n,d}$; thus, if one preperiodic point lies in the ramification locus, so must its entire Galois orbit.

The Galois orbit must map onto the Galois orbit of the corresponding periodic point, i.e., the periodic point y such that $f^r_{n,d}(x) = y$ and $f^s_{n,d}(x)$ is not periodic for any $s < r$. But this orbit consists of all periodic points of a fixed period b by Proposition 3.3. By specialisation to the d-power map $g_{n,d}$, we see that this is not possible:

the ramification locus of this map is actually preserved by the map and for no period b are all the periodic points of period b contained in the ramification locus. □

Theorem 3.6. *Let $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ be a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero. Then*:

(1) *If $X \subset \mathbb{P}^n_K$ is an irreducible subvariety such that $f^r(X) = X$ for some $r > 0$, then X is a point or $X = \mathbb{P}^n_K$.*

(2) *If $O_f(x)$ is infinite for $x \in \mathbb{P}^n(K)$, then $O_f(x)$ is Zariski-dense in $\mathbb{P}^n_K$.*

Proof. Since f is generic, we may identify f with $f_{n,d}$ defined above. Any specialisation of X to a subvariety of $\mathbb{P}^n$ defined over a finite extension of $k_{n,d}$ satisfies the same property as X, so we may assume that X is defined over a finite extension of degree m of $k_{n,d}$. Replacing X by the union of its Galois conjugates, we may assume that X is defined over $k_{n,d}$. Since $\dim(X) > 0$, it follows by [Fakhruddin 2003, Theorem 5.1] applied to f^r that X contains infinitely many periodic points of f. Since X is defined over $k_{n,d}$, it then follows from Proposition 3.3 that there is an infinite sequence of distinct integers b_i, $i = 1, 2, \ldots$, such that X contains all the periodic points of $f_{n,d}$ of period b_i for all i. Let $g_{n,d} : \mathbb{P}^n_k \to \mathbb{P}^n_k$ be, as before, the map that raises each coordinate to its d-th power. By specialisation and Lemma 2.2(2), we obtain a subvariety X' of $\mathbb{P}^n_k$ defined over k that contains all the periodic points of $g_{n,d}$ of period b_i for all i.

For $Z \subset \mathbb{A}^n_k$ any subvariety (with k any field), if there is a sequence of finite subsets $S_i \subset \mathbb{A}^n(\bar{k})$ such that $(S_i)^n \subset Z(\bar{k})$ for all i and $|S_i| \to \infty$, then $Z = \mathbb{A}^n_k$; this is well-known and follows, for example, from [Tao and Vu 2006, Theorem 9.2]. By the previous paragraph, for all i, X' contains all elements of $\mathbb{P}^n(\bar{k})$ of the form $[\xi_0, \ldots, \xi_{n-1}, 1]$ with $\xi_j \in \mathrm{Per}_{\phi_0}(b_i)$. Since $|\mathrm{Per}_{\phi_0}(b_i)| \to \infty$ with i, we must have $X' = \mathbb{P}^n_k$ and hence $X = \mathbb{P}^n_{k_{n,d}}$, proving (1).

To prove (2), we observe that the Zariski closure of $O_f(x)$ is mapped into itself by f. If it is infinite, it must contain a positive-dimensional subvariety X such that $f^r(X) = X$ for some $r > 0$. By (1), we must have $X = \mathbb{P}^n_K$. □

Remark 3.7. One can give a simpler proof of (1) using the Lefschetz trace formula rather than [Fakhruddin 2003, Theorem 5.1] if one only considers smooth X. Also, from Theorem 4.1, it follows that the elementary version of [Fakhruddin 2003, Theorem 5.1] asserting the density of preperiodic points suffices, but this introduces other dependencies.

4. The dynamical "Manin–Mumford" conjecture for generic endomorphisms

The original "Manin–Mumford" conjecture of Zhang asserted that, for any polarised endomorphism f of a projective variety X over a field K of characteristic 0, any subvariety Y of X containing a Zariski-dense set of preperiodic points is preperiodic.

This was known for abelian varieties and the multiplication-by-m maps but was later shown to be false in general, even if $X = \mathbb{P}^n_K$, by Ghioca and Tucker. Ghioca, Tucker and Zhang then proposed a modified conjecture that takes into account the action of f on the tangent space; see the article [Ghioca et al. 2011] for a discussion of the history, the statement of the modified version and some positive results.[2]

The following theorem implies that Zhang's conjecture, in its original form, holds for generic endomorphisms of $\mathbb{P}^n_K$:

Theorem 4.1. *For $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero, any infinite subset of $\mathbb{P}^n(K)$ consisting of f-preperiodic points is Zariski-dense in $\mathbb{P}^n_K$.*

As in the other results of this paper, one of the key ingredients of the proof is the Galois action on the set of periodic points. However, Proposition 3.3 alone does not suffice since we also need the transitivity of the monodromy action on the set of preperiodic points. It turns out that transitivity does not hold for the monodromy action on $\mathrm{Prep}_f(a, b)$ for f as in Theorem 3.2 and $d > 2$. Nevertheless, by considering a larger family of polynomials, we prove in Proposition 4.6 that the monodromy action on $\mathrm{Prep}_{f_{n,d}}(a, b)$ is indeed transitive for all a and b. This then allows us to use a specialisation argument to prove Theorem 4.1.

Fix $d > 1$. For $b > 0$, consider the polynomial $P_b(c) = \phi_c^b(0)$. The roots of $P_b(c)$ are exactly the parameters c so that 0 is a periodic point of period dividing b for the polynomial $\phi_c(z)$.

Lemma 4.2 (Gleason). *All roots of P_b are multiplicity-free.*

Proof. The proof given in [Douady and Hubbard 1985, Lemma 19.1] for $d = 2$ goes through for general d simply by replacing 2 by d. □

Lemma 4.3. *Fix $d > 1$. For $(c, \epsilon) \in \mathbb{A}^2$, let $\phi_{c,\epsilon} : \mathbb{A}^1 \to \mathbb{A}^1$ be given by $\phi_{c,\epsilon}(z) = z(z-\epsilon)^{d-1} + c$. Then the monodromy action on* $\mathrm{Prep}_{\phi_{c,\epsilon}}(1, b)$ *is transitive for all $b > 0$.*

Proof. It follows from Theorem 3.2 that the monodromy action on $\mathrm{Per}_{\phi_{c,\epsilon}}(b)$ is transitive for all $b > 0$, so it suffices to prove that for any $b > 0$ there exists $x \in \mathrm{Per}_{\phi_{c,\epsilon}}(b)$ and an element γ of the monodromy so that $\gamma(x) = x$ and γ cyclically permutes the $d-1$ elements of $\phi_{c,\epsilon}^{-1}(x)$ (so if $d = 2$, there is nothing to prove).

By Lemma 4.2 and a simple counting argument, it follows that, for any $b > 0$, there exists $c_b \in \mathbb{C}$ so that $0 \in \mathrm{Per}_{\phi_{c_b}}(b)$. Since 0 is a critical point of ϕ_{c_b}, it is of multiplicity 1 as an element of $\mathrm{Per}_{\phi_{c_b}}(b)$. It follows that, for $|\epsilon| \ll 0$, there exists $c_{b,\epsilon}$ close to c_b so that 0 is a periodic point of $\mathrm{Per}_{\phi_{c_{b,\epsilon},\epsilon}}(b)$ of multiplicity 1. By the

[2] It seems reasonable to expect the modified conjecture to hold even for quasipolarised endomorphisms.

definition of $\phi_{c,\epsilon}$, it then follows that ϵ is the unique element of $\mathrm{Prep}_{\phi_{c_{b,\epsilon},\epsilon}}(1, b)$ that is mapped to $\phi_{c_{b,\epsilon},\epsilon}(0) = c_{b,\epsilon}$ by $\phi_{c_{b,\epsilon},\epsilon}$.

We now consider the 1-parameter family of maps $\phi_{c,\epsilon}$ with ϵ such that $|\epsilon| \ll 0$ fixed. In a neighbourhood of $c_{b,\epsilon}$, the element $c_{b,\epsilon} \in \mathrm{Per}_{\phi_{c_{b,\epsilon},\epsilon}}(b)$ deforms uniquely as an element of $\mathrm{Per}_{\phi_{c,\epsilon}}(b)$. However, since the critical points of $\phi_{c,\epsilon}$ are independent of c, it follows that the element $\epsilon \in \mathrm{Prep}_{\phi_{c_{b,\epsilon},\epsilon}}(1, b)$ deforms to $d-1$ distinct elements of $\mathrm{Prep}_{\phi_{c,\epsilon}}(1, b)$ that are all mapped to the deformed periodic point above by $\phi_{c,\epsilon}$.

We claim that the monodromy action of a small loop around $c_{b,\epsilon}$ gives us the required element γ. Since $c_{b,\epsilon}$ deforms uniquely as a periodic point, the monodromy action of γ on this point is trivial as required. To prove that the second condition is satisfied, let $C \subset \mathbb{A}^2$ be the curve consisting of all points (z, c) so that z is a preperiodic point of preperiod 1 and period b of $\phi_{c,\epsilon}$ (with ϵ fixed as above). It suffices to prove that C is smooth at the point $(\epsilon, c_{b,\epsilon})$.

To see this, we consider the explicit equation for C. It is given in an neighbourhood of $(\epsilon, c_{b,\epsilon})$ by

$$\phi_{c,\epsilon}^{b+1}(z) - \phi_{c,\epsilon}(z) = 0. \tag{4-1}$$

To see that it is smooth at $(c_{b,\epsilon}, \epsilon)$, it suffices to substitute ϵ for z and check that the resulting polynomial in c, $\phi_{c,\epsilon}^{b+1}(\epsilon) - \phi_{c,\epsilon}(\epsilon)$, has $c_{b,\epsilon}$ as a root of multiplicity 1. However, from the definition of $\phi_{c,\epsilon}$ it follows that

$$\phi_{c,\epsilon}^{b+1}(\epsilon) - \phi_{c,\epsilon}(\epsilon) = \phi_{c,\epsilon}^{b+1}(0) - \phi_{c,\epsilon}(0),$$

so we may replace ϵ by 0. In a neighbourhood of the point $(0, c_{b,\epsilon})$, the curve given by (4-1) is smooth since it parametrises periodic points of period b and the periodic point 0 of $\phi_{c_{b,e},\epsilon}$ is of multiplicity 1. To show that the multiplicity of $c_{b,\epsilon}$ as a root of $\phi_{c,\epsilon}^{b+1}(0) - \phi_{c,\epsilon}(0)$ is 1, we may then specialise ϵ to 0, so it suffices to consider the multiplicity of c_b as a root of the polynomial $P_b(c) = \phi_c^b$. By Lemma 4.2, this multiplicity is indeed 1 as required. □

To prove the transitivity of the monodromy action on $\mathrm{Prep}_{\phi_{c,\epsilon}}(a, b)$ for $a > 1$, we shall need some results about Misiurewicz points. We refer the reader to [Lau and Schleicher 1994; Eberlein 1999] for the basic facts that we use below, which generalise results proved in [Douady and Hubbard 1984] in the case $d = 2$. Recall that $c_0 \in \mathbb{C}$ is called a Misiurewicz point if c_0 is a strictly preperiodic point of the map ϕ_{c_0}. By the results of [op. cit.], for any strictly preperiodic angle $\theta \in \mathbb{Q}/\mathbb{Z}$, there is a Misiurewicz point c_0 such that the parameter ray with angle θ lands at c_0. By [Eberlein 1999, Lemma 8.3], the preperiod of θ (with respect to multiplication by d) is equal to the preperiod of c_0 (with respect to ϕ_{c_0}) and the period of the kneading sequence of θ, $K(\theta)$, is equal to the period of c_0 (with respect to ϕ_{c_0}). For $\theta = 1/(d^a \cdot (d^b - 1))$, $a, b > 0$, the preperiod of θ is a and the period of $K(\theta)$ is b, so there exists a Misiurewicz point with any preperiod $a > 0$ and period b.

A point $\lambda \in \mathbb{C}$ is called parabolic of period b if it is the landing point of a parameter ray with angle θ that is periodic of period b. By results from [Douady and Hubbard 1984; Eberlein 1999], a parabolic point is never a Misiurewicz point.

To prove the transitivity of the monodromy action for $a > 1$, we shall need the following analogue of Lemma 4.2, due to Douady and Hubbard:

Lemma 4.4. *For a Misiurewicz point c_0 as above, the equation $\phi_c^{a+1+b}(0) - \phi_c^{a+1}(0) = 0$ has a simple root at $c = c_0$.*

Proof. The lemma is formulated and proved for $d = 2$ as Corollary 8.5 of [Douady and Hubbard 1984]; however, Proposition 8.5 of the same work holds for general d, and so the proof goes through if we substitute Theorems 8.1 and 8.2 of [Eberlein 1999] for Douady and Hubbard's Theorem 8.2. □

Lemma 4.5. *The monodromy action of the 2-dimensional family of polynomials $\phi_{c,\epsilon}$ on $\mathrm{Prep}_{\phi_{c,\epsilon}}(a, b)$ is transitive for all $a, b > 0$.*

Proof. We already know transitivity if $a = 1$. Thus, by induction, we may assume $a > 1$, and then it suffices to prove that for the 1-dimensional family of polynomials ϕ_c, there exists $x \in \mathrm{Per}_{\phi_c}(a - 1, b)$ and an element γ of the monodromy such that $\gamma(x) = x$ and γ induces a cyclic permutation on the d elements of $\mathrm{Per}_{\phi_c}(a - 1, b)$ comprising $\phi_c^{-1}(x)$.

Since 0 and ∞ are the only critical points of ϕ_λ, the preperiodic points for general λ are multiplicity-free. Let c be a Misiurewicz point of preperiod $a - 1$ and period b, and consider a small loop γ in the parameter plane around c such that all parabolic points of period b and all Misiurewicz points of preperiod a and period b are outside this loop. We note that the preperiodic points of ϕ_c of preperiod $\le a$ and period b are multiplicity-free since none of them are critical values and a Misiurewicz point is never a parabolic point. This remains true in a neighbourhood of c, so we may assume that this holds in a neighbourhood U of γ containing its interior; in particular, c deforms uniquely as a preperiodic point $c_\lambda \in \mathrm{Prep}_{\phi_\lambda}(a, b)$ as λ varies in this neighbourhood.

Since ϕ_c is totally ramified at 0, 0 is the unique element of $\mathrm{Prep}_{\phi_c}(a, b)$ mapping to c by ϕ_c. By construction, for any other $\lambda \in U$, there are d points of $\mathrm{Prep}_{\phi_\lambda}(a, b)$ mapping to c_λ and these points all come together at 0 as $\lambda \to c$. The set of these points in a neighbourhood of $(c, 0)$ is exactly the zero locus D of the polynomial $\phi_\lambda^{a+1+b}(z) - \phi_\lambda^{a+1}(z)$. By Lemma 4.4, the multiplicity of this after setting $z = 0$ is 1 at $(c, 0)$, so it follows that D must be smooth at this point. Since the inverse image of c in D is a single point, it follows that the map induced by the projection to the first factor is totally ramified of degree d at $(c, 0)$. Consequently, the monodromy around γ induces a cyclic permutation of order d on $\phi_\lambda^{-1}(c_\lambda)$. We may thus take $x = c_\lambda$ for any λ with $|\lambda| \ll 0$ to complete the proof. □

Proposition 4.6. *The Galois action on* $\mathrm{Prep}_{f_{n,d}}(a,b)$ *is transitive for all* $a, b > 0$.

Proof. Since $f_{n,d}$ is defined over $k_{n,d}$, the Galois action on $\mathrm{Prep}_{f_{n,d}}(a,b)$, $a, b \geq 1$, is compatible with the natural surjections

$$\mathrm{Prep}_{f_{n,d}}(a+1,b) \xrightarrow{f_{n,d}} \mathrm{Prep}_{f_{n,d}}(a,b).$$

Thus, by induction on a, it suffices to show that for any $a > 0$, there exists an element $x \in \mathrm{Prep}_{f_{n,d}}(a-1,b)$ such that, for any $y, y' \in \mathrm{Prep}_{f_{n,d}}(a,b)$ with $f_{n,d}(y) = f_{n,d}(y')$, there exists an element γ of the monodromy such that $\gamma(y) = y'$.

If $n = 1$, the claim follows immediately from Lemma 4.5, so in the following, we shall assume $n > 1$.

As before, we may assume that $k = \mathbb{C}$. The proof of transitivity is similar to that for the case of periodic points except that we replace the use of the maps ϕ_c by $\phi_{c,\epsilon}$. So for $0 \leq i \leq n$, consider the family of endomorphisms $f_i : \mathbb{P}^n \times \mathbb{A}^{2n} \to \mathbb{P}^n \times \mathbb{A}^{2n}$ defined by

$$\begin{aligned} &f_i(([x_0,\dots,x_n],(c_1,\dots,c_n,\epsilon_1,\dots,\epsilon_n))) \\ &\quad = \big(\big[x_0(x_0-\epsilon_1 x_i)^{d-1}+c_1x_i^d,\dots,x_{i-1}(x_{i-1}-\epsilon_i x_i)^{d-1}+c_ix_i^d, x_i^d, \\ &\qquad\qquad x_{i+1}(x_{i+1}-\epsilon_{i+1}x_i)^{d-1}+x_i^d,\dots,x_n(x_n-\epsilon_n x_i)^{d-1}+c_nx_i^d\big], \\ &\qquad (c_1,\dots,c_n,\epsilon_1,\dots,\epsilon_n)\big). \end{aligned}$$

On the open affine U_i given by $x_i \neq 0$, f_i is the product of the n polynomials ϕ_{c_j,ϵ_j}.

Let $g_{n,d}$ be the d-power map as before. Contrary to the case of periodic points, the map from the locus of preperiodic points to the base is not étale at preperiodic points of $g_{n,d}$ contained in its ramification locus. However, the map is étale at preperiodic points all of whose coordinates are nonzero, and this will suffice (except when $d = 2$) for our needs (compare the discussion of monodromy on page 589).

Suppose $a = 1$. Let ζ be a primitive (d^b-1)-th root of unity, and let $x' \in \mathrm{Per}_{g_{n,d}}(b)$ be the point $[\zeta, 1, \dots, 1, 1]$, and let $x = g_{n,d}(x')$. The preperiodic points y such that $g_{n,d}(y) = x$ are of the form $[\zeta\xi_1, \xi_2, \dots, \xi_n, 1]$, where all the ξ_i are d-th roots of unity and at least one of them is not equal to 1.

We now also assume that $d > 2$. Using the monodromy action of the family f_n and Lemma 4.3, it follows that we may assume that $\xi_i = \xi$, where ξ is a fixed d-th root of unity or $\xi_i = 1$.

Since $d > 1$, there exists a d-th root of unity ξ' such that $\xi' \neq 1, \xi$. Let $y = [\zeta\xi, 1, \dots, 1]$ and $y' = [\zeta\xi, \xi, \dots, \xi, 1, \dots, 1]$, where there are n_1 ξ's and n_2 1's with $n_1, n_2 > 0$. Each step of the following sequence of transformations is given either by multiplying through by a constant or by applying the monodromy of f_i for some i such that the i-th coordinate is equal to 1:

$$
\begin{aligned}
y = [\zeta\xi, 1, \dots, 1] &\to [1, \zeta^{-1}\xi^{-1}, \dots, \zeta^{-1}\xi^{-1}] \\
&\to [1, \zeta^{-1}\xi'^{-1}, \dots, \zeta^{-1}\xi'^{-1}, \zeta^{-1}\xi^{-1}, \dots, \zeta^{-1}\xi^{-1}] \\
&\to [\zeta\xi, \xi\xi'^{-1}, \dots, \xi\xi'^{-1}, 1 \dots, 1] \\
&\to [\zeta\xi, \xi, \dots, \xi, 1, \dots, 1] = y'.
\end{aligned}
$$

In the last transformation, we also use the fact that $\xi\xi'^{-1}$ is a d-th root of unity not equal to 1 so, like ξ and ξ', an element of $\mathrm{Prep}_{\phi_0}(1, 1)$.

Let $y'' = [\zeta, \xi, \dots, \xi, 1 \dots, 1]$, where there are n_1 ξ's and n_2 1's with $n_1, n_2 > 0$ as before. Since $d > 2$, there exists $\zeta' \in \mathrm{Per}_{\phi_0}(b)$ such that $\zeta\zeta'^{-1}$ also has period b. We then have a similar sequence of transformations:

$$
\begin{aligned}
y' = [\zeta\xi, \xi, \dots, \xi, 1, \dots, 1] &\to [1, \zeta^{-1}, \dots, \zeta^{-1}, \zeta^{-1}\xi^{-1}, \dots, \zeta^{-1}\xi^{-1}] \\
&\to [1, \zeta'\zeta^{-1}, \dots, \zeta\zeta^{-1}, \zeta^{-1}\xi^{-1}, \dots, \zeta^{-1}\xi^{-1}] \\
&\to [\zeta\xi, \zeta'\xi, \dots, \zeta'\xi, 1, \dots, 1] \\
&\to [\zeta\xi, \zeta'\xi', \dots, \zeta'\xi', 1, \dots, 1] \\
&\to [\zeta\xi(\zeta'\xi')^{-1}, 1, \dots, 1, (\zeta'\xi')^{-1}, \dots, (\zeta'\xi')^{-1}] \\
&\to [\zeta(\zeta'\xi)^{-1}, 1, \dots, 1, (\zeta'\xi)^{-1}, \dots, (\zeta'\xi)^{-1}] \\
&\to [\zeta, \zeta'\xi, \dots, \zeta'\xi, 1, \dots, 1] \\
&\to [1, \zeta^{-1}\zeta'\xi, \dots, \zeta^{-1}\zeta'\xi, \zeta^{-1}, \dots, \zeta^{-1}] \\
&\to [1, \zeta^{-1}\xi, \dots, \zeta^{-1}\xi, \zeta^{-1}, \dots, \zeta^{-1}] \\
&\to [\zeta, \xi, \dots, \xi, 1, \dots, 1] = y''.
\end{aligned}
$$

Here we have used that ζ, ζ', $\zeta\zeta'^{-1}$ and their inverses are in $\mathrm{Per}_{\phi_0}(b)$ and each one of these multiplied by ξ, ξ', $\xi\xi^{-1}$ or any of their inverses is an element of $\mathrm{Prep}_{\phi_0}(1, b)$. By symmetry, we then conclude that the desired transitivity holds in this case.

Now suppose $d = 2$. In this case, $\mathrm{Prep}_{\phi_0}(1, 1) = \{-1\}$, a singleton, so the above argument breaks down, and we will need to consider paths passing through elements of $\mathrm{Prep}_{g_{n,d}}(1, b)$, one of whose coordinates is 0. This is justified by Lemma 4.7 below.

As before, let $y = [-\zeta, \dots, -\zeta, 1]$ with ζ a primitive $(2^b - 1)$-th root of unity and $y' = [-\zeta, \dots, -\zeta, \zeta, \dots, \zeta, 1]$, where there is at least one $-\zeta$ and one ζ. We need to consider the cases $b = 1$ and $b > 1$ separately.

First suppose $b = 1$, so $\zeta = 1$. We then have the sequence of transformations

$$
\begin{aligned}
y = [-1, \dots, -1, 1] &\to [1, \dots, 1, -1] \to [1, \dots, 1, 0, -1] \\
&\to [-1, \dots, -1, 0, 1] \to [-1, \dots, -1, 1, 1].
\end{aligned}
$$

Repeating this procedure, we see that we can connect y to y', and by symmetry, the transitivity follows in this case.

Now suppose $b > 1$, so there exists $\zeta' \in \mathrm{Per}_{\phi_0}(b)$ so that, as before, $\zeta\zeta'^{-1} \in \mathrm{Per}_{\phi_0}(b)$. We then have a sequence of transformations

$$\begin{aligned} y = [-\zeta, \dots, -\zeta, 1] &\to [1, \dots, 1, -\zeta^{-1}] \to [1, \dots, 1, 0, -\zeta^{-1}] \\ &\to [-\zeta, \dots, -\zeta, 0, 1] \to [-\zeta, \dots, -\zeta, 1, 1] \\ &\to [1, \dots, 1, -\zeta^{-1}, -\zeta^{-1}] \to [1, \dots, 1, -\zeta'^{-1}, -\zeta^{-1}] \\ &\to [-\zeta^{-1}, \dots, -\zeta^{-1}, \zeta\zeta'^{-1}, 1] \to [-\zeta^{-1}, \dots, -\zeta^{-1}, \zeta, 1]. \end{aligned}$$

Repeating this procedure, we see that we can connect y to y', and then by symmetry, transitivity follows.

Finally, suppose $a > 1$. By induction, we can choose x to be an arbitrary point of preperiod $a-1$ and period b, so we let $x' = [\zeta, \zeta, \dots, \zeta, 1]$ where ζ is an element of $\mathrm{Prep}_{\phi_0}(a, 1)$ and $x = g_{n,d}(x)$. The points in $g_{n,d}^{-1}(x)$ are of the form $[\zeta\xi_1, \dots, \zeta\xi_n, 1]$ where ξ_1 is a d-th root of 1, so $\zeta\xi_i \in \mathrm{Prep}_{\phi_0}(a, 1)$. One sees that the monodromy acts transitively on $g_{n,d}^{-1}(x)$ simply by considering the monodromy of the family of maps f_n and applying Lemma 4.5. $\square$

Let $F : \mathbb{P}^n_k \times \mathrm{Mor}_{n,d/k} \to \mathbb{P}^n_k \to \mathrm{Mor}_{n,d/k}$ be the universal morphism of degree d, and consider $P_F(b+1, b) \subset \mathrm{Mor}_{n,d/k}$ (the notation is defined just before Lemma 2.2). The fibre of the projection map from $P_F(b+1, b)$ to $\mathrm{Mor}_{n,d/k}$ over any point $f \in \mathrm{Mor}_{n,d/k}$ consists of f-preperiodic points of preperiod at most 1 and period dividing b.

Lemma 4.7. *If $d = 2$ and $\mathrm{char}(k) \neq 2$, then $P_F(b+1, b)$ is smooth at any preperiodic point of $g_{n,2}$ of preperiod 1 and period b.*

Proof. We have $F(P_F(b+1, b)) = P_F(b)$. As we have already seen, $P_F(b)$ is smooth at all periodic points of $g_{n,d}$. Moreover, F is analytically locally at any point of $\mathrm{Prep}_{g_{n,d}}(1, b)$ with exactly one coordinate equal to 0, a cyclic cover of degree 2. Thus, to prove smoothness, it suffices to show that the discriminant of F intersects $P_F(b)$ transversely at any point of $\mathrm{Per}_{g_{n,d}}(b)$ with exactly one coordinate equal to 0.

To prove transversality, it then suffices to restrict to any subvariety of $\mathrm{Mor}_{n,d}$ and prove transversality for the induced subvarieties. By considering, say, the family f_i as in the proof of Proposition 3.3, where i is such that the i-th coordinate of the point under consideration is nonzero, we reduce to the case of the one parameter ϕ_c, $c \in \mathbb{A}^1_c$, and we need to prove transversality at the point $(z, c) = (0, 0)$. The discriminant locus is given by $z = 0$ (since $d = 2$) and the locus of fixed points by $z^2 + c - z = 0$, so the lemma follows. $\square$

Proof of Theorem 4.1. Let X be the Zariski closure of an infinite subset of preperiodic points. By Lemma 2.2(1), all preperiodic points of a quasipolarised map are defined over the algebraic closure of the base field, so we may assume without loss of generality that K is the algebraic closure of $k_{n,d}$ and X is defined over a finite extension of $k_{n,d}$. By replacing X by the union of its Galois conjugates, we may then assume that X is defined over $k_{n,d}$.

Since $\mathrm{Prep}_{f_{n,d}}(a, b)$ is finite for all a and b, there exists an infinite sequence of tuples (a_i, b_i), with $a_i \geq 0$ and $b_i > 0$, so that X contains a point $x_i \in \mathrm{Prep}_{f_{n,d}}(a_i, b_i)$ for all i. Since X is defined over $k_{n,d}$, it follows from Propositions 3.3 and 4.6 that $\mathrm{Prep}_{f_{n,d}}(a_i, b_i) \subset X$ for all i. We let X' be the Zariski closure of the specialisations, over the point in $\mathrm{Mor}_{n,d}$ corresponding to $g_{n,d}$, of the preperiodic points in $X(K)$. Since all preperiodic points lift to the generic fibre by Lemma 2.2, it follows that $X' \subset \mathbb{P}^n_k$ has the same properties as X but with respect to $\mathrm{Per}_{g_{n,d}}$.

For any i, the set of points in $\mathbb{P}^n_{\bar{k}}$ of the form $[\xi_0, \dots, \xi_{n_1}, 1]$ with $\xi_i \in \mathrm{Prep}_{\phi_0}(a_i, b_i)$ is contained in $\mathrm{Prep}_{g_{n,d}}(a_i, b_i)$ and hence in X'. As in the proof of Theorem 3.6, it then follows that $X' = \mathbb{P}^n_k$, so $X = \mathbb{P}^n_{k_{n,d}}$. □

Remark 4.8. Note that we do not use the full strength of the genericity hypothesis in the proofs of this section or of the previous one. It suffices to assume that the morphism under consideration corresponds to the generic point of an irreducibility subvariety of $\mathrm{Mor}_{d,n/k}$ that contains all the families $f_{c,\epsilon}$ and is smooth at $g_{n,d}$. Since all the $f_{c,\epsilon}$ are smooth and have dimension $2n$, there exist such subvarieties for all n with dimension independent of d.

5. The dynamical “Mordell–Lang” conjecture for generic endomorphisms

Let (X, f) be an algebraic dynamical system over a field K of characteristic 0; i.e., X is an algebraic variety and $f : X \to X$ is a morphism. The conjecture of Ghioca and Tucker [2009] asserts that, if $x \in X(K)$ and Y a subvariety of X are such that $O_f(x) \cap Y(K)$ is infinite, then there is a periodic subvariety Z of X with $Z \subset Y$ and $Z(K) \cap O_f(x) \neq \varnothing$. It has been proved when f is étale by Bell et al. [2010] and in a few other cases. It is not known in general if $X = \mathbb{P}^n_K$ and $\deg(f) > 1$; this was the original case investigated by Denis [1994], who proved the assertion under the assumption that $O_f(x) \cap Y(K)$ is large in a suitable sense.

For $(X, f) = (\mathbb{P}^n_K, f)$, with f a generic endomorphism, by Theorem 3.6, there are no nontrivial f-periodic subvarieties contained in $\mathbb{P}^n_K$, so the conjecture in this case is equivalent to the following:

Theorem 5.1. *Let $f : \mathbb{P}^n_K \to \mathbb{P}^n_K$ be a generic endomorphism of degree $d > 1$ over an algebraically closed field K of characteristic zero. For each $x \in \mathbb{P}^n(K)$, every infinite subset of $O_f(x)$ is Zariski-dense in $\mathbb{P}^n_K$.*

The idea of the proof is as follows. We first use specialisation to reduce to the case that K is a finite extension of $k_{n,d}$. We then show using a p-adic argument, for a prime p dividing d, that any Y such that $O_f(x) \cap Y$ is infinite must contain infinitely many periodic points, or there exists a prime q not diving d, such that x and f have specialisations $\bar{x}$ and $\bar{f}$ over $\mathbb{F}_q$ with $\bar{f}$ étale on the orbit of $\bar{x}$. Both these conditions lead to the conclusion that $Y = \mathbb{P}^n_K$, the first from Theorem 4.1 and second by using a result of Bell, Ghioca and Tucker, which we recall:

Lemma 5.2. *Let $L/\mathbb{Q}_p$ be a finite extension, $\pi : X \to \operatorname{Spec}(R)$ a smooth scheme of finite type over the ring of integers R of L and $f : X \to X$ a morphism over $\operatorname{Spec}(R)$. Suppose $x \in X(R)$ is such that f is étale on the orbit of x. If $Y \subset X$ is any closed subscheme with $Y \cap O_f(x)$ infinite, then Y_L contains a positive-dimensional periodic subvariety of X_L.*

Proof. If $L = \mathbb{Q}_p$, this is an immediate consequence of the results in [Bell et al. 2010]. For general L, it follows from the methods in the same work, if one replaces Theorem 3.3 therein by Theorem 7 of [Amerik 2011]. □

Lemma 5.3. *Theorem 5.1 for arbitrary extensions K of $k_{n,d}$ follows from the case of finite extensions.*

Proof. Without loss of generality, we may assume that $k = \mathbb{Q}$ and K is a finitely generated extension of $k_{n,d}$. Let $x \in \mathbb{P}^n(K)$, and let Y be a subvariety of $\mathbb{P}^n_K$ such that $O_f(x) \cap Y(K)$ is infinite.

Since K is finitely generated, there exists a smooth irreducible scheme M of finite type over $\mathbb{Z}$ with function field K and a dominant morphism $\pi : M \to \mathrm{Mor}_{n,d}$ inducing the inclusion $k_{n,d} \subset K$ on function fields. Let $\boldsymbol{f} : M \times \mathbb{P}^n \to M \times \mathbb{P}^n$ be the pullback of the universal morphism from $\mathrm{Mor}_{n,d} \times \mathbb{P}^n$, so $\boldsymbol{f}$ restricted to the generic fibre of the first projection is equal to f. Let $\boldsymbol{x}$ and $\boldsymbol{Y}$ be the Zariski closures of x and Y, respectively, in $M \times \mathbb{P}^n$. By shrinking M if necessary, we may assume that $\boldsymbol{x}$ and $\boldsymbol{Y}$ are flat over M.

Since p is a dominant finite-type morphism, there exists a point f' of M (which we think of as an endomorphism of $\mathbb{P}^n$ using π) mapping to the generic point of $\mathrm{Mor}_{n,d}$ and so that the residue field K' at f' is a finite extension of $k_{n,d}$. Let x' and Y' denote the fibres of $\boldsymbol{x}$ and $\boldsymbol{Y}$, respectively, over f'. If $O_{f'}(x')$ is infinite, then so is $O_{f'}(x') \cap Y'$; hence, by the condition on K', it would follow that $Y' = \mathbb{P}^n_{K'}$, which (by flatness) implies $Y = \mathbb{P}^n_K$.

If $O_{f'}(x')$ is finite, then x' is f'-preperiodic. Since f' is generic, it follows from Corollary 3.5 that $O_{f'}(x')$ does not intersect the ramification locus of f'. Let $|O_{f'}(x')| = n$, and let $\boldsymbol{Z}$ denote the Zariski closure of $\{x, f(x), \dots, f^{n-1}(x)\}$ in $M \times \mathbb{P}^n$. Let $\boldsymbol{R} \subset M \times \mathbb{P}^n$ be the ramification locus of $\boldsymbol{f}$, and consider the closed subset $\boldsymbol{R} \cap \boldsymbol{Z}$ of $M \times \mathbb{P}^n$. By the above, the fibre of this subset over f' is empty, so by the properness of $\mathbb{P}^n$, its projection in M is a proper closed subset. Replacing M by the complement of this subset, we may assume that $\boldsymbol{R} \cap \boldsymbol{Z} = \varnothing$.

Now let f'' be any closed point (which we again think of as an endomorphism) of M that lies in the closure of f'. Since M is of finite type over $\mathbb{Z}$, the residue field of f'' is a finite field F. Let x'' and Y'' denote the fibres of $\boldsymbol{x}$ and $\boldsymbol{Y}$, respectively, over f''. Since f'' is in the closure of f', x'' is in the closure of x'; hence, $|O_{f''}(x'')| \leq n$. Since $\boldsymbol{R} \cap \boldsymbol{Z} = \varnothing$, it follows that f'' is unramified at all points of $O_{f''}(x'')$. Let $W(F)$ be the ring of Witt vectors of F. Since M is smooth over $\mathbb{Z}$, by Hensel's lemma, the set of points in $M(W(F))$ that specialise to f'' is in bijection (after choosing local coordinates) with $W(F)^n$. The subset consisting of points that lie in a proper closed subscheme of M is a countable union of nowhere dense (in the adic topology) subsets. It follows by Baire's theorem that there exists a point in $M(W(F))$ specialising to f'' and not lying in any proper closed subscheme. Letting L be the quotient field of $W(K)$, it follows that the image of the induced map from $\operatorname{Spec}(L)$ to M must be the generic point. We thus get an inclusion of K into L, and we may apply Lemma 5.2 with $R = W(K)$ and $X = \operatorname{Spec}(W(F)) \times \mathbb{P}^n$ to the base change of $\boldsymbol{f}$, $\boldsymbol{x}$ and $\boldsymbol{Y}$ via the morphism $\operatorname{Spec}(W(F)) \to M$ to conclude using Theorem 3.6(1) that $Y = \mathbb{P}^n_K$. □

Lemma 5.4. *Let p be a prime and $g_{n,d,p}$ denote the endomorphism of $\mathbb{P}^n_{\mathbb{F}_p}$ given by raising each coordinate to its d-th power. Let $X \subset \mathbb{P}^n_{\overline{\mathbb{F}}_p}$ be a positive-dimensional subvariety. Then the set $\bigcup_{r\geq 0} g^r_{n,d,p}(X(\overline{\mathbb{F}}_p))$ contains periodic points of infinitely many distinct periods.*

Proof. Since $g_{n,d,p}$ preserves the standard decomposition of $\mathbb{P}^n$ as a disjoint union of affine spaces, by projecting to a suitable coordinate, we reduce to the statement for $n = 1$, in which case the statement is obvious. □

Remark 5.5. We expect that the lemma holds with $g_{n,d,p}$ replaced by an arbitrary quasipolarised morphism — or even more generally with some extra conditions on X — defined over a finite field, but this seems much harder to prove. However, for endomorphisms of abelian varieties the corresponding statement can indeed be proved.

Lemma 5.6. *Let $d > 1$ be an integer and p a prime such that $p \mid d$. Let ϕ be the morphism $\mathbb{A}^2 \to \mathbb{A}^2$ given by $\phi(x, c) = (x^d + c, c)$ over the field $\overline{\mathbb{F}}_p$.*

(a) *For $c \in \overline{\mathbb{F}}_p$, the monodromy action on the set of fixed points of ϕ_c is transitive.*

(b) *Let $X \subset \mathbb{A}^2$ be an irreducible subvariety of dimension* 1, *mapping dominantly to $\mathbb{A}^1$ via the second projection p_2. Assume that the intersection of X with the generic fibre of p_2 is not a preperiodic point of ϕ. Then the ϕ-periods of the points in $X(\overline{\mathbb{F}}_p)$* (*which are all preperiodic*) *are unbounded.*

Proof. We will use elementary intersection theory on $\mathbb{P}^1 \times Y$ with Y a smooth projective curve.

Let $X_n = \phi^n(X)$ for $n \geq 0$. By replacing X by $\phi^r(X)$ for some $r \geq 0$, we assume that the map $p_2 : X_n \to \mathbb{P}^1$ has degree e for all n.

For any integer $b > 0$, let P_b be the locus of points in $\mathbb{A}^2$ that are ϕ-periodic of period b. Since $p \mid d$, ϕ is inseparable on the fibres of p_2, so the graph of $\phi^b|_{\mathbb{P}^1 \times \{c\}}$ intersects the diagonal in $\mathbb{A}^2$ transversely for all $c \in \mathbb{A}^1$ and all b. It follows that P_b is a finite étale cover of $\mathbb{A}^1$ via p_2 and $P_b \cap P_{b'} = \varnothing$ for $b \neq b'$.

Let $\overline{P}_b$ be the closure of P_b in $\mathbb{P}^1 \times \mathbb{P}^1 \supset \mathbb{A}^1 \times \mathbb{A}^1 = \mathbb{A}^2$, and let $\overline{X}_n$ be the closure of X_n in $\mathbb{P}^1 \times \mathbb{P}^1$. The curve P_b is a subcurve of the curve Q_b in $\mathbb{A}^2$ with equation

$$(\ldots((x^d + c)^d + c)^d \ldots)^d + c - x = 0,$$

where we have b pairs of brackets. Replacing c by $1/c'$ in the above equation and multiplying through by ${c'}^{d^{(b-1)}}$, we get the equation

$$(\ldots((c'x^d + 1)^d + {c'}^{d^{b-1}-d^{b-2}})^d \ldots)^d + {c'}^{d^{b-1}-1} - {c'}^{d^{b-1}} x = 0.$$

It follows that the only point on all the $\overline{P}_b$ intersecting the fibre over $c = \infty$ is the point at infinity on this fibre and the support of $\overline{P}_b \cap \overline{P}_{b'}$ is equal to this point if $b \neq b'$. When $b = 1$, the equation is

$$c'x^d + 1 - c'x = 0.$$

One then sees that $\overline{P}_1$ is irreducible since the equation shows that it is smooth at the point (∞, ∞) and the closure in $\mathbb{P}^1 \times \mathbb{P}^1$ of any irreducible component of P_1 must contain this point; this proves (a).

If the ϕ-periods of all points in $X(\overline{\mathbb{F}}_p)$ are bounded, then there must exist some $b > 0$ so that $|\overline{X}_n \cap \overline{P}_b| \to \infty$ as $n \to \infty$ and, by assumption, $\overline{X}_n \not\subseteq \overline{P}_b$ for all b. Writing $[\overline{X}_n] = e[\{0\} \times \mathbb{P}^1] + a_n[\mathbb{P}^1 \times \{0\}]$ in $\mathrm{NS}(\mathbb{P}^1 \times \mathbb{P}^1)$, it follows that $a_n \to \infty$ as $n \to \infty$.

Let $\overline{X}'$ be the normalisation of $\overline{X}$, and denote by $\eta : \overline{X}' \to \mathbb{P}^1$ the composition of the normalisation map and the projection p_2. Let $x_1, x_2, \ldots, x_r$ be the points in $\overline{X}$ mapping to the point (∞, ∞) in $\mathbb{P}^1 \times \mathbb{P}^1$. If $r = 0$, it follows that $X \cap P_b \neq \varnothing$ for all b, so we may assume that $r > 0$. Let $b_1, b_2, \ldots, b_{r+1}$ be any distinct integers, and let $\gamma : Y \to \overline{X}'$ be a Galois cover such that there are components S_i of $(\mathrm{id} \times \eta\gamma)^{-1}(\overline{P}_{b_i})$ in $\mathbb{P}^1 \times Y$ that map isomorphically to Y via the second projection. Let Z be the section of this projection induced by the tautological section of $\mathbb{P}^1 \times \overline{X}'$, and let $Z_n = \psi^n(Z)$, where ψ is the map of $\mathbb{P}^1 \times Y$ induced by ϕ. We may write $[S_i] = [* \times Y] + s_i[\mathbb{P}^1 \times *]$ and $[Z_n] = [* \times Y] + a_n[\mathbb{P}^1 \times *]$ in $\mathrm{NS}(\mathbb{P}^1 \times Y)$ where $s_i \geq 0$ and $*$ denotes any point. It follows that the intersection number $S_i \cdot Z_n = s_i + a_n \to \infty$ as $n \to \infty$ for all $i = 1, 2, \ldots, r$.

Now consider the local intersection multiplicity $I_y(S_i, Z_n)$ of S_i and Z_n at a point $\infty \times y$, where $y \in Y$ is such that $\gamma(y) = x_j$ for some j. If this is bounded for all such y and all n, since $S_i \cdot Z_n \to \infty$, it would follow that for large n, S_i and Z_n must intersect in a point (z, y') such that $\eta\gamma(y') \in \mathbb{A}^1 \subset \mathbb{P}^1$, which implies that $X_n \cap P_{b_i} \neq \varnothing$.

Suppose this is not the case, so $I_y(S_i, Z_{n_l}) \to \infty$ for some infinite sequence $n_l \to \infty$. Since the S_i are all distinct smooth curves and there are only finitely many of them, it follows that $I_y(T, Z_{n_l})$ must remain bounded as $n_l \to \infty$, where T runs over all $S_{i'}$ for $i' \neq i$ and all of their Galois conjugates. Up to Galois conjugation, there are only r points y as above, so it follows that we must have that, for all large n_l, there exists $i_{n_l} \in \{1, 2, \dots, r+1\}$ so that $I_y(S_{i_{n_l}}, Z_{n_l})$ is bounded by an integer independent of n_l for all y as above. It follows that we must have $X_{n_l} \cap P_{b_{i_{n_l}}} \neq \varnothing$.

By choosing infinitely many disjoint sets of $r+1$ distinct integers $\{b_1, b_2, \dots, b_{r+1}\}$ as above, we see that $X_{n_b} \cap P_b \neq \varnothing$ for infinitely many distinct integers b (and n_b depending on b). Since all the P_b are disjoint, it follows that X contains preperiodic points of infinitely many distinct periods. □

Remark 5.7. We also expect this lemma to hold in much greater generality, e.g., for any 1-parameter family of maps defined over $\overline{\mathbb{F}}_p$.

The following lemma is the key to our construction of a periodic point in Y under the assumption that $O(x) \cap Y(K)$ is infinite:

Lemma 5.8. *Let $L/\mathbb{Q}_p$ be a finite extension, $\pi : X \to \mathrm{Spec}(R)$ a smooth projective scheme over the ring of integers R of L and $f : X \to X$ a quasipolarised morphism over $\mathrm{Spec}(R)$. Assume that the differential of f, df, is 0 on the special fibre of X. For any $x \in X(L) = X(R)$, let b be the period of the reduction $\bar{x}$ of x in the special fibre of X. Then for any integer $a \geq 0$, the sequence of points $f^{a+bn}(x)$ converges to a periodic point of $X(L)$ of period b.*

Proof. Replacing f by f^n and x by $f^{a'}(x)$, for any integer a' greater than the preperiod of $\bar{x}$, we may assume that $\bar{x}$ is a fixed point of f, and we then need to prove that $f^n(x)$ converges to a fixed point.

Since f is quasipolarised, by Lemma 2.2(2), $\bar{x}$ lifts to a fixed point y of f defined over a finite extension of L; by replacing L by this extension, we may assume that $y \in X(L)$.

Let A be the completion of the local ring of $\bar{x}$ on X. Since π is smooth, $A \cong R[[z_1, z_2, \dots, z_n]]$, where $n+1 = \dim(X)$. Using any such isomorphism, the set of points of $X(L)$ that specialise to $\bar{x}$ is identified with the set $(m_R)^n$, where m_R is the maximal ideal of R. We fix such an isomorphism, which we also assume identifies y with $(0, \dots, 0) \in (m_R)^n$.

Since $\bar{x}$ is a fixed point of f, f induces an endomorphism of A that, with respect to the chosen isomorphism, is given by an n-tuple of elements $(f_1, f_2, \dots, f_n)$ in the maximal ideal of the local ring $R[\![z_1, z_2, \ldots, z_n]\!]$. Moreover, since f fixes y, it follows that the constant term of each f_i is 0. Since df is assumed to be zero on the special fibre of X, it follows that the coefficients of the linear term of each f_i lies in m_R. For any $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in (m_R)^n$, let $|\lambda| = \max_i\{|\lambda_i|\}$. The conditions on the f_i imply that, for any such λ, $|f(\lambda)| < |\lambda|$ if $\lambda \neq (0, \ldots, 0)$. Since R is a discrete valuation ring, it follows that for any such λ we have that $f^n(\lambda) \to (0, 0, \ldots, 0)$ as $n \to \infty$; hence, $f^n(x) \to y$ as $n \to \infty$. □

Proof of Theorem 5.1. By Lemma 5.3, we may assume that K is a finite extension of $k_{n,d}$. Furthermore, we may assume without loss of generality that our base field $k = \mathbb{Q}$.

Let $x \in \mathbb{P}^n(K)$, and assume that Y is a subvariety defined over K such that $I = O(x) \cap Y(K)$ is infinite. Let X be the Zariski closure of the image of x in $\mathrm{Mor}_{n,d} \times \mathbb{P}^n_{\mathbb{Z}}$, and let χ denote the map $X \to \mathrm{Mor}_{n,d}$ induced by projection to the first factor.

Let p be a prime dividing d. Since $\mathrm{Mor}_{n,d}$ is smooth over $\mathrm{Spec}(\mathbb{Z})$, there is a map $g : \mathrm{Spec}(\mathbb{Z}_p) \to \mathrm{Mor}_{n,d}$ such that the generic point of $\mathrm{Spec}(\mathbb{Z}_p)$ maps to the generic point of $\mathrm{Mor}_{n,d}$ and the closed point maps to the point corresponding to $g_{d,n,p}$, the d-power map over $\mathbb{F}_p$. Since $p \mid d$, the differential of the endomorphism of $\mathbb{P}^n_{\mathbb{Z}_p}$ corresponding to g, which we also denote by g, is zero on the special fibre. Suppose the fibre $X_{g_{d,n,p}}$ of χ over $g_{d,n,p}$ is infinite. By Lemma 5.4, the set $\bigcup_{r\geq 0} g^r_{d,n,p}(X_{g_{d,n,p}}(\bar{\mathbb{F}}_p))$ contains infinitely many periodic points. By applying Lemma 5.8, we can lift all these periodic points to periodic points of $f_{n,d}$ contained in Y. It then follows from Theorem 4.1 that $Y = \mathbb{P}^n_K$. Thus, we may assume from now on that χ is finite over an open neighbourhood of $g_{n,d,p}$.

By replacing x by $f^r(x)$ for some large r, we may assume that $X_{g_{d,n,p}}$ contains a periodic point $x' = [x'_0, x'_1, \ldots, x'_n]$ with $x'_i \in \bar{\mathbb{F}}_p$. Since χ is finite in a neighbourhood of $g_{n,d,p}$, $\mathrm{Mor}_{d,n}$ is smooth, hence normal, and X is irreducible, it follows from the going-down theorem that if none of the $x'_i = 0$ then the fibre $X_{g_{n,d}}$ of χ over $g_{n,d}$ contains a point $\tilde{x}'$ lifting x'. By specialisation, it follows that for all large primes q the fibre of χ over $g_{n,d,q}$ contains a point all of whose coordinates are nonzero or, equivalently, not contained in the ramification locus of $g_{n,d,q}$. Since this locus is invariant under $g_{n,d,q}$, we may apply Lemma 5.2 to conclude the existence of a positive-dimensional periodic subvariety of Y, which, by Theorem 3.6, implies $Y = \mathbb{P}^n_K$.

We now use Lemma 5.6 to show that such an x' must exist, at least after replacing x by a Galois conjugate, or Y must contain infinitely many periodic points, both cases leading to the conclusion that $Y = \mathbb{P}^n_K$. Let x' be as above, and suppose that $x'_0 = 0$. Some other coordinate must be nonzero, so by symmetry, we may assume that $x'_n \neq 0$, and then by multiplying through by a scalar, we may assume $x'_n = 1$.

Consider the family of endomorphisms ψ_c of $\mathbb{P}^n_{\mathbb{F}_p}$ parametrised by $\mathbb{A}^1$ given by

$$\psi_c([x_0, x_1, \ldots, x_n]) = [x_0^d + cx_n^d, x_1^d, \ldots, x_n^d],$$

so $\psi_0 = g_{n,d,p}$. Note that, on the affine space given by the locus with $x_n \neq 0$, ψ_c is given in affine coordinates by $(z_0, z_1, \ldots, z_{n-1}) \mapsto (z_0^d + c, z_1^d, \ldots, z_{n-1}^d)$.

Let $S \subset \mathrm{Mor}_{n,d}$ be the subscheme corresponding to the family ψ_c. By the going-down theorem, there is an irreducible component T of $\chi^{-1}(S)$ mapping onto S and containing the point $(g_{n,d,p}, x')$. Let T' be the image of T in $\mathbb{P}^n_{\mathbb{F}_p}$ under the projection of X to $\mathbb{P}^n$. Since $x'_n = 1$, T' is not contained in the locus given by $x_n = 0$, so by projecting to the first n coordinates, we get a rational map ρ from T to $\mathbb{A}^n$.

Suppose the composition of ρ with the i-th projection is nonconstant for some i, $0 < i \leq n-1$. Since the action of ψ_c on the i-th coordinate doesn't depend on c, it follows that $T(\overline{\mathbb{F}}_p)$ must contain preperiodic points of arbitrarily large period. By Lemma 5.8 as before, we obtain infinitely many periodic points in Y, forcing $Y = \mathbb{P}^n_K$.

So suppose ρ composed with all the i-th projections are constant for $i > 0$, and let $\sigma : T \to \mathbb{A}^1 \times S$ be given by $(\pi_0\rho, \chi)$. By applying Lemma 5.6(b), it follows that if the image of T is not contained inside a preperiodic curve for the map ϕ (using the identification of S with $\mathbb{A}^1$) there must be ϕ preperiodic points in the image with unbounded period. By the construction of ψ_d, it follows that there are preperiodic points on T of unbounded period. As before, this implies that $Y = \mathbb{P}^n$.

The last case we need to consider is when the image of T lies in a preperiodic curve. By replacing x by an element in its orbit if necessary, we may assume that this image lies in the periodic locus. Now 0 is a fixed point of the map $z \mapsto z^d$, and the point $(0, 0)$ is contained in the image of T by construction. By Lemma 5.6(a), it follows that the point $(0, 1)$ is also in the image of T. We conclude that $X_{g_{d,n,p}}$ contains the periodic point $x'' = [1, x'_1, \ldots, x'_{n-1}, 1]$. By replacing x' with x'' and repeating the above argument if necessary, we conclude that Y contains infinitely many periodic points, in which case it must be $\mathbb{P}^n_k$, or $X_{g_{d,n,p}}$ contains a periodic point $x' = [x'_0, x'_1, \ldots, x'_n]$ with $x'_i \neq 0$ for all i. As we have already seen, this also implies that $Y = \mathbb{P}^n_K$, concluding the proof. □

Remark 5.9. Note that a statement similar to Remark 4.8 holds: it suffices to consider generic points of irreducible subschemes of $\mathrm{Mor}_{d,n}$ that contain all the families f_i and are smooth at the point $g_{n,d,p}$ for some prime p dividing d.

Acknowledgements

I thank Ekaterina Amerik, Amitava Bhattacharya, Bjorn Poonen, Joseph Silverman and Thomas Tucker for useful conversations and comments. I also thank the referee for his/her suggestions for improvements in the exposition.

References

[Amerik 2011] E. Amerik, “Existence of non-preperiodic algebraic points for a rational self-map of infinite order”, *Math. Res. Lett.* **18**:2 (2011), 251–256. MR 2012e:37182 Zbl 1241.14011

[Bell et al. 2010] J. P. Bell, D. Ghioca, and T. J. Tucker, “The dynamical Mordell–Lang problem for étale maps”, *Amer. J. Math.* **132**:6 (2010), 1655–1675. MR 2012a:37202 Zbl 1230.37112

[Bousch 1992] T. Bousch, *Sur quelques problemes de dynamique holomorphe*, Ph.D. thesis, Université de Paris-Sud, 1992.

[Denis 1994] L. Denis, “Géométrie et suites récurrentes”, *Bull. Soc. Math. France* **122**:1 (1994), 13–27. MR 95b:11065 Zbl 0795.14008

[Douady and Hubbard 1984] A. Douady and J. H. Hubbard, *Étude dynamique des polynômes complexes, I*, Publications Mathématiques d’Orsay **84**, Université de Paris-Sud, 1984. MR 87f:58072a Zbl 0552.30018

[Douady and Hubbard 1985] A. Douady and J. H. Hubbard, *Étude dynamique des polynômes complexes, II*, Publications Mathématiques d’Orsay **85**, Université de Paris-Sud, 1985. MR 87f:58072b Zbl 0571.30026

[Eberlein 1999] D. Eberlein, *Rational parameter rays of multibrot sets*, Master’s thesis, Technische Universität München, 1999.

[Fakhruddin 2003] N. Fakhruddin, “Questions on self maps of algebraic varieties”, *J. Ramanujan Math. Soc.* **18**:2 (2003), 109–122. MR 2004f:14038 Zbl 1053.14025

[Ghioca and Tucker 2009] D. Ghioca and T. J. Tucker, “Periodic points, linearizing maps, and the dynamical Mordell–Lang problem”, *J. Number Theory* **129**:6 (2009), 1392–1403. MR 2010i:37219 Zbl 1186.14047

[Ghioca et al. 2011] D. Ghioca, T. J. Tucker, and S. Zhang, “Towards a dynamical Manin–Mumford conjecture”, *Int. Math. Res. Not.* **2011**:22 (2011), 5109–5122. MR 2012i:37142 Zbl 1267.37110

[Lau and Schleicher 1994] E. Lau and D. Schleicher, “Internal addresses in the Mandelbrot set and irreducibility of polynomials”, preprint, Stony Brook University, 1994. arXiv math/9411238v1

[Morton 1998] P. Morton, “Galois groups of periodic points”, *J. Algebra* **201**:2 (1998), 401–428. MR 99a:12005 Zbl 0913.12005

[Tao and Vu 2006] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics **105**, Cambridge University Press, 2006. MR 2008a:11002 Zbl 1127.11002

Communicated by Joseph Silverman
Received 2012-11-30 Revised 2013-05-25 Accepted 2013-07-04

naf@math.tifr.res.in *School of Mathematics, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.609

The tame-wild principle for discriminant relations for number fields

John W. Jones and David P. Roberts

Consider tuples $(K_1, \ldots, K_r)$ of separable algebras over a common local or global number field F, with the K_i related to each other by specified resolvent constructions. Under the assumption that all ramification is tame, simple group-theoretic calculations give best possible divisibility relations among the discriminants of K_i/F. We show that for many resolvent constructions, these divisibility relations continue to hold even in the presence of wild ramification.

1. Overview

Let G be a finite group and let $\phi_1, \ldots, \phi_r$ be permutation characters of G. We say that a tuple $(K_1, \ldots, K_r)$ of separable algebras over a common ground field F has type $(G, \phi_1, \ldots, \phi_r)$ if for a joint splitting field K^{gal} one can identify $\text{Gal}(K^{\text{gal}}/F)$ with a subgroup of G such that the action of $\text{Gal}(K^{\text{gal}}/F)$ on $\text{Hom}_F(K_i, K^{\text{gal}})$ has character ϕ_i.

When F is a local or global number field, one has discriminants $\mathfrak{D}_{K_i/F}$ which are ideals in the ring of integers $\mathcal{O}_F$ of F. One can ask for the strongest divisibility relations among these discriminants which hold as $(K_1, \ldots, K_r)/F$ varies over all possibilities of a given type $(G, \phi_1, \ldots, \phi_r)$. This question has a simple group-theoretic answer if one restricts attention to tuples for which all ramification in each K_i/F is tame.

This paper focuses on the following phenomenon: *for many* $(G, \phi_1, \ldots, \phi_r)$*, the divisibility relations for tame* $(K_1, \ldots, K_r)/F$ *of type* $(G, \phi_1, \ldots, \phi_r)$ *hold also for arbitrary* $(K_1, \ldots, K_r)/F$ *of type* $(G, \phi_1, \ldots, \phi_r)$. In this case, we say that the tame-wild principle holds for $(G, \phi_1, \ldots, \phi_r)$. Our terminology "tame-wild principle" is intended to be reminiscent of the standard terminology "local-global principle": we are showing in this paper that simple tame computations can often solve a complicated wild problem, just as simple local computations can often solve a complicated global problem.

This work was partially supported by a grant from the Simons Foundation (#209472 to David Roberts).
MSC2010: primary 11S15; secondary 11S20, 11R32.
Keywords: number field, discriminant, ramification.

Section 2 provides an introductory example. Section 3 reviews some ramification theory centering on Artin characters, placing it in a framework which will be convenient for us. Section 4 states the tame-wild principle and gives two simple methods for proving instances of it.

If the tame-wild principle holds for a fixed G and any $(\phi_1, \dots, \phi_r)$ then we say it holds universally for G. Section 5 proves that the tame-wild principle holds universally for G in a small class of groups we call U-groups. Section 6 considers the remaining groups, called N-groups, finding that the tame-wild principle usually does not hold universally for them.

Sections 7 and 8 return to the more practical situation where one is given not only G but also a small list of naturally arising ϕ_i. Our theme is that the tame-wild principle is likely to hold, despite the negative results on N-groups. Section 7 focuses on comparing an arbitrary algebra K/F with its splitting field K^{gal}/F, proving that one of the two divisibility statements coming from the tame-wild principle holds for arbitrary G. Section 8 gives a collection of examples exploring the range of $(G, \phi_1, \dots, \phi_r)$ for which the tame-wild principle holds.

This paper is written with applications to tabulating number fields of small discriminant in mind. The topics in Section 2B, Section 7E, and Section 8A all relate to this application. Moreover, as we will make clear, the theory we present here still applies when permutation characters ϕ_i are replaced by general characters χ_i, and discriminants are correspondingly generalized to conductors. Applications to Artin L-functions of small conductor will be presented elsewhere.

2. An introductory example

In Section 2A, we provide an introductory instance of the tame-wild principle that we will revisit later to provide simple illustrations of general points. In Section 2B, we illustrate how this instance of the tame-wild principle gives an indirect but efficient way of solving a standard problem in tabulating number fields.

2A. *The tame-wild principle for* (S_5, ϕ_5, ϕ_6).

The Cayley–Weber type. For our introductory example, we take the type (S_5, ϕ_5, ϕ_6), where ϕ_5 is the character of the given degree five permutation representation, and ϕ_6 is the character of the degree six representation $S_5 \xrightarrow{\sim} \mathrm{PGL}_2(5) \subset S_6$. A pair of algebras (K_5, K_6) has type (S_5, ϕ_5, ϕ_6) exactly when K_6 is the Cayley–Weber resolvent of K_5 (see [Jensen et al. 2002, §2.3], for instance, for this notion). An explicit example over $\mathbb{Q}$ is given by $K_n = \mathbb{Q}[x]/f_n(x)$ with

$$f_5(x) = x^5 - 2x^4 + 4x^3 - 4x^2 + 2x - 4, \qquad D_5 = 2^8 3^4 5^1, \tag{2-1}$$

$$f_6(x) = x^6 - 2x^5 + 4x^4 - 4x^3 + 2x^2 - 4x - 6, \qquad D_6 = 2^{10} 3^4 5^3. \tag{2-2}$$

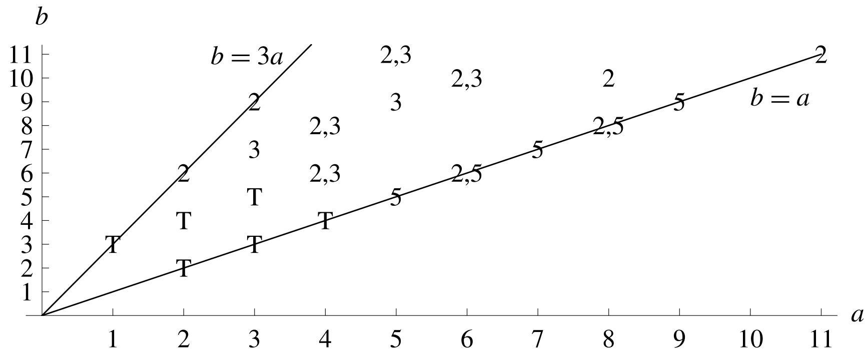


Figure 1. The complete list of pairs $(a, b) \neq (0, 0)$ which occur as $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$ for (S_5, ϕ_5, ϕ_6) over $\mathbb{Q}$. The pairs labeled T can occur with tame ramification, while the others can only occur for wild p-adic ramification as indicated.

In this example, the Galois group is all of S_5, discriminants $\mathfrak{D}_{K_n/\mathbb{Q}} = (D_n)$ are as indicated, and ramification is wild at 2 and tame at 3 and 5. We are concerned with exponent pairs $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$ on discriminants. Here $(a_2, b_2) = (8, 10)$, $(a_3, b_3) = (4, 4)$, $(a_5, b_5) = (1, 3)$, and otherwise $(a_{\mathfrak{p}}, b_{\mathfrak{p}}) = (0, 0)$.

All possibilities for $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$. Figure 1 gives all nonzero possibilities for $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$ over $\mathbb{Q}$. The fact that the tame list is complete is immediate from the general formalism of the next section. A brute force proof that the wild list is complete goes as follows: there are 113, 57, and 51 quintic algebras K_5 over $\mathbb{Q}_p$ for $p = 2$, 3, and 5 respectively [Jones and Roberts 2006]; for each, one can compute K_6 and thus the pairs $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$; the lists arising are the ones drawn in Figure 1. For larger number fields F, the list of possibilities for tame $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$ is exactly the same, but the list of possibilities for wild $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$ grows without bound.

The tame-wild principle. Figure 1 and the comment about general base fields F clearly illustrate two general phenomena about exponent vectors $(a_{\mathfrak{p}}, b_{\mathfrak{p}})$. First, in absolute terms, the exponents can be much larger in wild cases than they are in all tame cases. But second, in relative terms, one can hope that the ratios $a_{\mathfrak{p}}/b_{\mathfrak{p}}$ are quite similar in the wild and tame cases. We are interested in this paper only in the second phenomenon and so we systematically consider ratios.

In our example, the tame-wild principle is the statement that

$$\tfrac{1}{3} b_{\mathfrak{p}} \leq a_{\mathfrak{p}} \leq b_{\mathfrak{p}} \tag{2-3}$$

holds for all $(K_5, K_6)/F$ of type (S_5, ϕ_5, ϕ_6) and all primes $\mathfrak{p}$ of F. In other words, when $b_{\mathfrak{p}} \neq 0$ one must have $a_{\mathfrak{p}}/b_{\mathfrak{p}} \in \left[\frac{1}{3}, 1\right]$. We have summarized a proof that (2-3) holds when one restricts F to be $\mathbb{Q}$ or one of the $\mathbb{Q}_p$. We will see by a group-theoretic argument in Section 4C, not involving inspecting wild ramification at all, that (2-3) holds for general F. However the situation is subtle, as the analog of (2-3) holds for many (G, ϕ_n, ϕ_m) but not for all.

2B. ***Application to number field tabulation.*** The example of this section provides a convenient illustration of the application of tame-wild inequalities to number field tabulation. The right inequality of (2-3) globalizes to the divisibility relation $\mathfrak{D}_{K_5/F} \mid \mathfrak{D}_{K_6/F}$ which on the level of magnitudes becomes

$$|\mathfrak{D}_{K_5/F}| \leq |\mathfrak{D}_{K_6/F}|. \tag{2-4}$$

Consider the problem of finding sextic field extensions K_6/F with Galois group either $\mathrm{PSL}_2(5)$ or $\mathrm{PGL}_2(5)$. These all arise as Cayley–Weber resolvents of K_5/F with Galois group either A_5 or S_5. From (2-4), one sees that to find all K_6/F with $|\mathfrak{D}_{K_6/F}| \leq B$ it suffices to find all K_5/F with $|\mathfrak{D}_{K_5/F}| \leq B$, apply the Cayley–Weber resolvent, and keep those K_6/F with $|\mathfrak{D}_{K_6/F}| \leq B$. This indirect quintic method is enormously faster for large B, but the direct sextic method over $F = \mathbb{Q}$ was used in [Ford and Pohst 1992] and [Ford et al. 1998] for the $\mathrm{PSL}_2(5)$ and $\mathrm{PGL}_2(5)$ cases, respectively.

3. Character theory and discriminants

In this section, we review how Artin characters underlie discriminants. Each of the subsections introduces concepts and notation which play an important role in the rest of the paper. The notions we emphasize are slightly different from the most standard representation-theoretic notions. However they are appropriate here because all our characters are rational-valued.

3A. ***Class sets.*** Let G be a finite group. We say that two elements of G are *power-conjugate* if each is conjugate to a power of the other. Let $G^\sharp$ be the set of power-conjugacy classes. Thus one has a natural surjection $G \to G^\sharp$, with the fiber $C_\sigma \subset G$ above $\sigma \in G^\sharp$ being its set of representatives. The *order* $\bar{\sigma}$ of an element $\sigma \in G^\sharp$ is the order of a representing element $g \in C_\sigma$. Similarly the *power* σ^k of a class σ is the class of g^k for any representing element $g \in C_\sigma$.

When dealing with explicit examples, we most commonly indicate an element of $G^\sharp$ by giving its order and an extra identifying label, as in, e.g., 2B. To emphasize the role of order, we say that a class τ divides a class σ if some power of σ is τ. Thus divisibility of classes $\tau \mid \sigma$ implies divisibility of integers $\bar{\tau} \mid \bar{\sigma}$, but not conversely. In connection with divisibility, the quantity $[\sigma] = |C_\sigma|/\phi(\bar{\sigma})$ is useful, with $\phi(n) = |(\mathbb{Z}/n)^\times|$ the Euler ϕ-function. This quantity is integral because C_σ consists of $[\sigma]$ power-classes, each of size $\phi(\bar{\sigma})$. Alternatively, one can think of $G^\sharp$ as indexing conjugacy classes of cyclic subgroups of G, and then $[\sigma]$ is the number of cyclic subgroups of type σ.

Sections 5 and 6 systematically reason with class sets using diagrams based on the divisibility relation and the quantities $[\sigma]$. In general, G itself often recedes into the background of our considerations and the focus is on $G^\sharp$ and its inherited structures.

3B. *Characters.* Our calculations take place mainly in the ring $\mathbb{Q}(G^\sharp)$ of $\mathbb{Q}$-valued functions on $G^\sharp$. We also use the larger ring $\mathbb{R}(G^\sharp)$ of real-valued functions, so that we can use standard terms like cone, hull, and interval with their usual meaning. We make extensive use of the natural inner product on $\mathbb{Q}(G^\sharp)$, given by

$$(f_1, f_2) = \sum_{\sigma \in G^\sharp} \frac{|C_\sigma|}{|G|} f_1(\sigma) f_2(\sigma).$$

Important elements in $\mathbb{Q}(G^\sharp)$ for us include the characters ϕ_X of G-sets X. By definition, these characters are obtained by counting fixed points: $\phi_X(\sigma) = |X^g|$, for g any representative of σ. Both the identity class $e \in G^\sharp$ and the constant function $1 \in \mathbb{Q}(G^\sharp)$ usually play trivial roles in our situation. To efficiently remove these quantities from our attention, we define $G^{\sharp 0} = G^\sharp - \{e\}$ and let $\mathbb{Q}(G^\sharp)^0 \subset \mathbb{Q}(G^\sharp)$ be the orthogonal complement to 1.

The characters $\phi_{G/H}$ and a_H. Let H be a subgroup of G. Then the character of the G-set G/H is given by

$$\phi_{G/H}(\sigma) = \frac{|G||C_\sigma \cap H|}{|H||C_\sigma|}. \tag{3-1}$$

Taking $H = \{e\}$ gives the regular character ϕ_G with value $|G|$ at e and 0 elsewhere. We define the formal Artin character of H to then be the difference

$$a_H = \phi_G - \phi_{G/H}, \tag{3-2}$$

which lies in $\mathbb{Q}(G^\sharp)^0$. Here we use the adjective "formal" because often one talks about Artin characters only in the presence of fields, while currently we are in a purely group-theoretic setting.

The case that H is cyclic. The case where H is cyclic is particularly important to us. The generators of H all represent the same class $\tau \in G^\sharp$ and we use the alternative notation $a_\tau = a_H$, calling the a_τ tame characters for reasons which will be clear shortly in Section 3C.

To study the a_τ explicitly, it is convenient to make use of what we call *precharacters* $\hat{a}_\tau$, for $\tau \in G^\sharp$. By definition, $\hat{a}_e$ is the 0 function and otherwise $\hat{a}_\tau$ has two nonzero values:

$$\hat{a}_\tau(e) = |G|, \quad \hat{a}_\tau(\tau) = -\frac{|G|}{|C_\tau|}. \tag{3-3}$$

Tame characters and precharacters are related to each other via

$$a_\tau = \sum_{k|\bar\tau} \frac{\phi(\bar\tau/k)}{\bar\tau}\hat a_{\tau^k}, \quad \hat a_\tau = \sum_{k|\bar\tau} \frac{\bar\tau\mu(k)}{\phi(\bar\tau)k} a_{\tau^k}, \tag{3-4}$$

where μ is the Möbius μ-function taking values in $\{-1, 0, 1\}$. Thus, $a_e = \hat a_e = 0$, $a_\tau = \frac{\bar\tau-1}{\bar\tau}\hat a_\tau$ if $\bar\tau$ is prime, and otherwise a_τ and $\hat a_\tau$ are not proportional to each other. As τ ranges over $G^{\sharp 0}$, the $\hat a_\tau$ clearly form a basis for $\mathbb{Q}(G^\sharp)^0$. So the a_τ also form a basis for $\mathbb{Q}(G^\sharp)^0$

3C. *Artin characters.* Let F be a local or global number field. Let L/F be a Galois extension with Galois group identified with a subgroup of G. A permutation representation ρ of G gives an F-algebra K split by L. For $\mathfrak{p}$ a prime ideal of F, the discriminant exponent $c_\mathfrak{p}(K)$ depends only on the character $\phi \in \mathbb{Q}(G^\sharp)$ of ρ and in fact depends linearly on ϕ. The associated Artin character $a_{L/F,\mathfrak{p}}$ is the unique element of $\mathbb{Q}(G^\sharp)$ such that one has the general formula

$$c_\mathfrak{p}(K) = (a_{L/F,\mathfrak{p}}, \phi). \tag{3-5}$$

From $c_\mathfrak{p}(F) = 0$, one gets $(a_{L/F,\mathfrak{p}}, 1) = 0$ and so $a_{L/F,\mathfrak{p}} \in \mathbb{Q}(G^\sharp)^0$. One can completely compute $a_{L/F,\mathfrak{p}}$ by computing $c_\mathfrak{p}(K)$ for any $|G^{\sharp 0}|$ different K having characters which are linearly independent in $\mathbb{Q}(G^\sharp)/\mathbb{Q}$.

Before continuing, we note a subtlety that disappears in the Artin character formalism that we are reviewing. Namely, it can happen that nonisomorphic algebras K' and K'' give rise to the same permutation character ϕ. In this case K' and K'' are called arithmetically equivalent. They are indeed equivalent from the point of view of this paper, and any occurrence of K' can simply be replaced by K''.

An Artin conductor $a_{L/F,\mathfrak{p}}$ can be expressed directly in terms of inertia groups in their upper numbering as follows. Let $\mathfrak{P}$ be a prime of L above $\mathfrak{p}$ and let $I_{L/F,\mathfrak{P}} \subseteq \mathrm{Gal}(L/F) \subseteq G$ be the corresponding inertia group. Then one has rational numbers $1 \le s_1 < s_2 < \cdots < s_k$ and normal subgroups

$$I_{L/F,\mathfrak{P}} = I^{s_1} \supset I^{s_2} \supset \cdots \supset I^{s_k} \supset \{e\} \tag{3-6}$$

satisfying

$$a_{L/F,\mathfrak{p}} = \sum_{i=1}^{k} (s_i - s_{i-1}) a_{I^{s_i}}. \tag{3-7}$$

Here, for the sake of the conciseness of formulas, we put $s_0 = 0$. As a similar convention, we put $I^{s_{k+1}} = \{e\}$. The upper numbers s_i we are using here are called *slopes* in [Jones and Roberts 2006] and are designed to capture tame and wild ramification simultaneously; one has $s_i = u_i + 1$ where the u_i are the upper numbers used in the standard reference [Serre 1979].

If $s_1 = 1$ then I^{s_1}/I^{s_2} is cyclic of order prime to p. Otherwise, all the $I^{s_u}/I^{s_{u+1}}$ are abelian groups of exponent p. In particular, $I_{L/F,\mathfrak{P}}$ itself is a *p-inertial* group in the sense that it is an extension of a prime-to-p cyclic group by a p-group. In general, we say that a group is *inertial* if it is p-inertial for some prime p.

The prime $\mathfrak{p}$ is unramified in L/F if and only if $k = 0$ in which case $a_{L/F,\mathfrak{p}}$ is zero. The cases where $\mathfrak{p}$ is ramified but only tamely are those with $k = 1$ and $s_1 = 1$. In both these two settings, (3-7) becomes $a_{L/F,\mathfrak{p}} = a_\tau$ with τ being the class of any generator of any $I_{L/F,\mathfrak{P}}$. Thus the tame characters of the previous subsection are exactly the Artin characters which arise when ramification is tame.

3D. ***Bounds on Artin characters.*** Define cones in $\mathbb{R}(G^\sharp)^0$ spanned by characters or precharacters as follows:

$$\begin{aligned}
\text{the tame cone } \; & T_+(G) = \langle a_\tau \rangle, \\
\text{the wild cone } \; & W_+(G) = \langle a_{L/F,\mathfrak{p}} \rangle, \\
\text{the inertial cone } \; & \tilde{T}_+(G) = \langle a_I \rangle, \\
\text{the broad cone } \; & \hat{T}_+(G) = \langle \hat{a}_\tau \rangle.
\end{aligned}$$

The tame and broad cones are the simplest of these objects, as their generators are indexed by the small and explicit set $G^{\sharp 0}$. The inertial cone is also a purely group-theoretic object, although now more complicated as its generators are indexed by conjugacy classes of inertial subgroups. Finally, $W_+(G)$ is much more complicated in nature: its definition depends on the theory of p-adic fields, with $a_{L/F,\mathfrak{p}}$ running over all possible Artin characters, as above.

Our considerations in this section have established the following inclusions:

$$T_+(G) \subseteq W_+(G) \subseteq \tilde{T}_+(G) \subseteq \hat{T}_+(G). \tag{3-8}$$

The first inclusion holds because tame characters are special cases of Artin characters, the second by the expansion (3-7), and the third because all a_I take only positive values on $G^{\sharp 0}$.

4. The tame-wild principle

We begin in Section 4A by giving a formulation of the tame-wild principle in a somewhat abstract context, so that its motivation and structural features can be seen clearly. Next, Section 4B observes that the bounds from the previous section give techniques for group-theoretically proving instances of the tame-wild principle. Finally, Section 4C details one way of introducing coordinates to render everything explicit and Section 4D sketches alternative approaches.

4A. ***Abstract formulation.*** We seek settings where general ramification is governed by tame ramification. The statement that equality holds in $T_+(G) \subseteq W_+(G)$

is true for some G, in which case it is the ideal statement. For general G, we seek weaker statements in the same spirit. Accordingly, consider the orthogonal projection $a \mapsto a^V$ from $\mathbb{R}(G^\sharp)$ onto an arbitrary subspace $V \subseteq \mathbb{R}(G^\sharp)$. Let $T_+(G, V)$, $W_+(G, V)$, $\tilde{T}_+(G, V)$, and $\hat{T}(G, V)$ be the images of $T_+(G)$, $W_+(G)$, $\tilde{T}_+(G)$, and $\hat{T}(G)$, respectively.

Definition 4.1. Let G be a finite group and let $V \subseteq \mathbb{R}(G^\sharp)$ be a subspace. If equality holds in $T_+(G, V) \subseteq W_+(G, V)$, then we say the tame-wild principle holds for (G, V).

As V gets larger, the tame-wild principle for (G, V) becomes a stronger statement. If it holds when V is all of $\mathbb{R}(G^\sharp)$, then we say it holds universally for G.

An important aspect of our formalism is as follows. Given (G, V), consider inertial subgroups I of G. For each I, one has the subspace $V_I \subseteq \mathbb{R}(I^\sharp)$ consisting of pullbacks of functions in V under the natural map $I^\sharp \to G^\sharp$. Then the tame-wild principle holds for (G, V) if and only if it holds for all (I, V_I). In fact, while G typically arises as a global Galois group in our applications, whether or not the tame-wild principle holds for (G, V) is purely a question about local Galois extensions.

4B. ***Two proof methods.*** Projection turns the chain (3-8) into a chain of cones in V:

$$T_+(G, V) \subseteq W_+(G, V) \subseteq \tilde{T}_+(G, V) \subseteq \hat{T}_+(G, V). \tag{4-1}$$

As we will see, for many G all three inclusions are strict in the universal case $V = \mathbb{R}(G^\sharp)$. However strict inclusions can easily become equalities after projection, giving us elementary but quite effective proof techniques. Namely, the *broad method* for proving that the tame-wild principle holds for (G, V) is to show that equality holds in $T_+(G, V) \subseteq \hat{T}_+(G, V)$. The *inertial method* is to show that equality holds in $T_+(G, V) \subseteq \tilde{T}_+(G, V)$.

Applying the broad method gives the following simple result which we highlight because of its wide applicability:

Theorem 4.2. *Let G be a finite group and let V be a subspace of $\mathbb{R}(G^\sharp)$. Suppose that the broad cone $\hat{T}_+(G, V)$ is generated by the $\hat{a}_\tau^V$ with τ of prime order. Then $T_+(G, V) = \hat{T}_+(G, V)$ and the tame-wild principle holds for (G, V).*

Proof. For τ of prime order one has

$$\hat{a}_\tau = \frac{\bar{\tau}}{\bar{\tau} - 1} a_\tau,$$

as noted after (3-4). Thus $\hat{T}_+(G, V)$ is contained in $T(G, V)$ and so all four sets in (4-1) are the same. □

In general, the broad method is very easy to apply, while the harder inertial method can work when the broad method does not.

4C. ***Calculations with permutation characters.*** Let $\phi_1, \dots, \phi_r$ be permutation characters spanning V. Then we are exactly in the situation described in the introduction, and in this subsection we describe how one approaches the tame-wild principle in this particular coordinatization. We incorporate the ϕ_i into our notation in straightforward ways, for example by writing $(G, \phi_1, \dots, \phi_r)$ rather than (G, V).

Throughout this subsection, we illustrate the generalities by returning to the introductory example with $G = S_5$ and $V = \langle \phi_5, \phi_6 \rangle$. The very simple two-dimensional picture of V in Figure 1 serves as an adequate model for mental images of the general situation. In particular, we always think of the a_τ^V, $a_{L/F,\mathfrak{p}}$, a_I^V, and $\hat{a}_\tau$ as in the drawn V. We think of our various cones in the drawn V as well. On the other hand, it is not useful to draw the ϕ_i on these pictures. Rather, via the identification of V with its dual by the inner product, we think of the ϕ_i as coordinate functions on the drawn V.

Conductor vectors. The space V is identified with a subspace of $\mathbb{R}^r$, viewed but not always written as column vectors, via $v \mapsto (c_1, \dots, c_r)$ with $c_i = (v, \phi_i)$. For example, an Artin character $a_{L/F,\mathfrak{p}}^V$ becomes a vector of conductors as in the introduction:

$$c_{L/F,\mathfrak{p}} = (c_\mathfrak{p}(K_1), \dots, c_\mathfrak{p}(K_r)).$$

The main case is when the ϕ_i are linearly independent, so that V is all of $\mathbb{R}^r$. One can always work in this main case by picking a basis from among the ϕ_i.

Various matrices. Our approach to calculations centers on matrices. The $r \times G^{\sharp 0}$ *partition matrix* $P(G, \phi_1, \dots, \phi_r)$ has i-τ entry the cycle type $\lambda_\tau(\phi_i)$ of $\rho_i(g)$, where ρ_i is a permutation representation with character ϕ_i and $g \in G$ represents τ. Thus,

$$P(S_5, \phi_5, \phi_6) = \begin{pmatrix} 2111 & 221 & 311 & 41 & 5 & 32 \\ 222 & 2211 & 33 & 411 & 51 & 6 \end{pmatrix}. \tag{4-2}$$

Partition matrices are purely group-theoretic objects, but one can use fields in a standard way to help construct them. For example, the columns from left to right are the partitions obtained by factoring the pair $(f_5(x), f_6(x))$ from (2-1) and (2-2) modulo the primes 67, 211, 31, 13, 11, and 7 respectively.

One passes to the *tame matrix* $T(G, \phi_1, \dots, \phi_r)$ by replacing each partition $\lambda_\tau(\phi_i)$ by its conductor $c_\tau(\phi_i) = (a_\tau, \phi_i)$ — its degree minus its number of parts. Thus

$$T(S_5, \phi_5, \phi_6) = \begin{pmatrix} 1 & 2 & 2 & 3 & 4 & 3 \\ 3 & 2 & 4 & 3 & 4 & 5 \end{pmatrix}. \tag{4-3}$$

The *broad matrix* $\hat{T}(G, \phi_1, \dots, \phi_r)$ consists of what we call *preconductors*, the preconductor $\hat{c}_\tau(\phi_i) = (\hat{a}_\tau, \phi_i)$ being the degree of $\lambda_\tau(\phi_i)$ minus its number of ones.

Thus

$$\hat{T}(S_5, \phi_5, \phi_6) = \begin{pmatrix} 2 & 4 & 3 & 4 & 5 & 5 \\ 6 & 4 & 6 & 4 & 5 & 6 \end{pmatrix}. \tag{4-4}$$

Inertial matrices $\tilde{T}(G, \phi_1, \ldots, \phi_r)$ typically have more columns, because columns are indexed by conjugacy classes of inertial subgroups I. But an entry is just the *formal conductor* $c_I(\phi_i) = (a_I, \phi_i)$, this being the degree of ρ_i minus the number of orbits of $\rho_i(I)$, just as in the cyclic case. The cones $T_+ \subseteq \tilde{T}_+ \subseteq \hat{T}_+ \subset \mathbb{R}^r$ are then generated by the columns of the corresponding matrices T, $\tilde{T}$, and $\hat{T}$.

Inclusions $a^V \in T_+(G, V)$ in matrix terms. By dropping rows, we can assume that $\phi_1, \ldots, \phi_r$ span V and so $T = T(G, \phi_1, \ldots, \phi_r)$ has full rank r, as discussed above. In general, let $c \in \mathbb{R}^r$ be a column r-vector. For each r-element subset $J \subseteq G^{\sharp 0}$ for which the corresponding minor $T(J)$ is invertible, let $u(J) = (u(J)_\tau)_{\tau \in J}$ be the vector $T(J)^{-1}c$. Then $c = \sum_{\tau \in J} u(J)_\tau T_\tau$, with T_τ the τ-th column of T. Then c is in the tame cone T_+ if and only if there exists such a J with $u(J)_\tau \geq 0$ for all $\tau \in J$.

To prove that the tame-wild principle holds for $(G, \phi_1, \ldots, \phi_r)$ directly, one would have to show this positivity condition holds for all conductor vectors $c_{L/F,\mathfrak{p}}$. To show it via the inertial method, one has to show that it holds for all formal conductor vectors c_I. To show it holds via the broad method, one has to show that it holds for all preconductor vectors $\hat{c}_\tau$.

Projectivization. In the introductory example, we emphasized taking ratios of conductors, thereby removing the phenomenon that wild conductors are typically much larger than tame conductors, but keeping the phenomenon we are interested in. We can do this in the general case as well, assuming without loss of generality that ϕ_r comes from a faithful permutation representation so that the conductors $c_\tau(\phi_r)$ are strictly positive for all $\tau \in G^{\sharp 0}$. We projectivize $c = (c_1, \ldots, c_r)$ to $c' = (c'_1, \ldots, c'_{r-1})$ with $c'_i = c_i / c_r$.

Applying this projectivization process to columns gives the projective tame, inertial, and broad matrices respectively, each notationally indicated by a $'$. In our continuing introductory example, one has, very simply,

$$T'(S_5, \phi_5, \phi_6) = \begin{pmatrix} \frac{1}{3} & 1 & \frac{1}{2} & 1 & 1 & \frac{3}{5} \end{pmatrix}, \tag{4-5}$$

$$\hat{T}'(S_5, \phi_5, \phi_6) = \begin{pmatrix} \frac{1}{3} & 1 & \frac{1}{2} & 1 & 1 & \frac{5}{6} \end{pmatrix}. \tag{4-6}$$

In general, the τ-columns of $T'(G, \phi_1, \ldots, \phi_r)$ and $\hat{T}'(G, \phi_1, \ldots, \phi_r)$ agree if τ has prime order. Here they disagree only in the last column corresponding to the composite order 6.

Let $T'_+(G, \phi_1, \ldots, \phi_r)$ be the convex hull of the columns of $T'(G, \phi_1, \ldots, \phi_r)$ and define $W'_+(G, \phi_1, \ldots, \phi_r)$, $\tilde{T}'_+(G, \phi_1, \ldots, \phi_r)$ and $\hat{T}'_+(G, \phi_1, \ldots, \phi_r)$ to be the analogous hulls. Then (4-1) has its obvious analog at the level of hulls, and one can think about the broad method and the inertial method at this level. In the introductory example, (4-5) and (4-6) say that $T'_+(S_5, \phi_5, \phi_6) \subseteq \hat{T}'_+(S_5, \phi_5, \phi_6)$ is an equality because both sides are $\left[\frac{1}{3}, 1\right]$. Thus the tame-wild principle holds for (S_5, ϕ_5, ϕ_6).

The drop in dimension from r to $r-1$ has a number of advantages. As illustrated already by (4-5) and (4-6), it renders the $r=2$ case extremely concrete. As we will illustrate in Section 8A, it renders the $r=3$ case highly visible. In general, it lets one determine whether a^V is in $T'_+(G, \phi_1, \dots, \phi_r)$ by computation with $(r-1)\times(r-1)$ minors rather than $r \times r$ minors.

4D. ***Alternative approaches.*** Our abstract formulation of the tame-wild principle is designed to be very flexible. For example, say that a vector $v \in \mathbb{R}(G^\sharp)$ is *bad* if $(a_\tau, v) \geq 0$ for all $\tau \in G^{\sharp 0}$ but $(a_{L/F,\mathfrak{p}}, v) < 0$ for some Artin character $a_{L/F,\mathfrak{p}}$. The bad vectors form a union of cones in $\mathbb{R}(G^\sharp)$ and the tame-wild principle holds for (G, V) if and only if V misses all these cones. In this sense, the one-dimensional V spanned by bad v are essential cases, but these V are never spanned by permutation characters.

Sections 5 and 6 are in the universal setting $V = \mathbb{R}(G^\sharp)$, and we do not use ϕ_i at all. Sections 7 and 8 return to the permutation character setting described in Section 4C. In general, the systematic study of the tame-wild principle for a fixed G and varying V would be facilitated by the canonical basis of $\mathbb{Q}(G^\sharp)$ given by irreducible rational characters.

5. The universal tame-wild principle holds for U-groups

In Section 5A, we present a diagrammatic way of understanding class sets $G^\sharp$. Making use of this viewpoint, Section 5B gives the canonical expansion of a formal Artin character a_I as a sum of tame characters a_τ. Next, Section 5C introduces the notion of U-group and proves that the universal tame-wild principle holds for U-groups. However the class of U-groups is quite small, as discussed in Section 5D.

5A. ***Divisibility posets.*** For G a finite group, the set $G^\sharp$ is naturally a partially ordered set via the divisibility relation. We draw this divisibility poset in the standard way with an edge from σ down to τ of vertical length one if $\sigma^p = \tau$ for some prime p. With notation as in Section 3A, the natural weight $d(\tau, \sigma) = [\sigma]/[\tau]$ plays an important role, and we write it next to the edge whenever it is different from 1, considering this data as part of the divisibility poset.

The product of the edge weights from any vertex σ down to another τ is path-independent, being in fact just $d(\tau, \sigma) = [\sigma]/[\tau]$. Define integers $u_{G,\sigma}$ via

$$\sum_{\tau \mid \sigma} d(\tau, \sigma) u_{G,\sigma} = 1. \tag{5-1}$$

Thus $u_{G,\tau} = 1$ for maximal τ, and all the integers $u_{G,\tau}$ can be computed by downwards induction on the divisibility poset $G^\sharp$.

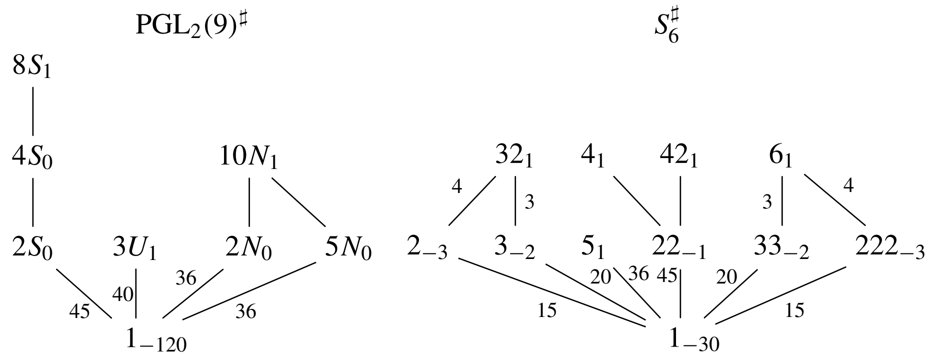


Figure 2. Two divisibility posets $G^\sharp$ with $u_{G,\tau}$ subscripted on $\tau \in G^\sharp$.

Figure 2 draws the divisibility posets $\mathrm{PGL}_2(9)^\sharp$ and $S_6^\sharp$, with each τ subscripted by its $u_{G,\tau}$. The case $\mathrm{PGL}_2(9)$ represents the general case $\mathrm{PGL}_2(p^f)$, with split-torus classes indexed by nonunital divisors of $p^f - 1$, a unipotent class pU, nonsplit-torus classes indexed by nonunital divisors of $p^f + 1$, and finally the identity class 1. The case S_6 represents the general case S_n, where classes are indexed by partitions of n, with 1s usually left unprinted.

In general, the largest edge weights on divisibility posets $G^\sharp$ tend to be on edges incident on the identity class. These edges do not play an important role for us and in the sequel we work instead with the divisibility poset associated to $G^{\sharp 0}$.

5B. ***Expansion of formal Artin characters.*** Divisibility posets for inertial groups I and the associated integers $u_{I,\sigma}$ are important to us because of the role they play in the following lemma.

Lemma 5.1. *Let G be a finite group, let I be a subgroup, and let $i : I^\sharp \to G^\sharp$ be the induced map. Then the expansion of the formal Artin character $a_I \in \mathbb{Q}(G^\sharp)^0$ in the basis $\{a_\tau\}_{\tau \in G^{\sharp 0}}$ can be read off from the divisibility poset $I^{\sharp 0}$ via the formula*

$$a_I = \frac{1}{|I|} \sum_{\sigma \in I^{\sharp 0}} u_{I,\sigma} [\sigma] \bar{\sigma} a_{i(\sigma)}. \tag{5-2}$$

Before proving the lemma, we explain the roles that various parts of (5-2) play in the sequel. The positive integer $[\sigma]\bar{\sigma}$ plays a very passive role: only the positivity of $[\sigma]\bar{\sigma}$ is used in the proof of Theorem 5.3; moreover, $[\sigma]\bar{\sigma}$ factors out in Section 6A and accordingly does not enter into Section 6B-Section 6E. The factor $|I|^{-1}$ is more important: while only its positivity enters into the proof of Theorem 5.3, it contributes to the index factor in (6-1) which enters significantly into the rest of Section 6. The part with the most important role is $u_{I,\sigma}$, as it is the possible negativity of $u_{I,\sigma}$ that can lead to failures of the tame-wild principle. Our use of the function i relegates the difference between I and G to the background, but one should note that for $\tau \in G^{\sharp 0}$ the actual coefficient of a_τ in (5-2) has $|i^{-1}(\tau)|$ terms.

Proof. First consider the case $I = G$. Then both sides of (5-2) are in $\mathbb{Q}(I^\sharp)^0$. The left side takes the value $a_I(\tau) = -1$ for all $\tau \in I^{\sharp 0}$. We thus need to evaluate the right side on an arbitrary $\tau \in I^{\sharp 0}$ and see that it simplifies to -1:

$$\begin{aligned}
\frac{1}{|I|}\sum_{\sigma\in I^{\sharp 0}} u_{I,\sigma}[\sigma]\bar{\sigma}a_\sigma(\tau) &= \frac{1}{|I|}\sum_{\tau|\sigma} u_{I,\sigma}[\sigma]\bar{\sigma}a_\sigma(\tau)\\
&= \frac{1}{|I|}\sum_{\tau|\sigma} u_{I,\sigma}[\sigma]\bar{\sigma}\left(-\frac{\phi(\bar{\tau})}{\bar{\sigma}}\frac{|I|}{|C_\tau|}\right)\\
&= -\sum_{\tau|\sigma} u_{I,\sigma}[\sigma]\frac{\phi(\bar{\tau})}{|C_\tau|}\\
&= -\sum_{\tau|\sigma} \frac{u_{I,\sigma}[\sigma]}{[\tau]}\\
&= -\sum_{\tau|\sigma} d(\tau,\sigma)u_{I,\sigma} = -1.
\end{aligned}$$

Here we have used formulas from Section 3A and Section 3B as well as the definition of $d(\tau, \sigma)$ and the defining property of the $u_{I,\sigma}$ from Section 5A. Finally the case of general G follows, by induction of both sides from I to G. □

5C. ***Applying the inertial method.*** Say that a class $\tau \in G^\sharp$ is a *U-class* if it divides exactly one maximal element σ of $G^\sharp$ and $d(\tau, \sigma) = 1$. Otherwise, say it is an *N-class*. Here U stands for unique and N for nonunique. The following three facts are immediate from the definition. First, a maximal class τ is always a U-class with $u_{G,\tau} = 1$. Second, other U-classes τ have $u_{G,\tau} = 0$. Third, a maximal N-class τ always has $u_{G,\tau} < 0$.

We divide all finite groups into two types, as follows.

Definition 5.2. A finite group is a *U-group* if every nonidentity element is contained in exactly one maximal cyclic subgroup. Otherwise it is an *N-group*.

It is immediate that a group G is a U-group if and only if all classes $\tau \in G^{\sharp 0}$ are U-classes. Thus from Figure 2, $\mathrm{PGL}_2(9)$ is a U-group while S_6 is an N-group. It follows easily from the definition that any subgroup of a U-group is itself a U-group. Similarly, any quotient of a U-group is a U-group [Suzuki 1950].

Via (3-4), the chain (3-8) completely collapses to the equality $T_+(G) = \hat{T}_+(G)$ if and only if all nonidentity elements in G have prime order. The following theorem is a subtler version of this idea.

Theorem 5.3. *Suppose G is a group such that all inertial subgroups of G are U-groups. Then one has $T_+(G) = \tilde{T}_+(G)$, and so the tame-wild principle holds universally for G.*

Proof. Let I be an arbitrary inertial subgroup. Since I is assumed to be a U-group, the associated integers $u_{I,\sigma}$ are nonnegative for all $\sigma \in I^{\sharp 0}$. For any $\tau \in G^{\sharp 0}$, the terms $|I|^{-1} u_{I,\sigma}[\sigma]\bar{\sigma}$ contributing to the coefficient of a_τ in Lemma 5.1 are all nonnegative. Hence the coefficient itself is nonnegative and so a_I is in the tame cone $T_+(G)$. Since the a_I generate the inertial cone $\tilde{T}_+(G)$, equality holds in $T_+(G) \subseteq \tilde{T}_+(G)$. □

In particular, the tame-wild principle holds for all U-groups. This is the main import of Theorem 5.3, as we are not aware of any group satisfying the hypothesis of Theorem 5.3 which is not itself a U-group.

5D. ***Classification of U-groups.*** Given Theorem 5.3, it is of interest to classify U-groups. This problem has been addressed in the literature, with Kontorovich [1939; 1940] referring to U-groups as completely decomposable groups, and Suzuki [1950] calling them groups with a complete partition. We give a summary of the classification situation here.

The condition to be a U-group is very restrictive, but it is easy to check that it includes many groups of small order. In particular, the following groups are U-groups: cyclic groups, dihedral groups, groups of prime exponent, and the Frobenius groups $F_p = C_p : C_{p-1}$. The last class is particularly important in our context, since an extension of a p-adic field of degree p has normal closure with Galois group a subgroup of F_p. If q is a prime power, the linear groups $\mathrm{PSL}_2(q)$ and $\mathrm{PGL}_2(q)$ are U-groups, so that in particular $S_4 \cong \mathrm{PGL}_2(3)$, $S_5 \cong \mathrm{PGL}_2(5)$ and $A_6 \cong \mathrm{PSL}_2(9)$ are all U-groups. There are more U-groups than those listed here, most of them being more general types of Frobenius groups.

The following observation is useful in understanding the nature of U-groups. In two settings, the extreme members of a class of groups are exactly the U-groups as follows. First, consider abelian p-groups of order p^n. Up to isomorphism, they correspond to partitions of n. The groups which are U-groups are the two extreme ones $(C_p)^n$ and C_{p^n}. Second, consider semidirect products $C_a :_\gamma C_b$ with a and b being relatively prime and $\gamma : C_b \to \mathrm{Aut}(C_a)$. If γ is trivial, then $C_a :_\gamma C_b \cong C_{ab}$ is a U-group. If γ is injective, then $C_a :_\gamma C_b$ is again a U-group, being of a nature similar to F_p above. Again, it is the intermediate cases which are N-groups: if γ is neither trivial nor injective then nontrivial elements in the kernel of γ are in more than one maximal cyclic subgroup.

6. The universal tame-wild principle usually fails for N-groups

In this section, we study the universal tame-wild principle for N-groups. In Section 6A, we give the canonical expansion of a general Artin character $a_{L/F,\mathfrak{p}}$ in terms of tame characters a_τ. In Section 6B, we list out N-groups of order pqr where p, q, and r are not necessarily distinct primes, finding six series. We show in Section 6C that the universal tame-wild principle generally fails for groups in the first four series. In Section 6D we take a close look at the quaternion group Q_8, which is the first group of the fifth series, finding failure again. On the other hand we show in Section 6E that the universal tame-wild principle holds for all groups in the sixth series. Finally, Section 6F explains how the negative results for small groups support the principle that most N-groups do not satisfy the universal tame-wild principle.

6A. ***Expansion of general Artin characters.*** Let $a_{L/F,\mathfrak{p}} \in \mathbb{Q}(G^\sharp)^0$ be an Artin character coming from an inertial subgroup $I \subseteq G$. Equation (3-7) expands $a_{L/F,\mathfrak{p}}$ in terms of formal Artin characters $a_{I^{s_i}}$ and Lemma 5.1 in turn expands each $a_{I^{s_i}}$ in terms of tame characters. Putting these two expansions together and replacing the divisibility posets $(I^{s_i})^{\sharp 0}$ with their images in $I^{\sharp 0}$ gives the following lemma.

Lemma 6.1. *Let G be a group, and let $a_{L/F,\mathfrak{p}} \in \mathbb{Q}(G^\sharp)^0$ be an Artin character with inertia group $I = I^{s_1} \supset I^{s_2} \supset \cdots$ as in* (3-6). *Let $i : I^\sharp \to G^\sharp$ be the induced map. Then one has the expansion*

$$a_{L/F,\mathfrak{p}} = \frac{1}{|I|} \sum_{\sigma \in I^{\sharp 0}} w_{L/F,\mathfrak{p},\sigma}[\sigma]\bar{\sigma} a_{i(\sigma)},$$

where

$$w_{L/F,\mathfrak{p},\sigma} = \sum_{i=1}^{k} (s_i - s_{i-1})[I : I^{s_i}] u_{I^{s_i},\sigma}. \tag{6-1}$$

While the lemma applies to the general situation, our focus in Sections 6B–6E is on the case $I = G$. Here $a_{L/F,\mathfrak{p}}$ is in the tame cone $T_+(G)$ if and only if $w_{L/F,\mathfrak{p},\tau} \geq 0$ for all $\tau \in G^{\sharp 0}$.

6B. ***Inertial N-groups of order pqr.*** Groups of order p or pq are U-groups. In the complete list of inertial groups of order pqr, in a rough sense about half of them are U-groups and the other half N-groups. For example, for a given prime p, there are two nonabelian groups of order p^3, the extra-special groups often denoted p_+^{1+2} and p_-^{1+2}. For p odd, p_+^{1+2} has exponent p and so is a U-group, while p_-^{1+2} is an N-group. Similarly the dihedral group $D_4 = 2_+^{1+2}$ is a U-group while the quaternion group $Q_8 = 2_-^{1+2}$ is an N-group.

In fact, it is easy to see that the inertial N-groups are as follows. Now p, q, and r are required to be different primes, with $q \mid p-1$ whenever $F_{p,q} = C_p : C_q$ is present:

1: The product $F_{p,q} \times C_r$.

2: The semidirect product $C_p : C_{q^2} \cong F_{p,q} *_{C_q} C_{q^2}$.

3: The abelian group $C_{pq} \times C_p \cong C_p \times C_p \times C_q$.

4: The product $F_{p,q} \times C_p$.

5: The extra-special group p_{-}^{1+2}.

6: The abelian group $C_{p^2} \times C_p$.

These groups I are all p-inertial groups, but not inertial groups for any other primes. Moreover, since all proper subgroups are U-groups, the universal tame-wild principle fails for I if and only if there exists a totally ramified local Galois extension L/F having $\mathrm{Gal}(L/F) \cong I$ with associated Artin character $a_{L/F,\mathfrak{p}}$ having a negative coefficient $w_{L/F,\mathfrak{p},\tau}$. Furthermore, in each case it turns out that there is exactly one N-class $\tau \in I^{\sharp 0}$. Only for this class τ could $w_{L/F,\mathfrak{p},\tau}$ possibly be negative, and this N-class is boxed in the displayed divisibility posets below.

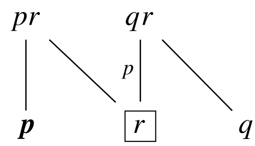
In general, let I be a p-inertial group. Then it is known that there indeed exists a totally ramified Galois extension of p-adic fields L/F with $\mathrm{Gal}(L/F)$ isomorphic to I. This fact for our particular I is essential to our proofs that the universal tame-wild principle does not hold. However it is easy to prove this fact for all the above I by direct exhibition of L/F. We will go into this level of detail only for the groups in Series 4 and Q_8 from Series 5, as here we need particular fields satisfying conditions on their wild ramification.

6C. *Negative results for four series.* Our first result concerns Series 1-4 and is negative:

Theorem 6.2. *$F_{3,2} \times C_3 \cong S_3 \times C_3$ satisfies the universal tame-wild principle, but otherwise the groups $F_{p,q} \times C_r$, $C_p : C_{q^2}$, $C_{pq} \times C_p$, and $F_{p,q} \times C_p$, do not.*

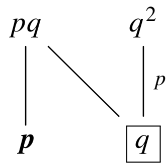
Proof. In the divisibility posets below, the wild classes, meaning the classes of p-power order, are put in boldface for further emphasis. For the first three series, the unique N-class τ has prime-to-p order and so we do not need to enter into an examination of wild slopes. In Series 4, τ has order p and bounds on wild slopes lead to the exception.

1. For the group $I = F_{p,q} \times C_r$, power-conjugacy classes are determined by their orders, and the divisibility poset $I^{\sharp 0}$ is



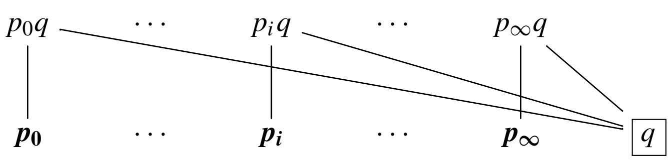
Equation (6-1) becomes $w_r = u_{I,r} = -p < 0$. So by the existence of totally ramified I-extensions as discussed in the previous subsection, the universal tame-wild principle does not hold for $I = F_{p,q} \times C_r$.

2. The group $I = C_p : C_{q^2}$ behaves very similarly. Again power-conjugacy classes are determined by their orders:



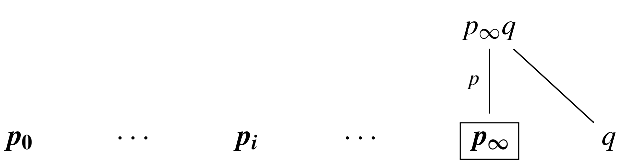
The key quantity $w_q = u_{I,q} = -p$ is again negative, so the universal tame-wild principle fails for $C_p : C_{q^2}$.

3. The group $I = C_p^2 \times C_q$ has a more complicated divisibility poset $I^{\sharp 0}$ but the behavior is otherwise similar. The classes of order p and the classes of order pq have the structure of projective lines over $\mathbb{F}_p$ in bijection with one another:



Once again $w_q = u_{I,q} = -p$ and so the universal tame-wild principle fails for $C_p^2 \times C_q$.

4. For $I = F_{p,q} \times C_p$ the divisibility poset $I^{\sharp 0}$ is disconnected:



Here p_0 and q lie in the factor $F_{p,q}$ while p_∞ lies in the factor C_p. The first term in (6-1) is $u_{I,p_\infty} = 1 - p$. However, now we must take into account how wild ramification contributes to the remaining terms. Let $s > 1$ be the slope associated to $F_{p,q}$ and let $c > 1$ be the slope associated to C_p. Since C_p is abelian, c must be integral and hence $c \geq 2$. On the other hand, s must have exact denominator q. Let

$m = \min(c, s)$, so that $I^m = C_p^2$ is the wild inertia group and $I^{\max(c,s)} \cong C_p$ is a higher inertia group. If $c > s$ then $(I^c)^{\sharp 0} = \{p_\infty\}$, while if $s > c$ then $(I^s)^{\sharp 0} = \{p_0\}$. Equation (6-1) becomes

$$w_{p_\infty} = \begin{cases} (1-p) + q(s-1) + qp(c-s) & \text{if } c > s, \\ (1-p) + q(c-1) & \text{if } s > c. \end{cases}$$

For $(p, q) = (3, 2)$, the general formula simplifies to

$$w_{3_\infty} = \begin{cases} 6c - 4s - 4 & \text{if } c > s, \\ 2c - 4 & \text{if } s > c. \end{cases}$$

Thus, using $c \geq 2$, one has $w_{3_\infty} \geq 0$ and so the universal tame-wild principle holds for $F_{3,2} \times C_3$.

There are many ways to produce an explicit instance with $w_{p_\infty} < 0$ for the remaining (p, q). We will present one in the setting $s > c = 2$ in which case $w_{p_\infty} = 1 + q - p$ is indeed negative. To get an $F_{p,q}$ extension, start with $x^p - p$, which gives a totally ramified $F_{p,p-1}$ extension of $\mathbb{Q}_p$ with wild slope best written in the form $1 + p/(p-1)$. Write $e = (p-1)/q$ and extend the ground field from $\mathbb{Q}_p$ to $F_e = \mathbb{Q}_p[\pi]/(\pi^e - p)$. Then $x^p - p$ has Galois group $F_{p,q}$ over F_e, with wild slope $1 + ep/(p-1)$, as tame base-change always scales slopes this way. But now $x^p - \pi x^{p-1} + \pi$ has wild slope 2 and, after perhaps replacing F_e by an unramified extension F, Galois group C_p [Amano 1971]. The splitting field of $(x^p - p)(x^p - \pi x^{p-1} + \pi)$ gives the desired extension L/F, showing that the universal tame-wild principle does not hold for $F_{p,q} \times C_p$. $\square$

6D. *Negative result for Q_8.* The fifth series, consisting of groups of the form p_-^{1+2}, is the most complicated. Here we treat only $2_-^{1+2} = Q_8$, getting a negative result.

Proposition 6.3. *The universal tame-wild principle fails for the quaternion group.*

Proof. The divisibility poset $Q_8^{\sharp 0}$, with unique N-class boxed as always, is

$$\begin{tikzcd} 4_i \arrow[dr, dash] & 4_j \arrow[d, dash] & 4_k \arrow[dl, dash]. \\ & \boxed{2} & \end{tikzcd}$$

The generic case has three distinct slopes. We seek only counterexamples and so we focus on the special case with two slopes $s_1 < s_2$, with s_1 occurring with multiplicity two. The key quantity (6-1) here becomes $w_2 = -2s_1 + 4(s_2 - s_1) = 4s_2 - 6s_1$. Thus one gets a counterexample to the universal tame-wild principle if and only if $s_2 < 1.5s_1$.

The table of octic 2-adic fields [Jones and Roberts 2008] available from the website of [Jones and Roberts 2006] then give four types of counterexamples in this context, after tame base-change from $\mathbb{Q}_2$ to its Galois extension F with ramification index t and residual field degree u:

#	c	slope content	I	D	s_1	s_2
2	10	$[1.\overline{3}, 1.\overline{3}, 1.5]_3^2$	$\mathrm{SL}_2(3)$	$\mathrm{GL}_2(3)$	2	2.5
2	16	$[2, 2, 2.5]^2$	Q_8	$\hat{Q}_8$	2	2.5
4	16	$[2, 2, 2.5]^4$	Q_8	8T17	2	2.5
8	22	$[2.\overline{6}, 2.\overline{6}, 3.5]_3^2$	$\mathrm{SL}_2(3)$	$\mathrm{GL}_2(3)$	6	8.5

(6-2)

Here in the first and last cases, we use the general conversion from slope content $[\dots\sigma_i\dots]_t^u$ over $\mathbb{Q}_2$ to slope content $[\dots s_i\dots]_1^1$ over F given by $s_i = 1 + t(\sigma_i - 1)$. A full treatment of the range of possible counterexamples could have Proposition 4.4 of [Fontaine 1971] as its starting point. □

Our counterexamples in Section 7D and Section 8B will be built from one of the two fields with slope content $[2, 2, 2.5]^2$. A point to note here is that $\mathbb{Q}_2$ does have totally ramified quaternionic extensions, in fact four of them, all with slope content $[2, 3, 4]$ [Jones and Roberts 2008]. However, these extensions do not give counterexamples to the universal tame-wild principle for Q_8. The fact that the first local counterexamples come from $\hat{Q}_8 = 8\mathrm{T}8$ extensions of $\mathbb{Q}_2$ plays a prominent role in our later global counterexamples.

6E. *Positive results for $C_{p^2} \times C_p$.* Here we prove that the N-groups in Series 6 always satisfy the universal tame-wild principle. Unlike most of our previous positive results, but like the exception $S_3 \times C_3$ of Section 6C, this result is not purely group-theoretic. Rather it depends on a close analysis of the possibilities for wild slopes. Said in a different way, the situation for these I is $T_+(I) = W_+(I) \subset \tilde{T}_+(I)$, so that the universal tame-wild principle holds, even though it is not provable by the inertial method.

Theorem 6.4. *The groups $C_{p^2} \times C_p$ satisfy the universal tame-wild principle.*

Proof. Let K_{p^2}/F be a cyclic extension of degree p^2 and slopes $s_1 < s_2$. Let K_p/F be a cyclic extension of degree p and slope t. Switch to the indexing scheme of [Serre 1979] via $s_i = 1 + v_i$ and $t = 1 + c$, so as to better align also with our reference [Fontaine 1971] and in particular make (6-5) below as simple as possible. There are three possibilities for how the slope filtration goes through the group:

$$
\begin{array}{ccc}
c < v_1 < v_2 & v_1 \le c < v_2 & v_1 < v_2 \le c \\
\begin{array}{l}
\mathbf{p^2} \\
\Big| {\scriptstyle p} \\
\boxed{\underline{\mathbf{p_0}}} \quad p_1 \quad \cdots \quad p_\infty,
\end{array}
&
\begin{array}{l}
p^2 \\
\Big| {\scriptstyle p} \\
\boxed{\underline{\mathbf{p_0}}} \quad \mathbf{p_1} \quad \cdots \quad \mathbf{p_\infty},
\end{array}
&
\begin{array}{l}
p^2 \\
\Big| {\scriptstyle p} \\
\boxed{\mathbf{p_0}} \quad \mathbf{p_1} \quad \cdots \quad \underline{\mathbf{p_\infty}}.
\end{array}
\end{array}
$$

Here, assuming all inequalities are strict, classes in the higher inertia group of order p^2 are put in bold and the classes in the higher inertia group of order p are furthermore underlined. If one has equality, the formulas below still apply.

As in the previous two subsections, only one w_τ from Lemma 6.1 could possibly be negative, and in this case it is $w = w_{p0}$. Equation (6-1) becomes

$$w = \begin{cases} (c+1)(1-p)+(v_2-v_1)p^2 & \text{if } c < v_1 < v_2, \\ (v_1+1)(1-p)+(c-v_1)p+(v_2-c)p^2 & \text{if } v_1 \le c < v_2, \\ (v_1+1)(1-p)+(v_2-v_1)p & \text{if } v_1 < v_2 \le c. \end{cases} \tag{6-3}$$

Let e be the ramification index of $F/\mathbb{Q}_2$ and put $B = e/(p-1)$. From the known behavior of cyclic degree p extensions, one has

$$1 \le c \le pB \quad \text{and} \quad 1 \le v_1 \le pB. \tag{6-4}$$

There are two regimes to consider: the geometric regime, where $v_1 < B$, and the arithmetic regime, where $v_1 \ge B$. One has

$$v_2 \ge pv_1 \quad \text{in the geometric regime,} \tag{6-5}$$

$$v_2 = v_1 + e \quad \text{in the arithmetic regime.} \tag{6-6}$$

These last two facts and other related information dating back to [Maus 1965] are conveniently available in [Fontaine 1971, Proposition 4.3].

The quantity e does not enter into the geometric inequality (6-5), and since we need to deal with arbitrary e the upper bounds in (6-4) are not available to us. This fact is the source of our terminology because the geometric case is now identified with the case where p-adic fields have been replaced by $\mathbb{F}_{p^f}((t))$, which have $e = \infty$. The worst case is always when $v_2 = pv_1$ and, in the second case, when c takes on its limiting bound v_2 as well. Substituting these worst cases into (6-3) and simplifying, one has

$$w \ge \begin{cases} (v_1+1)(1-p)+(pv_1-v_1)p^2 & \text{if } c < v_1 < v_2, \\ (v_1+1)(1-p)+(pv_1-v_1)p & \text{if } v_1 \le c < v_2, \\ (v_1+1)(1-p)+(pv_1-v_1)p & \text{if } v_1 < v_2 \le c. \end{cases}$$

With m equal to p^2, p, p in the three cases, one further simplifies by

$$w \ge (p-1)(-v_1 - 1 + v_1 m) = (p-1)((m-1)v_1 - 1) \ge 0.$$

Thus, in the geometric regime, w is never negative.

In the arithmetic regime the substitute (6-6) for (6-5) is simpler in that it is an equality, but now the upper bounds in (6-4) will need to be used. The substitution $v_2 = v_1 + e = v_1 + B(p-1)$ into (6-3) makes w factor, and we divide by the positive quantity $p-1$:

$$\frac{w}{p-1} = \begin{cases} Bp^2 - 1 - c & \text{if } c < v_1 < v_1 + e, \\ Bp^2 - v_1 - p(c - v_1) - 1 & \text{if } v_1 \le c < v_1 + e, \\ Bp - v_1 - 1 & \text{if } v_1 < v_1 + e \le c. \end{cases}$$

Using the bounds (6-4) one has

$$\frac{w}{p-1} \ge \begin{cases} Bp^2 - 1 - Bp = Bp(p-1) - 1 = ep - 1 \ge 1, \\ Bp^2 - Bp - p(e-1) - 1 = ep - p(e-1) - 1 = p - 1 \ge 1, \\ Bp - v_1 - 1 \le Bp - (Bp - e) - 1 = e - 1 \ge 0. \end{cases}$$

in the three cases. Thus here too $w \ge 0$. □

6F. ***From smaller to larger groups.*** Our final topic in this section is to promote our counterexamples from the small groups I to larger groups G that contain them. In general, let $I \subseteq G$ be an inclusion of groups and consider the induced map $i : I^\sharp \to G^\sharp$. Then the lack of injectivity of i can obstruct the promotion process. For example, consider Series 4 groups $I = F_{p,q} \times C_p$ and their product embedding into $G = S_{p^2}$. Then all $p+1$ classes in $I^\sharp$ of order p go to the single class in $S_{p^2}^\sharp$ indexed by the partition $p^p = p \cdots p$. To get the coefficient of a_{p^p} of the pushed-forward formal Artin conductor $a_I \in \mathbb{Q}(S_{p^2}^\sharp)^0$ one has to add the contributions of the fiber, as in Lemma 5.1. There are p contributions of $1/pq$ and one contribution of $(2-p)/pq$ for a total of $2/pq$. Equation (6-1) says that wild ramification can only increase this $2/pq$ to larger positive numbers, and so all pushed-forward Artin characters $a_{L/F,\mathfrak{p}}$ from I are in the tame cone $T_+(S_{p^2})$.

For Series 1–3 and also for Q_8, this complication does not arise because the unique N-class in $I^{\sharp 0}$ is the only class of its order. Hence the promotion process works:

Corollary 6.5. *Let G be a group containing a subgroup I of the form $F_{p,q} \times C_r$, $C_p : C_{q^2}$, $C_{pq} \times C_p$ or Q_8. Then G does not satisfy the universal tame-wild principle.*

Since there are so many possibilities for I, the hypothesis holds for many G. Moreover the fact that it holds for a given G is often easily verified. For example, when studying G one commonly has a list of maximal subgroups H, and one can often easily see that at least one I is in at least one of the H. As another example, the presence of $C_{pq} \times C_p$ can often be read off from the divisibility poset: suppose one has a class $\tau \in G^\sharp$ of order $\bar{\tau} = pq$ not dividing a class of order $p^2 q$ but such that p^2 divides the numerator of $|G|/|C_\tau|$. Then any representative g of τ lies in a group of type $C_{pq} \times C_p$. This criterion is satisfied particularly often for $p = 2$ and some odd prime q.

7. Comparing an algebra with its splitting field

In this section we return to a very concrete setting, considering types (G, ϕ_i, ϕ_r) where ϕ_i comes from a faithful permutation representation $i : G \subseteq S_n$ and ϕ_r is the regular character. Thus we are considering algebras $K = K_i$ of a specified Galois type compared with their splitting fields $K^{\mathrm{gal}} = K_r$.

In Section 7A we introduce explicit notation for comparing two algebras and in Section 7B we explain how it is sometimes best to highlight root discriminants $\mathfrak{d}$ rather than discriminants $\mathfrak{D}$. The tame-wild principle in the notation set up then takes the following form:

$$\mathfrak{d}_{K^{\mathrm{gal}}/F}^{\alpha(G,\phi_i,\phi_r)} \mid \mathfrak{d}_{K/F} \mid \mathfrak{d}_{K^{\mathrm{gal}}/F}^{\omega(G,\phi_i,\phi_r)}.$$

We observe in Section 7C that the right divisibility often trivially holds. In Section 7D, we give four examples where it holds nontrivially and one where it fails to hold. In Section 7E we show that the left divisibility always holds, and discuss applications to number field tabulation.

7A. ***Generalities.*** The case $r = 2$ of just two algebras deserves special attention for at least three reasons. First, hulls $T'_+(G, \phi_1, \phi_2) \subset \mathbb{R}^1$ are intervals while hulls for larger r can have up to $|G^{\sharp 0}|$ vertices. Second, the inequality for each face of any $T'_+(G, \phi_1, \ldots, \phi_r)$ also comes from some $T'_+(G, \psi_1, \psi_2)$ with the new characters ψ_j being certain sums of the old characters ϕ_i. Third, it is the case which applies most directly to number field tabulation.

To present results coming from $r = 2$ as explicitly as possible, we let $\alpha = \alpha(G, \phi_1, \phi_2)$ and $\omega = \omega(G, \phi_1, \phi_2)$ be the left and right endpoints of the interval $T'_+(G, \phi_1, \phi_2)$. The tame-wild principle says that all local exponents satisfy

$$\alpha c_{\mathfrak{p}}(K_2) \leq c_{\mathfrak{p}}(K_1) \leq \omega c_{\mathfrak{p}}(K_2). \tag{7-1}$$

In this $r = 2$ setting, the tame-wild principle breaks cleanly into two parts: the left and right tame-wild principles respectively say that the left and right inequalities in (7-1) always hold. Similarly, one has the perhaps larger inertial interval $[\tilde{\alpha}, \tilde{\omega}]$ and the perhaps even larger broad interval $[\hat{\alpha}, \hat{\omega}]$.

To transfer the additive inequalities (7-1) into the multiplicative language of divisibility, we make use of the following formalism. Note that the torsion-free group $\mathcal{I}$ of fractional ideals of a local or global number field F embeds into its tensor product over $\mathbb{Z}$ with $\mathbb{Q}$, a group we write as $\mathcal{I}^{\mathbb{Q}}$ to account for the fact that $\mathcal{I}$ is written multiplicatively. In $\mathcal{I}^{\mathbb{Q}}$, as our notation indicates, general rational exponents on ideals are allowed. Then (7-1) corresponds to

$$\mathfrak{D}_{K_2/F}^{\alpha} \mid \mathfrak{D}_{K_1/F} \mid \mathfrak{D}_{K_2/F}^{\omega}, \tag{7-2}$$

which makes sense for both local and global number fields. In this formalism, the relations of the introductory example take the form $\mathfrak{D}_{K_6/F}^{1/3} \mid \mathfrak{D}_{K_5/F} \mid \mathfrak{D}_{K_6/F}$.

7B. ***Mean-root normalization and the comparison interval.*** It is sometimes insightful to switch to a slightly different normalization. We call this normalization mean-root normalization, with "mean" capturing how additive quantities are renormalized and "root" capturing how multiplicative quantities are renormalized.

If K/F has degree n and discriminant $\mathfrak{D}_{K/F}$ then its *root discriminant* is by definition $\mathfrak{d}_{K/F} = \mathfrak{D}_{K/F}^{1/n}$. To make this shift in our formalism, we simply replace all permutation characters ϕ_i by the scaled-down quantities $\underline{\phi}_i = \phi_i/\phi_i(1)$. One has mean tame conductors $\underline{c}_\tau(\phi_i) = (a_\tau, \underline{\phi}_i)$ as well as their analogs $\underline{c}_I(\phi_i) = (a_I, \underline{\phi}_i)$ and $\hat{\underline{c}}_\tau(\phi_i) = (\hat{a}_\tau, \underline{\phi}_i)$. We always indicate this alternative convention by underlining. Thus the mean-root normalized tame hull for two characters indexed by dimension is $\underline{T}'_+(G, \phi_n, \phi_m) = [\underline{\alpha}, \underline{\omega}]$ where $\underline{\alpha} = m\alpha/n$ and $\underline{\omega} = m\omega/n$. The divisibility relation (7-2) becomes $\mathfrak{d}_{K_m/F}^{\underline{\alpha}} \mid \mathfrak{d}_{K_n/F} \mid \mathfrak{d}_{K_m/F}^{\underline{\omega}}$.

The *comparison interval* $[\underline{\alpha}, \underline{\omega}]$ just introduced supports an intuitive understanding of how ramification in K_n/F and K_m/F relate to each other. Suppose, for example, that $K_m/K_n/F$ is a tower of fields so that one has the standard divisibility relation

$$\mathfrak{d}_{K_n/F} \mid \mathfrak{d}_{K_m/F}. \tag{7-3}$$

Then, assuming K_m/F is actually ramified, the ratio

$$\frac{\log|\mathfrak{d}_{K_n/F}|}{\log|\mathfrak{d}_{K_m/F}|} \in [0, 1]$$

can be understood as the fraction of ramification in K_m/F which is seen already in K_n/F. If the corresponding tame-wild principle holds, then this quantity is guaranteed to be in $[\underline{\alpha}, \underline{\omega}]$.

The mean-root normalization introduces a sense of absolute scale, with the number one playing a prominent role, as illustrated by the preceding paragraph and the next three subsections. Assuming $\phi_n - \phi_m$ is not a constant, one always has strict inequality $\underline{\alpha} < \underline{\omega}$. The failure of resolvent constructions from (G, ϕ_n)-fields to (G, ϕ_m)-fields to preserve ordering by absolute discriminants is, roughly speaking, measured by the length of $[\underline{\alpha}, \underline{\omega}]$. For ϕ_n and ϕ_m coming from faithful transitive permutation representations, a very common situation is $\underline{\alpha} \le 1 \le \underline{\omega}$. This tendency gets stronger as n and m increase to $|G|$. For example, for $(A_5, \phi_{20}, \phi_{30})$ the partition matrix is

$$\begin{pmatrix} 2^{10} & 3^6 1^2 & 5^4 \\ 2^{14} 1^2 & 3^{10} & 5^6 \end{pmatrix},$$

and the comparison interval works out to $\left[\frac{9}{10}, \frac{15}{14}\right] \approx [0.90, 1.07]$.

7C. ***The right tame-wild principle often holds for (G, ϕ_i, ϕ_r).*** Applying (7-3) in our setting gives

$$\mathfrak{d}_{K/F} \mid \mathfrak{d}_{K^{\mathrm{gal}}/F} \tag{7-4}$$

when K is a field. This relation holds also when K is an algebra, as can be seen by expressing K as a product of fields and comparing each factor to the field K^{gal}.

The critical quantity is simply expressed as

$$\underline{\omega}(G, \phi_i, \phi_r) = \max_{\tau \in G^{\sharp 0}} \frac{\underline{c}_\tau(\phi_i)}{\underline{c}_\tau(\phi_r)}. \tag{7-5}$$

The denominator depends only on the order $\bar{\tau}$ of τ via $\underline{c}_\tau(\phi_r) = (\bar{\tau} - 1)/\bar{\tau}$. For the more complicated numerator, one has $\underline{c}_\tau(\phi_i) \leq (\bar{\tau} - 1)/\bar{\tau}$, with equality if and only if the partition $\lambda_\tau(\phi_i)$ has the form $\bar{\tau}^{n/\bar{\tau}} = \bar{\tau} \ldots \bar{\tau}$. A permutation is *semiregular* if all cycles have the same length. Therefore $\underline{\omega}(G, \phi_r, \phi_i) \leq 1$, with equality if and only if G contains a nonidentity element which is semiregular. Summarizing:

Proposition 7.1. *Let $G \subseteq S_n$ be a permutation group containing a nonidentity semiregular element, ϕ_i the given permutation character, and ϕ_r the regular character. Then the right tame-wild principle holds for (G, ϕ_i, ϕ_r) with*

$$\underline{\omega}(G, \phi_i, \phi_r) = 1.$$

However, this principle is nothing more than the classical statement that for any (K, K^{gal}) of type (G, ϕ_i, ϕ_r), one has $\mathfrak{d}_{K/F} \mid \mathfrak{d}_{K^{\mathrm{gal}}/F}$.

7D. ***Elusive groups.*** In the global setting, we are mainly interested in the case when K is a field and thus G is transitive. A transitive permutation group which does not contain a nonidentity semiregular element is called an *elusive group* [Cameron et al. 2002]. So Proposition 7.1 is the best statement for nonelusive transitive groups, but the situation needs to be investigated further for elusive groups.

Elusive groups are aptly named in that they are relatively rare. The smallest n for which S_n contains an elusive group is $n = 12$. There are five elusive groups in S_{12} up to permutation equivalence, listed in Table 1, all subgroups of the Mathieu group M_{11} in its transitive degree twelve realization 12T272. Here and in the sequel we use the T-notation for transitive permutation groups introduced in [Conway et al. 1998] and available online in several places, including [LMFDB 2013].

The following proposition treats these five groups.

Proposition 7.2. *The right tame-wild principle for (G, ϕ_i, ϕ_r) holds for the elusive groups* 12T46, 12T84, 12T181, *and* 12T272 *with $\underline{\omega}(G, \phi_i, \phi_r) = \frac{20}{21}$. Thus the strengthening $\mathfrak{d}_{K/F} \mid \mathfrak{d}_{K^{\mathrm{gal}}/F}^{20/21}$ of* (7-5) *holds for these groups. For* 12T47, *one has $\underline{\omega}(12\mathrm{T}47, \phi_i, \phi_r) = \frac{8}{9}$. Extensions (K_i, K_r) from* (7-6) *give an counterexample to the tame-wild principle over $\mathbb{Q}(\sqrt{-3}\,)$, but there is no counterexample over $\mathbb{Q}$.*

12T46 $\cong C_3^2 : Q_8$	$\surd$	$\surd$	$\surd$			$\surd$			
12T47 $\cong M_9$	$\surd$	$\surd$	$\surd$						
12T84 $\cong C_3^2 : \hat{Q}_8$	$\surd$	$\surd$	$\surd$		$\surd$	$\surd$			
12T181 $\cong M_{10}$	$\surd$	$\surd$	$\surd$	$\surd$		$\surd$			
12T272 $\cong M_{11}$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$		
τ	2A	3A	4A	5A	6A	8A	11A	Q_8	I
$\lambda_\tau(\phi_i)$	$2^4 1^4$	$3^3 1^3$	$4^2 2^2$	$5^2 1^2$	6321	84	(11)1	84	
$c_\tau(\phi_i)$	4	6	8	8	8	10	10	10	≤ 10
$\underline{c}_\tau(\phi_r)$	1/2	2/3	3/4	4/5	5/6	7/8	10/11	7/8	$(\lvert I\rvert-1)/\lvert I\rvert$
$\underline{c}'_\tau$	2/3	3/4	8/9	5/6	4/5	20/21	11/12	20/21	$\leq 20/21$

Table 1. Information used in the proof of Proposition 7.2.

Proof. The part below the line of Table 1 supports applying the broad and inertial methods for 12T272 $\cong M_{11}$. Thus the line labeled τ lists out the seven elements of 12T272$^{\sharp 0}$. The next two lines gives the corresponding dodecic partitions $\lambda_\tau(\phi_i)$ and conductors $c_\tau(\phi_i) = 12\underline{c}_\tau(\phi_i)$ respectively. The next lines give

$$\underline{c}_\tau(\phi_r) = \frac{c_\tau(\phi_r)}{|M_{11}|} \quad \text{and} \quad \underline{c}'_\tau = \frac{\underline{c}_\tau(\phi_i)}{\underline{c}_\tau(\phi_r)}.$$

Thus the comparison interval is $[\underline{\alpha}, \underline{\omega}] = \left[\frac{2}{3}, \frac{20}{21}\right]$.

One inertial subgroup of 12T272 is Q_8, which has orbit partition 84. As indicated in the second-to-last column of Table 1, its associated quantity is $\underline{c}'_{Q_8} = \frac{20}{21}$, which is the right endpoint of $\left[\frac{2}{3}, \frac{20}{21}\right]$. In general, the difficulty with the inertial method is that there can be many inertial N-subgroups I to inspect. However here we can treat them all at once as follows. Since none of the elusive groups from the complete list are themselves inertial groups, I must act intransitively and so $c_I(\phi_i) \leq 10$. Also $\underline{c}_I(\phi_r) = (|I|-1)/|I| \geq \frac{7}{8}$, since N-groups have order at least 8. So, as indicated by the table, $\underline{c}'_I \leq \frac{20}{21}$. Thus our initial case Q_8 was in fact the worst case, and the right tame-wild principle holds for (M_{11}, ϕ_i, ϕ_r).

The part of the table above the line gives the partitions which arise for all the G, as a subset of those that we have listed for 11T272. The smaller groups 12T46, 12T84, and 12T181 still have elements of cycle type 84, and so the same argument goes through for them, proving the tame-wild principle for (G, ϕ_i, ϕ_r) in these cases. Note that our uniform treatment of all the G uses that $\underline{c}_\tau(\phi_r)$ is independent of G; in contrast, the unnormalized quantity $c_\tau(\phi_r)$ depends on G.

A counterexample seems likely for $G =$ 12T47 because it contains Q_8 and its comparison interval is only $[\underline{\alpha}, \underline{\omega}] = \left[\frac{2}{3}, \frac{8}{9}\right]$, with $\frac{8}{9}$ being considerably less than $\frac{20}{21}$. However, as discussed in reference to (6-2), there are no quaternionic extensions of $\mathbb{Q}_2$ giving counterexamples to the universal tame-wild principle for Q_8. Other candidate I do not work either, and we are forced to leave $\mathbb{Q}$ as a ground field.

Our counterexample comes from fields $K_n = \mathbb{Q}[x]/f_n(x)$ with discriminants $D_n \in \mathbb{Z}$ and Galois groups $G_n = \mathrm{Gal}(K_n^{\mathrm{gal}}/\mathbb{Q})$ as follows:

n	$f_n(x)$	D_n	G_n
8	x^8+6x^4-3	$-2^{16}3^7$	$\hat{Q}_8$
9	$x^9-3x^8+18x^5+18x^4-27x+9$	$-2^{16}3^{15}$	9T19
12	$x^{12}-6x^{10}-4x^9+12x^7-36x^5+30x^4+8x^3-8$	$2^{22}3^{18}$	12T84

(7-6)

The overgroup $12\mathrm{T}84 \cong C_3^2 : \hat{Q}_8 \supset 12\mathrm{T}47$ was chosen because it contains not just Q_8 but also $\hat{Q}_8$. The nonic group 9T19 is a lower degree realization of 12T84, where the isomorphism with $C_3^2 : \hat{Q}_8$ is naturally realized. The field K_8 was chosen as a strong candidate from which to build a counterexample, because $\mathrm{Gal}(K_8^{\mathrm{gal}}/\mathbb{Q})$ is its own decomposition group with slope content $[2, 2, 2.5]^2$ as in (6-2). The field K_9 was extracted from the database [Jones and Roberts $\geq$ 2014] as a 9T19 field with K_8 as a resolvent, and then K_{12} was obtained from K_9 by resolvent calculations.

The splitting field K_{12}^{gal} contains $\mathbb{Q}(\sqrt{-3}\,)$ with $\mathrm{Gal}(K_{12}^{\mathrm{gal}}/\mathbb{Q}(\sqrt{-3}\,)) = 12\mathrm{T}47$ by construction. The root discriminant of K_{12}^{gal} is $2^2 3^{127/72}$, as computed by using the website of [Jones and Roberts 2006] to analyze ramification in $K_9/\mathbb{Q}$. Here the exponent 2 can be confirmed from a standard computation associated with the slope content $[2, 2, 2.5]^2$, namely $\frac{2}{8}+\frac{2}{4}+\frac{2.5}{2}=2$. On the other hand K_{12} has root discriminant $2^{22/12}3^{18/12}$. The quotient $\frac{22}{12}/2 = \frac{11}{12}$ is to the right of the root-normalized tame hull $[\underline{\alpha}, \underline{\omega}] = \left[\frac{2}{3}, \frac{8}{9}\right]$, giving a counterexample to the right tame-wild principle for $(12\mathrm{T}47, \phi_i, \phi_r)$ over $\mathbb{Q}(\sqrt{-3}\,)$. □

7E. *The left tame-wild principle always holds for (G, ϕ_i, ϕ_r).* The following theorem shows that an important part of the tame-wild principle holds for all finite groups G.

Theorem 7.3. *Let $G \subseteq S_n$ be any permutation group, ϕ_i the given permutation character, and ϕ_r the regular character. Let $\mathcal{F}(G)$ be the maximal number of fixed points of a nonidentity element of G. Then the left tame-wild principle holds for (G, ϕ_i, ϕ_r) with*

$$\underline{\alpha}(G, \phi_i, \phi_r) = 1 - \frac{\mathcal{F}(G)}{n}.$$

Thus for any (K, K^{gal}) of type (G, ϕ_i, ϕ_r), one has $\mathfrak{d}_{K^{\mathrm{gal}}/F}^{1-\mathcal{F}(G)/n} \mid \mathfrak{d}_{K/F}$.

Proof. We apply the broad method. Let τ be an arbitrary element of $G^{\sharp 0}$ and call its order t. Consider the corresponding column $\binom{\lambda}{\Lambda}$ of the partition matrix $P(G, \phi_i, \phi_r)$. Then $\lambda = \prod_{k=1}^{t} k^{m_k}$ is some partition of n and $\Lambda = t^{|G|/t}$ is the corresponding partition of $|G|$.

The projective matrices $T'(G, \phi_i, \phi_r)$ and $\hat{T}'(G, \phi_i, \phi_r)$ have just one row each. The entries in the τ column are respectively $c(\lambda)/c(\Lambda)$ and $\hat{c}(\lambda)/\hat{c}(\Lambda)$. Their difference is positive:

$$\begin{aligned}\frac{\hat{c}(\lambda)}{\hat{c}(\Lambda)} - \frac{c(\lambda)}{c(\Lambda)} &= \frac{\sum_{k=2}^{t} m_k k}{|G|} - \frac{\sum_{k=2}^{t} m_k (k-1)}{|G|(t-1)/t} \\ &= \frac{1}{|G|}\sum_{k=2}^{t} m_k \frac{k(t-1)}{t-1} - \frac{1}{|G|}\sum_{k=2}^{t} m_k \frac{t(k-1)}{t-1} \\ &= \frac{1}{|G|}\sum_{k=2}^{t} m_k \frac{(kt-k)-(kt-t)}{t-1} \\ &= \frac{1}{|G|}\sum_{k=2}^{t} m_k \frac{t-k}{t-1} \geq 0.\end{aligned}$$

Thus $\hat{c}(\lambda)/\hat{c}(\Lambda) \geq c(\lambda)/c(\Lambda)$, and so certainly all the $\hat{c}(\lambda)/\hat{c}(\Lambda)$ are at least

$$\alpha(G, \phi_i, \phi_r) = \min_{\tau \in G^{\sharp 0}} \frac{c_\tau(\lambda)}{c_\tau(\Lambda)}.$$

Thus the left tame-wild principle holds for (G, ϕ_i, ϕ_r), and moreover we can compute $\alpha(G, \phi_i, \phi_r)$ using $\hat{c}_\tau$ rather than c_τ, giving

$$\alpha(G, \phi_i, \phi_r) = \min_{\tau \in G^{\sharp 0}} \frac{\hat{c}_\tau(\lambda)}{\hat{c}_\tau(\Lambda)} = \min_{\tau \in G^{\sharp 0}} \frac{n - m_1}{|G|} = \frac{n - \mathcal{F}(G)}{|G|}.$$

Switching to the mean-root normalization gives $\underline{\alpha}(G, \phi_i, \phi_r) = 1 - \mathcal{F}(G)/n$. $\square$

Number field tabulation. For certain solvable transitive groups $G \subset S_n$, the techniques of [Jones and Wallington 2012] let one compute all degree n fields K of type G where $|\mathfrak{d}_{K^{\mathrm{gal}}/\mathbb{Q}}|$ is at most some constant β. Then the theorem just proved can be applied through its corollary $|\mathfrak{d}_{K^{\mathrm{gal}}/\mathbb{Q}}^{1-\mathcal{F}(G)/n}| \leq |\mathfrak{d}_{K/\mathbb{Q}}|$ to obtain all K with $|\mathfrak{d}_{K/\mathbb{Q}}|$ at most $B = \beta^{1-\mathcal{F}(G)/n}$. This computation is carried out in [Jones 2013] for the primitive nonic groups 9T9, 9T14, 9T15, 9T16, 9T23, and 9T27 to obtain the corresponding nonic fields with smallest absolute discriminant. This particular application served as the catalyst for the present paper.

8. Examples and counterexamples

The positive and negative results of the previous sections give one a good idea of the extent to which the tame-wild principle holds and how it can be applied. We now refine this picture, by considering various $(G, \phi_1, \ldots, \phi_r)$ of interest and determining whether the tame-wild principle holds. In Section 8A, we give examples illustrating the broad method and the inertial method. In Section 8B, we conclude by arguing that counterexamples to the tame-wild principle from pairs $(K_1, K_2)/\mathbb{Q}$ of number fields are not easily found, but present one such counterexample with Galois group 12T112 of order 192.

8A. ***The broad and inertial methods.*** We illustrate the two methods of Section 4B with positive results for three N-groups.

The broad method for $(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_8, \phi_{8a}, \phi_{8b})$. The group $\mathrm{Aff}_3(\mathbb{F}_2)$ provides a simple illustration of the broad method in the setting $r = 4$. It has five nontrivial small permutation representations $\rho_{7a}, \rho_{7b}, \rho_8, \rho_{8a}, \rho_{8b}$, with images the permutation groups 7T5, 7T5, 8T37, 8T48, 8T48. The first three representations are through the quotient $\mathrm{GL}_3(\mathbb{F}_2) \cong \mathrm{PGL}_2(\mathbb{F}_7)$ while the last two are faithful. The representations ρ_{7a} and ρ_{7b} share a common character ϕ_7. They are thus arithmetically equivalent and we call them identical twins. The representations ρ_{8a} and ρ_{8b} have different characters ϕ_{8a} and ϕ_{8b} and so we call them fraternal twins. The four characters ϕ_7, ϕ_8, ϕ_{8a}, and ϕ_{8b} are linearly independent.

Figure 3 first presents the partition matrix $P = P(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_8, \phi_{8a}, \phi_{8b})$ and the broad and tame matrices derived from it. For visualization purposes, it then drops consideration of ϕ_8. After this projection, it plots the columns of $\hat{T}'$ as +s and those of T' as •s. Since the +s are in the hull T'_+ of the •s, the tame-wild principle holds for $(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_{8a}, \phi_{8b})$. Working more algebraically, as described in Section 4C, one can verify the analogous convexity assertion in the presence of ϕ_8, giving the first sentence of the following result.

Proposition 8.1. *The tame-wild principle holds for* $(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_8, \phi_{8a}, \phi_{8b})$. *In particular, to find all* 8T48 *extensions with* $|\mathfrak{D}_{K_{8a}/F}| \leq B$, *one need look only at* 7T5 *extensions with* $|\mathfrak{D}_{K_7/F}| \leq B$ *and select from among the octic resolvents of their* 14T34 *quadratic overfields.*

The second sentence comes from an understanding of the algebraic meaning of Figure 3. Associate variables u, a, and b to ϕ_7, ϕ_{8a}, and ϕ_{8b} respectively. The four sides of the trapezoid $T'_+(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_{8a}, \phi_{8b})$ in the drawn $(u/b, a/b)$ plane correspond to the four faces of the cone $T_+(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_{8a}, \phi_{8b})$ in (u, a, b)-space. These four faces correspond to the four inequalities on local exponents on the left and they translate into divisibility relations among either local or global discriminants on the right:

$$u \leq a \leq u + b, \qquad \mathfrak{D}_{K_7/F} \mid \mathfrak{D}_{K_{8a}/F} \mid \mathfrak{D}_{K_7/F}\mathfrak{D}_{K_{8b}/F},$$
$$u \leq b \leq u + a, \qquad \mathfrak{D}_{K_7/F} \mid \mathfrak{D}_{K_{8b}/F} \mid \mathfrak{D}_{K_7/F}\mathfrak{D}_{K_{8a}/F}.$$

For tabulations of all extensions K_{8a}/F with $|\mathfrak{D}_{K_{8a}/F}|$ at most some bound B, the procedure referred to by the proposition is to look for all K_7/F with $|\mathfrak{D}_{K_7/F}| \leq B$,

τ	2A	2B	2C	3A	4A	4B	4C	6A	7A
	1^7	22111	22111	331	22111	421	421	331	7
P	1^8	2222	2222	3311	2222	44	44	3311	71
	2222	221111	2222	3311	44	4211	44	62	71
	2222	2222	221111	3311	44	44	4211	62	71
	0	4	4	6	4	6	6	6	7
$\hat{T}$	0	8	8	6	8	8	8	6	7
$+$	8	4	8	6	8	6	8	8	7
	8	8	4	6	8	8	6	8	7
	0	2	2	4	2	4	4	4	6
T	0	4	4	4	4	6	6	4	6
$\bullet$	4	2	4	4	6	4	6	6	6
	4	4	2	4	6	6	4	6	6

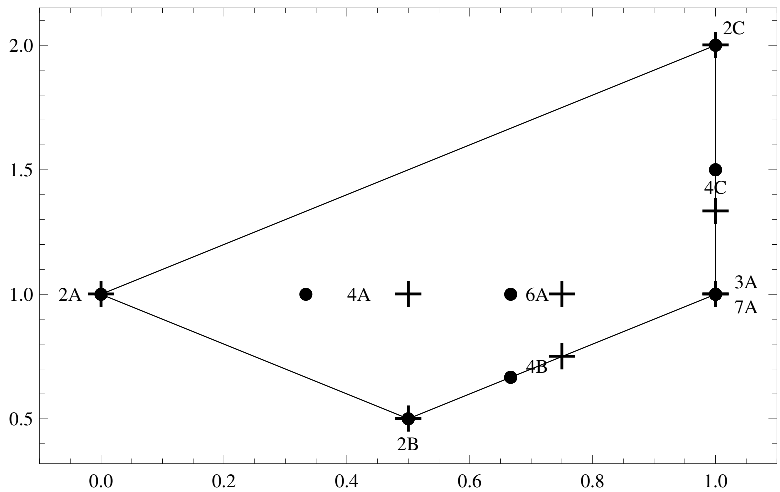


Figure 3. Top: The partition matrix, broad matrix, and tame matrix for $(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_8, \phi_{8a}, \phi_{8b})$. Bottom: the broad hull and tame hulls coinciding after removing ϕ_8 from consideration, proving the tame-wild principle for $(\mathrm{Aff}_3(\mathbb{F}_2), \phi_7, \phi_{8a}, \phi_{8b})$.

take suitable square roots to pass from 7T5 fields to 14T34 fields, and then use resolvents to obtain the desired 8T48 fields.

The inertial method for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$. The group S_6 has three faithful permutation representations of degree at most ten: two sextic ones ρ_{6a} and ρ_{6b} interchanged by the outer automorphism of S_6, and a decic one ρ_{10} coming from the exceptional isomorphism $S_6 \cong \mathrm{PSL}_2(\mathbb{F}_9).\mathrm{Gal}(\mathbb{F}_9/\mathbb{F}_3) = 10T32 \subset S_{10}$.

	2A	2B	2C	3A	3B	4A	4B	5A	6A	6B
P	222	21^4	2211	33	3111	42	411	51	6	321
	21^4	222	2211	3111	33	42	411	51	321	6
	$2^3 1^4$	$2^3 1^4$	$2^4 1^2$	3331	3331	4411	442	55	631	631
$\hat{T}$	6	2	4	6	3	6	4	5	6	5
$+$	2	6	4	3	6	6	4	5	5	6
	6	6	8	9	9	8	10	10	9	9
T	3	1	2	4	2	4	3	4	5	3
$\bullet$	1	3	2	2	4	4	3	4	3	5
	3	3	4	6	6	6	7	8	7	7

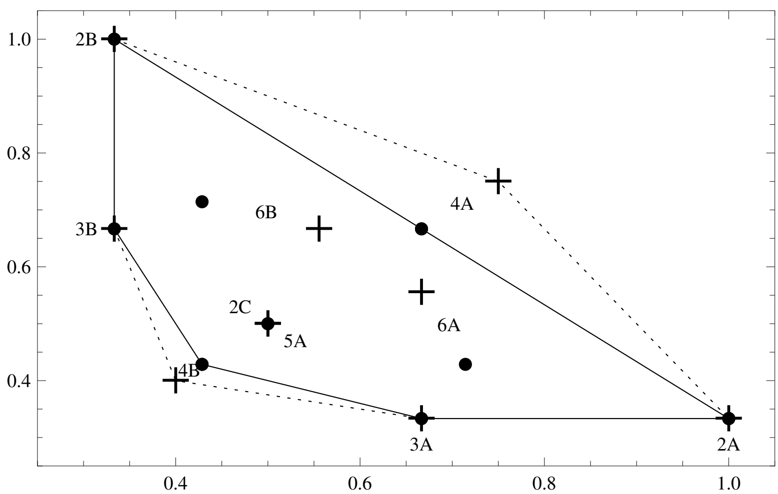


Figure 4. Top: The partition matrix, broad matrix, and tame matrix for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$. Bottom: the broad hull strictly containing the tame hull, showing that the broad method does not suffice to prove the tame-wild principle for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$.

Figure 4 presents our standard analysis of the situation. Since some +s are outside of the tame hull $T'_+(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$, the broad method does not suffice for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$. However after projection to the horizontal axis, the +s are indeed in the convex hull of the •s, so that the broad method establishes the tame-wild principle for $(S_6, \phi_{6a}, \phi_{10})$. Also the ratios a/b for the + points (a, b) are within the interval $\left[\frac{1}{3}, 3\right]$ formed by the ratios for the • points, proving the tame-wild principle for $(S_6, \phi_{6a}, \phi_{6b})$.

In fact, the tame-wild principle is true for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$ as follows. The only inertial subgroups not covered by previous considerations are $I_1 = D_4 \times C_2$

and the twin pair $(I_2, I_3) = (A_4 \times C_2, 6T6)$. The orbit partitions in the three cases are (42, 42, 442), (42, 6, 64), and (6, 42, 64). The associated conductor vectors are then (4, 4, 7), (4, 5, 8), and (5, 4, 8). Their projectivized versions are $(\frac{4}{7}, \frac{4}{7})$, $(\frac{1}{2}, \frac{5}{8})$, and $(\frac{5}{8}, \frac{1}{2})$. Since these points are visibly in $T'_+(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$, the tame-wild principle holds. We have given this argument to illustrate how the inertial method typically applies. However in this case the inertial groups I_2 and I_3 could also have been treated by using the techniques from Section 6, as in fact the universal tame-wild principle holds for $A_4 \times C_2$.

Summarizing, we have proved the first sentence:

Proposition 8.2. *The tame-wild principle holds for* $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$. *In particular to find all decic* S_6*-extensions with* $|\mathfrak{D}_{K_{10}/F}| \leq B$, *one need only look at sextic* S_6*-extensions with* $|\mathfrak{D}_{K_{6a}/F}| \leq B^{2/3}$ *and select from among their decic resolvents.*

For the second sentence, note first that the locations of the rightmost and highest points of the tame hull in Figure 4 respectively correspond to the equivalent statements $\mathfrak{D}_{K_{6a}/F} \mid \mathfrak{D}_{K_{10}/F}$ and $\mathfrak{D}_{K_{6b}/F} \mid \mathfrak{D}_{K_{10}/F}$. Each of these says that to find all decics with absolute discriminant $\leq B$, it suffices to look at all sextics up to that bound. A considerable improvement is to see that the long diagonal boundary between them corresponds to $\mathfrak{D}_{K_{6a}/F}\mathfrak{D}_{K_{6b}/F} \mid \mathfrak{D}^{4/3}_{K_{10}/F}$ which implies the statement.

The broad method for $(W(E_6), \phi_{27}, \phi_{36}, \phi_{40a}, \phi_{40b}, \phi_{45})$. As we have seen in Section 7E and by the earlier examples of this subsection, the broad method works well in the setting $r = 2$. As r increases, the difference between a_τ and $\hat{a}_\tau$ becomes more visible, and the broad method often fails even when the tame-wild principle is true, as we just saw for $(S_6, \phi_{6a}, \phi_{6b}, \phi_{10})$.

A clear illustration of the effectiveness of the broad method and its decay with increasing r comes from the Weyl group $W(E_6)$ of order $51840 = 2^6 3^4 5$ and the permutation characters $\phi_{27}, \phi_{36}, \phi_{40a}, \phi_{40b}, \phi_{45}$ corresponding to five maximal subgroups [Conway et al. 1985]. The broad method immediately shows that the tame-wild principle for $(W(E_6), \phi_u, \phi_v)$ holds for all ten possibilities for $\{u, v\}$. From ten pictures like Figures 3 and 4, now quite involved since $|W(E_6)^{\sharp 0}| = 24$, the broad method establishes the tame-wild principle in exactly four of the ten cases $(W(E_6), \phi_u, \phi_v, \phi_w)$ as follows.

Proposition 8.3. *For* $\{u, v, w\} = \{27, 36, 40a\}$, $\{27, 40a, 40b\}$, $\{36, 40a, 40b\}$, *and* $\{36, 40b, 45\}$, *the tame-wild principle holds for* $(W(E_6), \phi_u, \phi_v, \phi_w)$.

Pursuing this situation further with the inertial method would be harder, because $W(E_6)$ has many 2-inertial and 3-inertial subgroups.

8B. ***Best counterexamples.*** Let G be a group for which the universal tame-wild principle fails. Then there exists a vector $v \in \mathbb{Q}(G^\sharp)$ for which the tame-wild principle fails for $(G, \langle v \rangle)$. There are infinitely many solutions to $\phi_1 - \phi_2 \in \langle v \rangle$

with the ϕ_i permutation characters. So any failure of the universal tame-wild principle can be converted to a failure in the setting (G, ϕ_1, ϕ_2) of the introduction. By switching ϕ_1 and ϕ_2 if necessary, it can be converted to a failure of the left tame-wild principle for (G, ϕ_1, ϕ_2).

However these counterexamples are not guaranteed to have immediate bearing on our applications to tabulating number fields. All that is asserted by the failure of the principle for (G, ϕ_1, ϕ_2) is that there exists a pair of local extensions $(K_1, K_2)/F$ of the given type with

$$\mathfrak{D}_{K_2/F}^{\alpha(G,\phi_1,\phi_2)} \mid \mathfrak{D}_{K_1/F} \tag{8-1}$$

not holding. More directly relevant would be global counterexamples with the ϕ_i both coming from faithful transitive permutation representations and the extensions K_i/F full in the sense of each having Galois group $\mathrm{Gal}(K^{\mathrm{gal}}/F)$ all of G. More demanding still is to ask for counterexamples of this sort with $F = \mathbb{Q}$. Finally, one can seek examples for which even the weaker numerical statement

$$|\mathfrak{D}_{K_2/\mathbb{Q}}|^{\alpha(G,\phi_1,\phi_2)} \le |\mathfrak{D}_{K_1/\mathbb{Q}}| \tag{8-2}$$

fails. Examples of this explicit nature often do not exist for a given G, and even when they exist they can be hard to find. The rest of this subsection discusses the construction of global counterexamples built from one of the two local counterexamples with $I = Q_8$ with slope-content $[2, 2, \frac{5}{2}]^2$ from (6-2). There are several points of contact with Section 7D, but here we find counterexamples to (8-1) over $\mathbb{Q}$.

Inadequacy of $G = \hat{Q}_8$ *as a source of global counterexamples.* The group $\hat{Q}_8$ itself is not a source of global counterexamples of the sort we seek because it has only two transitive faithful permutation characters and the tame-wild principle holds for the corresponding type $(\hat{Q}_8, \phi_8, \phi_{16})$. To illustrate the best that can be done with this group, take

n	$f_n(x)$	D_n	G_n	$\lvert G_n\rvert$
8	x^8+6x^4-3	$-2^{16}3^7$	$\hat{Q}_8$	16
4	x^4+6x^2-3	$-2^6 3^3$	D_4	8
2	x^2+3	-3	C_2	2

(8-3)

The global and 2-adic Galois groups of $f_8(x)$ agree, and so one has this agreement for the resolvents $f_4(x)$ and $f_2(x)$ as well. The Galois groups G_n and the field discriminants D_n are as indicated. The fields $K_n = \mathbb{Q}[x]/f_n(x)$ belong to transitive characters ϕ_8, ϕ_4, and ϕ_2 of $\hat{Q}_8$.

Figure 5 is an analog of Figure 1, but now for $(\hat{Q}_8, \phi_8+\phi_2, \phi_8+\phi_4)$. The algebra pair $(K_8 \times K_2, K_8 \times K_4)/\mathbb{Q}$ yields the exponent pair $(a_2, b_2) = (16, 22)$ which is just outside the tame cone. So this pair of algebras indeed contradicts (8-1), but we are seeking counterexamples among pairs of fields.

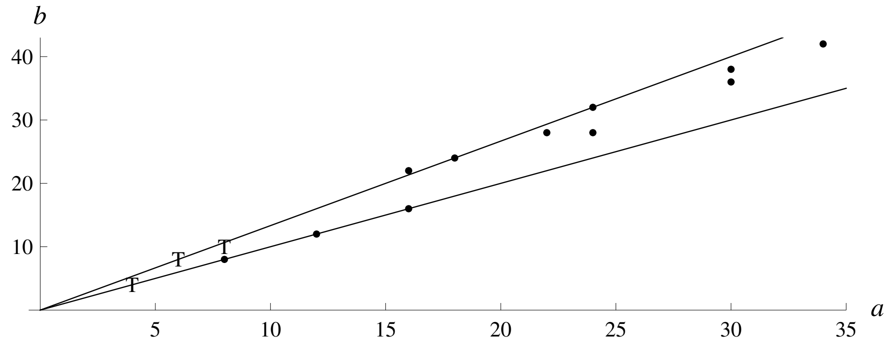


Figure 5. An analog of the introductory Figure 1 for the type $(\hat{Q}_8, \phi_8+\phi_2, \phi_8+\phi_4)$. The points are exactly all the possibilities for exponent pairs (a_2, b_2) from wild 2-adic ramification over $\mathbb{Q}_2$, and (16, 22) is just outside the tame cone.

Failure of the inertial method for $(M_{12}, \phi_{12a}, \phi_{12b})$. To get a better global counterexample corresponding to the same local counterexample, we need to replace $\hat{Q}_8$ by larger groups G containing it. An initial key observation is that the quaternion group Q_8 is the four-point stabilizer of the Mathieu group $M_{12} \subset S_{12}$ of order $12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$ in its natural action, and also one has $Q_8 \subset \hat{Q}_8 \subset M_{12}$. On the one hand, the given character ϕ_{12a} of the Mathieu group has decomposition $\phi_8+\phi_2+2$ when restricted to $\hat{Q}_8$. On the other hand, there is a twin dodecic character ϕ_{12b} coming from the outer involution of M_{12}; its restriction to $\hat{Q}_8$ decomposes as $\phi_8+\phi_4$.

Further group-theoretic facts are necessary for this situation to give number fields as desired. First, the partition matrix and projective tame matrix of $(M_{12}, \phi_{12a}, \phi_{12b})$ are as follows:

τ	2A	2B	3A	3B	4A	4B	5A	6A	6B	8A	8B	10A	11AB	Q_8
$\lambda_\tau(\phi_{12a})$	2^6	$2^4 1^4$	$3^3 1^3$	3^4	$4^2 2^2$	$4^2 1^4$	$5^2 1^2$	6^2	6321	84	821^2	(10)2	(11)1	81^4
$\lambda_\tau(\phi_{12b})$	2^6	$2^4 1^4$	$3^3 1^3$	3^4	$4^2 1^4$	$4^2 2^2$	$5^2 1^2$	6^2	6321	821^2	84	(10)2	(11)1	84
c'_τ	1	1	1	1	$\frac{8}{6}$	$\frac{6}{8}$	1	1	1	$\frac{10}{8}$	$\frac{8}{10}$	1	1	$\frac{7}{10}$

Thus, in the language introduced in Section 8A, extensions $(K_{12a}, K_{12b})/F$ of full type $(M_{12}, \phi_{12a}, \phi_{12b})$ are fraternal twins, this being necessary for our purposes. But they are near-identical in the sense that the interval $[\alpha, \omega]$ is small, being $\left[\frac{3}{4}, \frac{4}{3}\right] = [0.75, 1.\overline{33}]$ here, rather than the intervals $\left[\frac{1}{2}, 2\right]$ and $\left[\frac{1}{3}, 3\right]$ seen in Section 8A for $\mathrm{Aff}_3(\mathbb{F}_2)$-twins and S_6-twins respectively. The orbit partitions of Q_8 are as indicated above, yielding $c_{Q_8}(\phi_{12a})/c_{Q_8}(\phi_{12b}) = \frac{7}{10} = 0.70$ which is outside the interval $[0.75, 1.\overline{33}]$. Thus the inertial method for proving the tame-wild principle fails here.

Failure of the tame-wild principle for $(M_{12}, \phi_{12a}, \phi_{12b})$. Computing with the slopes $[2, 2, 2.5]^2$, the 8s in the last column above give discriminant exponent $3\cdot 2+4\cdot 2.5 = 16$ while the 4 gives the discriminant exponent $3 \cdot 2 = 6$. So the ratio of wild conductors is $\frac{16}{22} = 0.\overline{72}$, which is still outside the interval $[0.75, 1.\overline{33}]$. Thus the tame-wild principle itself fails for $(M_{12}, \phi_{12a}, \phi_{12b})$.

Smaller groups. We have looked in several places, including the two-parameter family of [Malle 2000], for twin pairs (K_{12a}, K_{12b}) of M_{12} fields with the needed quaternionic 2-adic behavior. We did not find any, and so we consider smaller groups as follows as potential sources of counterexamples:

$$\begin{array}{ccccccccccccc} & & \mathrm{Aff}_2(\mathbb{F}_3) & & P\Gamma L_2(\mathbb{F}_9) & & & & & & & & \\ & & \wr\| & & \wr\| & & & & & & & & \\ \hat{Q}_8 & \subset & M_9.2 & \subset & M_{10}.2 & & & & & & & & \\ \cup & & \cup & & \cup & & & & & & & & \\ Q_8 & \subset & M_9 & \subset & \boxed{M_{10}} & \subset & \boxed{M_{11}} & \subset & \boxed{M_{12}} & \supset & \boxed{T} & \supset & P. \end{array}$$

The four groups in the middle are boxed to stress that they appear in Proposition 8.4 below.

Proceeding from M_{12} to the left, the groups M_{11} and M_{10} contain $\hat{Q}_8$, since $\hat{Q}_8$ has orbit partition 8211. Thus, using $0.\overline{72} \notin [0.75, 1.\overline{33}]$ exactly as above, the tame-wild principle fails also for $(M_{11}, \phi_{11}, \phi_{12b})$ and $(M_{10}, \phi_{10}, \phi_{12b})$. Here the transitive permutation groups in question are respectively (11T6, 12T272) and (10T31, 12T181). The analog of Figure 5 for M_{10} and M_{11} has the same tame cone, but more dots. For M_{12} there are many more dots, and a symmetry appears with the cone doubling so that its bounding lines have slope $\frac{3}{4}$ and $\frac{4}{3}$ rather than 1 and $\frac{4}{3}$.

Moving further leftward to M_9 and Q_8 relates our current discussion to our earlier counterexamples. For M_9, the transitive groups are (9T14, 12T47). However now $\hat{Q}_8$ is not contained in M_9 and so we do not have counterexamples over $\mathbb{Q}$. However the counterexample for $(M_9, \phi_{12b}, \phi_{72})$ over $\mathbb{Q}(\sqrt{-3}\,)$ from Section 7D also gives a counterexample for $(M_9, \phi_9, \phi_{12b})$, as always because the projectivized wild Artin conductor $0.\overline{72}$ is not in the tame hull $[0.75, 1.\overline{33}]$. Finally for $\hat{Q}_8$ itself we recover (8-3), now interpreted as an intransitive counterexample for $(Q_8, \phi_8, \phi_8 + \phi_4)$ over $\mathbb{Q}(\sqrt{-3}\,)$.

The extended groups $M_9.2$ and $M_{10}.2$ corresponding to the pairs (9T19, 12T84) and (10T35, 12T220) are natural candidates to support examples over $\mathbb{Q}$ because they contain $\hat{Q}_8$. However they have orbit partitions 921 and (10)2 as subgroups of M_{12}. Computation in the column headed by 8B then has to be adjusted, with the 2 in 821^2 removed. The conductor ratio is then $\frac{7}{10}$ rather than $\frac{8}{10}$ and in fact the inertial method above works to prove the tame-wild principle for $(M_9.2, \phi_9, \phi_{12b})$ and $(M_{10}.2, \phi_{10}, \phi_{12b})$. This phenomenon illustrates the fundamental difficulty in promoting local nontransitive counterexamples to global transitive ones with a larger group. While wild Artin conductor ratios, here $0.\overline{72}$ stay the same, tame hulls increase, here from [0.75, 1] for $\hat{Q}_8$ itself to $\left[\frac{7}{10}, \frac{7}{6}\right] = [0.70, 1.1\overline{6}]$ for $M_{10}.2$.

There are other good candidates for global Galois groups. The 2-Sylow subgroup P of M_{12} of order 2^6 is not good for us, because neither the given orbit decomposition nor its twin is transitive, both having orbit partition 84. However an overgroup T of order $2^6 \cdot 3 = 192$ is good, with (ϕ_{12a}, ϕ_{12b}) remaining a fraternal pair of type (12T112, 12T112). Our computations have shown that the tame-wild principle fails for (12T112, ϕ_{12a}, ϕ_{12b}).

Number fields. Constructing number fields with nonsolvable Galois groups and prescribed ramification remains a difficult problem despite the increasing attention it has been receiving recently. Just as we have not found M_{12} fields with the appropriate quaternionic ramification, we have also not found M_{11} or M_{10} fields.

In contrast, it is relatively easy to build solvable fields step by step, and we have found many explicit pairs $(K_{12a}, K_{12b})/\mathbb{Q}$ providing counterexamples to the tame-wild principle for (12T112, ϕ_{12a}, ϕ_{12b}). One such, with tame ramification at the prime number $q = 277$, is

$$f_{12a}(x) = x^{12} + 223x^{10} + 14856x^8 + 1784qx^6 + 38160qx^4 + 1712q^2x^2 + 9216q^2,$$
$$f_{12b}(x) = x^{12} + 202x^8 + 49qx^4 + 4q^2.$$

The discriminants are $D_{12a} = 2^{16}277^8$ and $D_{12b} = 2^{22}277^6$, with the tame prime 277 having ramification partitions $\mu_{12a} = 4^22^2$ and $\mu_{12b} = 4^21^4$.

By design, $\mathfrak{D}_{K_{12b}}^{0.75} \nmid \mathfrak{D}_{K_{12a}}$. However the tame ramification at 277 completely overwhelms the wild ramification at 2 in terms of magnitudes, and easily $|\mathfrak{D}_{K_{12b}}|^{.75} \leq |\mathfrak{D}_{K_{12a}}|$. Indeed $|\mathfrak{D}_{K_{12b}}|^{1.15} \approx |\mathfrak{D}_{K_{12a}}|$. To improve upon the counterexample (f_{12a}, f_{12b}), one would like examples with $D_{12a} = 2^{16}(p_1 \ldots p_k)^6$ and $D_{12b} = 2^{22}(p_1 \ldots p_k)^8$ so that (8-2) is contradicted as well. However no such counterexamples exist with $G =$ 12T112, as the subgroup Q_8 together with all elements of type 4^21^4 generate an index two subgroup of type 12T63 and this subgroup does not contain $\hat{Q}_8$. Partially summarizing:

Proposition 8.4. *The tame-wild principle for* $(G, \phi_{12a}, \phi_{12b})$ *fails for the groups* $G = M_{12}, M_{11}, M_{10}$ *and* 12T112. *For* $G =$ 12T112 *the pair of number fields* (K_{12a}, K_{12b}) *contradicts the divisibility statement* (8-1), *but no pair with* $G =$ 12T112 *contradicts the numerical statement* (8-2).

The group-theoretic argument for 12T112 does not apply to the three M_n and we expect that there exist pairs (K_{12a}, K_{12b}) for them contradicting not only (8-1) but also (8-2). In general, a closer analysis of the exact range of applicability of the tame-wild principle would be interesting.

Acknowledgments

We thank Nigel Boston and Griff Elder for helpful conversations.

References

[Amano 1971] S. Amano, "Eisenstein equations of degree p in a $\mathfrak{p}$-adic field", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **18** (1971), 1–21. MR 46 #7201 Zbl 0231.12019

[Cameron et al. 2002] P. J. Cameron, M. Giudici, G. A. Jones, W. M. Kantor, M. H. Klin, D. Marušič, and L. A. Nowitz, "Transitive permutation groups without semiregular subgroups", *J. London Math. Soc.* (2) **66**:2 (2002), 325–333. MR 2003f:20001 Zbl 1015.20001

[Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups*, Oxford University Press, Eynsham, 1985. MR 88g:20025 Zbl 0568.20001

[Conway et al. 1998] J. H. Conway, A. Hulpke, and J. McKay, "On transitive permutation groups", *LMS J. Comput. Math.* **1** (1998), 1–8. MR 99g:20011 Zbl 0920.20001

[Fontaine 1971] J.-M. Fontaine, "Groupes de ramification et représentations d'Artin", *Ann. Sci. École Norm. Sup.* (4) **4** (1971), 337–392. MR 44 #6648 Zbl 0232.12006

[Ford and Pohst 1992] D. Ford and M. Pohst, "The totally real A_5 extension of degree 6 with minimum discriminant", *Experiment. Math.* **1**:3 (1992), 231–235. MR 93k:11108 Zbl 0773.11067

[Ford et al. 1998] D. Ford, M. Pohst, M. Daberkow, and N. Haddad, "The S_5 extensions of degree 6 with minimum discriminant", *Experiment. Math.* **7**:2 (1998), 121–124. MR 2000c:11182 Zbl 0921.11052

[Jensen et al. 2002] C. U. Jensen, A. Ledet, and N. Yui, *Generic polynomials: Constructive aspects of the inverse Galois problem*, Mathematical Sciences Research Institute Publications **45**, Cambridge University Press, 2002. MR 2004d:12007 Zbl 1042.12001

[Jones 2013] J. W. Jones, "Minimal solvable nonic fields", *LMS J. Comput. Math.* **16** (2013), 130–138. MR 3065803

[Jones and Roberts 2006] J. W. Jones and D. P. Roberts, "A database of local fields", *J. Symbolic Comput.* **41**:1 (2006), 80–97. Database at http://math.asu.edu/~jj/localfields. MR 2006k:11230 Zbl 1140.11350

[Jones and Roberts 2008] J. W. Jones and D. P. Roberts, "Octic 2-adic fields", *J. Number Theory* **128**:6 (2008), 1410–1429. MR 2009d:11163 Zbl 1140.11056

[Jones and Roberts $\geq$ 2014] J. W. Jones and D. P. Roberts, "A database of number fields", submitted. Database at http://hobbes.la.asu.edu/NFDB.

[Jones and Wallington 2012] J. W. Jones and R. Wallington, "Number fields with solvable Galois groups and small Galois root discriminants", *Math. Comp.* **81**:277 (2012), 555–567. MR 2012e:11182 Zbl 06006889

[Kontorovitch 1939] P. Kontorovitch, "Sur la représentation d'un groupe fini sous la forme d'une somme directe de sous-groupes, I", *Rec. Math. [Mat. Sbornik] N.S.* **5 (47)**:2 (1939), 289–296. In Russian. MR 2,3e Zbl 0022.31304

[Kontorovitch 1940] P. Kontorovitch, "Sur la représentation d'un groupe fini sous la forme d'une somme directe de sous-groupes, II", *Rec. Math. [Mat. Sbornik] N.S.* **7 (49)**:1 (1940), 27–33. In Russian. MR 2,4a Zbl 0023.01205

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

[Malle 2000] G. Malle, "Multi-parameter polynomials with given Galois group: Algorithmic methods in Galois theory", *J. Symbolic Comput.* **30**:6 (2000), 717–731. MR 2002a:12007 Zbl 0967.12005

[Maus 1965] E. Maus, *Existenz* $\mathfrak{p}$*-adischer Zahlkörper zu vorgegebenem Verzweigungsverhalten*, Ph.D. thesis, Hamburg, 1965.

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, New York, 1979. MR 82e:12016 Zbl 0423.12016

[Suzuki 1950] M. Suzuki, "On the finite group with a complete partition", *J. Math. Soc. Japan* **2** (1950), 165–185. MR 13,907b Zbl 0040.00703

Communicated by John Henry Coates
Received 2012-12-13 Revised 2013-09-11 Accepted 2013-10-21

jj@asu.edu *School of Mathematical and Statistical Sciences, Arizona State University, P.O. Box 871804, Tempe, AZ 85287-1804, United States*

roberts@morris.umn.edu *Division of Science and Mathematics, University of Minnesota - Morris, Morris, MN 56267, United States*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.647

Linear forms in logarithms and integral points on higher-dimensional varieties

Aaron Levin

We apply inequalities from the theory of linear forms in logarithms to deduce effective results on S-integral points on certain higher-dimensional varieties when the cardinality of S is sufficiently small. These results may be viewed as a higher-dimensional version of an effective result of Bilu on integral points on curves. In particular, we prove a completely explicit result for integral points on certain affine subsets of the projective plane. As an application, we generalize an effective result of Vojta on the three-variable unit equation by giving an effective solution of the polynomial unit equation $f(u, v) = w$, where u, v, and w are S-units, $|S| \leq 3$, and f is a polynomial satisfying certain conditions (which are generically satisfied). Finally, we compare our results to a higher-dimensional version of Runge's method, which has some characteristics in common with the results here.

1. Introduction

The problem of proving effective results in Diophantine questions is one of the most pervasive and basic problems in number theory. Already in the case of curves, the fundamental finiteness theorems for integral points and rational points (Siegel's theorem and Faltings' theorem, respectively) are not known in an effective way, that is, in general there is no known algorithm to provably compute the finite sets in the conclusion of either theorem. In certain special cases, however, effective techniques have been developed for computing integral or rational points on curves. The most general and widely used effective methods for integral points on curves come from the theory of linear forms in logarithms, developed originally by Baker [1975]. In higher dimensions, effective techniques have not received much attention. A natural first step towards proving higher-dimensional effective results consists of taking the known effective techniques for curves and applying them, to the extent possible, to the higher-dimensional situation. In [Levin 2008], some progress towards this goal was achieved by formulating a higher-dimensional version of an effective

Research partially supported by NSF grant DMS-1102563.
MSC2010: primary 11G35; secondary 11J86, 11D61.
Keywords: integral points, unit equation, linear forms in logarithms, Runge's method.

method of Runge for computing integral points on curves. In this article we will consider the theory of linear forms in logarithms and applications to integral points on higher-dimensional varieties.

One of the few directions in which progress has been made on the study of integral points on higher-dimensional varieties involves varieties which, roughly speaking, have many components at infinity (see, for example, [Autissier 2009; 2011; Corvaja et al. 2009; Corvaja and Zannier 2004; 2006; Levin 2008; 2009]). The results given here also fit into this framework. We prove the following effective result for integral points on higher-dimensional varieties.

Theorem 1. *Let X be a nonsingular projective variety defined over a number field k. Let $D_1, \dots, D_n$ be effective ample divisors on X defined over k. Let $D = \sum_{i=1}^n D_i$. Let $m \le n$ be a positive integer such that for all subsets $I \subset \{1, \dots, n\}$, $|I| = m$, the set $\bigcap_{i \in I} (\operatorname{Supp} D_i)(\bar{k})$ consists of finitely many points. Suppose that for each point $P \in (\operatorname{Supp} D)(\bar{k})$, there exists a nonconstant rational function $\phi \in k(X)$ satisfying $P \notin \operatorname{Supp} \phi$ and $\operatorname{Supp} \phi \subset \operatorname{Supp} D$. Let S be a set of places of k containing the archimedean places with*

$$(m-1)|S| < n.$$

Let R be a set of S-integral points on $X \setminus D$. Suppose that $X, D_1, \dots, D_n, D, R, S, k$ satisfy (*) *in Section 3.2. Then R is contained in an effectively computable proper closed subset Z of X.*

To make the meaning of "effective" precise, we have assumed in the theorem that one can compute certain natural quantities described in Section 3.2. An explicit description of the higher-dimensional part of Z is given in Theorem 14. If $X = C$ is a curve, then Theorem 1 (with $m = 1$) is easily seen to be equivalent to the following theorem of Bilu.

Theorem 2 [Bilu 1995]. *Let $C \subset \mathbb{A}^n$ be an affine curve defined over a number field k. Suppose that there exist two everywhere nonvanishing regular functions on C with multiplicatively independent images in $k(C)^*/k^*$. For any finite set of places S of k containing the archimedean places, the set $C(\mathbb{O}_{k,S})$ is finite and effectively computable.*

Thus, Theorem 1 may be viewed as a higher-dimensional generalization of Bilu's theorem. We note that, as mentioned in [Bilu 1995], when combined with finite covers and the Chevalley–Weil theorem, Theorem 2 appears to be responsible for all known "universally effective" results on integral points on curves (results valid for all number fields k and finite sets of places S).

As an easy consequence of Theorem 1, we obtain the following result for integral points on surfaces.

Corollary 3. *Let X be a nonsingular projective surface defined over a number field k. Let $D_1, \dots, D_n$ be ample effective divisors on X, defined over k, that generate a subgroup of $\operatorname{Pic}(X)$ of rank r and pairwise do not have any common components. Let $D = \sum_{i=1}^n D_i$. Suppose that the intersection of the supports of any $n-r$ of the divisors D_i is empty. Let S be a set of places of k containing the archimedean places with*

$$|S| < n.$$

Let R be a set of S-integral points on $X \setminus D$. Suppose that $X, D_1, \dots, D_n, D, R, S, k$ satisfy (*) *in Section 3.2. Then R is contained in an effectively computable proper closed subset Z of X.*

The requirement, in the above results, that the number of components at infinity be large relative to the cardinality of S appears prominently in Runge's method [Levin 2008] as well. We will compare our results with a higher-dimensional version of Runge's method in Section 5.

As an application of our result on surfaces, we prove an effective result on two-variable polynomials that take on S-unit values at S-unit arguments when $|S| \le 3$.

Corollary 4. *Let $f \in k[x, y]$ be a polynomial of degree $d > 0$ such that $f(0, 0) \neq 0$ and x^d and y^d have nonzero coefficients in f. Let S be a finite set of places of k containing the archimedean places with $|S| \le 3$. Then the set of solutions to*

$$f(u, v) = w, \quad u, v, w \in \mathbb{O}_{k,S}^*,$$

consists of a finite effectively computable set and a finite number of infinite families of solutions where one of u, v, or w is constant.

The infinite families of solutions are explicitly described in Corollary 21 in Section 4.

Taking $f(x, y)$ to be an appropriate affine linear polynomial, we find that Corollary 4 generalizes an effective result of Vojta on the three-variable S-unit equation with $|S| \le 3$.

Theorem 5 [Vojta 1983]. *Let k be a number field, S a finite set of places of k containing the archimedean places, and $a_1, a_2, a_3 \in k^*$. If $|S| \le 3$, then the set of solutions to the equation*

$$a_1u_1 + a_2u_2 + a_3u_3 = 1, \quad u_1, u_2, u_3 \in \mathbb{O}_{k,S}^*,$$

with

$$\sum_{i \neq j} a_i u_i \neq 0, \quad j = 1, 2, 3,$$

is finite and effectively computable.

We note that versions of Theorem 5 with $k=\mathbb{Q}$ were also proved by Mo and Tijdeman [1992] and Skinner [1990]. Ineffectively, versions of Corollary 4 and Theorem 5 can be proven without any assumption on the (finite) cardinality of S. For Theorem 5, this is a special case of a well-known result on unit equations, proved independently by Evertse [1984] and van der Poorten and Schlickewei [1982]. In the case of Corollary 4, this is an easy consequence of a result of Vojta [1987, Corollary 2.4.3] and the proof of Corollary 4. The ineffectivity here comes ultimately from usage of the Schmidt subspace theorem.

More generally, Vojta proved the following result for systems of unit equations.

Theorem 6 (Vojta). *Let m and n be positive integers with $n>m$. Let (a_{ij}) be an $m\times n$ matrix with elements in a number field k such that no $m+1$ distinct columns of the matrix have rank less than m, and such that no column is identically zero. Assume further that S is a finite set of places of k, containing the archimedean places, satisfying*

$$(n-m-2)|S|<n.$$

Then the set of solutions to the system of unit equations

$$a_{i1}u_1+\cdots+a_{in}u_n=0,\quad 1\le i\le m,\quad u_1,\dots,u_n\in\mathbb{O}_{k,S}^*,$$

can be effectively determined.

More precisely, viewing a solution in Theorem 6 as a point in $\mathbb{P}^{n-1}$, the set of solutions to a system of equations as in Theorem 6 lies in finitely many proper linear subspaces of $\mathbb{P}^{n-1}$, and these solutions may be explicitly described and parametrized. In work to appear, Bennett (personal communication) has improved the inequality on $|S|$ in Theorem 6 to $(n-m-1)|S|<2n$. In particular, Bennett's methods allow one to extend Theorem 5 to four-variable unit equations, that is, to effectively solve the unit equation

$$a_1u_1+a_2u_2+a_3u_3+a_4u_4=1,\quad u_1,u_2,u_3,u_4\in\mathbb{O}_{k,S}^*,$$

where $a_1,a_2,a_3,a_4\in k^*$ and $|S|\le 3$. It would be interesting to determine the extent to which Bennett's methods may be applied to gain a similar improvement to the results presented here.

In Section 7, we prove a completely explicit version of Corollary 3 when $X=\mathbb{P}^2$ is the projective plane.

Theorem 7. *Let k be a number field of degree δ and discriminant Δ. Let $C_1,\dots,C_n$ be distinct curves over k in $\mathbb{P}^2$ such that the intersection of any $n-1$ of the curves is empty. Let S be a set of places of k containing the archimedean places with $s=|S|<n$. Then the set of integral points $\left(\mathbb{P}^2\setminus\bigcup_{i=1}^n C_i\right)(\mathbb{O}_{k,S})$ is contained in an effectively computable proper Zariski closed subset Z of $\mathbb{P}^2$. Explicitly, let*

$d_i = \deg C_i$, $d = \max_i d_i$, $h = \max_i h(C_i)$, and $N = \max_{v\in S} N(v)$. Let C_i be defined by $f_i \in k[x, y, z]$, $i = 1, \ldots, n$. Let $T = \bigcup_{i\neq j}(C_i \cap C_j)(\bar{k})$, and for each point $P \in T$, let $I_P = \{i : P \notin C_i\}$. For $P \in T$, let

$$\Phi_P = \left\{ \frac{f_i^{d_j}}{f_j^{d_i}} : i, j \in I_P \right\}.$$

Then Z may be taken to consist of the union of the finite set of points

$$\big\{P \in X(k) :$$
$$h(P) < 2^{20s+4\delta+75} d^{6s+34} \delta^{5s+8\delta-3} s^{4s+2} N^{d^2} (\log^* N)^{2s} |\Delta|^{3/2} (\log^* |\Delta|)^{3\delta} (h+1)\big\}$$

and the Zariski closure Z' of the set

$$\bigcup_{P\in T} \bigcap_{\phi\in\Phi_P} \{Q \in X(\bar{k}) : \phi(Q) = \phi(P)\}.$$

Being more interested in the general shape of the explicit height bound in the theorem, we have made no effort here to obtain the best possible explicit bound coming from the proof of Theorem 7 (and indeed, carefully following the proof gives a superior, but more cumbersome, expression).

Finally, we give a brief sketch of the proof of Theorem 1. The proof is a generalization of the proofs of Bilu's and Vojta's results (Theorem 2 and Theorem 5). Let R be a set of S-integral points on $X \setminus D$, as in Theorem 1, and let $P \in R$. Let $T \subset X(\bar{k})$ be the finite set of points contained in the support of m or more divisors D_i. Using the assumption on the cardinality of S, the pigeonhole principle implies that for some point $Q \in T$ and $v \in S$, P is v-adically close to Q. Our hypotheses then provide us with a nonconstant rational function $\phi \in k(X)$ with zeros and poles only in $\operatorname{Supp} D \setminus \{Q\}$. Since $P \in R$, $\phi(P)$ is essentially an S-unit, and $\phi(P)$ is v-adically close to $\phi(Q)$. Now assuming that $\phi(P) \neq \phi(Q)$ (this is where a higher-dimensional exceptional set may appear), we apply a Baker-type inequality to conclude that $\phi(P)$, and hence P, must have height bounded by an explicit constant.

2. Notation and definitions

Let k be a number field and let S be a finite set of places of k containing the archimedean places. We use $\mathcal{O}_k$, $\mathcal{O}_{k,S}$, and $\mathcal{O}_{k,S}^*$ to denote the ring of integers of k, ring of S-integers of k, and group of S-units of k, respectively. Throughout, we let $\delta = [k : \mathbb{Q}]$ be the degree of k, Δ the (absolute) discriminant of k, R_k the regulator of k, and R_S the S-regulator.

Recall that we have a canonical set M_k of places (or absolute values) of k consisting of one place for each prime ideal $\mathfrak{p}$ of $\mathcal{O}_k$, one place for each real embedding $\sigma : k \to \mathbb{R}$, and one place for each pair of conjugate embeddings $\sigma, \overline{\sigma} : k \to \mathbb{C}$. For $v \in M_k$, we define

$$N(v) = \begin{cases} 2 & \text{if } v \text{ is archimedean,} \\ N(\mathfrak{p}) & \text{if } v \text{ corresponds to the prime } \mathfrak{p}, \end{cases}$$

where $N(\mathfrak{p}) = |\mathbb{O}_k/\mathfrak{p}|$ is the norm of $\mathfrak{p}$. We normalize our absolute values so that $|p|_v = 1/p$ if v corresponds to $\mathfrak{p}$ and $\mathfrak{p}$ lies above a rational prime p, and $|x|_v = |\sigma(x)|$ if v corresponds to an embedding σ. For $v \in M_k$, let k_v denote the completion of k with respect to v. We set

$$\|x\|_v = |x|_v^{[k_v:\mathbb{Q}_v]/[k:\mathbb{Q}]}.$$

A fundamental equation is the product formula

$$\prod_{v \in M_k} \|x\|_v = 1,$$

which holds for all $x \in k^*$.

For x a positive real number we let

$$\log^* x = \max\{\log x, 1\},$$
$$\epsilon_v(x) = \begin{cases} x & \text{if } v \text{ is archimedean,} \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\epsilon_v'(x) = \epsilon_v(x)^{[k_v:\mathbb{Q}_v]/[k:\mathbb{Q}]}.$$

We note that

$$\prod_{v \in M_k} \epsilon_v'(x) = x.$$

In this notation, for $v \in M_k$ and $x, y \in k$ we have the inequalities

$$|x+y|_v \le \epsilon_v(2) \max\{|x|_v, |y|_v\},$$
$$\|x+y\|_v \le \epsilon_v'(2) \max\{\|x\|_v, \|y\|_v\}.$$

For $v \in M_k$ and $\alpha \in k$, we define the local height

$$h_v(\alpha) = \log \max\{\|\alpha\|_v, 1\}$$

and the height

$$h(\alpha) = \sum_{v \in M_k} h_v(\alpha).$$

We will frequently make the identification $\mathbb{P}^1(k) = k \cup \{\infty\}$. More generally, for a point $P = (x_0, \dots, x_n) \in \mathbb{P}^n(k)$, we have the absolute logarithmic height

$$h(P) = \sum_{v \in M_k} \log \max\{\|x_0\|_v, \dots, \|x_n\|_v\}.$$

Note that this is independent of the number field k and the choice of coordinates $x_0, \dots, x_n \in k$.

For a polynomial $f \in k[x_1, \dots, x_n]$ and $v \in M_k$, we let $|f|_v$ denote the maximum of the absolute values of the coefficients of f with respect to v. We define $\|f\|_v$ similarly. We define the height of a nonzero polynomial by

$$h(f) = \sum_{v \in M_k} \log \|f\|_v.$$

This is the same as the height of the point in projective space whose coordinates are given by the coefficients of f. If $\phi : \mathbb{P}^n \to \mathbb{P}^m$ is a rational map, where $\phi = (f_0, \dots, f_m)$ and $f_0, \dots, f_m \in k[x_0, \dots, x_n]$ are polynomials with no common factor, then we define

$$h(\phi) = \sum_{v \in M_k} \log \max_i \|f_i\|_v.$$

Let D be a hypersurface in $\mathbb{P}^n$ defined by a homogeneous polynomial $f \in k[x_0, \dots, x_n]$ of degree d. We define

$$h(D) = h(f).$$

For $v \in M_k$ and $P = (x_0, \dots, x_n) \in \mathbb{P}^n(k) \setminus D$, $x_0, \dots, x_n \in k$, we define the local height function

$$h_{D,v}(P) = \log \frac{\|f\|_v \max_i \|x_i\|_v^d}{\|f(P)\|_v}. \tag{1}$$

Note that this definition is independent of the choice of the defining polynomial f and the choice of the coordinates for P. We let $h_D(P) = (\deg D)h(P)$. By the product formula, if $P \in \mathbb{P}^n(k) \setminus D$, then $\sum_{v \in M_k} h_{D,v}(P) = h_D(P)$.

If $P = (x_0, \dots, x_n)$, $Q = (y_0, \dots, y_n) \in \mathbb{P}^n(k)$, $x_i, y_i \in k$, $P \neq Q$, and $v \in M_k$, we define

$$h_{Q,v}(P) = \log \frac{\max_i \|x_i\|_v \max_i \|y_i\|_v}{\max_{i,j} \|x_i y_j - x_j y_i\|_v}.$$

Much more generally, one can associate a height to any closed subscheme of a projective variety. We give here a quick summary of the relevant properties of such heights and refer the reader to [Silverman 1987] for the general theory and details.

Let Y be a closed subscheme of a projective variety X, both defined over k. For $v \in M_k$, one can associate a local height function $h_{Y,v} : X(k) \setminus Y \to \mathbb{R}$, well-defined up to $O(1)$, and a global height function h_Y, well-defined up to $O(1)$, which is a sum of appropriate local height functions. If $Y = D$ is an effective (Cartier) divisor (which we will frequently identify with the associated closed subscheme), these height functions agree with the usual height functions associated to divisors. Local height functions satisfy the following properties: if Y and Z are two closed subschemes of X, defined over k, and $v \in M_k$, then up to $O(1)$,

$$\begin{aligned} h_{Y\cap Z,v} &= \min\{h_{Y,v}, h_{Z,v}\}, \\ h_{Y+Z,v} &= h_{Y,v} + h_{Z,v}, \\ h_{Y,v} &\le h_{Z,v} && \text{if } Y \subset Z, \\ h_{Y,v} &\le c h_{Z,v} && \text{if } \operatorname{Supp} Y \subset \operatorname{Supp} Z, \end{aligned}$$

for some constant $c > 0$, where $\operatorname{Supp} Y$ denotes the support of Y. If $\phi : W \to X$ is a morphism of projective varieties, then

$$h_{Y,v}(\phi(P)) = h_{\phi^*Y,v}(P) \quad \text{for all } P \in W(k) \setminus \phi^*Y.$$

Here, $Y \cap Z$, $Y + Z$, $Y \subset Z$, and ϕ^*Y are defined in terms of the associated ideal sheaves (see [Silverman 1987]). Global height functions satisfy similar properties (except the first property above, which becomes $h_{Y\cap Z} \le \min\{h_Y, h_Z\} + O(1)$).

Let D be a divisor on a nonsingular projective variety X. For a nonzero rational function $\phi \in \bar{k}(X)$, we let $\operatorname{div}(\phi)$ denote the divisor associated to ϕ. We let $\operatorname{Supp} D$ denote the support of D and $\operatorname{Supp}\phi = \operatorname{Supp}\operatorname{div}(\phi)$. Let

$$L(D) = \{\phi \in \bar{k}(X) : \operatorname{div}(\phi) + D \ge 0\}$$

and $h^0(D) = \dim H^0(X, \mathcal{O}(D)) = \dim L(D)$. If $h^0(nD) = 0$ for all $n > 0$, then we let $\kappa(D) = -\infty$. Otherwise, we define the dimension of D to be the integer $\kappa(D)$ such that there exist positive constants c_1 and c_2 with

$$c_1 n^{\kappa(D)} \le h^0(nD) \le c_2 n^{\kappa(D)}$$

for all sufficiently divisible $n > 0$. We define a divisor D on X to be big if $\kappa(D) = \dim X$.

Let D be an effective divisor on X and $h_D = \sum_{v\in M_k} h_{D,v}$ a height function associated to D. A set of points $R \subset X(k) \setminus D$ is called a set of S-integral points on $X \setminus D$ if there exist constants c_v, $v \in M_k$, such that $c_v = 0$ for all but finitely many v, and for all $v \notin S$,

$$h_{D,v}(P) \le c_v$$

for all $P \in R$. This is well-defined, independent of how we write $X \setminus D$ [Vojta 1987, Corollary 1.4.2, Theorem 1.4.11]. There are other essentially equivalent definitions of integrality (see, for example, [ibid., Proposition 1.4.7]), but since our main tools involve heights, this will be the most natural definition for our purposes.

Let Z be a closed subset of $\mathbb{P}^n$ defined over k. Let S be a finite set of places of k containing the archimedean places. In this case there is a natural set of integral points on $\mathbb{P}^n \setminus Z$. We define $(\mathbb{P}^n \setminus Z)(\mathbb{O}_{k,S})$ to be the set of points $P \in \mathbb{P}^n(k)$ such that the Zariski closures of P and Z in $\mathbb{P}^n_{\mathbb{O}_k}$ do not meet over any $v \notin S$. Equivalently, if D is an effective divisor on $\mathbb{P}^n$, using the local height functions defined in (1) one easily finds that

$$\begin{aligned}(\mathbb{P}^n \setminus D)(\mathbb{O}_{k,S}) &= \{P \in \mathbb{P}^n(k) \setminus D : h_{D,v}(P) = 0, \forall v \in M_k \setminus S\}\\ &= \Big\{P \in \mathbb{P}^n(k) \setminus D : \sum_{v \in S} h_{D,v}(P) = (\deg D) h(P)\Big\}.\end{aligned}$$

3. General results

For the purpose of clarifying our later proofs, we first collect together various elementary facts about heights.

3.1. *Heights.* Throughout, we let X be a nonsingular projective variety defined over a number field k. We first recall the Northcott property for heights associated to ample divisors.

Lemma 8. *Let D be an ample divisor on X and $c \in \mathbb{R}$. Then the set of points $\{P \in X(k) : h_D(P) < c\}$ is finite.*

More generally, finiteness holds for points of $X(\bar{k})$ of bounded degree and bounded ample height. Every height is bounded by a multiple of an ample height [Vojta 1987, Proposition 1.2.9(f)].

Lemma 9. *Let A and D be divisors on X with A ample. Then there exists a positive integer N such that*

$$h_D(P) < N h_A(P) + O(1)$$

for all $P \in X(\bar{k})$.

The next two lemmas give relations between the height of a point and its image under a rational map.

Lemma 10. *Let $\phi \in k(X)$ and let $P_1, \dots, P_q \in X(k) \setminus \operatorname{Supp} \phi$. Let S be a finite set of places of k. Then*

$$\sum_{i=1}^{q} \sum_{v \in S} h_{P_i,v}(P) < \sum_{i=1}^{q} \sum_{v \in S} h_{\phi(P_i),v}(\phi(P)) + O(1)$$

for all $P \in X(k) \setminus \operatorname{Supp} \phi$ such that $\phi(P) \neq \phi(P_i)$, $i = 1, \dots, q$.

Proof. For an appropriate blow-up $\pi : \tilde{X} \to X$, where π is an isomorphism on $\pi^{-1}(X \setminus \operatorname{Supp} \phi)$, ϕ extends to a morphism $\tilde{\phi} : \tilde{X} \to \mathbb{P}^1$ such that $\tilde{\phi} = \phi \circ \pi$

on $\pi^{-1}(X \setminus \operatorname{Supp}\phi)$. For a point $P \in X(k) \setminus \operatorname{Supp}\phi$, we let $\tilde{P} = \pi^{-1}(P)$. Let $P \in X(k) \setminus \operatorname{Supp}\phi$ be such that $\phi(P) \neq \phi(P_i)$, $i = 1, \ldots, q$. By functoriality of heights,

$$\sum_{i=1}^{q} \sum_{v \in S} h_{\tilde{\phi}^*\phi(P_i),v}(\tilde{P}) = \sum_{i=1}^{q} \sum_{v \in S} h_{\phi(P_i),v}(\tilde{\phi}(\tilde{P})) + O(1).$$

Since $\tilde{P}_i$ is in the support of $\tilde{\phi}^*\phi(P_i)$, we have

$$\sum_{i=1}^{q} \sum_{v \in S} h_{\tilde{P}_i,v}(\tilde{P}) < \sum_{i=1}^{q} \sum_{v \in S} h_{\tilde{\phi}^*\phi(P_i),v}(\tilde{P}) + O(1).$$

Now the lemma follows from the above two equations, noting that $\tilde{\phi}(\tilde{P}) = \phi(P)$ and by functoriality, $h_{\tilde{P}_i,v}(\tilde{P}) = h_{\pi^*P_i,v}(\tilde{P}) = h_{P_i,v}(P) + O(1)$. □

Lemma 11. *Let D be an effective divisor on X and let $\phi \in k(X)$ be a rational function with every pole contained in* Supp D. *Then for some constant $c > 0$,*

$$h(\phi(P)) < ch_D(P) + O(1)$$

for all $P \in X(\bar{k}) \setminus \operatorname{Supp}\phi$.

Proof. We use the same notation as in the proof of Lemma 10. Let $P \in X(\bar{k}) \setminus \operatorname{Supp}\phi$. By functoriality,

$$h(\phi(P)) = h(\tilde{\phi}(\tilde{P})) = h_{\tilde{\phi}^*\infty}(\tilde{P}) + O(1).$$

Since $\operatorname{Supp}\tilde{\phi}^*\infty \subset \operatorname{Supp}\pi^*D$, there exists a constant $c > 0$ such that

$$h_{\tilde{\phi}^*\infty}(\tilde{P}) < ch_{\pi^*D}(\tilde{P}) + O(1).$$

By functoriality again, $h_{\pi^*D}(\tilde{P}) = h_D(P) + O(1)$ and the result follows. □

The next lemma is crucial in our later proofs.

Lemma 12. *Let $E_1, \ldots, E_m$ be effective divisors on X, defined over k, such that $\bigcap_{i=1}^{m} E_i$ consists of a finite number of points, all defined over k. Let $v \in M_k$. Then there exists a positive integer N such that*

$$\min_i h_{E_i,v}(P) \leq N \sum_{Q \in \bigcap_{i=1}^{m} E_i(k)} h_{Q,v}(P) + O(1),$$

for all $P \in X(k) \setminus \bigcup_i E_i$.

Proof. If $\bigcap_{j=1}^{m} \operatorname{Supp} E_i = \varnothing$, then in fact

$$\min\{h_{D_{i_1},v}(P), \ldots, h_{D_{i_m},v}(P)\} \leq c$$

for some constant c. This is well known and follows, for instance, from formal properties of heights since in this case $\min\{h_{D_{i_1},v}, \dots, h_{D_{i_m},v}\}$ is a local height associated to the trivial divisor.

Otherwise, let N be a positive integer such that $\bigcap_{i=1}^m E_i \subset N \sum_{Q \in \bigcap_{i=1}^m E_i(k)} Q$. Then by properties of heights,

$$\min_i h_{E_i,v}(P) = h_{\bigcap_{i=1}^m E_i,v}(P) + O(1) \le N \sum_{Q \in \bigcap_{i=1}^m E_i(k)} h_{Q,v}(P) + O(1)$$

for all $P \in X(k) \setminus \bigcup_i E_i$. □

Finally, we record two basic facts about integral points that follow from the definitions and basic properties of height functions (see also [Vojta 1987, Lemma 1.4.6]).

Lemma 13. *Let $D_1, \dots, D_n$ be effective divisors on X, defined over k, and let $D = \sum_{i=1}^n D_i$. Let S be a finite set of places of k containing the archimedean places and let R be a set of S-integral points on $X \setminus D$. Then*

$$\sum_{v \in S} h_{D_i,v}(P) = h_{D_i}(P) + O(1), \quad i = 1, \dots, n,$$

for all $P \in R$. If $\phi \in k(X)$ with $\operatorname{Supp} \phi \subset \operatorname{Supp} D$, *then there exists a finite set of places T of k such that*

$$\phi(P) \in \mathbb{O}_{k,T}^*$$

for all $P \in R$.

3.2. ***Results.*** Let X be a nonsingular projective variety defined over a number field k. Let $D_1, \dots, D_n$ be effective ample divisors on X, defined over k, and set $D = \sum_{i=1}^n D_i$. Let S be a finite set of places of k containing the archimedean places and R a set of S-integral points on $X \setminus D$. We need a hypothesis asserting that one can effectively compute the height relations of the last section. We say that $X, D_1, \dots, D_n, D, R, S, k$ satisfy (*) if there are height functions associated to $D_1, \dots, D_n, D$ and points of X such that:

(1) The finite set in Lemma 8 is effectively computable for D and any $c \in \mathbb{R}$.

(2) The positive integer N and $O(1)$ in Lemma 9 are effectively computable for D and $A = D_i$, $i = 1, \dots, n$.

(3) The $O(1)$ in Lemma 10 is effectively computable for S, any $\phi \in k(X)$ with $\operatorname{Supp} \phi \subset \operatorname{Supp} D$, and any set of points $\{P_1, \dots, P_q\} \subset X(k) \setminus \operatorname{Supp} \phi$.

(4) The $O(1)$ in Lemma 11 is effectively computable for D and any ϕ.

(5) The positive integer N and $O(1)$ in Lemma 12 are effectively computable for any $v \in S$ and subset $\{E_1, \dots, E_m\} \subset \{D_1, \dots, D_n\}$.

(6) The finite set T and $O(1)$ in Lemma 13 are effectively computable for R, $D_1, \dots, D_n$, D, and any ϕ.

(7) The above remain true upon replacing k by a finite extension of k and S by any finite set of places containing the set of places lying above places of S.

Additionally, we assume that we can compute in $\operatorname{Pic}(X)$ as follows:

(8) All of the relations between the images of $D_1, \dots, D_n$ in $\operatorname{Pic}(X)$ are effectively computable, and for any principal divisor E supported on $D_1, \dots, D_n$ one can effectively compute a rational function $\phi \in k(X)$ with $\operatorname{div}(\phi) = E$.

Examples of varieties where (*) is satisfied (for any reasonably defined R) include curves, projective space, and more generally projective subvarieties of $\mathbb{P}^N$ where the divisors D_i are hypersurface sections. We explicitly work out the case $X = \mathbb{P}^2$ in Section 7. For curves, key algorithms include the computation of Riemann–Roch spaces [Schmidt 1991] and relations amongst points in the Jacobian [Masser 1988].

The main result of this section is a slightly more explicit version of Theorem 1.

Theorem 14. *Let X be a nonsingular projective variety defined over a number field k. Let $D_1, \dots, D_n$ be effective ample divisors on X defined over k. Let $D = \sum_{i=1}^n D_i$. Let $m \le n$ be a positive integer such that for all subsets $I \subset \{1, \dots, n\}$, $|I| = m$, the set $\bigcap_{i \in I} (\operatorname{Supp} D_i)(\bar{k})$ consists of finitely many points. Suppose that for each point $P \in (\operatorname{Supp} D)(\bar{k})$, there exists a nonconstant rational function $\phi \in k(X)$ satisfying $P \notin \operatorname{Supp} \phi$ and $\operatorname{Supp} \phi \subset \operatorname{Supp} D$. Let S be a set of places of k containing the archimedean places with*

$$(m-1)|S| < n.$$

Let R be a set of S-integral points on $X \setminus D$. Suppose that $X, D_1, \dots, D_n, D, R, S, k$ satisfy (*) in Section 3.2. Then R is contained in an effectively computable proper closed subset Z of X. Explicitly, for a point $P \in X(\bar{k})$, let

$$\Phi_P = \{\phi \in k(X)^* : P \notin \operatorname{Supp} \phi, \operatorname{Supp} \phi \subset \operatorname{Supp} D\}$$

and let

$$T = \bigcup_{\substack{I \subset \{1, \dots, n\} \\ |I| = m}} \bigcap_{i \in I} (\operatorname{Supp} D_i)(\bar{k}).$$

Then we may take Z to consist of a finite effectively computable set of points together with the Zariski closure of the set

$$\bigcup_{P \in T} \bigcap_{\phi \in \Phi_P} \{Q \in X(\bar{k}) : \phi(Q) = \phi(P)\}.$$

Remark 15. As is typical, more generally one could replace ample with big in the theorem by modifying the theorem slightly (e.g., increasing the exceptional set Z to account for the base loci of certain divisors).

If $D_1, \ldots, D_{r+1}$ are nontrivial effective divisors on a variety X that generate a subgroup of $\mathrm{Pic}(X)$ of rank r and pairwise do not have any common components, then there exists a nonconstant rational function ϕ on X with all zeros and poles contained in the support of $\sum_{i=1}^{r+1} D_i$. Using this fact to construct appropriate rational functions ϕ in Theorem 14, we immediately obtain the following corollary.

Corollary 16. *Let X be a nonsingular projective variety defined over a number field k. Let $D_1, \ldots, D_n$ be ample effective divisors on X, defined over k, that generate a subgroup of* $\mathrm{Pic}(X)$ *of rank r and pairwise do not have any common components. Let $D = \sum_{i=1}^{n} D_i$. Let $m \leq n$ be a positive integer such that for all subsets $I \subset \{1, \ldots, n\}$, $|I| = m$, the set $\bigcap_{i \in I}(\mathrm{Supp}\, D_i)(\bar{k})$ consists of finitely many points. Suppose that the intersection of the supports of any $n - r$ of the divisors D_i is empty. Let S be a set of places of k containing the archimedean places with*

$$(m-1)|S| < n.$$

Let R be a set of S-integral points on $X \setminus D$. Suppose that $X, D_1, \ldots, D_n, D, R, S, k$ satisfy (*). *Then R is contained in an effectively computable proper closed subset Z of X.*

Of particular interest is the case where X is a surface.

Corollary 17. *Let X be a nonsingular projective surface over a number field k. Let $D_1, \ldots, D_n$ be ample effective divisors on X, defined over k, that generate a subgroup of* $\mathrm{Pic}(X)$ *of rank r and pairwise do not have any common components. Suppose that the intersection of the supports of any $n - r$ of the divisors D_i is empty. Let S be a set of places of k containing the archimedean places with*

$$|S| < n.$$

Let R be a set of S-integral points on $X \setminus D$. Suppose that $X, D_1, \ldots, D_n, D, R, S, k$ satisfy (*). *Then R is contained in an effectively computable proper closed subset Z of X. Let*

$$T = \bigcup_{i \neq j} (D_i \cap D_j)(\bar{k}),$$

and let Φ_P be as in Theorem 14. Then we may take Z to consist of a finite effectively computable set of points together with the Zariski closure of the set

$$\bigcup_{P \in T} \bigcap_{\phi \in \Phi_P} \{Q \in X(\bar{k}) : \phi(Q) = \phi(P)\}.$$

3.3. ***Proofs.*** The key tool in this section is the main theorem from the theory of linear forms in logarithms, which we now state in the language of heights (see Theorem 24 for a completely explicit version).

Theorem 18. *Let k be a number field and S a finite set of places of k containing the archimedean places. Let $v \in M_k$, $\alpha \in k^*$, and $\epsilon > 0$. Then there exists an effective constant C such that*

$$h_{\alpha,v}(x) \leq \epsilon h(x) + C$$

for all $x \in \mathbb{O}_{k,S}^*$, $x \neq \alpha$.

We note that with an ineffective constant C, the theorem follows easily from Roth's theorem. Before proving Theorem 14, we prove a result which can be regarded as a higher-dimensional version of Theorem 18.

Theorem 19. *Let X be a nonsingular projective variety defined over a number field k and let D be an effective divisor on X defined over k. Let $\phi \in k(X)$ be a nonconstant rational function with* $\operatorname{Supp}\phi \subset \operatorname{Supp} D$. *Let S be a finite set of places of k and R a set of S-integral points on $X \setminus D$. Suppose that X, D, R, S, k satisfy* (3), (4), (6), (7) *of Section 3.2. Let $P_1, \dots, P_q \in X(k) \setminus \operatorname{Supp}\phi$ and $\epsilon > 0$. Then*

$$\sum_{i=1}^{q} \sum_{v \in S} h_{P_i,v}(P) < \epsilon h_D(P) + O(1)$$

for all $P \in R \setminus Z$, where Z is the proper closed subset of X defined as the Zariski closure of the set

$$\{P \in X(\bar{k}) : \phi(P) = \phi(P_i) \text{ for some } i \in \{1, \dots, q\}\}.$$

Here, as well as elsewhere, the implicit constant in the $O(1)$ is an effective constant.

Proof. By Lemma 13, since R is a set of S-integral points on $X \setminus D$, without loss of generality, after enlarging S we can assume that $\phi(P) \in \mathbb{O}_{k,S}^*$ for all $P \in R$. Then by Theorem 18,

$$\sum_{i=1}^{q} \sum_{v \in S} h_{\phi(P_i),v}(\phi(P)) < \epsilon h(\phi(P)) + O(1)$$

for all $P \in R \setminus Z$. By Lemma 10,

$$\sum_{i=1}^{q} \sum_{v \in S} h_{P_i,v}(P) < \sum_{i=1}^{q} \sum_{v \in S} h_{\phi(P_i),v}(\phi(P)) + O(1)$$

for all $P \in X(k) \setminus (Z \cup \operatorname{Supp}\phi)$. By Lemma 11,

$$\epsilon h(\phi(P)) < \epsilon c h_D(P) + O(1)$$

for some positive constant c and all $P \in X(k) \setminus \operatorname{Supp} D$. Replacing ϵ by ϵ/c and combining the above inequalities yields

$$\sum_{i=1}^{q}\sum_{v\in S} h_{P_i,v}(P) < \epsilon h_D(P) + O(1)$$

for all $P \in R \setminus Z$. □

We now prove Theorem 14.

Proof of Theorem 14. By Lemma 13, since R is a set of S-integral points on $X \setminus D$, we have

$$\sum_{v\in S} h_{D_i,v}(P) = h_{D_i}(P) + O(1), \quad i = 1, \dots, n,$$

for all $P \in R$. Let $P \in R$. Then for each i, there exists a place $v \in S$ such that $h_{D_i,v}(P) \geq (1/|S|)h_{D_i}(P) + O(1)$. Since $(m-1)|S| < n$, there exists a place $v \in S$ and distinct elements $i_1, i_2, \dots, i_m \in \{1, \dots, n\}$ such that

$$\min\{h_{D_{i_1},v}(P), \dots, h_{D_{i_m},v}(P)\} \geq \frac{1}{|S|}\min_j h_{D_{i_j}}(P) + O(1).$$

By Lemma 9, there exists a positive integer N such that

$$h_D(P) \leq N h_{D_i}(P) + O(1)$$

for all i and all $P \in X(\bar{k})$. So for $P \in R$,

$$\min\{h_{D_{i_1},v}(P), \dots, h_{D_{i_m},v}(P)\} \geq \frac{1}{N|S|} h_D(P) + O(1).$$

The theorem is then a consequence of the following lemma.

Lemma 20. *Let m, X, $D_1, \dots, D_n$, R, S, k be as in the hypotheses of Theorem 14. Let $\epsilon > 0$, $v \in S$, and let $i_1, \dots, i_m \in \{1, \dots, n\}$ be distinct integers. Then the set of points*

$$\{P \in R : \min\{h_{D_{i_1},v}(P), \dots, h_{D_{i_m},v}(P)\} > \epsilon h_D(P)\}$$

is contained in an effectively computable proper closed subset Z of X. For $P \in X(\bar{k})$, let Φ_P be the set from Theorem 14 and let

$$T = \bigcap_{j=1}^{m} (\operatorname{Supp} D_{i_j})(\bar{k}).$$

Then we may take Z to consist of a finite effectively computable set of points together with the Zariski closure of the set

$$\bigcup_{P\in T}\bigcap_{\phi\in\Phi_P}\{Q \in X(\bar{k}) : \phi(Q) = \phi(P)\}.$$

Proof. If L is a finite extension of k and w is a place of L lying above v, then we can define a local height function

$$h_{D_i,w}(P) = \frac{[L_w : k_v]}{[L : k]} h_{D_i,v}(P)$$

for all $P \in X(k) \setminus D_i$. It follows that without loss of generality, after replacing k by a finite extension of k and v by a place lying above v, we may assume that every point $P \in X(\bar{k})$ in the intersection $\bigcap_{j=1}^m \operatorname{Supp} D_{i_j}$ is defined over k (note that by hypothesis this intersection consists of a finite number of points).

If $\bigcap_{j=1}^m \operatorname{Supp} D_{i_j} = \varnothing$, then by Lemma 12,

$$\min\{h_{D_{i_1},v}(P), \ldots, h_{D_{i_m},v}(P)\} \le C$$

for some effective constant C. In this case, the lemma follows immediately from the fact that since D is ample, the set of points $\{P \in X(k) : h_D(P) < C/\epsilon\}$ is finite.

Suppose now that $\bigcap_{j=1}^m \operatorname{Supp} D_{i_j} \neq \varnothing$, in which case it consists of a finite number q of points. By Lemma 12, there exists a positive integer N such that

$$\min_j h_{D_{i_j},v}(P) \le N \sum_{Q \in \bigcap_{j=1}^m D_{i_j}(k)} h_{Q,v}(P) + O(1)$$

for all $P \in X(k) \setminus \bigcup_j D_{i_j}$.

Let $Q \in \bigcap_{j=1}^m D_{i_j}(k)$. Note that Φ_Q is a monoid under multiplication, generated by k^* and finitely many rational functions in $k(X)^*$. Let $\epsilon > 0$. Since R is a set of S-integral points on $X \setminus D$, applying Theorem 19 multiple times yields the inequality

$$h_{Q,v}(P) < \frac{\epsilon}{2Nq} h_D(P) + O(1)$$

for all $P \in R \setminus Z_Q$, where Z_Q is the Zariski closure of the set

$$\bigcap_{\phi \in \Phi_Q} \{P \in X(\bar{k}) : \phi(P) = \phi(Q)\}.$$

Summing over all points in $\bigcap_{j=1}^m D_{i_j}(k)$, we obtain

$$\min_j h_{D_{i_j},v}(P) \le N \sum_{Q \in \bigcap_{j=1}^m D_{i_j}(k)} h_{Q,v}(P) + O(1) < \frac{\epsilon}{2} h_D(P) + C$$

for all $P \in R \setminus Z$, where

$$Z = \bigcup_{Q \in \bigcap_{j=1}^m D_{i_j}(k)} Z_Q$$

and C is an effectively computable constant. So if $P \in R \setminus Z$ satisfies

$$\min_j h_{D_{i_j},v}(P) > \epsilon h_D(P),$$

then $h_D(P) < \frac{2}{\epsilon} C$. It follows that

$$\left\{P \in R : \min_j h_{D_{i_j},v}(P) > \epsilon h_D(P)\right\} \subset Z \cup \left\{P \in X(k) : h_D(P) < \frac{2}{\epsilon} C\right\},$$

where Z is a proper Zariski closed subset of X and the last set on the right is finite. □

4. An application to polynomial unit equations

We prove a complete version of Corollary 4 from the introduction.

Corollary 21. *Let $f \in k[x, y]$ be a polynomial of degree d such that $f(0,0) = c_0 \neq 0$ and x^d and y^d have nonzero coefficients c_x and c_y in f, respectively. Let S be a set of places of k containing the archimedean places with $|S| \leq 3$. Then the set of solutions to*

$$f(u, v) = w, \quad u, v, w \in \mathcal{O}_{k,S}^*,$$

consists of a finite effectively computable set and a finite number of infinite families of solutions where one of u, v, or w is constant. Let

$$T_1 = \{a \in \mathcal{O}_{k,S}^* : (x - a) \mid (f(x, y) - c_y y^d), c_y \in \mathcal{O}_{k,S}^*\},$$
$$T_2 = \{a \in \mathcal{O}_{k,S}^* : (y - a) \mid (f(x, y) - c_x x^d), c_x \in \mathcal{O}_{k,S}^*\},$$
$$T_3 = \{a \in \mathcal{O}_{k,S}^* : (y - ax) \mid (f(x, y) - c_0), c_0 \in \mathcal{O}_{k,S}^*\}.$$

Then the infinite families of solutions are

$$(u, v, w) = (a, t, c_y t^d), \quad t \in \mathcal{O}_{k,S}^*, \quad \text{for each } a \in T_1,$$
$$(u, v, w) = (t, a, c_x t^d), \quad t \in \mathcal{O}_{k,S}^*, \quad \text{for each } a \in T_2,$$
$$(u, v, w) = (t, at, c_0), \quad t \in \mathcal{O}_{k,S}^*, \quad \text{for each } a \in T_3.$$

Proof. It will be convenient to work with the homogenized polynomial

$$F(x, y, z) = z^d f(x/z, y/z).$$

Consider $\mathbb{P}^2$ with homogeneous coordinates (x, y, z) and let D_1, D_2, D_3, D_4 be the curves defined by $x = 0$, $y = 0$, $z = 0$, and $F(x, y, z) = 0$, respectively. Let $D = \sum_{i=1}^4 D_i$. Let

$$R = \{(u, v, 1) \in \mathbb{P}^2(k) : u, v, f(u, v) \in \mathcal{O}_{k,S}^*\}.$$

Then $R \subset (\mathbb{P}^2 \setminus D)(\mathcal{O}_{k,S})$. Let $\{i, j, k\} = \{1, 2, 3\}$ and $P \in (D_i \cap D_4)(\bar{k})$. Then, using the notation of Corollary 17, the Zariski closure of $\bigcap_{\phi \in \Phi_P} \{Q \in X(\bar{k}) : \phi(Q) = \phi(P)\}$ is a line through P and the unique point of $D_j \cap D_k$. Now let $P_1 = (1, 0, 0)$,

$P_2 = (0, 1, 0)$, and $P_3 = (0, 0, 1)$, so that $\{P_i\} = \bigcap_{j \in \{1,2,3\} \setminus \{i\}} D_j(\bar{k})$. Let Z_i be the Zariski closure of

$$\bigcap_{\phi \in \Phi_{P_i}} \{Q \in X(\bar{k}) : \phi(Q) = \phi(P_i)\}$$

for $i = 1, 2, 3$. Since $(F(x, y, z)/x^d)(P_1) = c_x$, $(F(x, y, z)/y^d)(P_2) = c_y$, and $(F(x, y, z)/z^d)(P_3) = c_0$, it follows that we have the equations

$$Z_1 : F(x, y, z) = c_x x^d, \quad Z_2 : F(x, y, z) = c_y y^d, \quad Z_3 : F(x, y, z) = c_0 z^d.$$

Let Z be the closed subset of $\mathbb{P}^2$ consisting of all lines connecting points of $(D_i \cap D_j)(\bar{k})$ with points of $(D_k \cap D_l)(\bar{k})$, where $\{i, j, k, l\} = \{1, 2, 3, 4\}$, together with the closed subsets Z_1, Z_2, and Z_3. Then it follows from Corollary 17 that $R \setminus Z$ consists of a finite effectively computable set of points (in fact, an explicit height bound for points in this set follows from Theorem 7).

Now let C be a geometrically irreducible curve in Z. If C is not defined over k and C' is any nontrivial conjugate of C over k, then $C(k) \subset (C \cap C')(\bar{k})$, a finite effectively computable set. In particular, $R \cap C$ is finite and effectively computable. Assume now that C is defined over k. Then $R \cap C$ is a set of integral points on $C \setminus (C \cap D)$. Consider the rational functions on C given by $\phi_1 = (x/z)|_C$ and $\phi_2 = (y/z)|_C$. The functions ϕ_1 and ϕ_2 have zeros and poles only in $C \cap D$. If ϕ_1 and ϕ_2 are multiplicatively independent modulo k^*, then Bilu's Theorem 2 implies that $R \cap C$ is finite and effectively computable. Suppose now that this is not the case. Then this easily implies that C is given by an equation $x^m y^{n-m} = az^n$, $x^m z^{n-m} = ay^n$, or $y^m z^{n-m} = ax^n$ for some nonnegative integers m and n and $a \in k^*$. Suppose first that $n \geq 2$. Then C is a component of Z_1, Z_2, or Z_3. Suppose that, say, C is given by $x^m y^{n-m} = az^n$ and is a component of Z_1. Then $F(x, y, z) = c_x x^d + (x^m y^{n-m} - az^n) g(x, y, z)$ for some homogeneous polynomial $g(x, y, z) \in k[x, y, z]$. Since C is geometrically irreducible and $n \geq 2$, we must have $0 < m < n$. But from the form of $F(x, y, z)$ we then see that y^d cannot have a nonzero coefficient in $F(x, y, z)$, contradicting our assumptions. The other possible cases are similar and we conclude that $n = 1$. So C is defined by a linear form $x - az$, $y - az$, or $y - ax$, for some $a \in k^*$.

Suppose that C is defined by $x - az = 0$. If $R \cap C \neq \varnothing$ then $a \in \mathbb{O}_{k,S}^*$, which we now assume. Since y^d must have a nonzero coefficient in $F(x, y, z)$, it follows that C cannot be an irreducible component of Z_1 or Z_3. If C is an irreducible component of Z_2, then $f(x, y) = c_y y^d + (x - a) g(x, y)$ for some polynomial $g(x, y) \in k[x, y]$. If C connects a point of $D_i \cap D_j$ with a point of $D_k \cap D_l$, where $\{i, j, k, l\} = \{1, 2, 3, 4\}$, then it must be that C connects the unique point of $D_1 \cap D_3$ with a point of $D_2 \cap D_4$. If C intersects D in more than two points over $\bar{k}$, then it follows easily again from Theorem 2 that $R \cap C$ is finite and effectively computable. So suppose that

$|(C \cap D)(\bar{k})| \le 2$, in which case $|(C \cap D)(\bar{k})| = 2$. Then the fact that $C \cap D_4$ consists of a single point contained in D_2 implies that $f(x, y) = c_y y^d + (x-a)g(x, y)$ for some polynomial $g(x, y) \in k[x, y]$. So in any case, $f(x, y) = c_y y^d + (x-a)g(x, y)$ for some polynomial $g(x, y) \in k[x, y]$. Now $C \cap R \neq \varnothing$ implies that $c_y \in \mathbb{O}^*_{k,S}$, and in this case we find that $C \cap R = \{(a, t, 1) : t \in \mathbb{O}^*_{k,S}\}$, leading to the infinite family of solutions $(u, v, w) = (a, t, c_y t^d)$, where $t \in \mathbb{O}^*_{k,S}$.

The cases where C is defined by $y - az$ or $y - ax$ are similar, and we are led to the classification of the infinite families in the theorem. □

5. Comparison with Runge's method

An old method of Runge yields effective finiteness for the set of integral points on certain curves. In its most basic form, Runge proved:

Theorem 22 [Runge 1887]. *Let $f \in \mathbb{Q}[x, y]$ be an absolutely irreducible polynomial of total degree n. Let f_0 denote the leading form of f, that is, the sum of the terms of total degree n in f. Suppose that f_0 factors as $f_0 = g_0 h_0$, where $g_0, h_0 \in \mathbb{Q}[x, y]$ are nonconstant relatively prime polynomials. Then the set of solutions to*

$$f(x, y) = 0, \quad x, y \in \mathbb{Z},$$

is finite and effectively computable.

We will state a general higher-dimensional version of Runge's method from [Levin 2008] (see [Bombieri 1983] for earlier work on curves). Before stating a higher-dimensional version, we give some definitions which allow for varying sets of places and number fields. It will be more convenient here to use a definition of integrality involving regular functions. Let V be a variety (not necessarily projective or affine) defined over a number field k. Let s be a positive integer. We call a set $R \subset V(\bar{k})$ a set of s-integral points on V if for every point $P \in R$ there exists a set of places S_P of $k(P)$, containing the archimedean places of $k(P)$, such that $|S_P| \le s$ and for every regular function $\phi \in \bar{k}(V)$ on V there exists a nonzero constant $c_\phi \in k^*$, independent of P, such that $|c_\phi \phi(P)|_v \le 1$ for all places v of $k(P)$ not in S_P (extending each place v of $k(P)$ to $\bar{k}$ in some fixed way). With these definitions, we have a higher-dimensional version of Runge's theorem:

Theorem 23. *Let X be a nonsingular projective variety defined over a number field k. Let $D = \sum_{i=1}^r D_i$ be a divisor on X, with $D_1, \ldots, D_r$ effective divisors defined over k. Suppose that the intersection of any $m + 1$ of the supports of the divisors D_i is empty. Let s be a positive integer satisfying*

$$ms < r.$$

Let R be a set of s-integral points on $X \setminus D$. Suppose that for every regular function

$\phi \in \bar{k}(X)$ on $X \setminus D$, the constant c_ϕ in the definition of s-integral is effectively computable with respect to R. Suppose also that one can effectively compute a basis of $L(nD_i)$ for all $n > 0$ and all i. Then the following statements hold.

(a) *If $\kappa(D_i) > 0$ for all i, then R is contained in an effectively computable proper Zariski closed subset $Z \subset X$.*

(b) *If D_i is big for all i, then there exists an effectively computable proper Zariski closed subset $Z \subset X$, independent of R, such that the set $R \setminus Z$ is finite (and effectively computable).*

(c) *If D_i is ample for all i, then R is finite and effectively computable.*

We now briefly discuss some of the advantages and disadvantages of the higher-dimensional Runge method as compared to our results here. To begin, in some respects the conditions on the divisors D_i in Theorem 23 are weaker than the conditions required in Theorem 14. The divisors in Theorem 23 are not required to be ample or big (though one still needs $\kappa(D_i) > 0$) and furthermore there is no linear equivalence condition present in Theorem 23. On the other hand, for the necessary rational functions ϕ to exist in Theorem 14, it is necessary that the subgroup of the Picard group generated by the divisors D_i not be too large (this condition is more explicitly present in Corollary 16). Another advantage of Theorem 23 is that the result is uniform in $|S|$, giving degeneracy of integral points even as S and k vary subject to an appropriate inequality. This is also, however, a limitation of Theorem 23, as many results are simply not true in this generality (e.g., the unit equation $u + v = 1$ likely has infinitely many solutions in rational S-units, $|S| \leq 3$, since for instance there are expected to be infinitely many Mersenne primes). It is not apparent from the statement of Theorem 23, but when Runge's method applies it also gives much smaller bounds than techniques coming from Baker's theorem.

As compared to Theorem 23, we note that the intersection condition on the divisors D_i in Theorem 14 is much weaker, especially on surfaces. For instance, the intersection condition on the divisors in Corollary 17 allows for highly degenerate configurations of the divisors D_i. Finally, we note that even in cases where the divisors D_i are in general position, the crucial inequality involving $|S|$ in Theorem 14 is superior to the inequality in Theorem 23. This is particularly notable in the case of surfaces, where the superior inequality on $|S|$ is crucial, for instance, in proving Corollary 4.

6. Effective inequalities

In preparation for the next section, we recall several needed inequalities and prove them here.

6.1. ***Linear forms in logarithms.*** The deepest effective result we need is from the theory of linear forms in logarithms. We give a statement in terms of local heights, based on an inequality of Bérczes, Evertse, and Győry [Bérczes et al. 2009].

Theorem 24. *Let k be a number field of degree δ and let G be a finitely generated multiplicative subgroup of k^* of rank $t > 0$. Let $\alpha \in k^*$ and $v \in M_k$. Let $0 < \epsilon < 1$. Then if $x \in G$, $x \neq \alpha$, we have*

$$h_{\alpha,v}(x) \leq \epsilon h(x) + c_1(\epsilon, k, G, v, \alpha) + \log 2,$$

where

$$c_1(\epsilon, k, G, v, \alpha) = 6.4 \frac{c_2(\delta, t)N(v)}{\epsilon \log N(v)} Q_G \max\{h(\alpha), 1\} \max\left\{\log \frac{c_2(\delta, t)N(v)}{\epsilon}, \log^* Q_G\right\},$$

$$c_2(\delta, t) = 36(16e\delta)^{3t+5}(\log^* \delta)^2.$$

Here, if G is a finitely generated multiplicative subgroup of $\overline{\mathbb{Q}}$ of rank $t > 0$, then we let Q_G be the minimum value of

$$h(u_1) \cdots h(u_t),$$

where $u_1, \ldots, u_t$ are generators for G modulo the roots of unity in G. We let $Q_S = Q_{\mathbb{O}^*_{k,S}}$.

Proof. Let $x \in G$, $x \neq \alpha$. Suppose first that

$$h_{\alpha,v}(x) \leq \epsilon h(x) + h_v(\alpha) + h_v(1/\alpha) + h_v(2).$$

Then, using that $h(\alpha) = h(1/\alpha)$, we have

$$h_{\alpha,v}(x) \leq \epsilon h(x) + 2h(\alpha) + \log 2 \leq \epsilon h(x) + c_1(\epsilon, k, G, v, \alpha) + \log 2,$$

as $2h(\alpha)$ is easily bounded by $c_1(\epsilon, k, G, v, \alpha)$. Suppose now that

$$h_{\alpha,v}(x) > \epsilon h(x) + h_v(\alpha) + h_v(1/\alpha) + h_v(2).$$

We have

$$\begin{aligned} h_{\alpha,v}(x) &= \log \frac{\max\{\|\alpha\|_v, 1\} \max\{\|x\|_v, 1\}}{\|x-\alpha\|_v} \\ &= \log \frac{\max\{\|\alpha\|_v, 1\} \max\left\{\left\|\frac{x}{\alpha}\right\|_v, \left\|\frac{1}{\alpha}\right\|_v\right\}}{\left\|\frac{x}{\alpha} - 1\right\|_v}. \end{aligned}$$

Now

$$\left\|\frac{x}{\alpha}\right\|_v = \left\|\left(\frac{x}{\alpha} - 1\right) + 1\right\|_v \leq \epsilon_v'(2) \max\left\{\left\|\frac{x}{\alpha} - 1\right\|_v, 1\right\}.$$

It follows that

$$h_{\alpha,v}(x) \le \log \frac{\epsilon'_v(2)\max\{\|\alpha\|_v, 1\}\max\left\{\left\|\frac{x}{\alpha}-1\right\|_v, \left\|\frac{1}{\alpha}\right\|_v, 1\right\}}{\left\|\frac{x}{\alpha}-1\right\|_v}.$$

If

$$\max\left\{\left\|\frac{x}{\alpha}-1\right\|_v, \left\|\frac{1}{\alpha}\right\|_v, 1\right\} = \left\|\frac{x}{\alpha}-1\right\|_v,$$

then $h_{\alpha,v}(x) \le \log \epsilon'_v(2)\max\{\|\alpha\|_v, 1\} \le h_v(\alpha) + h_v(2)$, contradicting our assumptions. Then we must have

$$\begin{aligned} h_{\alpha,v}(x) &\le \log \frac{\epsilon'_v(2)\max\{\|\alpha\|_v, 1\}\max\left\{\left\|\frac{1}{\alpha}\right\|_v, 1\right\}}{\left\|\frac{x}{\alpha}-1\right\|_v} \\ &\le h_v(2) + h_v(\alpha) + h_v(1/\alpha) - \log\left\|\frac{x}{\alpha}-1\right\|_v. \end{aligned}$$

So

$$\log\left\|\frac{x}{\alpha}-1\right\|_v < -\epsilon h(x).$$

By [Bérczes et al. 2009, Theorem 4.2], this implies that

$$h(x) \le c_1(\epsilon, k, G, v, \alpha).$$

Now we note that by Lemma 32, proved later in this section, for any $x \in k$, $x \ne \alpha$,

$$h_{\alpha,v}(x) \le \log 2 + \sum_{v \in M_k} h_{\alpha,v}(x) = \log 2 + h(x).$$

Thus,

$$h_{\alpha,v}(x) \le \epsilon h(x) + c_1(\epsilon, k, G, v, \alpha) + \log 2. \qquad \square$$

6.2. ***Hilbert's Nullstellensatz.*** We will need an effective version of Hilbert's Nullstellensatz. We use the following version, due to Masser and Wüstholz [1983].

Theorem 25 (effective Hilbert's Nullstellensatz). *Let k be a number field and let $p_1, \dots, p_m, q \in \mathcal{O}_k[x_1, \dots, x_n]$ be polynomials of degree at most $d \ge 1$ such that q vanishes at all common zeros of $p_1, \dots, p_m$ in $\mathbb{A}^n(\bar{k})$. Then there exists a positive integer $M \le (8d)^{2^n}$ and polynomials $a_1, \dots, a_m \in \mathcal{O}_k[x_1, \dots x_n]$ of degrees at most $(8d)^{2^n+1}$, such that*

$$aq^M = a_1 p_1 + \dots + a_m p_m$$

for some nonzero element $a \in \mathcal{O}_k$. Furthermore, if

$$h_\infty = \log \max_{\substack{v \in M_k \\ v \mid \infty}} \{|p_1|_v, \dots, |p_m|_v, |q|_v\},$$

then

$$\log \max_{\substack{v \in M_k \\ v \mid \infty}} \{|a_1|_v, \dots, |a_m|_v, |a|_v\} \le (8d)^{2^{n+1}-1}(h_\infty + 8d \log 8d).$$

Remark 26. Applying the theorem appropriately to $\mathbb{A}^n$, it's clear that the same result holds for homogeneous polynomials $p_1, \ldots, p_m, q \in \mathbb{O}_k[x_1, \ldots, x_n]$ such that q vanishes at all common zeros of $p_1, \ldots, p_m$ in $\mathbb{P}^{n-1}(\bar{k})$. Furthermore, in this case one can clearly choose $a_1, \ldots, a_m$ to be homogeneous polynomials with $\deg a_i = M \deg q - \deg p_i$.

6.3. ***Arithmetic Bézout.*** We will make use of the following arithmetic Bézout theorem for curves in $\mathbb{P}^2$, which is essentially a special case of a general arithmetic Bézout theorem of Philippon [1995].

Theorem 27. *Let C_1 and C_2 be distinct curves in $\mathbb{P}^2$ over $\overline{\mathbb{Q}}$. Then*

$$\sum_{P \in (C_1 \cap C_2)(\overline{\mathbb{Q}})} h(P) \leq (\deg C_1) h(C_2) + (\deg C_2) h(C_1) + 4(\deg C_1)(\deg C_2).$$

Proof. We will denote the height used by Philippon [1995] by h_{Ph}. By [ibid., Proposition 4],

$$\sum_{P \in (C_1 \cap C_2)(\overline{\mathbb{Q}})} h_{\mathrm{Ph}}(P) \leq (\deg C_1) h_{\mathrm{Ph}}(C_2) + (\deg C_2) h_{\mathrm{Ph}}(C_1).$$

From [ibid., p. 347], for $i = 1, 2$ we have

$$h_{\mathrm{Ph}}(C_i) = h_{\mathrm{Ph}}(f_i) + \frac{\deg C_i}{2}$$

and from the definitions of the heights, easy estimates give

$$h_{\mathrm{Ph}}(f_i) \leq h(f_i) + \left(\log 2 + \tfrac{3}{4}\right) \deg C_i = h(C_i) + \left(\log 2 + \tfrac{3}{4}\right) \deg C_i.$$

So

$$h_{\mathrm{Ph}}(C_i) \leq h(C_i) + 2 \deg C_i \quad \text{for } i = 1, 2.$$

Finally, we note that if P is a point in $\mathbb{P}^n$, then $h_{\mathrm{Ph}}(P)$ is the usual height $h(P)$ except that at the archimedean places one uses the ℓ^2-norm. In particular, $h(P) \leq h_{\mathrm{Ph}}(P)$. Combining the above inequalities gives the result. □

6.4. ***Units and regulators.*** Let k be a number field of degree δ and discriminant Δ. We will use the following bound on the product of the class number and the regulator, proven by Lenstra [1992, Theorem 6.5].

Lemma 28. *Suppose that $k \neq \mathbb{Q}$. Let r_2 denote the number of complex places of k and let $C = (2/\pi)^{r_2} \sqrt{|\Delta|}$. We have*

$$h_k R_k \leq \frac{C (\log C)^{\delta - 1 - r_2} (\delta - 1 + \log C)^{r_2}}{(\delta - 1)!}.$$

Let S be a finite set of places of k containing the archimedean places. Recall that Q_S is the minimum value of

$$h(u_1)\cdots h(u_{s-1}),$$

where $s=|S|$ and $u_1,\dots,u_{s-1}$ are generators of $\mathbb{O}^*_{k,S}$ modulo roots of unity. For the S-regulator and Q_S, Bugeaud and Győry [1996, Lemmas 1 and 3] gave the bounds

$$R_S \le h_k R_k \prod_{v\in S\setminus S_\infty} \log N(v),$$
$$Q_S \le \frac{((s-1)!)^2}{2^{s-2}\delta^{s-1}} R_S.$$

More crudely, letting $s=|S|$ and $N=\max_{v\in S} N(v)$, we have the estimates

$$\begin{aligned} R_k \le h_k R_k &\le \frac{\sqrt{|\Delta|}(\frac{1}{2}\log|\Delta|)^{\delta-1-r_2}(\delta-1+\frac{1}{2}\log|\Delta|)^{r_2}}{(\delta-1)!} \\ &\le \frac{\sqrt{|\Delta|}(\delta-1+\frac{1}{2}\log|\Delta|)^{\delta-1}}{(\delta-1)!} \le \frac{\sqrt{|\Delta|}2^{\delta-1}\delta^{\delta-1}(\log^*|\Delta|)^{\delta-1}}{(\delta-1)!} \\ &\le \delta^\delta\sqrt{|\Delta|}(\log^*|\Delta|)^{\delta-1} \end{aligned}$$

and

$$R_S \le \delta^\delta (\log^* N)^{s-\delta/2}\sqrt{|\Delta|}(\log^*|\Delta|)^{\delta-1}, \tag{2}$$
$$Q_S \le 2^{2-s}s^{2s-4}\delta^{\delta-s+1}(\log^* N)^{s-\delta/2}\sqrt{|\Delta|}(\log^*|\Delta|)^{\delta-1}. \tag{3}$$

6.5. *Points in projective space.* We first recall an inequality of Silverman relating the height of a point in projective space and the discriminant of its field of definition.

Theorem 29 [Silverman 1984, Theorem 2]. *Let k be a number field of degree δ and discriminant Δ. Let $P\in\mathbb{P}^n(k)$. Then*

$$\frac{\log|\Delta|}{\delta} \le (2\delta-2)h(P)+\log\delta.$$

For a number field k and finite set of places S of k containing the archimedean places, define

$$c_3(k,S)=\begin{cases} 0 & \text{if } \delta=1 \text{ or } s=1,\\ \dfrac{2s!\,s^{s+\frac{1}{2}}R_S}{(\log\delta/6\delta^3)^{s-2}} & \text{otherwise,}\end{cases}$$

where $s=|S|$. If S_∞ denotes the set of archimedean places of k, then we let $c_3(k)=c_3(k,S_\infty)$.

The next lemma describes certain choices of coordinates for a point in projective space.

Lemma 30. *Let k be a number field of degree δ, s the number of archimedean places of k, and $P \in \mathbb{P}^n(k)$.*

(a) *There exists a choice of homogeneous coordinates $P = (x_0, \dots, x_n)$ such that $x_0, \dots, x_n \in \mathbb{O}_k$ and for any $v \in M_k$,*

$$\begin{aligned} \frac{1}{s}h(P) - c_3(k) \le \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} \\ \le \frac{1}{s}h(P) + \frac{1}{2\delta s}\log|\Delta| + c_3(k) \qquad \text{if } v \mid \infty, \\ -\frac{1}{2\delta}\log|\Delta| \le \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} \le 0 \quad \text{if } v \nmid \infty. \end{aligned}$$

(b) *There exists a choice of homogeneous coordinates $P = (x_0, \dots, x_n)$ such that $x_0, \dots, x_n \in \mathbb{O}_k$ and for any $v \in M_k$,*

$$\begin{aligned} 0 \le \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} \\ \le (2\delta + 1)h(P) + \log\delta \qquad \text{if } v \mid \infty, \\ -\delta^2 h(P) - \frac{\delta}{2}\log\delta \le \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} \le 0 \qquad \text{if } v \nmid \infty. \end{aligned}$$

Proof. Let S_∞ denote the set of archimedean places of k. The case $k = \mathbb{Q}$ follows immediately by choosing $x_0, \dots, x_n$ to be integers with $\gcd(x_0, \dots, x_n) = 1$ and $P = (x_0, \dots, x_n)$. We assume from now on that $\delta > 1$. Let $P = (x_0, \dots, x_n)$ be some choice of homogeneous coordinates with $x_0, \dots, x_n \in \mathbb{O}_k$. Let I be the ideal of $\mathbb{O}_k$ generated by $x_0, \dots, x_n$. From the Minkowski bound, the ideal class of I contains an (integral) ideal with norm $\le \sqrt{|\Delta|}$. Thus, after rescaling $x_0, \dots, x_n$, we may assume that the norm of I satisfies $N(I) \le \sqrt{|\Delta|}$. From the definition of the height, we have

$$h(P) = \sum_{v \in M_k} \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} = \sum_{v \in S_\infty} \log\max_i \|x_i\|_v - \frac{1}{\delta}\log N(I).$$

So

$$h(P) \le \sum_{v \in S_\infty} \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} \le h(P) + \frac{1}{2\delta}\log|\Delta|.$$

We first consider (a). The case $s = |S_\infty| = 1$ is immediate from the above, so we assume from now on that $s > 1$. Consider the image of the unit group $\mathbb{O}_k^*$ via the logarithmic map $\lambda : \mathbb{O}_k^* \mapsto \mathbb{R}^s$, $\lambda(u) = (\log\|u\|_v)_{v \in S_\infty}$. The image is a lattice in the hyperplane of $\mathbb{R}^s$ defined by $\sum_{v \in S_\infty} x_v = 0$. From [Hajdu 1993, p. 5], there exists a fundamental domain of this lattice with diameter $\le 2s!s^{s+\frac{1}{2}}R_k/(\log\delta/6\delta^3)^{s-2}$. Let $c = \sum_{v \in S_\infty} \log\max\{\|x_0\|_v, \dots, \|x_n\|_v\}$ and consider the vector

$$\boldsymbol{v} = (\log\max\{\|x_0\|_v, \dots, \|x_n\|_v\} - c/s)_{v \in S_\infty}.$$

Then there exists a unit $u \in \mathcal{O}_k^*$ such that

$$|\boldsymbol{v} - \lambda(u)| \le \frac{2s!\,s^{s+\frac{1}{2}} R_k}{(\log\delta/6\delta^3)^{s-2}} = c_3(k).$$

Therefore, for every $v \in S_\infty$,

$$\left|\log\max\{\|u^{-1}x_0\|_v, \ldots, \|u^{-1}x_n\|_v\} - c/s\right| \le c_3(k)$$

and

$$\frac{1}{s}h(P) - c_3(k) \le \log\max\{\|u^{-1}x_0\|_v, \ldots, \|u^{-1}x_n\|_v\} \le \frac{1}{s}h(P) + \frac{1}{2\delta s}\log|\Delta| + c_3(k).$$

Note that if $v \nmid \infty$, we also have

$$-\frac{1}{2\delta}\log|\Delta| \le -\frac{1}{\delta}\log N(I) \le \log\max\{\|x_0\|_v, \ldots, \|x_n\|_v\} \le 0.$$

We now prove (b). From our earlier choice of coordinates, we have in particular

$$\sum_{v\in S_\infty}\log\|x_0\|_v = \frac{1}{\delta}\log|N_{\mathbb{Q}}^k(x_0)| \le h(P) + \frac{1}{2\delta}\log|\Delta|.$$

Then after scaling by $N_{\mathbb{Q}}^k(x_0)/x_0 \in \mathcal{O}_k$, we may take $P = (x_0, \ldots, x_n)$ where $x_0 \in \mathbb{Z}$,

$$\frac{1}{\delta}\log|x_0| \le h(P) + \frac{1}{2\delta}\log|\Delta| \le \delta h(P) + \tfrac{1}{2}\log\delta$$

by Theorem 29, and $x_1, \ldots, x_n \in \mathcal{O}_k$. Let $v \in S_\infty$. Then $\log\max_i\|x_i\|_v \ge 0$ and

$$\begin{aligned}\log\max\{\|x_0\|_v, \ldots, \|x_n\|_v\} &= h(P) - \sum_{w\in M_k\setminus\{v\}}\log\max\{\|x_0\|_w, \ldots, \|x_n\|_w\}\\ &\le h(P) - \sum_{w\in M_k\setminus\{v\}}\log\|x_0\|_w \le h(P) + \log\|x_0\|_v\\ &\le h(P) + 2(\delta h(P) + \tfrac{1}{2}\log\delta) \le (2\delta+1)h(P) + \log\delta.\end{aligned}$$

We also clearly have

$$-\delta^2 h(P) - \frac{\delta}{2}\log\delta \le -\log|x_0| \le \log\max\{\|x_0\|_v, \ldots, \|x_n\|_v\} \le 0$$

if v is nonarchimedean. □

We also need the following result from the main theorem of [Hajdu 1993], which is closely related to the previous lemma.

Theorem 31. *Let k be a number field of degree δ and let S be a finite set of places of k containing the archimedean places. Let $\alpha \in k$. Then we can write*

$$\alpha = \beta u,$$

where $u \in \mathbb{O}_{k,S}^*$ *and*

$$h(\beta) < sc_3(k, S) + \sum_{v \notin S} h_v(\alpha) + \sum_{v \in S} \log \|\alpha\|_v < sc_3(k, S) + \sum_{v \notin S} h_v(\alpha) + h_v(1/\alpha).$$

The last inequality follows from the product formula. This result is actually only proven in [ibid.] for S-integers α, but the same proof given there yields the result above.

We note the estimates

$$c_3(k, S) \leq 2^{4s} s^{2s} \delta^{3s+\delta-6} \sqrt{|\Delta|} (\log^* |\Delta|)^{\delta-1} (\log^* N)^{s-\delta/2}, \tag{4}$$

$$c_3(k) \leq 2^{4\delta} \delta^{6\delta-6} \sqrt{|\Delta|} (\log^* |\Delta|)^{\delta-1}. \tag{5}$$

6.6. *Miscellaneous elementary estimates.* We have the following lower bound for heights on $\mathbb{P}^1$.

Lemma 32. *Let S be a set of places of a number field k. Let $P, Q \in \mathbb{P}^1(k)$, $P \neq Q$. Then*

$$\sum_{v \in S} h_{Q,v}(P) \geq -\log 2.$$

Proof. Let $P = (x_1, y_1)$, $Q = (x_2, y_2)$, $x_1, x_2, y_1, y_2 \in k$. Then

$$\begin{aligned} h_{Q,v}(P) &= \log \frac{\max\{\|x_1\|_v, \|y_1\|_v\} \max\{\|x_2\|_v, \|y_2\|_v\}}{\|x_1 y_2 - x_2 y_1\|_v} \\ &\geq \log \frac{\max\{\|x_1\|_v, \|y_1\|_v\} \max\{\|x_2\|_v, \|y_2\|_v\}}{\epsilon_v'(2) \max\{\|x_1 y_2\|_v, \|x_2 y_1\|_v} \\ &\geq \log \frac{\max\{\|x_1\|_v, \|y_1\|_v\} \max\{\|x_2\|_v, \|y_2\|_v\}}{\epsilon_v'(2) \max\{\|x_1\|_v, \|y_1\|_v\} \max\{\|x_2\|_v, \|y_2\|_v\}} \\ &\geq -\log \epsilon_v'(2). \end{aligned}$$

Therefore, $\sum_{v \in S} h_{Q,v}(P) \geq -\log 2$. □

We need an estimate on the height of a product of polynomials [Hindry and Silverman 2000, Proposition B.7.4].

Lemma 33. *Let k be a number field. Let $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ be polynomials and let $f = f_1 \cdots f_m$. Then for any $v \in M_k$,*

$$|f|_v \leq \epsilon_v \left(\prod_{i=2}^{m} 2^{\deg f_i} \right) \prod_{i=1}^{m} |f_i|_v.$$

In particular,

$$h(f) \leq \sum_{i=1}^{m} h(f_i) + \left(\sum_{i=2}^{m} \deg f_i \right) \log 2.$$

For maps between projective spaces, we have the following height inequality [Hindry and Silverman 2000, p. 181].

Lemma 34. *Let $\phi : \mathbb{P}^n \to \mathbb{P}^m$ be a rational map of degree d defined over $\overline{\mathbb{Q}}$. Then*

$$h(\phi(P)) \le dh(P) + h(\phi) + \log \binom{n+d}{n}$$

for all $P \in \mathbb{P}^n(\overline{\mathbb{Q}})$ where ϕ is defined.

We also need an elementary estimate for polynomials in two variables.

Lemma 35. *Let k be a number field. Let $f \in k[x, y]$ be a polynomial of degree d and let $v \in M_k$. Let $a, b, x, y \in k$ and suppose that $|x-a|_v, |y-b|_v \le 1$. Then*

$$|f(x, y)-f(a, b)|_v \le \epsilon_v((d+2)^4 2^d)|f|_v \max\{|a|_v, |b|_v, 1\}^d \max\{|x-a|_v, |y-b|_v\}.$$

Proof. Let $f(x, y) = \sum c_{ij} x^i y^j$. Looking at the Taylor series for $f(x, y)$ around (a, b) and applying the triangle inequality, we find

$$\begin{aligned} |f(x, y) - f(a, b)|_v &\le \left| \sum_{m,n,m+n>0} \left(\frac{\partial^{m+n} f}{\partial x^m \partial y^n} \right)(a, b) \frac{(x-a)^m (y-b)^n}{m!n!} \right|_v \\ &\le \epsilon_v\left(\binom{d+2}{2}\right) \max_{m,n} \left| \frac{1}{m!n!} \left(\frac{\partial^{m+n} f}{\partial x^m \partial y^n} \right)(a, b) \right|_v \max\{|x-a|_v, |y-b|_v\}. \end{aligned}$$

Since

$$\begin{aligned} \left| \frac{1}{m!n!} \left(\frac{\partial^{m+n} f}{\partial x^m \partial y^n} \right)(a, b) \right|_v &= \left| \sum c_{ij} \binom{i}{m} \binom{j}{n} a^{i-m} b^{j-n} \right|_v \\ &\le \epsilon_v\left(\binom{d+2}{2}\right) \max_{i,j} |c_{ij}|_v \left| \binom{i}{m} \binom{j}{n} \right|_v |a^{i-m} b^{j-n}|_v \\ &\le \epsilon_v\left(\binom{d+2}{2} 2^d\right) |f|_v \max\{|a|_v, |b|_v, 1\}^d, \end{aligned}$$

we have

$$\begin{aligned} &|f(x, y) - f(a, b)|_v \\ &\qquad \le \epsilon_v((d+2)^4 2^d)|f|_v \max\{|a|_v, |b|_v, 1\}^d \max\{|x-a|_v, |y-b|_v\}. \qquad \square \end{aligned}$$

Finally, we prove an explicit version of Lemma 10 when $X = \mathbb{P}^2$.

Lemma 36. *Let k be a number field and let $\phi \in k(\mathbb{P}^2)$ be a rational function of degree d on $\mathbb{P}^2$. Let $P, Q \in \mathbb{P}^2(k) \setminus \operatorname{Supp} \phi$ and $T \subset M_k$. Suppose that $\phi(P) \ne \phi(Q)$. Then*

$$\sum_{v \in T} h_{Q,v}(P) \le \sum_{v \in T} h_{\phi(Q),v}(\phi(P)) + (2d+2)h(Q) + 8\log(d+2) + (2d+4)\log 2.$$

Proof. Let $\phi = f_1/f_2$, where $f_1, f_2 \in \mathbb{O}_k[x, y, z]$ are homogeneous polynomials of degree d. Let $Q = (x_0, y_0, z_0)$, $P = (x, y, z)$, and $\alpha = \phi(Q)$. From the definitions,

$$h_{Q,v}(P) = \log \frac{\max\{\|x_0\|_v, \|y_0\|_v, \|z_0\|_v\} \max\{\|x\|_v, \|y\|_v, \|z\|_v\}}{\max\{\|z_0x - x_0z\|_v, \|z_0y - y_0z\|_v, \|x_0y - y_0x\|_v\}},$$

$$h_{\alpha,v}(\phi(P)) = \log \frac{\max\{\|\alpha\|_v, 1\} \max\{\|f_1(x, y, z)\|_v, \|f_2(x, y, z)\|_v\}}{\|f_1(x, y, z) - \alpha f_2(x, y, z)\|_v}.$$

Without loss of generality, after permuting the variables, we can assume that $z_0 \neq 0$ and $Q = (x_0, y_0, 1)$. If $z = 0$, then

$$\begin{aligned} h_{Q,v}(P) &= \log \frac{\max\{\|x_0\|_v, \|y_0\|_v, 1\} \max\{\|x\|_v, \|y\|_v\}}{\max\{\|x\|_v, \|y\|_v, \|x_0y - y_0x\|_v\}} \\ &\leq \log \max\{\|x_0\|_v, \|y_0\|_v, 1\}. \end{aligned}$$

So

$$\sum_{v \in T} h_{Q,v}(P) \leq \sum_{v \in T} \log \max\{\|x_0\|_v, \|y_0\|_v, 1\} \leq h(Q).$$

Then using Lemma 32, in this case we have

$$\sum_{v \in T} h_{Q,v}(P) \leq \sum_{v \in T} h_{\alpha,v}(P) + h(Q) + \log 2.$$

Suppose now that $z \neq 0$, in which case we can take $P = (x, y, 1)$, for some $x, y \in k$.

First suppose that

$$\max\{|x - x_0|_v, |y - y_0|_v\} < \frac{1}{\epsilon_v((d+2)^4 2^{d+1})} \min_{j=1,2} \frac{|f_j(x_0, y_0, 1)|_v}{|f_j|_v \max\{|x_0|_v, |y_0|_v, 1\}^d}.$$

In particular, $\max\{|x - x_0|_v, |y - y_0|_v\} \leq 1$. Let $F(u, v) = f_1(u, v, 1) - \alpha f_2(u, v, 1)$. Note that $\deg F \leq d$. From the definition of α, $F(x_0, y_0) = 0$. Then by Lemma 35, with $a = x_0$, $b = y_0$, we have

$$\begin{aligned} &|F(x, y)|_v \\ &\quad \leq \epsilon_v((d+2)^4 2^d) |F|_v \max\{|x_0|_v, |y_0|_v, 1\}^d \max\{|x - x_0|_v, |y - y_0|_v\}. \end{aligned} \tag{6}$$

For $j = 1, 2$, using Lemma 35 again, we find, if v is archimedean,

$$\begin{aligned} &|f_j(x, y, 1)|_v \\ &\geq |f_j(x_0, y_0, 1)|_v - (d+2)^4 2^d |f_j|_v \max\{|x_0|_v, |y_0|_v, 1\}^d \max\{|x - x_0|_v, |y - y_0|_v\} \\ &\geq \tfrac{1}{2} |f_j(x_0, y_0, 1)|_v. \end{aligned}$$

By the same reasoning, if v is nonarchimedean we have

$$|f_j(x, y, 1) - f_j(x_0, y_0, 1)|_v < |f_j(x_0, y_0, 1)|_v,$$

and so

$$|f_j(x, y, 1)|_v = |f_j(x_0, y_0, 1)|_v \quad \text{for } j = 1, 2.$$

Then in any case,

$$|f_j(x, y, 1)|_v \geq \frac{1}{\epsilon_v(2)} |f_j(x_0, y_0, 1)|_v \quad \text{for } j = 1, 2.$$

Since $\max\{|x - x_0|_v, |y - y_0|_v\} \leq 1$, we also have

$$\max\{|x|_v, |y|_v, 1\} \leq \epsilon_v(2) \max\{|x_0|_v, |y_0|_v, 1\}.$$

Then

$$\begin{aligned} h_{Q,v}(P) &= \log \frac{\max\{\|x_0\|_v, \|y_0\|_v, 1\} \max\{\|x\|_v, \|y\|_v, 1\}}{\max\{\|x - x_0\|_v, \|y - y_0\|_v, \|x_0 y - y_0 x\|_v\}} \\ &\leq 2 \log \max\{\|x_0\|_v, \|y_0\|_v, 1\} + \log \epsilon'_v(2) - \log \max\{\|x - x_0\|_v, \|y - y_0\|_v\} \end{aligned}$$

and

$$\begin{aligned} h_{\alpha,v}(\phi(P)) &= \log \frac{\max\{\|\alpha\|_v, 1\} \max\{\|f_1(x, y, 1)\|_v, \|f_2(x, y, 1)\|_v\}}{\|f_1(x, y, 1) - \alpha f_2(x, y, 1)\|_v} \\ &= \log \max_{j=1,2} \|f_j(x, y, 1)\|_v + \log \max\{\|\alpha\|_v, 1\} - \log \|F(x, y)\|_v \\ &\geq \log \max_{j=1,2} \|f_j(x_0, y_0, 1)\|_v + \log \max\{\|\alpha\|_v, 1\} - \epsilon'_v(\log(d+2)^4 2^{d+1}) \\ &\quad - \log \|F\|_v - d \log \max\{\|x_0\|_v, \|y_0\|_v, 1\} - \log \max\{\|x - x_0\|_v, \|y - y_0\|_v\} \end{aligned}$$

by (6). We can write $|F|_v = |f_1 - \alpha f_2|_v \leq \epsilon_v(2) \max\{|f_1|_v, |f_2|_v\} \max\{|\alpha|_v, 1\}$. So

$$\begin{aligned} h_{\alpha,v}(\phi(P)) \geq \log \max_{j=1,2} \|f_j(x_0, y_0, 1)\|_v - \epsilon'_v(\log(d + 2)^4 2^{d+2}) - \log \max_{j=1,2} \|f_j\|_v \\ - d \log \max\{\|x_0\|_v, \|y_0\|_v, 1\} - \log \max\{\|x - x_0\|_v, \|y - y_0\|_v\}. \end{aligned}$$

Note that

$$|f_j(x_0, y_0, 1)|_v \leq \epsilon_v\left(\binom{d+2}{2}\right) |f_j|_v \max\{|x_0|_v, |y_0|_v, 1\}^d \quad \text{for } j = 1, 2.$$

This implies that

$$\begin{aligned} &\sum_{v \in T} \log \frac{\max\{\|f_1(x_0, y_0, 1)\|_v, \|f_2(x_0, y_0, 1)\|_v\}}{\max\{\|f_1\|_v, \|f_2\|_v\} \max\{\|x_0\|_v, \|y_0\|_v, 1\}^d} \\ &\qquad \geq \sum_{j=1}^{2} \sum_{v \in M_k} \log \frac{\|f_j(x_0, y_0, 1)\|_v}{\|f_j\|_v \max\{\|x_0\|_v, \|y_0\|_v, 1\}^d} - 2 \log \epsilon'_v\left(\binom{d+2}{2}\right) \\ &\qquad \geq -2dh(Q) - 4 \log(d + 2) \end{aligned}$$

by the product formula. So

$$\sum_{v\in T} h_{\alpha,v}(\phi(P)) \geq -2dh(Q) - 8\log(d+2) - (d+2)\log 2 - \sum_{v\in T}\log\max\{\|x-x_0\|_v, \|y-y_0\|_v\}.$$

Then

$$\sum_{v\in T} h_{Q,v}(P) \leq \sum_{v\in T} h_{\alpha,v}(\phi(P)) + (2d+2)h(Q) + 8\log(d+2) + (d+3)\log 2.$$

Finally, suppose that

$$\max\{|x-x_0|_v, |y-y_0|_v\} \geq C_v,$$

where

$$C_v = \frac{1}{\epsilon_v((d+2)^4 2^{d+1})}\min\left\{\frac{|f_1(x_0,y_0,1)|_v}{|f_1|_v\max\{|x_0|_v,|y_0|_v,1\}^d}, \frac{|f_2(x_0,y_0,1)|_v}{|f_2|_v\max\{|x_0|_v,|y_0|_v,1\}^d}\right\}.$$

As noted before, $C_v \leq 1$. Then one easily finds that

$$\begin{aligned}\frac{\max\{|x|_v,|y|_v,1\}}{\max\{|x-x_0|_v,|y-y_0|_v\}} &= \frac{\max\{|(x-x_0)+x_0|_v, |(y-y_0)+y_0|_v, 1\}}{\max\{|x-x_0|_v,|y-y_0|_v\}} \\ &\leq \frac{\epsilon_v(2)\max\{|(x-x_0)|_v, |(y-y_0)|_v, |x_0|_v, |y_0|_v, 1\}}{\max\{|x-x_0|_v,|y-y_0|_v\}} \\ &\leq \frac{\epsilon_v(2)\max\{|x_0|_v,|y_0|_v,1\}}{C_v}.\end{aligned}$$

So

$$\begin{aligned}h_{Q,v}(P) &= \frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}\log\frac{\max\{|x_0|_v,|y_0|_v,1\}\max\{|x|_v,|y|_v,1\}}{\max\{|x-x_0|_v,|y-y_0|_v,|x_0y-y_0x|_v\}} \\ &\leq \frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}\log\epsilon_v(2)\max\{|x_0|_v,|y_0|_v,1\}^2/C_v \\ &\leq \frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}(2\log\max\{|x_0|_v,|y_0|_v,1\} + \log\epsilon_v(2) - \log C_v).\end{aligned}$$

Then using Lemma 32, we find

$$\sum_{v\in T} h_{Q,v}(P) \leq \sum_{v\in T} h_{\alpha,v}(\phi(P)) + 2h(Q) + 2\log 2 - \sum_{v\in T}\frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}\log C_v.$$

Since

$$\begin{aligned}\sum_{v\in T}\frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}\log C_v &\geq \sum_{v\in M_k}\frac{[k_v:\mathbb{Q}_v]}{[k:\mathbb{Q}]}\log C_v\\ &\geq \sum_{j=1}^{2}\sum_{v\in M_k}\log\frac{\|f_j(x_0,y_0,1)\|_v}{\epsilon'_v((d+2)^4 2^{d+1})\|f_j\|_v\max\{\|x_0\|_v,\|y_0\|_v,1\}^d}\\ &\geq -8\log(d+2)-(2d+2)\log 2-2dh(Q),\end{aligned}$$

where we have used the product formula in the last line, we obtain

$$\sum_{v\in T}h_{Q,v}(P)\leq\sum_{v\in T}h_{\alpha,v}(\phi(P))+(2d+2)h(Q)+(2d+4)\log 2+8\log(d+2).\ \square$$

7. Explicit results for $\mathbb{P}^2$

In this section we give a proof of Theorem 7. The proof will follow the proof in Section 3.3, except that we will give explicit estimates at each step. We begin with an explicit version of Theorem 19.

Theorem 37. *Let k be a number field of degree δ and discriminant Δ. Let S be a finite set of places of k, containing the archimedean places, of cardinality s. Let C_1 and C_2 be distinct curves over k in $\mathbb{P}^2$ defined by homogeneous polynomials $f_1, f_2 \in \mathcal{O}_k[x,y,z]$, respectively, of degrees d_1 and d_2, respectively. Let $d = \max\{d_1,d_2\}$ and $\phi = f_1^{d_2}(x,y,z)/f_2^{d_1}(x,y,z)$, a rational function on $\mathbb{P}^2$. Let $Q\in\mathbb{P}^2(\bar{k})\setminus(C_1\cup C_2)$ and let $\delta'=[k(Q):\mathbb{Q}]$. Let $w\in M_{k(Q)}$ and let $0<\epsilon<1$. Then for all $P\in(\mathbb{P}^2\setminus(C_1\cup C_2))(\mathcal{O}_{k,S})$, either*

$$h_{Q,w}(P)\leq\epsilon h(P)+c_4(\epsilon,k,S,w,Q,C_1,C_2) \tag{7}$$

or

$$\phi(P)=\phi(Q),$$

where

$$\begin{aligned}c_4 &= (2d^2+2)h(Q)+10\log(d^2+2)+(2d^2+7)\log 2\\ &\qquad +\frac{1}{d}\,(h(C_1)+h(C_2))+\frac{\log|\Delta|}{\delta d^2}+\frac{2\delta}{d^2}c_3(k)+c_5,\\ c_5 &= 6.4(d^2c_2(\delta',s)/\epsilon)\frac{N(w)}{\log N(w)}c_6c_7\max\{\log((d^2c_2(\delta',s)/\epsilon)N(w)),\log^* c_6\},\\ c_6 &= Q_S\Big(1+sc_3(k,S)+\frac{1}{\delta}\log|\Delta|\Big),\\ c_7 &= d^2h(Q)+dh(C_1)+dh(C_2)+\frac{1}{\delta}\log|\Delta|+2\delta c_3(k)+2d^2\log 2+2\log(d^2+2).\end{aligned}$$

In particular, (7) *holds for all* $P \in (\mathbb{P}^2 \setminus (C_1 \cup C_2))(\mathcal{O}_{k,S})$ *outside of an effectively computable finite union of plane curves* Z.

Proof. Let I_1 and I_2 be the ideals of $\mathcal{O}_k$ generated by the coefficients of $g_1 = f_1^{d_2}$ and $g_2 = f_2^{d_1}$, respectively. We rescale g_1 and g_2 as in Lemma 30(a) and its proof (viewing the coefficients of the polynomials as giving points in projective space). In particular, $N(I_1), N(I_2) \leq \sqrt{|\Delta|}$. Let $\phi = g_1/g_2$ and let $P \in (\mathbb{P}^2 \setminus (C_1 \cup C_2))(\mathcal{O}_{k,S})$. Then it follows from the definitions that we have an equality of fractional ideals $\phi(P)\mathcal{O}_k = (I_1/I_2)J$, where J is a fractional ideal supported on the primes in S. By Theorem 31, we can write $\phi(P) = \beta u$, where $u \in \mathcal{O}_{k,S}^*$ and

$$h(\beta) \leq sc_3(k, S) + \frac{1}{\delta}\log N(I_1) + \frac{1}{\delta}\log N(I_2) \leq sc_3(k, S) + \frac{1}{\delta}\log|\Delta|.$$

Let $\alpha = \phi(Q)$ and suppose that $\phi(P) \neq \alpha$. By Theorem 24, substituting ϵ/d^2 for ϵ and taking G to be the multiplicative group generated by β and $\mathcal{O}_{k,S}^*$, we have the inequality

$$h_{\alpha,w}(\phi(P)) \leq \frac{\epsilon}{d^2}h(\phi(P)) + c_1\Big(\frac{\epsilon}{d^2}, k(Q), G, w, \alpha\Big) + \log 2.$$

Note that $\deg\phi \leq d^2$. By Lemma 36,

$$\begin{aligned} h_{Q,w}(P) &\leq h_{\alpha,w}(\phi(P)) + (2d^2+2)h(Q) + 8\log(d^2+2) + (2d^2+4)\log 2 \\ &\leq \frac{\epsilon}{d^2}h(\phi(P)) + (2d^2+2)h(Q) + 8\log(d^2+2) \\ &\qquad\qquad + (2d^2+5)\log 2 + c_1\Big(\frac{\epsilon}{d^2}, k(Q), G, w, \alpha\Big). \end{aligned}$$

By Lemma 34,

$$h(\phi(P)) \leq d^2h(P) + h(\phi) + \log\binom{d^2+2}{2} \leq d^2h(P) + h(\phi) + 2\log(d^2+2).$$

Let s_∞ be the number of archimedean places of k. By Lemma 30(a) and the construction of g_1 and g_2,

$$\begin{aligned} h(\phi) &= \sum_{v\in M_k}\log\max\{\|g_1\|_v, \|g_2\|_v\} \leq \sum_{\substack{v\in M_k \\ v|\infty}}\log\max\{\|g_1\|_v, \|g_2\|_v\} \\ &\leq s_\infty\Big(\frac{1}{s_\infty}h(f_1^{d_2}) + \frac{1}{2\delta s_\infty}\log|\Delta| + c_3(k) + \frac{1}{s_\infty}h(f_2^{d_1}) + \frac{1}{2\delta s_\infty}\log|\Delta| + c_3(k)\Big) \\ &\leq h(f_1^{d_2}) + h(f_2^{d_1}) + \frac{1}{\delta}\log|\Delta| + 2\delta c_3(k). \end{aligned}$$

By Lemma 33,

$$h(f_1^{d_2}) + h(f_2^{d_1}) \leq dh(f_1) + dh(f_2) + 2d^2\log 2 = dh(C_1) + dh(C_2) + 2d^2\log 2.$$

So

$$h(\phi) \le dh(C_1) + dh(C_2) + \frac{1}{\delta}\log|\Delta| + 2\delta c_3(k) + 2d^2\log 2.$$

Then

$$\begin{aligned}
&h_{Q,w}(P)\\
&< \epsilon h(P) + (2d^2+2)h(Q) + 10\log(d^2+2) + (2d^2+7)\log 2\\
&\qquad + \frac{1}{d}(h(C_1)+h(C_2)) + \frac{1}{\delta d^2}\log|\Delta| + \frac{2\delta}{d^2}c_3(k) + c_1\Big(\frac{\epsilon}{d^2}, k(Q), G, w, \alpha\Big).
\end{aligned}$$

Finally, we can estimate the last term using

$$Q_G \le Q_S \max\{h(\beta), 1\} \le Q_S\Big(1 + sc_3(k, S) + \frac{1}{\delta}\log|\Delta|\Big)$$

and, using Lemma 34 again,

$$\begin{aligned}
h(\alpha) &\le d^2h(Q) + h(\phi) + 2\log(d^2+2)\\
&\le d^2h(Q) + dh(C_1) + dh(C_2) + \frac{1}{\delta}\log|\Delta|\\
&\qquad\qquad + 2\delta c_3(k) + 2d^2\log 2 + 2\log(d^2+2). \qquad \square
\end{aligned}$$

Proof of Theorem 7. Let $d_i = \deg C_i$, $i = 1, \dots, n$. Let $P \in \big(\mathbb{P}^2 \setminus \bigcup_{i=1}^n C_i\big)(\mathbb{O}_{k,S})$. Then

$$\sum_{v\in S} h_{C_i,v}(P) = d_i h(P) \quad \text{for } i = 1, \dots, n.$$

So for each i, there exists a place $v \in S$ such that $h_{C_i,v}(P) \ge (1/s)h(P)$. Since $s < n$, there exists a place $v \in S$ and distinct elements $i, j \in \{1, \dots, n\}$ such that

$$\min\{h_{C_i,v}(P), h_{C_j,v}(P)\} \ge \frac{1}{s}h(P).$$

The theorem is then a consequence of the following lemma.

Lemma 38. *Let k be a number field of degree δ and discriminant Δ. Let S be a finite set of places of k, containing the archimedean places, of cardinality s. Let $C_1, \dots, C_n \subset \mathbb{P}^2$ be distinct curves over k such that at most $n-2$ of the curves C_i intersect at any point of $\mathbb{P}^2(\bar{k})$. Let $d_i = \deg C_i$, $d = \max_i d_i$, $h = \max_i h(C_i)$, and $N = \max_{v\in S} N(v)$. Let Z' be the set from Theorem 7. Let $0 < \epsilon < 1$ and $v \in S$. Then any point $P \in \big(\mathbb{P}^2 \setminus \bigcup_{i=1}^n C_i\big)(\mathbb{O}_{k,S})$ with*

$$\min\{h_{C_1,v}(P), h_{C_2,v}(P)\} \ge \epsilon h(P)$$

satisfies either $P \in Z'$ or

$$h(P) < 2^{20s+4\delta+75} d^{6s+34} \delta^{5s+8\delta-3} s^{4s-1} N^{d^2} (\log^* N)^{2s} |\Delta|^{3/2} (\log^* |\Delta|)^{3\delta} (h+1)/\epsilon^3.$$

Proof. Let

$$(C_1 \cap C_2)(\bar{k}) = \{Q_1, \ldots, Q_r\} \subset \mathbb{P}^2(\bar{k})$$

and let $Q_i = (x_i, y_i, z_i)$, $x_i, y_i, z_i \in \mathcal{O}_{k(Q_i)}$, $i = 1, \ldots, r$, where $r \leq d^2$. Let $L = k(Q_1, \ldots, Q_r)$. We note that $[k(Q_i) : k] \leq d^2$ for all i. Let C_i be defined by $f_i \in \mathcal{O}_k[x, y, z]$, $i = 1, \ldots, n$, and let

$$h_\infty = \log \max_{\substack{w \in M_L \\ w \mid \infty}} \left\{ |f_1|_w, |f_2|_w, \max \left| \prod_{i=1}^r g_i \right|_w \right\},$$

where the max is taken over all possible choices of

$$g_i \in \{z_i x - x_i z, z_i y - y_i z, x_i y - y_i x\} \subset \mathcal{O}_L[x, y, z], \quad i = 1, \ldots, r.$$

Now fix a choice of $g_i \in \{z_i x - x_i z, z_i y - y_i z, x_i y - y_i x\}$, $i = 1, \ldots, r$. Since $\prod_{i=1}^r g_i$ vanishes at all the points Q_i, by the effective Hilbert Nullstellensatz (see Remark 26), there exists a positive integer M, homogeneous polynomials $a_1, a_2 \in \mathcal{O}_L[x, y, z]$ with $\deg a_1 = rM - \deg f_1$, $\deg a_2 = rM - \deg f_2$, and a constant $a \in \mathcal{O}_L$ such that

$$f_1(x, y, z) a_1(x, y, z) + f_2(x, y, z) a_2(x, y, z) = a \left(\prod_{i=1}^r g_i \right)^M$$

and

$$M \leq (8d)^8,$$

$$\log \max_{\substack{w \in M_L \\ w \mid \infty}} \{|a_1|_w, |a_2|_w, |a|_w\} \leq (8d)^{15} (h_\infty + 8d \log 8d).$$

Let w be a place of L lying above v (we will choose a specific such w later). Let $x, y, z \in k$. It follows that there exists a_1, a_2, a, and M, as above, such that

$$\begin{aligned}
&\left(\prod_{i=1}^r \max\{|z_i x - x_i z|_w, |z_i y - y_i z|_w, |x_i y - y_i x|_w\} \right)^M \\
&\qquad = \frac{1}{|a|_w} |f_1(x, y, z) a_1(x, y, z) + f_2(x, y, z) a_2(x, y, z)|_w \\
&\qquad \leq 2 \max\{|f_1(x, y, z) a_1(x, y, z)|_w, |f_2(x, y, z) a_2(x, y, z)|_w\} / |a|_w \\
&\qquad \leq 2(rM)^2 \max_{i=1,2} \left\{ |f_i(x, y, z)|_w |a_i|_w \max\{|x|_w, |y|_w, |z|_w\}^{rM - \deg f_i} \right\} / |a|_w.
\end{aligned}$$

So

$$\begin{aligned}
&\left(\prod_{i=1}^r \frac{\max\{|z_i x - x_i z|_w, |z_i y - y_i z|_w, |x_i y - y_i x|_w}{\max\{|x|_w, |y|_w, |z|_w\}} \right)^M \\
&\qquad\qquad \leq \frac{2(rM)^2}{|a|_w} \max\{|a_1|_w, |a_2|_w\} \max_{i=1,2} \frac{|f_i(x, y, z)|_w}{\max\{|x|_w, |y|_w, |z|_w\}^{\deg f_i}}.
\end{aligned}$$

Let $r_{w/v} = [L : k]/[L_w : k_v]$. Taking logarithms, rearranging, and using the definitions and inequalities above, we find

$$\begin{aligned}\epsilon h(P) &\le \min\{h_{C_1,v}(P), h_{C_2,v}(P)\} = r_{w/v}\min\{h_{C_1,w}(P), h_{C_2,w}(P)\}\\ &\le Mr_{w/v}\sum_{i=1}^{r} h_{Q_i,w}(P) - Mr_{w/v}\sum_{i=1}^{r}\log\max\{\|x_i\|_w, \|y_i\|_w, \|z_i\|_w\}\\ &\qquad + \log 2(rM)^2 + r_{w/v}\log\max\{\|f_1\|_w, \|f_2\|_w\}\\ &\qquad + r_{w/v}\log\max\{\|a_1\|_w, \|a_2\|_w\} - r_{w/v}\log\|a\|_w. \qquad (8)\end{aligned}$$

Let $Q_l \in (C_1 \cap C_2)(\bar{k})$. Then by assumption, there exists $i, j \in \{1, \dots, n\}$, $i \ne j$, such that $Q_l \notin C_i \cup C_j$. Let w_l be the place of $k(Q_l)$ lying below w and let $r_{w_l/v} = [k(Q_l) : k]/[k(Q_l)_{w_l} : k_v]$. Let Φ_P and Z' be as in Theorem 7.

By Theorem 37, either

$$P \in \bigcap_{\phi\in\Phi_{Q_l}} \{Q \in X(\bar{k}) : \phi(Q) = \phi(Q_l)\} \subset Z'$$

or

$$\begin{aligned}r_{w/v}h_{Q_l,w}(P) &= r_{w_l/v}h_{Q_l,w_l}(P)\\ &\qquad < \frac{\epsilon}{2rM}h(P) + \max_{i,j} c_4\Big(\frac{\epsilon}{2rMr_{w_l/v}}, k, S, w_l, Q_l, C_i, C_j\Big)\end{aligned}$$

for all $P \in \big(\mathbb{P}^2 \setminus \bigcup_{i=1}^{n} C_i\big)(\mathbb{O}_{k,S}) = \bigcap_{i,j}(\mathbb{P}^2 \setminus (C_i \cup C_j))(\mathbb{O}_{k,S})$.

Suppose now that $P \notin Z'$. Summing over all points in $C_1 \cap C_2$, we obtain

$$Mr_{w/v}\sum_{l=1}^{r} h_{Q_l,w}(P) < \frac{\epsilon}{2}h(P) + \sum_{l=1}^{r}\max_{i,j} c_4\Big(\frac{\epsilon}{2rMr_{w_l/v}}, k, S, w_l, Q_l, C_i, C_j\Big).$$

Substituting into (8) we find that

$$\begin{aligned}h(P) < \frac{2}{\epsilon}\Bigg(&\sum_{l=1}^{r}\max_{i,j} c_4\Big(\frac{\epsilon}{2rMr_{w_l/v}}, k, S, w_l, Q_l, C_i, C_j\Big)\\ &+ \log 2(rM)^2 + r_{w/v}\log\max_{i=1,2}\|f_i\|_w + r_{w/v}\log\max_{i=1,2}\|a_i\|_w\\ &- Mr_{w/v}\sum_{i=1}^{r}\log\max\{\|x_i\|_w, \|y_i\|_w, \|z_i\|_w\} - r_{w/v}\log\|a\|_w\Bigg). \qquad (9)\end{aligned}$$

We now estimate all the terms on the right-hand side. The dominant term, which comes from the first sum above, is

$$\sum_{l=1}^{r}\max_{i,j} c_5\Big(\frac{\epsilon}{2rMr_{w_l/v}}, k, S, w_l, Q_l, C_i, C_j\Big). \qquad (10)$$

We estimate this term first. We note that by (2), (3), and (4),

$$c_6(k,S) \le 2^{3s+3} s^{4s-3} \delta^{2s+2\delta-5} (\log^* N)^{2s-\delta} |\Delta| (\log^* |\Delta|)^{2\delta-2}.$$

Then

$$\begin{aligned} 2rMr_{w_l/v} d^2 c_2(d^2\delta, s) N(w_l) c_6(k,S) \\ &= 72rMd^2 r_{w_l/v} N(w_l) (16ed^2\delta)^{3s+5} (\log^* d^2\delta)^2 c_6(k,S) \\ &\le 2^{20s+61} d^{6s+28} \delta^{5s+2\delta+2} s^{4s-3} N^{d^2} (\log^* N)^{2s-\delta} |\Delta| (\log^* |\Delta|)^{2\delta-2}. \end{aligned}$$

Simple estimates then also give

$$\log 2rMr_{w_l/v} d^2 c_2(d^2\delta, s) N(w_l) c_6(k,S)/\epsilon \le 2^7 s^2 d^2 (\log^* N)(\log^* |\Delta|)/\epsilon.$$

We have $\sum_{l=1}^r h(Q_l) \le d(h(C_1) + h(C_2) + 4d) \le 4d^2(h+1)$, by Theorem 27, and

$$\begin{aligned} \sum_{l=1}^r \max_{i,j} c_7\left(k, Q_l, C_i, C_j\right) \\ &\le \sum_{l=1}^r (d^2 h(Q_l) + 2dh + 2d^2 \log 2 + 2\log(d^2+2) + \frac{1}{\delta} \log|\Delta| + 2\delta c_3(k)) \\ &\le 2^{4\delta+2} \delta^{6\delta-5} \sqrt{|\Delta|} (\log^* |\Delta|)^\delta d^4 (h+1). \end{aligned}$$

Then (10) is bounded by

$$2^{20s+4\delta+73} d^{6s+34} \delta^{5s+8\delta-3} s^{4s-1} N^{d^2} (\log^* N)^{2s} |\Delta|^{3/2} (\log^* |\Delta|)^{3\delta} (h+1)/\epsilon^2. \quad (11)$$

In the remainder of the proof, we will show that the sum of the remaining elements in the parentheses on the right-hand side of (9) can also be bounded by this quantity. Thus, we find that

$$h(P) \le 2^{20s+4\delta+75} d^{6s+34} \delta^{5s+8\delta-3} s^{4s-1} N^{d^2} (\log^* N)^{2s} |\Delta|^{3/2} (\log^* |\Delta|)^{3\delta} (h+1)/\epsilon^3,$$

proving the lemma.

First, we handle the remaining terms coming from the first sum in (9):

$$\begin{aligned} \sum_{l=1}^r (2d^2+2) h(Q_l) + 10r \log(d^2+2) + (2d^2+7) r \log 2 \\ &\quad + \frac{r}{d}\left(h(C_1) + h(C_2)\right) + \frac{r}{\delta d^2} \log|\Delta| + \frac{2\delta r}{d^2} c_3(k) \\ &\le 4(2d^2+2) d^2 (h+1) + 10d^2 \log(d^2+2) + (2d^2+7) d^2 \\ &\quad + 2dh + \frac{1}{\delta} \log|\Delta| + 2\delta 2^{4\delta} \delta^{6\delta} \sqrt{|\Delta|} (\log^* |\Delta|)^\delta \\ &\le 55d^4(h+1) + 2^{4\delta+2} \delta^{6\delta+1} \sqrt{|\Delta|} (\log^* |\Delta|)^\delta. \end{aligned}$$

We now bound h_∞, after making some appropriate choices. Choose f_1, f_2, and $Q_l = (x_l, y_l, z_l)$, $l = 1, \dots, r$, as in Lemma 30(b). Then

$$\begin{aligned}
-\Big((d^2\delta)^2 h(Q_l) + \frac{(d^2\delta)}{2}\log(d^2\delta)\Big) &\le \log\max\{\|x_l\|_{w_l}, \|y_l\|_{w_l}, \|z_l\|_{w_l}\} \\
&\le (2d^2\delta+1)h(Q_l) + \log d^2\delta, \\
\log\max_{\substack{v\in M_k\\ v|\infty}} |f_1|_v &\le (2\delta+1)\delta h(C_1) + \delta\log\delta \le 3\delta^2(h+1), \\
\log\max_{\substack{v\in M_k\\ v|\infty}} |f_2|_v &\le (2\delta+1)\delta h(C_2) + \delta\log\delta \le 3\delta^2(h+1).
\end{aligned}$$

Let $g_l \in \{z_l x - x_l z, z_l y - y_l z, x_l y - y_l x\}$, $i = 1, \dots, r$. If $w \in M_L$, $w|\infty$, then

$$\begin{aligned}
\log\Big|\prod_{l=1}^r g_l\Big|_w &\le \log 2^r \prod_{l=1}^r |g_l|_w \le r\log 2 + \log\sum_{l=1}^r |g_l|_w \\
&\le r\log 2 + \sum_{l=1}^r \log\max\{|x_l|_{w_l}, |y_l|_{w_l}, |z_l|_{w_l}\} \\
&\le d^2\log 2 + \sum_{l=1}^r \big((2d^2\delta+1)d^2\delta h(Q_l) + d^2\delta\log d^2\delta\big) \\
&\le d^2\log 2 + 4(2d^2\delta+1)d^4\delta(h+1) + d^4\delta\log d^2\delta.
\end{aligned}$$

Since $w \in M_L$ was arbitrary, we have

$$\log\max_{\substack{w\in M_L\\ w|\infty}} \Big|\prod_{l=1}^r g_l\Big|_w \le d^2\log 2 + 4(2d^2\delta+1)d^4\delta(h+1) + d^4\delta\log d^2\delta.$$

It follows easily that

$$h_\infty \le 14d^6\delta^2(h+1).$$

Then from the above, we have

$$\begin{aligned}
\log 2(rM)^2 &\le \log 2d^4(8d)^{16} \le 2^6 d, \\
r_{w/v}\log\max\{\|f_1\|_w, \|f_2\|_w\} &\le \log\max\{|f_1|_w, |f_2|_w\} \le 4\delta^2(h+1), \\
r_{w/v}\log\max\{\|a_1\|_w, \|a_2\|_w\} &\le \log\max\{|a_1|_w, |a_2|_w\} \le (8d)^{15}(h_\infty + 8d\log 8d) \\
&\le 2^{51}d^{21}\delta^2(h+1).
\end{aligned}$$

We also find

$$
\begin{aligned}
-Mr_{w/v}\sum_{l=1}^{r}\log\max\{\|x_l\|_w,\,\|y_l\|_w,\,\|z_l\|_w\}& \\
&= -M\sum_{l=1}^{r} r_{w_l/v}\log\max\{\|x_l\|_{w_l},\,\|y_l\|_{w_l},\,\|z_l\|_{w_l}\}\\
&\le (8d)^8 d^2\Bigl((d^2\delta)^2\sum_{l=1}^{r} h(Q_l)+\frac{d^4\delta}{2}\log(d^2\delta)\Bigr)\\
&\le 2^{24}d^{10}\Bigl(4d^6\delta^2(h+1)+\frac{d^4\delta}{2}\log(d^2\delta)\Bigr)\\
&\le 2^{27}d^{16}\delta^2(h+1).
\end{aligned}
$$

From the product formula and the fact that $a \in O_L$, we have the inequality

$$
\begin{aligned}
-\sum_{\substack{w'\in M_L\\ w'|v}}\log\|a\|_{w'} = \sum_{\substack{w'\in M_L\\ w'\nmid v}}\log\|a\|_{w'} &\le \sum_{\substack{w'\in M_L\\ w'|\infty}}\max\{\log\|a\|_{w'},0\}\\
&\le (8d)^{15}(h_\infty+8d\log 8d)\le 2^{51}d^{21}\delta^2(h+1).
\end{aligned}
$$

Since L/k is Galois, there are exactly $r_{w/v}$ places $w' \in M_L$ with $w'|v$. Therefore, there exists a place $w' \in M_L$ with $w'|v$ and $-\log\|a\|_{w'} \le (1/r_{w/v})2^{51}d^{36}\delta^2(h+1)$. Choosing now $w = w'$, we have

$$-r_{w/v}\log\|a\|_w \le 2^{51}d^{21}\delta^2(h+1).$$

Summing all of the inequalities above, we find that, as claimed, the remaining terms in (9) are easily bounded by (11). □

Acknowledgments

The author would like to thank Mike Bennett for sharing and discussing his results.

References

[Autissier 2009] P. Autissier, "Géométries, points entiers et courbes entières", *Ann. Sci. Éc. Norm. Supér.* (4) **42**:2 (2009), 221–239. MR 2010j:14044 Zbl 1173.14016

[Autissier 2011] P. Autissier, "Sur la non-densité des points entiers", *Duke Math. J.* **158**:1 (2011), 13–27. MR 2012j:14033 Zbl 1217.14020

[Baker 1975] A. Baker, *Transcendental number theory*, Cambridge University Press, 1975. MR 54 #10163 Zbl 0297.10013

[Bérczes et al. 2009] A. Bérczes, J.-H. Evertse, and K. Győry, "Effective results for linear equations in two unknowns from a multiplicative division group", *Acta Arith.* **136**:4 (2009), 331–349. MR 2009j:11054 Zbl 1234.11034

[Bilu 1995] Y. Bilu, "Effective analysis of integral points on algebraic curves", *Israel J. Math.* **90**:1-3 (1995), 235–252. MR 96e:11082 Zbl 0840.11028

[Bombieri 1983] E. Bombieri, "On Weil's "théorème de décomposition"", *Amer. J. Math.* **105**:2 (1983), 295–308. MR 84g:14029 Zbl 0516.12009

[Bugeaud and Győry 1996] Y. Bugeaud and K. Győry, "Bounds for the solutions of unit equations", *Acta Arith.* **74**:1 (1996), 67–80. MR 97b:11045 Zbl 0861.11023

[Corvaja and Zannier 2004] P. Corvaja and U. Zannier, "On integral points on surfaces", *Ann. of Math.* (2) **160**:2 (2004), 705–726. MR 2005m:11118 Zbl 1146.11035

[Corvaja and Zannier 2006] P. Corvaja and U. Zannier, "On the integral points on certain surfaces", *Int. Math. Res. Not.* **2006** (2006), Art. ID 98623. MR 2006m:14030 Zbl 1145.14023

[Corvaja et al. 2009] P. Corvaja, A. Levin, and U. Zannier, "Integral points on threefolds and other varieties", *Tohoku Math. J.* (2) **61**:4 (2009), 589–601. MR 2011a:11125 Zbl 1250.11066

[Evertse 1984] J.-H. Evertse, "On sums of S-units and linear recurrences", *Compositio Math.* **53**:2 (1984), 225–244. MR 86c:11045 Zbl 0547.10008

[Hajdu 1993] L. Hajdu, "A quantitative version of Dirichlet's S-unit theorem in algebraic number fields", *Publ. Math. Debrecen* **42**:3-4 (1993), 239–246. MR 94e:11118 Zbl 0798.11051

[Hindry and Silverman 2000] M. Hindry and J. H. Silverman, *Diophantine geometry: An introduction*, Graduate Texts in Mathematics **201**, Springer, New York, 2000. MR 2001e:11058 Zbl 0948.11023

[Lenstra 1992] H. W. Lenstra, Jr., "Algorithms in algebraic number theory", *Bull. Amer. Math. Soc.* (*N.S.*) **26**:2 (1992), 211–244. MR 93g:11131 Zbl 0759.11046

[Levin 2008] A. Levin, "Variations on a theme of Runge: Effective determination of integral points on certain varieties", *J. Théor. Nombres Bordeaux* **20**:2 (2008), 385–417. MR 2009j:11104 Zbl 1179.11018

[Levin 2009] A. Levin, "Generalizations of Siegel's and Picard's theorems", *Ann. of Math.* (2) **170**:2 (2009), 609–655. MR 2010k:11116 Zbl 1250.11067

[Masser 1988] D. W. Masser, "Linear relations on algebraic groups", pp. 248–262 in *New advances in transcendence theory* (Durham, 1986), edited by A. Baker, Cambridge Univ. Press, 1988. MR 89j:11065 Zbl 0656.10031

[Masser and Wüstholz 1983] D. W. Masser and G. Wüstholz, "Fields of large transcendence degree generated by values of elliptic functions", *Invent. Math.* **72**:3 (1983), 407–464. MR 85g:11060 Zbl 0516.10027

[Mo and Tijdeman 1992] D. Z. Mo and R. Tijdeman, "Exponential Diophantine equations with four terms", *Indag. Math.* (*N.S.*) **3**:1 (1992), 47–57. MR 93d:11035 Zbl 0765.11018

[Philippon 1995] P. Philippon, "Sur des hauteurs alternatives, III", *J. Math. Pures Appl.* (9) **74**:4 (1995), 345–365. MR 97a:11098 Zbl 0878.11025

[van der Poorten and Schlickewei 1982] A. J. van der Poorten and H. P. Schlickewei, "The growth conditions for recurrence sequences", Macquarie Univ. Math. Rep. 82-0041, Macquarie University, 1982.

[Runge 1887] C. Runge, "Über ganzzahlige Lösungen von Gleichungen zwischen zwei Veränderlichen", *J. Reine Angew. Math.* **100** (1887), 425–435. JFM 19.0076.03

[Schmidt 1991] W. M. Schmidt, "Construction and estimation of bases in function fields", *J. Number Theory* **39**:2 (1991), 181–224. MR 93b:11079 Zbl 0764.11046

[Silverman 1984] J. H. Silverman, "Lower bounds for height functions", *Duke Math. J.* **51**:2 (1984), 395–403. MR 87d:11039 Zbl 0579.14035

[Silverman 1987] J. H. Silverman, "Arithmetic distance functions and height functions in Diophantine geometry", *Math. Ann.* **279**:2 (1987), 193–216. MR 89a:11066 Zbl 0607.14013

[Skinner 1990] C. M. Skinner, "On the Diophantine equation $ap^x + bq^v = c + dp^z q^w$", *J. Number Theory* **35**:2 (1990), 194–207. MR 91h:11021 Zbl 0703.11017

[Vojta 1983] P. A. Vojta, *Integral points on varieties*, Ph.D. thesis, Harvard University, 1983, Available at http://search.proquest.com/docview/303162150.

[Vojta 1987] P. Vojta, *Diophantine approximations and value distribution theory*, Lecture Notes in Mathematics **1239**, Springer, Berlin, 1987. MR 91k:11049 Zbl 0609.14011

Communicated by Joseph Silverman
Received 2013-04-14 Revised 2013-09-02 Accepted 2013-10-02

adlevin@math.msu.edu *Department of Mathematics, Michigan State University, 619 Red Cedar Road, East Lansing, MI 48824, United States*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.689

Lefschetz theorem for abelian fundamental group with modulus

Moritz Kerz and Shuji Saito

We prove a Lefschetz hypersurface theorem for abelian fundamental groups allowing wild ramification along some divisor. In fact, we show that isomorphism holds if the degree of the hypersurface is large relative to the ramification along the divisor.

1. Statement of main results

Lefschetz hyperplane theorems represent an important technique in the study of Grothendieck's fundamental group $\pi_1(X)$ of an algebraic variety X (we omit base points for simplicity). Roughly speaking, one gets an isomorphism of the form

$$\iota_{Y/X} : \pi_1(Y) \xrightarrow{\sim} \pi_1(X)$$

for a suitable hypersurface section $Y \to X$ if $\dim(X) \geq 3$. Purely algebraic Lefschetz theorems for projective varieties satisfying certain regularity assumptions were developed in [SGA 2 1968]. The case of nonproper varieties X and Y is more intricate because one needs a precise control of the ramification at the infinite locus. We show in the present note that for the abelian quotient of the fundamental group a Lefschetz hyperplane theorem does in fact hold. Our basic technical ingredient is the higher-dimensional ramification theory of Brylinski, Kato and Matsuda, which is recalled in Section 2. We expect that there is a noncommutative analog of our Lefschetz theorem, which should have applications to ℓ-adic representations of fundamental groups, especially over finite fields as studied in [Esnault and Kerz 2012].

To formulate our main result, let X be a normal variety over a perfect field k, and let $U \subset X$ be an open subset such that $X \setminus U$ is the support of an effective Cartier divisor on X. Let D be an effective Cartier divisor on X with support in $X \setminus U$. We introduce the abelian fundamental group $\pi_1^{\mathrm{ab}}(X, D)$ as a quotient of $\pi_1^{\mathrm{ab}}(U)$ classifying abelian étale coverings of U with ramification bounded by D. More precisely, for an integral curve $Z \subset U$, let Z^N be the normalization of the closure of Z in X with $\phi_Z : Z^N \to X$, the natural map. Let $Z_\infty \subset Z^N$ be the finite set

MSC2010: primary 14H30; secondary 14E22.
Keywords: fundamental group, ramification, Lefschetz theorem .

of points x such that $\phi_Z(x) \notin U$. Then $\pi_1^{\mathrm{ab}}(X, D)$ is defined as the Pontryagin dual of the group $\mathrm{fil}_D H^1(U)$ of continuous characters $\chi : \pi_1^{\mathrm{ab}}(U) \to \mathbb{Q}/\mathbb{Z}$ such that, for any integral curve $Z \subset U$, its restriction $\chi|_Z : \pi_1^{\mathrm{ab}}(Z) \to \mathbb{Q}/\mathbb{Z}$ satisfies the following inequality of Cartier divisors on Z^N:

$$\sum_{y \in Z_\infty} \mathrm{art}_y(\chi|_Z)[y] \le \phi_Z^* D,$$

where $\mathrm{art}_y(\chi|_Z) \in \mathbb{Z}_{\ge 0}$ is the Artin conductor of $\chi|_Z$ at $y \in Z_\infty$ and $\phi_Z^* D$ is the pullback of D by the natural map $\phi_Z : Z^N \to X$.

Such a global measure of ramification in terms of curves has been first considered by Deligne and Laumon; see [Laumon 1981].

Now assume that X is smooth projective over k (we fix a projective embedding) and that $C = X \setminus U$ is a simple normal crossing divisor. Let Y be a smooth hypersurface section such that $Y \times_X C$ is a reduced simple normal crossing divisor on Y, and write $\deg(Y)$ for the degree of Y with respect to the fixed projective embedding of X. Set $E = Y \times_X D$. Then one sees from the definition that the map $Y \cap U \to U$ induces a natural map

$$\iota_{Y/X} : \pi_1^{\mathrm{ab}}(Y, E) \to \pi_1^{\mathrm{ab}}(X, D).$$

Our main theorem says:

Theorem 1.1. *Assume that Y is* sufficiently ample *for (X, D) (see Definition 3.1). If $d := \dim(X) \ge 3$, $\iota_{Y/X}$ is an isomorphism. If $d = 2$, $\iota_{Y/X}$ is surjective.*

The prime-to-p part of the theorem is due to [Schmidt and Spieß 2000], where $p = \mathrm{ch}(k)$. Below we see that Y is sufficiently ample if $\deg(Y) \gg 0$.

Corollary 1.2. *Let X be a normal proper variety over a finite field k. Then $\pi_1^{\mathrm{ab}}(X, D)^0$ is finite, where*

$$\pi_1^{\mathrm{ab}}(X, D)^0 = \mathrm{Ker}\big(\pi_1^{\mathrm{ab}}(X, D) \to \pi_1^{\mathrm{ab}}(\mathrm{Spec}(k))\big).$$

Proof. In case X and $X \setminus U$ satisfy the assumption of Theorem 1.1, the corollary follows from the corresponding statement for curves. The finiteness in the curves case is a consequence of class field theory. For the general case, one can take by [de Jong 1996] an alteration $f : X' \to X$ such that X' and $X' \setminus U'$ with $U' = f^{-1}(U)$ satisfy the assumption of Theorem 1.1. Then the assertion follows from the fact that the map $f_* : \pi_1^{\mathrm{ab}}(U') \to \pi_1^{\mathrm{ab}}(U)$ has a finite cokernel. □

Corollary 1.2 can also be deduced from [Raskind 1995, Theorem 6.2]. It has recently been generalized to the noncommutative setting by Deligne; see [Esnault and Kerz 2012].

Theorem 1.1 is a central ingredient in our paper [Kerz and Saito 2013]. There we use it to construct a reciprocity isomorphism between a Chow group of zero cycles with modulus and the abelian fundamental group with bounded ramification. In fact, Theorem 1.1 allows us to restrict to surfaces in this construction.

2. Review of ramification theory

First we review local ramification theory. Let K denote a henselian discrete valuation field of $\mathrm{ch}(K) = p > 0$ with the ring $\mathcal{O}_K$ of integers and residue field κ. Let π be a prime element of $\mathcal{O}_K$ and $\mathfrak{m}_K = (\pi) \subset \mathcal{O}_K$ the maximal ideal. By the Artin–Schreier–Witt theory, we have a natural isomorphism for $s \in \mathbb{Z}_{\geq 1}$,

$$\delta_s : W_s(K)/(1-F)W_s(K) \xrightarrow{\sim} H^1(K, \mathbb{Z}/p^s\mathbb{Z}), \tag{2-1}$$

where $W_s(K)$ is the ring of Witt vectors of length s and F is the Frobenius. We have the Brylinski–Kato filtration indexed by integers $m \geq 0$

$$\mathrm{fil}_m^{\log} W_s(K) = \{(a_{s-1}, \dots, a_1, a_0) \in W_s(K) \mid p^i v_K(a_i) \geq -m\},$$

where v_K is the normalized valuation of K. In this paper, we use its nonlog version introduced by Matsuda [1997]:

$$\mathrm{fil}_m W_s(K) = \mathrm{fil}_{m-1}^{\log} W_s(K) + V^{s-s'} \mathrm{fil}_m^{\log} W_{s'}(K),$$

where $s' = \min\{s, \mathrm{ord}_p(m)\}$. We define ramification filtrations on $H^1(K) := H^1(K, \mathbb{Q}/\mathbb{Z})$ as

$$\mathrm{fil}_m^{\log} H^1(K) = H^1(K)\{p'\} \oplus \bigcup_{s\geq 1} \delta_s(\mathrm{fil}_m^{\log} W_s(K)) \quad (m \geq 0),$$
$$\mathrm{fil}_m H^1(K) = H^1(K)\{p'\} \oplus \bigcup_{s\geq 1} \delta_s(\mathrm{fil}_m W_s(K)) \quad (m \geq 1),$$

where $H^1(K)\{p'\}$ is the prime-to-p part of $H^1(K)$. We note that this filtration is shifted by one from the filtration of Matsuda [1997, Definition 3.1.1]. We also let $\mathrm{fil}_0 H^1(K)$ be the subgroup of all unramified characters.

Definition 2.1. For $\chi \in H^1(K)$, we denote the minimal m with $\chi \in \mathrm{fil}_m H^1(K)$ by $\mathrm{art}_K(\chi)$ and call it the Artin conductor of χ.

We have the following facts (cf. [Kato 1989; Matsuda 1997]):

Lemma 2.2. (1) $\mathrm{fil}_1 H^1(K)$ *is the subgroup of tamely ramified characters.*

(2) $\mathrm{fil}_m H^1(K) \subset \mathrm{fil}_m^{\log} H^1(K) \subset \mathrm{fil}_{m+1} H^1(K)$.

(3) $\mathrm{fil}_m H^1(K) = \mathrm{fil}_{m-1}^{\log} H^1(K)$ *if* $(m, p) = 1$.

The structure of graded quotients

$$\mathrm{gr}_m H^1(K) = \mathrm{fil}_m H^1(K)/\mathrm{fil}_{m-1} H^1(K) \quad (m > 1)$$

is described as follows. Let Ω^1_K be the absolute Kähler differential module, and put

$$\mathrm{fil}_m\,\Omega^1_K = \mathfrak{m}_K^{-m} \otimes_{\mathbb{O}_K} \Omega^1_{\mathbb{O}_K}.$$

We have an isomorphism

$$\mathrm{gr}_m\,\Omega^1_K = \mathrm{fil}_m\,\Omega^1_K / \mathrm{fil}_{m-1}\,\Omega^1_K \simeq \mathfrak{m}_K^{-m}\Omega^1_{\mathbb{O}_K} \otimes_{\mathbb{O}_K} \kappa. \tag{2-2}$$

We have the maps

$$F^s d : W_s(K) \to \Omega^1_K, \quad (a_{s-1}, \dots, a_1, a_0) \mapsto \sum_{i=0}^{s-1} a_i^{p^i-1} da_i,$$

and one can check $F^s d(\mathrm{fil}_n\,W_s(K)) \subset \mathrm{fil}_n\,\Omega^1_K$.

Theorem 2.3 [Matsuda 1997]. *The maps $F^s d$ factor through δ_s and induce a natural map*

$$\mathrm{fil}_n\,H^1(K) \to \mathrm{fil}_n\,\Omega^1_K,$$

which induces for $m > 1$ an injective map (called the refined Artin conductor for K)

$$\mathrm{art}_K : \mathrm{gr}_n\,H^1(K) \hookrightarrow \mathrm{gr}_n\,\Omega^1_K. \tag{2-3}$$

Next we review global ramification theory. Let X and C be as in the introduction, and fix a Cartier divisor D with $|D| \subset C$. We recall the definition of $\pi_1^{\mathrm{ab}}(X, D)$. We write $H^1(U)$ for the étale cohomology group $H^1(U, \mathbb{Q}/\mathbb{Z})$, which is identified with the group of continuous characters $\pi_1^{\mathrm{ab}}(U) \to \mathbb{Q}/\mathbb{Z}$.

Definition 2.4. We define $\mathrm{fil}_D\,H^1(U)$ to be the subgroup of $\chi \in H^1(U)$ satisfying this condition. For all integral curves $Z \subset X$ not contained in C, its restriction $\chi|_Z : \pi_1^{\mathrm{ab}}(Z) \to \mathbb{Q}/\mathbb{Z}$ satisfies the following inequality of Cartier divisors on Z^N:

$$\sum_{y \in Z_\infty} \mathrm{art}_y(\chi|_Z)[y] \le \phi_Z^* D,$$

where $\mathrm{art}_y(\chi|_Z) \in \mathbb{Z}_{\ge 0}$ is the Artin conductor of $\chi|_Z$ at $y \in Z_\infty$ and $\phi_Z^* D$ is the pullback of D by the natural map $\phi_Z : Z^N \to X$. Define

$$\pi_1^{\mathrm{ab}}(X, D) = \mathrm{Hom}(\mathrm{fil}_D\,H^1(U), \mathbb{Q}/\mathbb{Z}), \tag{2-4}$$

endowed with the usual profinite topology of the dual.

For the rest of this section, we assume that X is smooth and C is a simple normal crossing. Let I be the set of generic points of C, and let $C_\lambda = \overline{\{\lambda\}}$ for $\lambda \in I$. Write

$$D = \sum_{\lambda \in I} m_\lambda C_\lambda. \tag{2-5}$$

For $\lambda \in I$, let K_λ be the henselization of $K = k(X)$ at λ. Note that K_λ is a henselian discrete valuation field with residue field $k(C_\lambda)$.

Proposition 2.5. *We have*

$$\mathrm{fil}_D H^1(U) = \mathrm{Ker}\Big(H^1(U) \to \bigoplus_{\lambda \in I} H^1(K_\lambda)/\, \mathrm{fil}_{m_\lambda} H^1(K_\lambda)\Big).$$

Proof. This is a consequence of ramification theory developed in [Kato 1989; Matsuda 1997]. See [Kerz and Saito 2013, Corollary 2.7] for a proof. □

Proposition 2.6. *Fix $\lambda \in I$ such that $m_\lambda > 1$ in* (2-5). *The refined Artin conductor* art_{K_λ} (*cf. Theorem 2.3*) *induces a natural injective map*

$$\mathrm{art}_{C_\lambda} : \mathrm{fil}_D H^1(U)/\, \mathrm{fil}_{D-C_\lambda} H^1(U) \hookrightarrow H^0(C_\lambda, \Omega^1_X(D) \otimes_{\mathcal{O}_X} \mathcal{O}_{C_\lambda}),$$

which is compatible with pullback along maps $f : X' \to X$ of smooth varieties with the property that $f^{-1}(C)$ is a reduced simple normal crossing divisor.

Proof. This follows from the integrality result [Matsuda 1997, Corollary 4.2.2] of the refined Artin conductor. □

Proposition 2.6 motivates us to introduce the following log-variant of $\mathrm{fil}_D H^1(U)$:

Definition 2.7. We define $\mathrm{fil}^{\log}_D H^1(U)$ as

$$\mathrm{fil}^{\log}_D H^1(U) = \mathrm{Ker}\Big(H^1(U) \to \bigoplus_{\lambda \in I} H^1(K_\lambda)/\, \mathrm{fil}^{\log}_{m_\lambda} H^1(K_\lambda)\Big).$$

Lemma 2.8. (1) $\mathrm{fil}_C H^1(U)$ *is the subgroup of tamely ramified characters.*

(2) $\mathrm{fil}_D H^1(U) \subset \mathrm{fil}^{\log}_D H^1(U) \subset \mathrm{fil}_{D+C} H^1(U)$.

(3) $\mathrm{fil}_D H^1(U) = \mathrm{fil}^{\log}_{D-C} H^1(U)$ *if $(m_\lambda, p) = 1$ for all $\lambda \in I$.*

Proof. This is a direct consequence of Lemma 2.2. □

3. Proof of the main theorem

Let X be a smooth projective variety over a perfect field of characteristic $p > 0$ and $C \subset X$ a reduced simple normal crossing divisor on X. Let $j : U = X \setminus C \subset X$ be the open immersion. We use the same notation as in the last part of the previous section. Take an effective Cartier divisor

$$D = \sum_{\lambda \in I} m_\lambda C_\lambda \quad \text{with } m_\lambda \geq 0.$$

Let $I' = \{\lambda \in I \mid p \mid m_\lambda\}$, and put

$$D' = \sum_{\lambda \in I'} (m_\lambda + 1) C_\lambda + \sum_{\lambda \in I \setminus I'} m_\lambda C_\lambda.$$

Let Y be a smooth hypersurface section such that $Y \times_X C$ is a reduced simple normal crossing divisor on Y.

Definition 3.1. (1) Assuming $\dim(X) \geq 3$, we say that Y is sufficiently ample for (X, D) if the following conditions hold:

(A1) $H^i(X, \Omega^d_X(-\Xi + Y)) = 0$ for any effective Cartier divisor $\Xi \leq D$ and for $i = d, d-1, d-2$.

(A2) For any $\lambda \in I'$, we have

$$H^0(C_\lambda, \Omega^1_X(D' - Y) \otimes \mathcal{O}_{C_\lambda}) = H^0(C_\lambda, \mathcal{O}_{C_\lambda}(D' - Y)) = H^1(C_\lambda, \mathcal{O}_{C_\lambda}(D' - 2Y)) = 0.$$

(2) Assuming $\dim(X) = 2$, we say that Y is sufficiently ample for (X, D) if the following condition holds:

(B) $H^i(X, \Omega^d_X(-\Xi + Y)) = 0$ for any effective Cartier divisor $\Xi \leq D$ and for $i = 1, 2$.

We remark that there is an integer N such that any smooth Y of degree $\geq N$ is sufficiently ample for (X, D).

Theorem 1.1 is a direct consequence of the following:

Theorem 3.2. *Let Y be sufficiently ample for (X, D). Write $E = Y \times_X D$.*

(1) *Assuming $d := \dim(X) \geq 3$, we have isomorphisms*

$$\mathrm{fil}_D H^1(U) \xrightarrow{\sim} \mathrm{fil}_E H^1(U \cap Y) \quad \textit{and} \quad \mathrm{fil}^{\log}_D H^1(U) \xrightarrow{\sim} \mathrm{fil}^{\log}_E H^1(U \cap Y).$$

(2) *Assuming $d = 2$, we have injections*

$$\mathrm{fil}_D H^1(U) \hookrightarrow \mathrm{fil}_E H^1(U \cap Y) \quad \textit{and} \quad \mathrm{fil}^{\log}_D H^1(U) \hookrightarrow \mathrm{fil}^{\log}_E H^1(U \cap Y).$$

For an abelian group M, we let $M\{p'\}$ denote the prime-to-p torsion part of M.

Lemma 3.3. (1) *Assuming $d := \dim(X) \geq 3$, we have an isomorphism*

$$\mathrm{fil}_D H^1(U)\{p'\} \xrightarrow{\sim} \mathrm{fil}_E H^1(U \cap Y)\{p'\}$$

and the same isomorphism for $\mathrm{fil}^{\log}_D$.

(2) *Assuming $d = 2$, we have an injection*

$$\mathrm{fil}_D H^1(U)\{p'\} \hookrightarrow \mathrm{fil}_E H^1(U \cap Y)\{p'\}$$

and the same injection for $\mathrm{fil}^{\log}_D$.

Proof. Noting

$$\mathrm{fil}_D H^1(U)\{p'\} = \mathrm{fil}_C H^1(U)\{p'\} = \mathrm{fil}^{\log}_C H^1(U)\{p'\} = \mathrm{fil}^{\log}_D H^1(U)\{p'\},$$

this follows from the tame case of Theorem 1.1 due to [Schmidt and Spieß 2000]. □

By the above lemma, Theorem 3.2 is reduced to the following:

Theorem 3.4. *Let the assumption be as in Theorem 3.2. Take an integer $n > 0$.*

(1) *Assuming $d := \dim(X) \geq 3$, we have isomorphisms*

$$\mathrm{fil}_D\, H^1(U)[p^n] \xrightarrow{\sim} \mathrm{fil}_E\, H^1(U \cap Y)[p^n]$$

and the same isomorphism for $\mathrm{fil}_D^{\log}$.

(2) *Assuming $d = 2$, we have an injection*

$$\mathrm{fil}_D\, H^1(U)[p^n] \hookrightarrow \mathrm{fil}_E\, H^1(U \cap Y)[p^n]$$

and the same injection for $\mathrm{fil}_D^{\log}$.

In what follows, we consider an effective Cartier divisor with $\mathbb{Z}[1/p]$-coefficient:

$$D = \sum_{\lambda \in I} m_\lambda C_\lambda, \quad m_\lambda \in \mathbb{Z}[1/p]_{\geq 0}.$$

We put

$$[D] = \sum_{\lambda \in I} [m_\lambda] C_\lambda \quad \text{with } [m_\lambda] = \max\{i \in \mathbb{Z} \mid i \leq m_\lambda\}$$

and $\mathcal{F}(\pm D) = \mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{O}_X(\pm[D])$ for an $\mathcal{O}_X$-module. For D as above, let $\mathrm{fil}_D^{\log} W_n\mathcal{O}_X$ be the subsheaf of $j_* W_n \mathcal{O}_U$ of local sections

$$\underline{a} \in W_n\mathcal{O}_U \quad \text{such that } \underline{a} \in \mathrm{fil}_{m_\lambda}^{\log} W_n(K_\lambda) \text{ for any } \lambda \in I,$$

where $\mathrm{fil}_{m_\lambda}^{\log} W_n(K_\lambda) := \mathrm{fil}_{[m_\lambda]}^{\log} W_n(K_\lambda)$ is defined in Section 2 for the henselization K_λ of $K = k(X)$ at λ. We note

$$\mathcal{O}_X(D) = \mathrm{fil}_D^{\log} W_n\mathcal{O}_X \quad \text{for } n = 1.$$

The following facts are easily checked:

- The Frobenius F induces $F : \mathrm{fil}_{D/p}^{\log} W_n\mathcal{O}_X \to \mathrm{fil}_D^{\log} W_n\mathcal{O}_X$.
- The Verschiebung V induces $V : \mathrm{fil}_D^{\log} W_{n-1}\mathcal{O}_X \to \mathrm{fil}_D^{\log} W_n\mathcal{O}_X$.
- The restriction R induces $R : \mathrm{fil}_D^{\log} W_n\mathcal{O}_X \to \mathrm{fil}_{D/p}^{\log} W_{n-1}\mathcal{O}_X$.
- The following sequence is exact:

$$0 \to \mathcal{O}_X(D) \xrightarrow{V^{n-1}} \mathrm{fil}_D^{\log} W_n\mathcal{O}_X \xrightarrow{R} \mathrm{fil}_{D/p}^{\log} W_{n-1}\mathcal{O}_X \to 0. \tag{3-1}$$

We define an object $(\mathbb{Z}/p^n\mathbb{Z})_{X|D}$ of the derived category $D^b(X)$ of bounded complexes of étale sheaves on X:

$$(\mathbb{Z}/p^n\mathbb{Z})_{X|D} = \mathrm{Cone}\big(\mathrm{fil}_{D/p}^{\log} W_n\mathcal{O}_X \xrightarrow{1-F} \mathrm{fil}_D^{\log} W_n\mathcal{O}_X\big)[-1].$$

We have a distinguished triangle in $D^b(X)$:

$$(\mathbb{Z}/p^n\mathbb{Z})_{X|D} \to \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X \xrightarrow{1-F} \mathrm{fil}^{\log}_{D} W_n\mathbb{O}_X \xrightarrow{+} . \tag{3-2}$$

Lemma 3.5. *There is a distinguished triangle*

$$(\mathbb{Z}/p\mathbb{Z})_{X|D} \to (\mathbb{Z}/p^n\mathbb{Z})_{X|D} \to (\mathbb{Z}/p^{n-1}\mathbb{Z})_{X|D/p} \xrightarrow{+} .$$

Proof. The lemma follows from the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{O}_X(D/p) \arrow[r, "V^{n-1}"] \arrow[d, "1-F"] & \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X \arrow[r, "R"] \arrow[d, "1-F"] & \mathrm{fil}^{\log}_{D/p^2} W_{n-1}\mathbb{O}_X \arrow[r] \arrow[d, "1-F"] & 0 \\
0 \arrow[r] & \mathbb{O}_X(D) \arrow[r, "V^{n-1}"] & \mathrm{fil}^{\log}_{D} W_n\mathbb{O}_X \arrow[r, "R"] & \mathrm{fil}^{\log}_{D/p} W_{n-1}\mathbb{O}_X \arrow[r] & 0
\end{tikzcd}$$

□

Lemma 3.6. *There is a canonical isomorphism*

$$\mathrm{fil}^{\log}_D H^1(U)[p^n] \simeq H^1(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}).$$

Proof. Noting that the restriction of $(\mathbb{Z}/p^n\mathbb{Z})_{X|D}$ to U is $\mathbb{Z}/p^n\mathbb{Z}$ on U, we have the localization exact sequence

$$H^1(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \to H^1(U, \mathbb{Z}/p^n\mathbb{Z}) \to H^2_C(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}). \tag{3-3}$$

For the generic point λ of C_λ, (3-2) gives us an exact sequence

$$\begin{aligned} H^1_\lambda(X, \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X) \xrightarrow{1-F} H^1_\lambda(X, \mathrm{fil}^{\log}_{D} W_n\mathbb{O}_X) \\ \to H^2_\lambda(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \to H^2_\lambda(X, \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X). \end{aligned}$$

By [Grothendieck 1967, Corollary 3.10] and (3-1), we have

$$H^i_\lambda(X, \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X) = H^i_\lambda(X, \mathrm{fil}^{\log}_{D} W_n\mathbb{O}_X) = 0 \quad \text{for } i \geq 2$$

and

$$H^1_\lambda(X, \mathrm{fil}^{\log}_{D/p} W_n\mathbb{O}_X) \simeq W_n(K_\lambda)/\, \mathrm{fil}^{\log}_{m_\lambda/p} W_n(K_\lambda),$$
$$H^1_\lambda(X, \mathrm{fil}^{\log}_{D} W_n\mathbb{O}_X) \simeq W_n(K_\lambda)/\, \mathrm{fil}^{\log}_{m_\lambda} W_n(K_\lambda).$$

Thus, we get

$$H^2_\lambda(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \simeq H^1(K_\lambda)[p^n]/\, \mathrm{fil}^{\log}_{m_\lambda} H^1(K_\lambda)[p^n].$$

Hence, Lemma 3.6 follows from (3-3) and the injectivity of

$$H^2_C(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \to \bigoplus_{\lambda\in I} H^2_\lambda(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}).$$

This injectivity is a consequence of:

Claim 3.7. *For $x \in C$ with* $\dim(\mathcal{O}_{X,x}) \geq 2$, *we have*

$$H^2_x(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) = 0.$$

By Lemma 3.5, it suffices to show Claim 3.7 in case $n = 1$. Triangle (3-2) gives us an exact sequence

$$H^1_x(X, \mathcal{O}_X(D)) \to H^2_x(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) \to H^2_x(X, \mathcal{O}_X(D/p)) \xrightarrow{1-F} H^2_x(X, \mathcal{O}_X(D)).$$

If $\dim(\mathcal{O}_{X,x}) > 2$, $H^1_x(X, \mathcal{O}_X(D)) = 0$ and $H^2_x(X, \mathcal{O}_X(D/p)) = 0$ by [Grothendieck 1967, Corollary 3.10], which implies $H^2_x(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) = 0$ as desired.

We now assume $\dim(\mathcal{O}_{X,x}) = 2$. Let $(\mathbb{Z}/p\mathbb{Z})_X$ denote the constant sheaf $\mathbb{Z}/p\mathbb{Z}$ on X, and put

$$\mathcal{F}_{X|D} = \operatorname{Coker}\bigl(\mathcal{O}_X(D/p) \xrightarrow{1-F} \mathcal{O}_X(D)\bigr).$$

Note that $\mathcal{F}_{X|D} = 0$ for $D = 0$. By definition, we have a distinguished triangle

$$(\mathbb{Z}/p\mathbb{Z})_X \to (\mathbb{Z}/p\mathbb{Z})_{X|D} \to \mathcal{F}_{X|D} \xrightarrow{+} .$$

By [SGA 1 1971, Exposé X, Théorèm 3.1], we have $H^2_x(X, (\mathbb{Z}/p\mathbb{Z})_X) = 0$. Hence, we are reduced to showing

$$H^2_x(X, \mathcal{F}_{X|D}) = 0. \tag{3-4}$$

Without loss of generality, we can assume that D has integral coefficients. We prove (3-4) by induction on multiplicities of D reducing to the case $D = 0$. Fix an irreducible component C_λ of C with the multiplicity $m_\lambda \geq 1$ in D, and put $D' = D - C_\lambda$. We have a commutative diagram with exact rows and columns

$$\begin{tikzcd}
 & (\mathbb{Z}/p\mathbb{Z})_X \arrow[d] & (\mathbb{Z}/p\mathbb{Z})_X \arrow[d] & & \\
0 \arrow[r] & \mathcal{O}_X(D'/p) \arrow[r] \arrow[d, "1-F"] & \mathcal{O}_X(D/p) \arrow[r] \arrow[d, "1-F"] & \mathcal{L} \arrow[r] \arrow[d, "F"] & 0 \\
0 \arrow[r] & \mathcal{O}_X(D') \arrow[r] & \mathcal{O}_X(D) \arrow[r] & \mathcal{O}_{C_\lambda}(D) \arrow[r] & 0
\end{tikzcd}$$

Here $\mathcal{O}_{C_\lambda}(D) = \mathcal{O}_X(D) \otimes \mathcal{O}_{C_\lambda}$, and $\mathcal{L} = \mathcal{O}_{C_\lambda}(D/p)$ if $p \mid m_\lambda$, and $\mathcal{L} = 0$ otherwise. Thus, we get short exact sequences

$$0 \to \mathcal{F}_{X|D'} \to \mathcal{F}_{X|D} \to \mathcal{O}_{C_\lambda}(D) \to 0 \qquad \text{if } p \nmid m_\lambda,$$
$$0 \to \mathcal{F}_{X|D'} \to \mathcal{F}_{X|D} \to \mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(D/p)^p \to 0 \quad \text{if } p \mid m_\lambda.$$

We may assume $H^2_x(X, \mathcal{F}_{X|D'}) = 0$ by the induction hypothesis. Hence, (3-4) follows from

$$H^2_x(C_\lambda, \mathcal{O}_{C_\lambda}(D)) = 0, \tag{3-5}$$

$$H^2_x(C_\lambda, \mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(E)^p) = 0, \tag{3-6}$$

where we put $E = [D/p]$. We may assume $x \in C_\lambda$ so that $\dim(\mathcal{O}_{C_\lambda,x}) = 1$ by the assumption $\dim(\mathcal{O}_{X,x}) = 2$. Equation (3-5) is a consequence of [Grothendieck 1967, Corollary 3.10]. In view of an exact sequence

$$0 \to \mathcal{O}_{C_\lambda}(pE)/\mathcal{O}_{C_\lambda}(E)^p \to \mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(E)^p \to \mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(pE) \to 0,$$

(3-6) follows from

$$H^2_x(C_\lambda, \mathcal{O}_{C_\lambda}(pE)/\mathcal{O}_{C_\lambda}(E)^p) = 0 \quad \text{and} \quad H^2_x(C_\lambda, \mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(pE)) = 0.$$

The first assertion follows from [Grothendieck 1967, Corollary 3.10] noting that $\mathcal{O}_{C_\lambda}(pE)/\mathcal{O}_{C_\lambda}(E)^p$ is a locally free $\mathcal{O}^p_{C_\lambda}$-module. The second assertion holds since $\mathcal{O}_{C_\lambda}(D)/\mathcal{O}_{C_\lambda}(pE)$ is supported in a proper closed subscheme T of C_λ and x is a generic point of T if $x \in T$. This completes the proof of Lemma 3.6. □

Proof of Theorem 3.4. In view of the above results, the assertions for $\mathrm{fil}^{\log}_D$ of Theorem 3.4(1) and (2) follow from the following:

Theorem 3.8. *Let the assumption be as in Theorem 3.2. The natural map*

$$H^1(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \to H^1(Y, (\mathbb{Z}/p^n\mathbb{Z})_{Y|D})$$

is an isomorphism for $d := \dim(X) \geq 3$, *and it is injective for* $d = 2$.

Proof. By Lemma 3.5, we have a commutative diagram:

$$\begin{tikzcd}
0 \arrow[d] & 0 \arrow[d] \\
H^1(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) \arrow[r] \arrow[d] & H^1(Y, (\mathbb{Z}/p\mathbb{Z})_{Y|E}) \arrow[d] \\
H^1(X, (\mathbb{Z}/p^n\mathbb{Z})_{X|D}) \arrow[r] \arrow[d] & H^1(Y, (\mathbb{Z}/p^n\mathbb{Z})_{Y|D}) \arrow[d] \\
H^1(X, (\mathbb{Z}/p^{n-1}\mathbb{Z})_{X|D/p}) \arrow[r] \arrow[d] & H^1(Y, (\mathbb{Z}/p^{n-1}\mathbb{Z})_{Y|E/p}) \arrow[d] \\
H^2(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) \arrow[r] & H^2(Y, (\mathbb{Z}/p\mathbb{Z})_{Y|E})
\end{tikzcd}$$

The theorem follows by the induction on n from the following lemma. □

Lemma 3.9. *Let the assumption be as in Theorem 3.2.*

(1) *Assuming $d \geq 3$, the natural map*

$$H^i(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) \to H^i(Y, (\mathbb{Z}/p\mathbb{Z})_{Y|E})$$

is an isomorphism for $i = 1$ and injective for $i = 2$.

(2) *Assuming $d = 2$, the natural map*

$$H^1(X, (\mathbb{Z}/p\mathbb{Z})_{X|D}) \to H^1(Y, (\mathbb{Z}/p\mathbb{Z})_{Y|E})$$

is injective.

Proof. We define an object $\mathcal{K}$ of $D^b(X)$:

$$\mathcal{K} = \mathrm{Cone}\big(\mathcal{O}_X(D/p - Y) \xrightarrow{1-F} \mathcal{O}_X(D - Y)\big)[-1].$$

By the commutative diagram with exact horizontal sequences

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_X(D/p - Y) \arrow[r] \arrow[d, "1-F"] & \mathcal{O}_X(D/p) \arrow[r] \arrow[d, "1-F"] & \mathcal{O}_Y(E/p) \arrow[r] \arrow[d, "1-F"] & 0 \\
0 \arrow[r] & \mathcal{O}_X(D - Y) \arrow[r] & \mathcal{O}_X(D) \arrow[r] & \mathcal{O}_Y(E) \arrow[r] & 0
\end{tikzcd}$$

we have a distinguished triangle in $D^b(X)$:

$$\mathcal{K} \to (\mathbb{Z}/p\mathbb{Z})_{X|D} \to (\mathbb{Z}/p\mathbb{Z})_{Y|E} \xrightarrow{+} .$$

Hence, it suffices to show $H^i(X, \mathcal{K}) = 0$ for $i = 1, 2$ in case $d \geq 3$ and $H^1(X, \mathcal{K}) = 0$ in case $d = 2$. We have an exact sequence

$$\begin{aligned} H^0(\mathcal{O}_X(D - Y)) \to H^1(X, \mathcal{K}) \to H^1(\mathcal{O}_X(D/p - Y)) \\ \to H^1(\mathcal{O}_X(D - Y)) \to H^2(X, \mathcal{K}) \to H^2(\mathcal{O}_X(D/p - Y)). \end{aligned}$$

By Serre duality, for a divisor Ξ on X, we have

$$H^i(X, \mathcal{O}_X(\Xi - Y)) = H^{d-i}(X, \Omega_X^d(-\Xi + Y))^\vee.$$

Thus, the desired assertion follows from Definition 3.1(A1) and (B). □

It remains to deduce the assertions for fil_D of Theorem 3.4(1) and (2) from those for $\mathrm{fil}_D^{\log}$. Let D' be as in the beginning of this section and $E' = D' \times_X Y$. Noting that the multiplicities of D' are prime to p, we have by Lemma 2.8(3)

$$\mathrm{fil}_{D'} H^1(U) = \mathrm{fil}^{\log}_{D'-C} H^1(U) \quad \text{and} \quad \mathrm{fil}_{E'} H^1(U \cap Y) = \mathrm{fil}^{\log}_{E'-C\cap Y} H^1(U \cap Y).$$

Thus, the assertions for $\mathrm{fil}^{\log}_{D'-C}$ of Theorem 3.4 imply that for $\mathrm{fil}_{D'}$. Since $\mathrm{fil}_D \subset \mathrm{fil}_{D'}$, it immediately implies the injectivity of

$$\mathrm{fil}_D\, H^1(U) \to \mathrm{fil}_E\, H^1(U \cap Y).$$

It remains to deduce its surjectivity from that of

$$\mathrm{fil}_{D'}\, H^1(U) \to \mathrm{fil}_{E'}\, H^1(U \cap Y)$$

assuming $d \geq 3$. For this it, suffices to show the injectivity of

$$\mathrm{fil}_{D'}\, H^1(U)/\,\mathrm{fil}_D\, H^1(U) \to \mathrm{fil}_{E'}\, H^1(U \cap Y)/\,\mathrm{fil}_E\, H^1(U \cap Y).$$

By Proposition 2.6, we have a commutative diagram

$$\begin{tikzcd}
\mathrm{fil}_{D'}\, H^1(U)/\,\mathrm{fil}_D\, H^1(U) \arrow[r, hook] \arrow[d] & \bigoplus_{\lambda \in I'} H^0(C_\lambda, \Omega^1_X(D') \otimes_{\mathcal{O}_X} \mathcal{O}_{C_\lambda}) \arrow[d] \\
\mathrm{fil}_{E'}\, H^1(U \cap Y)/\,\mathrm{fil}_E\, H^1(U \cap Y) \arrow[r, hook] & \bigoplus_{\lambda \in I'} H^0(C_\lambda \cap Y, \Omega^1_Y(D') \otimes_{\mathcal{O}_Y} \mathcal{O}_{C_\lambda \cap Y})
\end{tikzcd}$$

Thus, we are reduced to showing the injectivity of the right vertical map. Putting $\mathscr{L} = \mathrm{Ker}(\Omega^1_X \to i_*\Omega^1_Y)$ where $i : Y \subset X$, the assertion follows from

$$H^0(C_\lambda, \mathscr{L}(D') \otimes_{\mathcal{O}_X} \mathcal{O}_{C_\lambda}) = 0.$$

Note that we used the fact that Y and C_λ intersect transversally. We have an exact sequence

$$0 \to \Omega^1_X(-Y) \to \mathscr{L} \to \mathcal{O}_X(-Y) \otimes \mathcal{O}_Y \to 0.$$

From this, we get an exact sequence

$$0 \to \Omega^1_X(D' - Y) \otimes_{\mathcal{O}_X} \mathcal{O}_{C_\lambda} \to \mathscr{L}(D') \otimes_{\mathcal{O}_X} \mathcal{O}_{C_\lambda} \to \mathcal{O}_{C_\lambda}(D' - Y) \otimes \mathcal{O}_{C_\lambda \cap Y} \to 0.$$

We also have an exact sequence

$$0 \to \mathcal{O}_{C_\lambda}(D' - 2Y) \to \mathcal{O}_{C_\lambda}(D' - Y) \to \mathcal{O}_{C_\lambda}(D' - Y) \otimes \mathcal{O}_{C_\lambda \cap Y} \to 0.$$

Therefore, the desired assertion follows from Definition 3.1(A2). This completes the proof of Theorem 3.4. $\square$

References

[Esnault and Kerz 2012] H. Esnault and M. Kerz, "A finiteness theorem for Galois representations of function fields over finite fields (after Deligne)", *Acta Math. Vietnam.* **37**:4 (2012), 531–562. MR 3058662 Zbl 06133191

[Grothendieck 1967] A. Grothendieck, *Local cohomology*, Lecture Notes in Mathematics **41**, Springer, Berlin, 1967. MR 37 #219 Zbl 0185.49202 [de Jong 1996] A. J. de Jong, "Smoothness, semi-stability and alterations", *Inst. Hautes Études Sci. Publ. Math.* **83** (1996), 51–93. MR 98e:14011 Zbl 0916.14005

[Kato 1989] K. Kato, "Swan conductors for characters of degree one in the imperfect residue field case", pp. 101–131 in *Algebraic K-theory and algebraic number theory* (Honolulu, 1987), edited by M. R. Stein and R. K. Dennis, Contemp. Math. **83**, Amer. Math. Soc., Providence, RI, 1989. MR 90g:11164 Zbl 0716.12006

[Kerz and Saito 2013] M. Kerz and S. Saito, "Chow group of 0-cycles with modulus and higher dimensional class field theory", preprint, 2013. arXiv 1304.4400

[Laumon 1981] G. Laumon, "Semi-continuité du conducteur de Swan (d'après P. Deligne)", pp. 173–219 in *The Euler–Poincaré characteristic*, edited by J.-L. Verdier, Astérisque **83**, Soc. Math. France, Paris, 1981. MR 83g:14007 Zbl 0504.14013

[Matsuda 1997] S. Matsuda, "On the Swan conductor in positive characteristic", *Amer. J. Math.* **119**:4 (1997), 705–739. MR 99h:14022 Zbl 0928.14017

[Raskind 1995] W. Raskind, "Abelian class field theory of arithmetic schemes", pp. 85–187 in *K-theory and algebraic geometry: connections with quadratic forms and division algebras* (Santa Barbara, 1992), part 1, edited by B. Jacob and A. Rosenberg, Proc. Sympos. Pure Math. **58**, Amer. Math. Soc., Providence, RI, 1995. MR 96b:11089 Zbl 0832.19004

[Schmidt and Spieß 2000] A. Schmidt and M. Spieß, "Singular homology and class field theory of varieties over finite fields", *J. Reine Angew. Math.* **527** (2000), 13–36. MR 2001m:11108 Zbl 0961.14013

[SGA 1 1971] A. Grothendieck, *Séminaire de Géométrie Algébrique du Bois Marie 1960–1961: revêtements étales et groupe fondamental*, Lecture Notes in Mathematics **224**, Springer, Berlin, 1971. MR 50 #7129 Zbl 0234.14002

[SGA 2 1968] A. Grothendieck, *Séminaire de Géométrie Algébrique du Bois-Marie 1962: cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux*, Advanced Studies in Pure Mathematics **2**, North-Holland, Amsterdam, 1968. MR 57 #16294 Zbl 0197.47202

Communicated by Jean-Louis Colliot-Thélène
Received 2013-04-27 Revised 2013-11-12 Accepted 2013-12-10

moritz.kerz@mathematik.uni-regensburg.de
Naturwissenschaftliche Fakultät I – Mathematik, Universität Regensburg, 93040 Regensburg, Germany

sshuji@msb.biglobe.ne.jp *Interactive Research Center of Science, Graduate School of Science and Engineering, Tokyo Institute of Technology, Ookayama, Meguro, Tokyo 152-8551, Japan*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.703

Localization of spherical varieties

Friedrich Knop

We prove some fundamental structural results for spherical varieties in arbitrary characteristic. In particular, we study Luna's two types of localization and use them to analyze spherical roots, colors, and their interrelation. At the end, we propose a preliminary definition of a p-spherical system.

1. Introduction

Let G be a connected reductive group defined over an algebraically closed ground field k of arbitrary characteristic p. A normal G-variety X is called *spherical* if a Borel subgroup B of G has an open orbit in X. In characteristic zero, there exists by now an extensive body of research on spherical varieties culminating in a complete classification [Luna and Vust 1983; Luna 2001; Losev 2009; Cupit-Foutou 2010; Bravi and Pezzini 2011a; 2011b; 2011c].

In positive characteristic, much less work has been done. Most papers dealing with spherical varieties in positive characteristic are restricted to particular examples (like flag or symmetric varieties) or other special classes of spherical varieties (like varieties obtained by reduction mod p).

This paper is part of a program to develop a general theory of spherical varieties in arbitrary characteristic, possibly also leading to a classification. In this sense, the old paper [Knop 1991] on a characteristic-free approach to spherical embeddings is already part of the program.

A crucial portion of Luna's theory of spherical varieties depends on Akhiezer's classification [1983] of spherical varieties of rank 1. In this paper, we present results which are independent of that classification. On the other hand, in the companion paper [Knop 2013], we determine all spherical varieties of rank 1 in arbitrary characteristic and present results whose proofs depend (so far) on it.

More precisely, in this paper we recover most of Luna's results [1997] on the "big cell". We start by generalizing Luna's fundamental relations for the colors of a spherical variety. At this point, we introduce additional data needed to describe a spherical variety in positive characteristic. These are certain p-powers $q_{D,\alpha}$, where

MSC2010: primary 14M27; secondary 14L30, 20G15, 14M17, 20G05.
Keywords: spherical varieties.

α is a simple root and D is a color "moved by α". Our exposition of this part is different from (and, we think, simpler than) Luna's, and seems to be new even in characteristic zero.

Next, we define the notion of spherical roots as properly normalized normal vectors to the valuation cone. Luna's method of viewing them as weights attached to a wonderful variety does not generalize.

Next, we consider Luna's construction [1997], called *localization at S*. Basically, it consists of analyzing the open Białynicki-Birula cell with respect to a dominant 1-parameter subgroup of G. Our results are more general than Luna's, even in characteristic zero, since Luna restricts his attention to wonderful varieties, while we formulate everything for so-called toroidal varieties. From this, we derive Luna's important result that the colors are, to a large extent, already determined by the spherical roots.

Then we consider a construction called *localization at Σ*. This procedure amounts to analyzing G-orbits of a toroidal variety. Since this technique is mostly classical, only the proof for the behavior of type-(a) colors is new. Unfortunately, our results remain somewhat incomplete, since it is unknown whether orbit closures in toroidal spherical embeddings are normal or not.

We use localization at Σ to prove the important nonpositivity result Corollary 6.6. Unlike Luna's proof, which uses Wasserman's tables [1996] of rank-2 varieties, our proof is conceptual.

Finally, in Section 7 we attempt to generalize Luna's notion of a spherical system to positive characteristic. This is a combinatorial structure describing the roots and the colors of a spherical variety. As additional data we propose the p-powers $q_{\alpha,D}$ mentioned above, and we hope that, at least for $p \neq 2, 3$, these data are enough to describe a spherical variety. As for the axioms, we restrict ourselves to those which immediately generalize axioms in characteristic zero. Conditions which are only meaningful in positive characteristic (like bounding the denominators of the pairings $\delta_D(\alpha)$) are deferred to future work.

So additional axioms will have to be added on at a later stage.

Notation. In the entire paper, the ground field k is algebraically closed. Its characteristic exponent is denoted by p, that is, $p = 1$ if $\operatorname{char} k = 0$ and $p = \operatorname{char} k$ otherwise. The group G is connected reductive, $B \subseteq G$ is a Borel subgroup, and $T \subseteq B$ is a maximal torus. Let $\Xi(T) = \Xi(B)$ be its character group. The set of simple roots with respect to B is denoted by $S \subset \Xi(T)$.

A rational function f on X is B-semiinvariant if there is a character $\chi_f \in \mathscr{X}(B) = \mathscr{X}(T)$ such that $f(b^{-1}x) = \chi_f(b) f(x)$ for all $b \in B$ and generic $x \in X$. If X is spherical, the character χ_f determines f up to a nonzero scalar. Let $\Xi(X) \subseteq \Xi(T)$ be the set of characters of the form χ_f. It is a finitely generated abelian group whose rank is called the rank of X. We also use $\Xi_{\mathbb{Q}}(X) := \Xi(X) \otimes \mathbb{Q}$ and $\Xi_p(X) := \Xi(X) \otimes \mathbb{Z}_p$ with $\mathbb{Z}_p := \mathbb{Z}[1/p]$.

2. Colors

Many properties of a spherical variety are determined by two sets of data and their interrelation: colors and valuations. We start with colors. Our results generalize those of [1997] in characteristic zero, but the approach is different. We do not use compactifications, but use the completeness of flag varieties instead.

Let X be a spherical G-variety with group of characters $\Xi(X)$, and let

$$N_{\mathbb{Q}}(X) = \mathrm{Hom}(\Xi(X), \mathbb{Q}). \tag{2-1}$$

A *color* of X is an irreducible divisor which is B- but not G-invariant. Every color D produces an element $\delta_D \in N_{\mathbb{Q}}(X)$ by

$$\delta_D(\chi_f) := v_D(f) \tag{2-2}$$

for all B-semiinvariants f. Here v_D is the valuation of $k(X)$ attached to D. The color D is, in general, not uniquely determined by δ_D.

Since X is spherical, we can choose a point $x_0 \in X$ such that Bx_0 is open and dense in X. Let $\Delta(X)$ be the set of colors of X. Since every color intersects the open G-orbit Gx_0, we have $\Delta(X) = \Delta(Gx_0) = \Delta(G/H)$, where $H = G_{x_0}$ is the isotropy subgroup scheme of x_0. We start by recalling a well known formula for the number of colors.

Proposition 2.1. *Let G be a semisimple group and $H \subseteq G$ a spherical subgroup. Then*

$$\#\Delta(G/H) = \mathrm{rk}\, G/H + \mathrm{rk}\, \Xi(H). \tag{2-3}$$

Proof. We compute the Picard group in two different ways. Set $X = G/H$. First, we have an exact sequence

$$\Xi(G) \to \mathrm{Pic}^G X \to \mathrm{Pic}\, X. \tag{2-4}$$

The group on the left is trivial since G is semisimple. The cokernel of the homomorphism on the right is torsion by [Sumihiro 1974]. On the other hand, $\mathrm{Pic}^G X = \Xi(H)$. Thus $\mathrm{rk}\,\mathrm{Pic}\, X = \mathrm{rk}\, \Xi(H)$. Now let $X_0 = Bx_0 \subseteq X$ be the open B-orbit. Then the colors are the irreducible components of $X \setminus X_0$. Thus we have an exact sequence

$$k^\times = \mathbb{O}(X)^\times \to \mathbb{O}(X_0)^\times \to \mathbb{Z}^a \to \mathrm{Pic}\, X \to \mathrm{Pic}\, X_0 = 0, \tag{2-5}$$

where $a = \#\Delta(X)$. By definition, $\mathrm{rk}\, \mathbb{O}(X_0)^\times / k^\times = \mathrm{rk}\, X$. Thus $\mathrm{rk}\,\mathrm{Pic}\, X = a - \mathrm{rk}\, X$. $\square$

Given a simple root $\alpha \in S$, one can construct colors as follows: let $P_\alpha \subseteq G$ be the minimal parabolic subgroup corresponding to α. Then $P_\alpha x_0$ is an open B-stable subset of X which, according to [Knop 1995b, Lemma 3.2], decomposes into at most three B-orbits. One of them is the open B-orbit Bx_0; the others are of codimension 1 in X, and hence their closures are colors. We say that these colors are *moved by* α. Clearly, this just means that $P_\alpha D \neq D$. In particular, every color is moved by some (not necessarily unique) simple root.

A more precise description is as follows. Let $H_\alpha := (P_\alpha)_{x_0} = H \cap P_\alpha$ such that $P_\alpha x_0 = P_\alpha / H_\alpha$. Then the B-orbits in $P_\alpha x_0$ correspond to H_α^{red}-orbits in $B\backslash P_\alpha \cong \boldsymbol{P}^1$. Let $\overline{H}_\alpha$ denote the image of H_α^{red} in Aut $\boldsymbol{P}^1 \cong \mathrm{PGL}(2)$. Then, up to conjugation, there are four possibilities for $\overline{H}_\alpha$:

Type of α	$\overline{H}_\alpha$	colors
(p)	G_0	—
(b)	S_0U_0	D
(a)	T_0	D, D'
$(2a)$	N_0	D

(2-6)

Here $G_0 = \mathrm{PGL}(2)$. The subgroups B_0, U_0, and T_0 of G_0 are a Borel subgroup, a maximal unipotent subgroup, and a maximal torus, respectively. Moreover, $S_0 \subseteq T_0$ and $N_0 = N_{G_0}(T_0)$. Thus, the set of simple roots decomposes as a disjoint union according to their type:

$$S = S^{(p)} \cup S^{(b)} \cup S^{(a)} \cup S^{(2a)}. \tag{2-7}$$

Observe that $\alpha \in S^{(p)}$ if and only if the open B-orbit Bx_0 is P_α-invariant. Thus, $S^{(p)}$ is the set of simple roots of the parabolic P_X, the stabilizer of the open B-orbit.

Let D be a color moved by α. Then the morphism

$$\varphi_{D,\alpha} : P_\alpha \times^B D \to X \tag{2-8}$$

is generically finite. Its separable degree is 1, that is, $\varphi_{D,\alpha}$ is bijective if and only if α is of type (b) or (a). It is 2 for α of type $(2a)$. The inseparable degree of $\varphi_{D,\alpha}$ will be denoted by $q_{D,\alpha} \in p^{\mathbb{N}}$.

Example. Assume $p > 3$ and let $P \subseteq G$ be a subgroup scheme which contains ${}^-B$, the Borel subgroup which is opposite to B. Wenzel [1993; 1994] showed that such subgroup schemes are classified by functions $f : S \to \mathbb{Z}_{\geq 0} \cup \{\infty\}$, where $f(\alpha)$ can be defined as the supremum of the set of all $n \in \mathbb{Z}_{\geq 0}$ such that P contains the n-th Frobenius kernel of P_α. See also [Haboush and Lauritzen 1993] for a simplified account. Now let $X = G/P$, a complete homogeneous G-variety. Then the following is easy to see: A simple root $\alpha \in S$ is of type (p) if and only if $f(\alpha) = \infty$. All other simple roots are of type (b) and they all move a different divisor D_α. Moreover, $q_{D_\alpha,\alpha} = p^{f(\alpha)}$. In particular, this shows that in this example the numbers $q_{D,\alpha}$ may be arbitrary p-powers.

To formulate the following permanence property we renormalize δ_D as follows:

$$\delta_D^{(\alpha)} := q_{D,\alpha}\delta_D. \tag{2-9}$$

Lemma 2.2. *Let $\pi : X_1 \to X_0$ be a finite surjective equivariant morphism between spherical G-varieties, let E be a color of X_1, and let $D = \pi(E)$ be its image in X_0. Let, moreover, $\alpha \in S$ be a simple root moving E (and D). Then $\delta_D^{(\alpha)} = \delta_E^{(\alpha)}$.*

Proof. We consider first the case that α is of type (a) or (b) for X_0. Then its type for X_1 is the same. Moreover, both $\varphi_{D,\alpha}$ and $\varphi_{E,\alpha}$ are bijective and, as an equality of divisors, $\pi^{-1}(D) = qE$, where q is some p-power. Thus $\delta_E = q\delta_D$. Now consider the diagram

$$\begin{array}{ccc} P_\alpha \times^B \pi^{-1}(D) & \longrightarrow & X_1 \\ \downarrow & & \downarrow \\ P_\alpha \times^B D & \xrightarrow{\varphi_{D,\alpha}} & X_0 \end{array} \tag{2-10}$$

It is cartesian, and hence both horizontal arrows have the same (inseparable) degree, namely $q_{D,\alpha}$. On the other hand, the top arrow has degree $q\, q_{E,\alpha}$. Hence

$$\delta_D^{(\alpha)} = q_{D,\alpha}\delta_D = qq_{E,\alpha}\delta_D = q_{E,\alpha}\delta_E = \delta_E^{(\alpha)}. \tag{2-11}$$

Now assume that α is of type $(2a)$ for X_0. Then there are two cases. If $\pi^{-1}(D)^{\mathrm{red}}$ is irreducible then α is of type $(2a)$ for X_1, as well. Moreover, the degree of both horizontal arrows is now $2q_{D,\alpha}$. From here one argues as above.

The second case is when α is of type (a) for X_0. Then $\pi^{-1}(D)^{\mathrm{red}} = E_1 \cup E_2$ has two components. As divisors, we have $\pi^{-1}(D) = q_1E_1 + q_2E_2$. Thus $\delta_{E_1} = q_1\delta_D$ and $\delta_{E_2} = q_2\delta_D$. Moreover, as above, we get

$$2\delta_D^{(\alpha)} = 2q_{D,\alpha}\delta_D = q_1q_{E_1,\alpha}\delta_D + q_2q_{E_2,\alpha}\delta_D = \delta_{E_1}^{(\alpha)} + \delta_{E_2}^{(\alpha)}. \tag{2-12}$$

Now we claim that actually $\delta_{E_1}^{(\alpha)} = \delta_{E_2}^{(\alpha)}$, which would prove our assertion.

To prove the claim, we may assume that $X_0 = G/H_0$ and $X_1 = G/H_1$ are homogeneous. Moreover, the cases proved above allow replacement of H_0 and H_1 by H_0^{red} and H_1^{red}, respectively. We can even replace H_1 by its connected component of unity since E cannot split any further (otherwise D would split into more than two components). Then H_1 is normal in H_0 and π is the quotient by the finite group $\Gamma := H_0/H_1$. Since Γ acts transitively on the connected components of the fibers of π, there is an element $g \in \Gamma$ which maps E_1 to E_2, which proves the claim. □

For any simple root α, let $\alpha^r \in N_{\mathbb{Q}}(X)$ be the restriction of $\alpha^\vee$ to $\Xi_{\mathbb{Q}}(X)$:

$$\alpha^r(\chi) = \langle \chi, \alpha^\vee \rangle \quad \text{for all } \chi \in \Xi_{\mathbb{Q}}(X). \tag{2-13}$$

Proposition 2.3. *Fix a simple root $\alpha \in S$. Then the following relations hold*:

$$
\begin{array}{ll}
\textit{Type of } \alpha & \\
\hline
(p) & \alpha^r = 0 \\
(b) & \delta_D^{(\alpha)} = \alpha^r \\
(a) & \delta_D^{(\alpha)} + \delta_{D'}^{(\alpha)} = \alpha^r \\
 & \alpha \in \Xi_p(X),\, \delta_D^{(\alpha)}(\alpha) = \delta_{D'}^{(\alpha)}(\alpha) = 1 \\
(2a) & \delta_D^{(\alpha)} = \frac{1}{2}\alpha^r
\end{array}
\tag{2-14}
$$

Proof. Let $X = G/H$ be homogeneous and put $X_1 = G/H^{0,\mathrm{red}}$. We claim that it suffices to prove the assertions for X_1. Indeed, if α is of type (p) for X, then the same holds for X_1 and the claim follows from $\Xi_{\mathbb{Q}}(X) = \Xi_{\mathbb{Q}}(X_1)$. If α is not of type (p) for X and X_1, then the claim follows immediately from Lemma 2.2 if α has the same type for X and X_1. Otherwise, α is of type $(2a)$ for X (moving one color D) and of type (a) for X_1 (moving two colors E_1, E_2). But then Lemma 2.2 implies

$$\delta_D^{(\alpha)} = \tfrac{1}{2}\big(\delta_{E_1}^{(\alpha)} + \delta_{E_2}^{(\alpha)}\big) = \tfrac{1}{2}\alpha^r, \tag{2-15}$$

proving the claim.

Thus we may assume that H is connected and reduced. Then consider the diagram

$$
\begin{tikzcd}
 & G \arrow[dl, "p_1"'] \arrow[dr, "p_2"] & \\
X = G/H & & B\backslash G =: \mathscr{F}
\end{tikzcd}
\tag{2-16}
$$

Both morphisms p_1 and p_2 are smooth with connected fibers. Therefore, an irreducible B-stable divisor $D \subset X$ corresponds to an irreducible H-stable divisor $E \subset \mathscr{F}$. Moreover, any B-semiinvariant rational function f on X corresponds to an H-invariant rational section s of the homogeneous line bundle $\mathscr{L}_\chi$ (with $\chi = \chi_f$) on $\mathscr{F}$. Furthermore, (D, f) is related to (E, s) by

$$v_E(s) = v_D(f) = \delta_D(\chi). \tag{2-17}$$

Now consider the $\boldsymbol{P}^1$-fibration $\pi : \mathscr{F} = B\backslash G \to \mathscr{F}_\alpha := P_\alpha\backslash G$. Moreover, let $y \in \mathscr{F}$ be in the open H-orbit and let $F \cong \boldsymbol{P}^1$ be the fiber through y.

Assume first that α is of type (p). Then the open B-orbit in X is P_α-stable, which translates into the open H-orbit in $\mathscr{F}$ being the preimage of an open set in $P_\alpha\backslash G$. But then $\mathscr{L}_\chi$ is a pull-back from $P_\alpha\backslash G$, which implies $\langle \chi_f, \alpha^\vee\rangle = 0$.

Now assume that $\alpha \in S^{(b)}$. Then E is the only H-invariant divisor mapping dominantly onto $\mathscr{F}_\alpha$. Moreover, since $E \cap F$ consists of a single point, the map $E \to \mathscr{F}_\alpha$ is generically bijective, and hence purely inseparable. Its degree is $q_{D,\alpha}$.

Thus we get

$$\langle\chi,\alpha^\vee\rangle=\deg\mathscr{L}_\chi|_F=(s)\cdot F=v_E(s)E\cdot F=q_{D,\alpha}\delta_D(\chi)=\delta_D^{(\alpha)}(\chi),\tag{2-18}$$

proving the assertion.

If $\alpha\in S^{(a)}$, then there are two divisors E, E' mapping generically injectively to $\mathscr{F}_\alpha$ with degree $q_{D,\alpha}$ and $q_{D',\alpha}$, respectively. Then

$$\langle\chi,\alpha^\vee\rangle=(s)\cdot F=\big(v_E(s)E+v_{E'}(s)E'\big)\cdot F=q_{D,\alpha}\delta_D(\chi)+q_{D',\alpha}\delta_{D'}(\chi).\tag{2-19}$$

Now we prove $\alpha\in\Xi_p(X)$. By construction, there is an equivariant morphism $P_\alpha x_0\twoheadrightarrow\mathrm{PGL}(2)/\tilde{H}$ with $\tilde{H}^{\mathrm{red}}=T_0$. Thus the pull-back of any nonconstant B_0-semiinvariant is a B-semiinvariant with character $q_0\alpha$ for some p-power q_0. The H_α-linearization of $\mathscr{L}_{q_0\alpha}|_F$ factors through a PGL(2)-linearization. One reason is, for example, that $\mathscr{L}_{-\alpha}$ is the relative canonical bundle of the fibration π. This implies that $s|_F$ has two zeroes of the same multiplicity on F. Hence $\delta_D^{(\alpha)}(\alpha)=\delta_{D'}^{(\alpha)}(\alpha)$, and therefore both are equal to 1.

Finally, assume that $\alpha\in S^{(2a)}$. Then there is one divisor E mapping generically $2:1$ to $\mathscr{F}_\alpha$. The degree of inseparability of this map is $q_{D,\alpha}$. Then $E\cdot F=2q_{D,\alpha}$, and therefore

$$\langle\chi,\alpha^\vee\rangle=(s)\cdot F=v_E(s)E\cdot F=2q_{D,\alpha}\delta_D(\chi)=2\delta_D^{(\alpha)}(\chi),\tag{2-20}$$

as claimed. □

We note the following consequence:

Corollary 2.4. *Let $p\neq 2$ and let $\alpha\in S$ be of type $(2a)$. Then $\alpha\notin\Xi_p(X)$ and $\langle\chi,\alpha^\vee\rangle$ is even for all $\chi\in\Xi(X)$.*

Proof. We keep the notation of the proof of Proposition 2.3. Let $N_0=\langle s_0\rangle T_0$ and let $n\in\mathbb{Z}$ with $n\alpha\in\Xi(X)$. Then s_0 acts on the T_0-invariant section of $\mathscr{L}_{q_0\alpha}|_F$ by multiplication with $(-1)^n$. Hence n is even and $\alpha\notin\Xi_p(X)$. The rest follows directly from Proposition 2.3. □

Now we analyze the case where a color is moved by more than one simple root.

Lemma 2.5. *Let D be a color which is moved by two distinct simple roots α_1 and α_2. Then either $\alpha_1,\alpha_2\in S^{(b)}$ or $\alpha_1,\alpha_2\in S^{(a)}$. In the latter case, let D' and D'' be the second color moved by α_1 and α_2, respectively. Then $D'\neq D''$.*

Proof. Clearly, neither α_1 nor α_2 is of type (p). Recall from [Knop 1995b, §2] that any B-orbit on X has a rank attached to it. Moreover, if $\alpha\in S$ moves the color D, then $\operatorname{rank}D=\operatorname{rank}X$ if α is of type (b), and $\operatorname{rank}D=\operatorname{rank}X-1$ in case α is of type (a) or $(2a)$ [Knop 1995b, §2 and Lemma 3.2]. This entails that α_1, α_2 are either both of type (b) or both of type (a) or $(2a)$.

Suppose they are both of type $(2a)$. Then, since $\alpha_1 \in \Xi_{\mathbb{Q}}(X)$,

$$0 < q_{D,\alpha_1}^{-1} = q_{D,\alpha_1}^{-1}\, \tfrac{1}{2}\alpha_1^r(\alpha_1) = \delta_D(\alpha_1) = q_{D,\alpha_2}^{-1}\, \tfrac{1}{2}\alpha_2^r(\alpha_1) \le 0. \tag{2-21}$$

Similarly, suppose α_1 is of type (a) and α_2 is of type $(2a)$. Then

$$0 < q_{D,\alpha_1}^{-1} = \delta_D(\alpha_1) = q_{D,\alpha_2}^{-1} \tfrac{1}{2}\alpha_2^r(\alpha_1) \le 0. \tag{2-22}$$

This finishes the proof of the first part.

Now let both α_1 and α_2 be of type (a) and suppose $D' = D''$. Then

$$0 < \frac{q_{D,\alpha_2}}{q_{D,\alpha_1}} + \frac{q_{D',\alpha_2}}{q_{D',\alpha_1}} = q_{D,\alpha_2}\delta_D(\alpha_1) + q_{D',\alpha_2}\delta_{D'}(\alpha_1) = \alpha_2^r(\alpha_1) \le 0. \qquad \square$$

Examples. 1. Let $G = \mathrm{SL}(2) \times \mathrm{SL}(2)$ and $H = \mathrm{SL}(2)$ embedded into G via $\mathrm{id} \times F_q$, where F_q is a Frobenius morphism. Then $X := G/H$ has only one color D. Moreover, both simple roots α_1, α_2 are of type (b), and D is moved by both of them. Furthermore, $q_{D,\alpha_1} = 1$ and $q_{D,\alpha_2} = q$, which shows that $q_{D,\alpha}$ may depend on α.

2. Let $G = \mathrm{SL}(3)$, let q be a p-power, and let H be the subgroup consisting of the matrices

$$\begin{pmatrix} t^{q+2} & & \\ & t^{q-1} & \\ & & t^{-2q-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & x & y \\ & 1 & x^q \\ & & 1 \end{pmatrix} \quad \text{with } t \in \boldsymbol{G}_m \text{ and } x, y \in \boldsymbol{G}_a. \tag{2-23}$$

Then both simple roots are of type (a) and there are three colors D_0, D_1, D_2 where α_i moves D_0 and D_i. Furthermore, $q_{D_1,\alpha_1} = q_{D_0,\alpha_2} = q_{D_2,\alpha_2} = 1$, while $q_{D_0,\alpha_1} = q$. The values $\delta_D(\alpha_i)$ are given by the following table:

	δ_{D_0}	δ_{D_1}	δ_{D_2}
α_1	q^{-1}	1	$-1-q^{-1}$
α_2	1	$-q-1$	1

(2-24)

So indeed $q\delta_{D_0} + \delta_{D_1} = \alpha_1^r$ and $\delta_{D_0} + \delta_{D_2} = \alpha_2^r$.

Both examples show that G contains for $p \ge 2$ infinitely many conjugacy classes of self-normalizing spherical subgroups, a phenomenon which does not occur in characteristic zero.

Remark. The Lemma shows, in particular, that one can assign unambiguously a type to any color. Thereby, one gets a decomposition

$$\Delta(X) = \Delta^{(b)}(X) \cup \Delta^{(a)}(X) \cup \Delta^{(2a)}(X). \tag{2-25}$$

3. Spherical roots

For a spherical variety X, let $\mathcal{V}(X)$ be the set of G-invariant $\mathbb{Q}$-valued valuations of the field $k(X)$. The map

$$\mathcal{V}(X) \to N_{\mathbb{Q}}(X) = \mathrm{Hom}(\Xi(X), \mathbb{Q}) : v \mapsto (\chi_f \mapsto v(f)) \tag{3-1}$$

is injective [Knop 1991, Corollary 1.8]. According to [ibid., Corollary 5.3], it identifies $\mathcal{V}(X)$ with a finitely generated convex rational cone inside $N_{\mathbb{Q}}(X)$ which contains the image of the antidominant Weyl chamber under the projection $\mathrm{Hom}(\Xi(T), \mathbb{Q}) \twoheadrightarrow N_{\mathbb{Q}}(X)$. This can be phrased in terms of the dual cone $\mathcal{V}(X)^\vee$ of $\mathcal{V}(X)$: it is a finitely generated rational convex cone in $\Xi_{\mathbb{Q}}(X)$ with $\mathcal{V}(X) = (\mathcal{V}(X)^\vee)^\vee$ and $-\mathcal{V}(X)^\vee \subseteq \mathbb{Q}_{\geq 0}S$, where $\mathbb{Q}_{\geq 0}S$ is the cone generated by the simple roots of G. In particular, $-\mathcal{V}(X)^\vee$ is a pointed cone. Thus, it has a canonical set of generators:

Definition 3.1. An element $\sigma \in \Xi_{\mathbb{Q}}(X)$ is called a *spherical root of* X if

- $\mathbb{Q}_{\geq 0}\sigma$ is an extremal ray of $-\mathcal{V}(X)^\vee$ (thus $\sigma \in \mathbb{Q}S$) and
- σ is a primitive element of $\mathbb{Z}S \cap \Xi_p(X)$.

The set of spherical roots is denoted by $\Sigma(X)$.

Clearly, each extremal ray of $-\mathcal{V}(X)^\vee$ contains a unique spherical root. Moreover, the spherical roots determine the valuation cone via

$$\mathcal{V}(X) = \{v \in N_{\mathbb{Q}}(X) : \sigma(v) \leq 0 \text{ for all } \sigma \in \Sigma(X)\}, \tag{3-2}$$

and are in bijection with faces of codimension 1 of $\mathcal{V}(X)$.

The normalization for a spherical root is chosen such that the following statement holds:

Lemma 3.2. *Let $\varphi : X_1 \to X_2$ be a morphism of spherical varieties which is either purely inseparable or a quotient by a central subgroup scheme of G. Then $\Sigma(X_1) = \Sigma(X_2)$.*

Proof. If φ is purely inseparable, then clearly $\mathcal{V}(X_1) = \mathcal{V}(X_2)$ and $\Xi_p(X_1) = \Xi_p(X_2)$. Hence $\Sigma(X_1) = \Sigma(X_2)$.

Now let φ be the quotient by $A \subseteq Z(G)$ (which might be positive dimensional). Then

$$\Xi_p(X_2) = \{\chi \in \Xi_p(X_1) : \chi|_A = 1\}. \tag{3-3}$$

Since roots of G are trivial on A, this implies

$$\mathbb{Z}S \cap \Xi_p(X_2) = \mathbb{Z}S \cap \Xi_p(X_1). \tag{3-4}$$

Also, $N_{\mathbb{Q}}(X_2)$ is a quotient of $N_{\mathbb{Q}}(X_1)$ and $\mathcal{V}(X_1)$ is the preimage of $\mathcal{V}(X_2)$ (this follows, for example, from [Knop 1991, Theorem 6.1]). Then $\Sigma(X_2) = \Sigma(X_1)$. $\square$

4. Localization at S

There are two types of constructions, called localization at S and at Σ, respectively, which both allow reduction of a spherical variety to a simpler one. In this section we describe localization at S, which, in characteristic zero, was first introduced by Luna [1997]. To this end, we first recall and prove some properties of the Białynicki-Birula decomposition [1976] of a $\boldsymbol{G}_m$-variety.

Let X be a complete normal $\boldsymbol{G}_m$-variety. Then for any $x \in X$, the limit

$$\pi(x) := \lim_{t\to 0} t \cdot x \in X \tag{4-1}$$

exists and is a $\boldsymbol{G}_m$-fixed point. Thus, letting F be the set of connected components of the fixed point set $X^{\boldsymbol{G}}_m$, we get a set partition of X by putting

$$X_Z := \{x \in X : \pi(x) \in Z\}. \tag{4-2}$$

These are the Białynicki-Birula cells which are indexed by F. Except when X is smooth or projective, they are, in general, not very well behaved. One cell is always good, though:

Proposition 4.1. *Let X be a complete normal $\boldsymbol{G}_m$-variety. Then there is a unique connected component S of $X^{\boldsymbol{G}}_m$ (the* source *of X) such that X_S is open. Moreover, the map $\pi_S := \pi_S : X_S \to S$ is affine and a categorical quotient by $\boldsymbol{G}_m$. In particular, the source S is irreducible and normal.*

Proof. The statement is well known. For example, it follows from the theory in [Białynicki-Birula and Święcicka 1982]: Let S be the source, that is, the connected component of $X^{\boldsymbol{G}}_m$ such that $\pi(x) \in S$ for $x \in X$ generic. By [Białynicki-Birula and Święcicka 1982, Proposition 2.3], the set $A^+ = \{S\}$ defines a sectional set. Now the assertion is [Białynicki-Birula and Święcicka 1982, Theorem 1.5]. □

Lemma 4.2. *Let X be as above. Then the general fibers of π_S are irreducible and generically reduced.*

Proof. Since X_S is normal, the generic fiber of π_S is geometrically unibranch [Grothendieck 1965, 6.15.6]. Since all irreducible components contain the $\boldsymbol{G}_m$-fixed point, it follows that the generic fiber is geometrically irreducible. Thus, there is an open subset of S over which all fibers are also geometrically irreducible [Jouanolou 1983, Theorem 4.10].

The second property follows from the fact that π_S is a categorical quotient. This entails that $k(S)$ is separable inside $k(X_S)$. Therefore, π_S is smooth generically on X_S. □

There is a second well-behaved cell. For this, let $X^- := X$ but with the opposite $\boldsymbol{G}_m$-action: $t * x := t^{-1} \cdot x$. Then the source T of X^- is called the *sink of X*.

It is characterized by the fact that $\pi(x) \in T$ implies $x \in T$. Thus, $T = X_T$ is a Białynicki-Birula cell by itself. Symmetrically, S is the sink of X^- and therefore $X_S^- = S$.

Now assume that X is a G-variety for some connected reductive group G and that the $\boldsymbol{G}_m$-action is induced by a 1-parameter subgroup $\lambda : \boldsymbol{G}_m \to G$. Then we put $X^\lambda := S$ and $X_\lambda := X_S$ and $\pi_\lambda = \pi_S$. Observe that X^λ is not the entire fixed point set of $\lambda(\boldsymbol{G}_m)$ but only a very special component.

Let $G^\lambda := C_G(\lambda(\boldsymbol{G}_m))$ be the fixed point set under the conjugation action and let

$$G_\lambda := \bigl\{g \in G : \pi_G(g) := \lim_{t\to 0} \lambda(t) g \lambda(t)^{-1} \text{ exists}\bigr\}. \tag{4-3}$$

Then G_λ is a parabolic subgroup with Levi complement G^λ and the map

$$\pi_G : G_\lambda \to G^\lambda$$

is the natural homomorphism with kernel G_λ^u, the unipotent radical of G_λ. The following lemma is well known; for example, the proof given in [Luna 1997] carries over verbatim to positive characteristic.

Lemma 4.3. *The open cell X_λ is G_λ-invariant and $\pi_\lambda : X_\lambda \to X^\lambda$ is G_λ-equivariant where G_λ acts on X^λ via $\pi_G : G_\lambda \to G^\lambda$, that is, $\pi_\lambda(gx) = \pi_G(g)\pi_\lambda(x)$ for all $g \in G_\lambda$ and $x \in X_\lambda$. Moreover, X^λ consists of fixed points for the opposite unipotent radical ${}^-G_\lambda^u$.*

We next provide a link between closed orbits in X and closed orbits in X^λ:

Lemma 4.4. *Let X be a complete normal G-variety and let $Z \subseteq X^\lambda$ be a closed G^λ-orbit. Then $GZ \subseteq X$ is a closed G-orbit with $GZ \cap X^\lambda = Z$.*

Proof. Since Z is complete and homogeneous, it contains a unique fixed point z for ${}^-B^\lambda$, the Borel subgroup opposite to B^λ. Since z is in the source, it is a ${}^-G_\lambda^u$-fixed point. So z is fixed by ${}^-B = {}^-B^\lambda\, {}^-G_\lambda^u$, which implies that $Gz = GZ$ is a complete, and hence closed, G-orbit of X. Moreover, since λ is dominant, we have $z \in (GZ)^\lambda$ and therefore $GZ \cap X^\lambda = (Gz)^\lambda = G^\lambda z = Z$. $\square$

Recall that a complete spherical G-variety X is called *toroidal* if no color of X contains a G-orbit.

Proposition 4.5. *Let λ be a dominant 1-parameter subgroup and let X be a complete toroidal spherical variety. Then X^λ is a complete toroidal spherical G^λ-variety.*

Proof. This is basically [Luna 1997, Proposition 1.4]. We recall the proof and check that it is characteristic-free.

First of all, completeness of X^λ is clear, while irreducibility and normality were obtained in Proposition 4.1. Moreover, the dominance of λ implies $B \subseteq G_\lambda$. Hence the open B-orbit of X is contained in X_λ. Its image in X^λ is an open B^λ-orbit. Thus, X^λ is spherical as well.

Now let $D \subseteq X^\lambda$ be a color containing the closed G^λ-orbit Z. Then $\pi_\lambda^{-1}(D)$ contains a unique component E' which maps onto D (by Lemma 4.2 and the fact that D meets the open G^λ-orbit). Then E, the closure of D in X, is a color which contains Z. Since GZ is a closed G-orbit by Lemma 4.4, we have $GZ = \overline{BZ}$, and therefore $GZ \subseteq E$. It follows that E is G-invariant since X is toroidal. From this we get that $D = \pi_\lambda(E)$ is G^λ-invariant, in contradiction to D being a color. $\square$

A toroidal spherical variety X determines a (pointed) fan $\mathcal{F} = \mathcal{F}(X)$ in $N_\mathbb{Q}(X)$ whose support is $\mathcal{V}(X)$. More precisely, for any G-orbit $Y \subseteq X$, the invariant valuations whose center in X contain Y form a cone $\mathcal{C}_Y \subseteq \mathcal{V}(X)$. Then $\mathcal{F}$ is the collection of cones of the form $\mathcal{C}_Y$.

The fan $\mathcal{F}$ is precisely the piece of information needed to reconstruct X from its open G-orbit X_0. In fact, X is the compactification of X_0 corresponding to the colored fan $(\mathcal{F}, \varnothing)$. See [Knop 1991] for more details.

Let λ be a dominant 1-parameter subgroup and let X be a complete toroidal spherical variety with associated fan $\mathcal{F}$. Then λ induces, via restriction to $\Xi_\mathbb{Q}(X)$, an element $\lambda^r \in N_\mathbb{Q}(X)$. The dominance of λ implies $-\lambda^r \in \mathcal{V}(X)$. Thus we can consider the fan

$$\tilde{\mathcal{F}}^\lambda := \{\mathcal{C} + \mathbb{Q}_{\geq 0}\lambda : -\lambda^r \in \mathcal{C} \in \mathcal{F}\}. \tag{4-4}$$

One may think of $\tilde{\mathcal{F}}^\lambda$ as the restriction of $\mathcal{F}$ to a neighborhood of $-\lambda^r$. Visibly, this fan is not pointed, since all its members contain the line $\mathbb{Q}\lambda^r$. More precisely, let $\mathcal{C}(\lambda)$ be the unique cone in $\mathcal{F}$ such that $-\lambda^r$ is contained in its relative interior. Then $V(\lambda) := \langle \mathcal{C}(\lambda)\rangle_\mathbb{Q} = \mathcal{C}(\lambda) + \mathbb{Q}_{\geq 0}\lambda^r$ is the unique element of $\tilde{\mathcal{F}}^\lambda$ which is a subspace. Thus,

$$\mathcal{F}^\lambda := \{\mathcal{C}/V(\lambda) : \mathcal{C} \in \tilde{\mathcal{F}}^\lambda\}. \tag{4-5}$$

is a pointed fan which lives in the vector space $N_\mathbb{Q}(X)/V(\lambda)$ and is called the *localization of $\mathcal{F}$ at λ.*

Theorem 4.6. *Let λ and X as in Proposition 4.5. Then $\Xi(X^\lambda) = \Xi(X) \cap V(\lambda)^\perp$, $N_\mathbb{Q}(X^\lambda) = N_\mathbb{Q}(X)/V(\lambda)$, and $\mathcal{F}(X^\lambda) = \mathcal{F}(X)^\lambda$.*

Proof. Let z be an arbitrary ${}^-B$-fixed point in X. Then z corresponds to the complete orbit Gz and therefore to a cone $\mathcal{C}_{Gz}$ of maximal dimension in $\mathcal{F}(X)$. Let P be the parabolic which is opposite to the reduced stabilizer G_z^{red}. Then the local structure theorem [Knop 1993, Satz 1.2] asserts the existence of a normal affine T-variety $\overline{A}$ and a T-equivariant morphism $\varphi_0 : \overline{A} \to X$ such that the morphism

$$\varphi : P_u \times \overline{A} \to X : (u, a) \mapsto u\varphi_0(a) \tag{4-6}$$

is finite onto an open neighborhood of z in X. Moreover, the torus T has an open orbit A in $\overline{A}$ such that $\Xi_{\mathbb{Q}}(A) = \Xi_{\mathbb{Q}}(X)$. The embedding $A \hookrightarrow \overline{A}$ corresponds to the cone $\mathscr{C}_{Gz}$. In particular, $\overline{A}$ contains a unique T-fixed point, denoted by 0, such that $\varphi_0(0) = z$.

From this we see that z lies in the source of λ on X if and only if λ has a source in $P_u \times \overline{A}$. This is automatically the case for P_u since λ is dominant and acts by conjugation. The fixed point set is $P_u^\lambda = P_u \cap G^\lambda$. On the other hand, we have

$$f_\chi(\lambda(t)a) = t^{-\lambda^r(\chi)} f_\chi(a). \tag{4-7}$$

For the limit when $t \to 0$ to exist for all $a \in \overline{A}$, it is necessary and sufficient that $-\lambda^r(\chi) \geq 0$ for all $\chi \in \Xi_{\mathbb{Q}}(\overline{A})$ with $f_\chi \in \mathbb{O}(\overline{A})$. This condition boils down to $-\lambda^r \in \mathscr{C}_{Gz}$. In that case, one readily checks that the fixed point set $\overline{A}^\lambda$ is the closure of the orbit corresponding to the face $\mathscr{C}(\lambda)$ of $\mathscr{C}_{Gz}$. The restricted morphism

$$\varphi^\lambda : P_u^\lambda \times \overline{A}^\lambda \to X^\lambda \tag{4-8}$$

describes the local structure of X^λ in a neighborhood of z.

From this we already infer that $\Xi_{\mathbb{Q}}(X^\lambda) = \Xi_{\mathbb{Q}}(\overline{A}^\lambda) = V(\lambda)^\perp$. We claim that $\Xi(X^\lambda) = \Xi_{\mathbb{Q}}(X^\lambda) \cap \Xi(X)$, which then proves the assertion $\Xi(X^\lambda) = \Xi(X) \cap V(\lambda)^\perp$. In fact, only "$\supseteq$" is an issue. To prove it, let $\chi \in \Xi_{\mathbb{Q}}(X^\lambda) \cap \Xi(X)$. Then there are $n \in \mathbb{Z}_{>0}$ and rational semiinvariants f_χ on X and $f_{n\chi}$ on X^λ such that $f_\chi^n = \pi_\lambda^* f_{n\chi}$. Let $X' \subseteq X^\lambda$ be the open subset on which $f_{n\chi}$ is regular. Then the normality of X implies that f_χ is regular on $\pi_\lambda^{-1}(X')$. Since f_χ is also λ-invariant, we conclude that f_χ pushes down to a rational function on X^λ, which shows $\chi \in \Xi(X^\lambda)$, as claimed. The equality $N_{\mathbb{Q}}(X^\lambda) = N_{\mathbb{Q}}(X)/V(\lambda)$ follows immediately.

Finally, we compute the fan $\mathscr{F}(X^\lambda)$. Clearly, it suffices to determine its cones $\mathscr{C}$ of maximal dimension corresponding to closed orbits. Lemma 4.4 and the discussion above show that the closed G^λ-orbits in X^λ correspond precisely to those closed G-orbits Gz in X such that $-\lambda^r \in \mathscr{C}_{Gz}$. In that case, it is easy to check that the toroidal embedding $A^\lambda \hookrightarrow \overline{A}^\lambda$ corresponds to the cone $(\mathscr{C}_{Gz} + \mathbb{Q}_{\geq 0}\lambda)/V(\lambda)$. But these are precisely the cones of maximal dimension in $\mathscr{F}^\lambda$, which shows $\mathscr{F}(X^\lambda) = \mathscr{F}(X)^\lambda$. □

Corollary 4.7. *Let λ and X be as above. Then*

$$\Sigma(X^\lambda) = \Sigma(X) \cap V(\lambda)^\perp = \Sigma(X) \cap \lambda^\perp. \tag{4-9}$$

Proof. The valuation cone $\mathscr{V}(X^\lambda)$ equals the support of $\mathscr{F}(X)^\lambda$. Its codimension-1 faces are, by construction, the codimension-1 faces of $\mathscr{V}(X)$ which contain $\mathscr{C}(\lambda)$. From $\Xi_p(X^\lambda) = \Xi_p(X) \cap V(\lambda)^\perp$ we get $\Sigma(X^\lambda) = \Sigma(X) \cap V(\lambda)^\perp$. The second equality follows from the fact that $\langle \sigma, \mathscr{V}(X)\rangle \geq 0$ for all $\sigma \in \Sigma(X)$. Hence $\langle \sigma, V(\lambda)\rangle = 0$ if and only if $\langle \sigma, \mathscr{C}(\lambda)\rangle = 0$ if and only if $\langle \sigma, \lambda\rangle = 0$. □

Proposition 4.8. *Let λ and X be as above. Then $GX^\lambda = \overline{Y}_0$, where $Y_0 \subseteq X$ is the G-orbit with $\mathscr{C}_{Y_0} = \mathscr{C}(\lambda)$. Moreover, for any G-orbit $Y \subseteq X$,*

$$\mathscr{C}(\lambda) \subseteq \mathscr{C}_Y \iff X^\lambda \cap Y \neq \varnothing, \tag{4-10}$$

$$\mathscr{C}(\lambda) \supseteq \mathscr{C}_Y \iff X^\lambda \subseteq \overline{Y}. \tag{4-11}$$

Proof. First note that X^λ is stable under ${}^-B = {}^-B^\lambda\,{}^-G^u_\lambda$. Hence GX^λ is closed, and hence an orbit closure $\overline{Y_0}$, in X. Choose a closed orbit Gz in $\overline{Y_0}$. The orbits of X which contain Gz in their closure correspond precisely to the T-orbits in the slice $\overline{A}$. It follows from (4-8) that in this way, $\overline{Y_0}$ corresponds to $\overline{A}^\lambda$, which shows $\mathscr{C}_{Y_0} = \mathscr{C}(\lambda)$, as claimed.

The two equivalences follow easily: we have $X^\lambda \cap Y \neq \varnothing$ if and only if $Y \subseteq GX^\lambda = \overline{Y_0}$ if and only if $\mathscr{C}_Y \supseteq \mathscr{C}(\lambda)$; and $X^\lambda \subseteq \overline{Y}$ if and only if $\overline{Y_0} = GX^\lambda \subseteq \overline{Y}$ if and only if $\mathscr{C}(\lambda) \supseteq \mathscr{C}_Y$. $\square$

Next we compute the colors of X^λ. Let $D \subseteq X^\lambda$ be a color and let D_0 be its restriction to the open G^λ-orbit. Then $\pi_\lambda^{-1}(D_0)$ is irreducible by Lemma 4.2. Hence its closure $D^* \subset X$ is a B-stable prime divisor.

Proposition 4.9. *Let λ and X be as above.*

a) *Let $\alpha \in S^\lambda := S(G^\lambda) = S \cap \lambda^\perp$ and let D be a color of X^λ. Then*
   i) *α has the same type for X^λ as it has for X,*
   ii) *δ_D is the restriction of δ_{D^*} to $\Xi_{\mathbb{Q}}(X^\lambda) \subseteq \Xi_{\mathbb{Q}}(X)$, and*
   iii) *D is moved by α if and only if D^* is moved by α. In that case $q_{D,\alpha} = q_{D^*,\alpha}$.*

b) *The map $D \mapsto D^*$ is a bijection between the set of colors of X^λ and the set of colors of X which are moved by some $\alpha \in S^\lambda$.*

Proof. The first part of iii) follows from $P_\alpha \subseteq G_\lambda$ and the equivariance of π_λ. Then b) is an immediate consequence. For ii), recall that $\pi_\lambda^{-1}(D_0)$ is even a reduced divisor by Lemma 4.2. Thus, $v_{D^*}(\pi_\lambda^* f) = v_D(f)$ for all $f \in k(X^\lambda)$. Moreover, there is a commutative diagram

$$\begin{array}{ccc}
P_\alpha \times^B D^* & \xrightarrow{\varphi_{D^*,\alpha}} & X \\
\uparrow & & \uparrow \\
P_\alpha \times^B \pi_\lambda^{-1}(D_0) & \longrightarrow & X_\lambda \\
\downarrow & & \downarrow{\scriptstyle \pi_\lambda} \\
P_\alpha \times^B D_0 & \longrightarrow & X^\lambda \\
\downarrow & & \| \\
P_\alpha^\lambda \times^{B^\lambda} D & \xrightarrow{\varphi^\lambda_{D,\alpha}} & X^\lambda
\end{array} \tag{4-12}$$

where the middle square is cartesian and the injections are open embeddings. It follows that $\varphi^{\lambda}_{D,\alpha}$ and $\varphi_{D^*,\alpha}$ have the same inseparable degree, showing the second part of iii). Both morphisms also have the same separable degree (1 or 2). Thus, α is of type $(2a)$ for X^{λ} if and only if it is of type $(2a)$ for X, proving part of i). The other types (p), (b), and (a) are distinguished by the number of colors (0, 1, and 2, respectively) moved by α. Thus, the rest of i) follows from iii). □

5. The interrelation of roots and colors

In practice, the 1-parameter subgroup λ has to be chosen diligently.

Lemma 5.1. *Let X be a complete toroidal spherical G-variety and $S' \subseteq S$. Then there is a 1-parameter subgroup λ such that*

a) *$S(G^{\lambda}) = S'$ and*

b) *the connected center of G^{λ} acts trivially on X^{λ}.*

Proof. Let $F \subseteq N_{\mathbb{Q}}(T) := \mathrm{Hom}(\Xi(T), \mathbb{Q})$ be the open face of the Weyl chamber defined by $\alpha = 0$ for all $\alpha \in S'$ and $\alpha > 0$ for $\alpha \in S \setminus S'$. Then $\lambda \in F$ is equivalent to a) (such λ are called *adapted* to S').

Now consider the projection $\pi : N_{\mathbb{Q}}(T) \twoheadrightarrow N_{\mathbb{Q}}(X)$. Since $\pi(F) \subseteq \mathcal{V}(X)$, the fan $\mathcal{F}$ associated to X induces a complete fan $\mathcal{F}'$ on $\pi(F)$. Let $F^0 := F \setminus \bigcup \pi^{-1}(\mathcal{C})$, where $\mathcal{C}$ runs through all $\mathcal{C} \in \mathcal{F}'$ with $\dim \mathcal{C} < \dim \pi(F)$, which is obviously a dense open subset of F. We claim that $\lambda \in F^0$ ensures the second property b).

Indeed, $\pi(\lambda)$ lies by construction in the relative interior of a maximal dimensional cone $\mathcal{C}'$ of $\mathcal{F}'$. This implies that $\pi(F) \subseteq V'$, where V' is the span of $\mathcal{C}'$. Now let $\mathcal{C} \in \mathcal{F}$ be minimal with $\mathcal{C}' \subseteq \mathcal{C}$. Then $\pi(\lambda)$ is also in the relative interior of $\mathcal{C}$. Thus, the subspace V spanned by $\mathcal{C}$ contains $\pi(F)$. Hence $\langle F \rangle_{\mathbb{Q}} \subseteq \pi^{-1}(V)$, which implies b). □

If the fan is changed, one can do better.

Lemma 5.2. *Let X_0 be a homogeneous spherical G-variety and $S' \subseteq S$. Then there is a 1-parameter subgroup λ and a toroidal compactification X of X_0 such that*

a) *$S(G^{\lambda}) = S'$ and*

b) *$\Xi_{\mathbb{Q}}(X^{\lambda}) = \langle \Sigma(X^{\lambda}) \rangle_{\mathbb{Q}}$.*

Proof. Same construction as above, but this time we choose $\mathcal{F}$ such that $\dim \mathcal{C}$ is as large as possible, that is, the dimension of the smallest face of $\mathcal{V}$ which contains $\mathcal{C}'$. This means precisely b). □

Remark. A rather trivial application of the last lemma is when $S' = S$. Then X^{λ} has the same roots and colors as X_0 but $\Xi_{\mathbb{Q}}(X^{\lambda})$ is spanned by $\Sigma(X)$.

Following Luna [1997], an important application of this technique is:

Proposition 5.3. *Let $\alpha \in S$ be a simple root. If $p \neq 2$, then*:

a) $\alpha \in \Sigma(X)$ *if and only if* α *is of type* (a). *Thus* $S^{(a)} = S \cap \Sigma(X)$.

b) $2\alpha \in \Sigma(X)$ *if and only if* α *is of type* $(2a)$. *Thus* $S^{(2a)} = S \cap \frac{1}{2}\Sigma(X)$.

If $p = 2$, *then*:

c) $\alpha \in \Sigma(X)$ *if and only if* α *is of type* (a) *or* $(2a)$. *Thus* $S^{(a)} \cup S^{(2a)} = S \cap \Sigma(X)$.

Proof. Without loss of generality, we may replace X by a toroidal compactification of its open G-orbit. The corresponding fan is denoted by $\mathcal{F}$. Choose λ as in Lemma 5.1 with $S' := \{\alpha\}$. Then G^λ acts on X^λ only via a semisimple quotient G_0 of rank 1. Let G_0/H_0 be the open G_0-orbit in X^λ.

Assume first $p \neq 2$. Then

$$\alpha \in S^{(a)}(X) \iff \alpha \in S^{(a)}(X^\lambda) \iff H_0^{\mathrm{red}} \sim T_0 \iff \alpha \in \Sigma(X^\lambda) \iff \alpha \in \Sigma(X).$$

This proves a).

The argument for b) is analogous, with T_0 replaced by $N(T_0)$. Finally, for $p = 2$ we argue with $\Sigma(G_0/T_0) = \Sigma(G_0/N(T_0)) = \{\alpha\}$. ☐

Remarks. 1. The proposition shows that in case $p \neq 2$, the type of the simple roots can be recovered from $S^{(p)}$ and $\Sigma(X)$ as follows:

$$\alpha \in S \text{ is of type } \begin{cases} (p) & \text{if } \alpha \in S^{(p)}, \\ (a) & \text{if } \alpha \in \Sigma(X), \\ (2a) & \text{if } 2\alpha \in \Sigma(X), \\ (b) & \text{otherwise.} \end{cases} \tag{5-1}$$

This way, all colors can be recovered, but some might appear multiple times (see Lemma 2.5). For colors of type (b), that behavior is controlled by $\Sigma(X)$ as well (see Proposition 5.4 below).

2. For $p = 2$ and $\alpha \in S^{(2a)}$, it is tempting to define the corresponding spherical root to be 2α instead of α. This would make parts a) and b) of Proposition 5.3 work uniformly in all characteristics. We opted against this procedure. The main reason is that otherwise spherical roots would not be roots (possibly not simple) of some root system. Example: For $p = 2$ and $X = \mathrm{SL}(3)/\mathrm{SO}(3)$, the two roots are α_1 and $\alpha_1 + \alpha_2$ (see [Schalke 2011]), which are visibly contained in an A_2-root system. On the other hand, the root α_1 is of type $(2a)$ and the set $\{2\alpha_1, \alpha_1 + \alpha_2\}$ is not part of any root system.

Proposition 5.4. *For two distinct simple roots* $\alpha_1, \alpha_2 \in S^{(b)}$, *there is equivalence between*:

a) α_1 *and* α_2 *move the same color* D.

b) *α_1 and α_2 are orthogonal to each other and $q_1\alpha_1 + q_2\alpha_2 \in \Sigma(X)$ for two p-powers q_1, q_2 (one of which is necessarily equal to* 1).

If these conditions hold, we have

$$q_1^{-1}\alpha_1^r = q_2^{-1}\alpha_2^r, \tag{5-2}$$

and D is not moved by any other simple root.

Proof. Again replace X by a toroidal compactification and choose λ as in Lemma 5.1 with $S' = \{\alpha_1, \alpha_2\}$. Now G^λ acts on X^λ via a semisimple group G_0 of rank 2. The simple roots of G_0 are α_1 and α_2, and both are of type (b) with respect to X^λ.

Assume first b). Then G_0 is of type $\mathsf{A}_1\mathsf{A}_1$. An easy inspection of its subgroups shows that X^λ has a spherical root of the given form if and only if its open orbit is isogenous to $\mathrm{SL}(2) \times \mathrm{SL}(2)/(F_{q_1} \times F_{q_2})\,\mathrm{SL}(2)$ (with F_q = Frobenius morphism of $\mathrm{SL}(2)$). In that case one checks that a) holds for X^λ and therefore for X.

Conversely assume a). Then X^λ has precisely one color D which is moved by both simple roots. Thus, (2-14) implies that a relation like (5-2) holds. In particular, the rank of X^λ is 1.

Now one could use the classification of spherical varieties of rank ≤ 1 in [Knop 2013] and conclude that G_0 is of type $\mathsf{A}_1\mathsf{A}_1$ having a root of the given form. A self-contained argument goes as follows. We may assume that $X = G/H$ is homogeneous, where H is reduced and connected. The color and the half-line $\mathcal{V}(X)$ lie opposite to each other. By [Knop 1991, Theorem 6.7], the variety X is affine, and thus H is reductive. Since there is only one color, (2-3) implies that H is semisimple. Moreover, the dimension formula [Knop 1991, Theorem 6.6] shows that $\dim H = 3, 4, 5, 7$ for $G = \mathsf{A}_1\mathsf{A}_1, \mathsf{A}_2, \mathsf{B}_2, \mathsf{G}_2$, respectively. This shows that G is isogenous to $\mathrm{SL}(2) \times \mathrm{SL}(2)$ and that $H \cong \mathrm{SL}(2)$ is embedded diagonally using the Frobenius morphisms. The assertion b) follows.

Formula (5-2) follows immediately from (2-14). Finally, assume $\alpha_3 \in S$ moves D as well. Then $\alpha_3 \in S^{(b)}$ (Lemma 2.5). Thus, by the above, α_3 would be orthogonal to α_1 and α_2. Moreover, $q_1'\alpha_1 + q_3\alpha_3 \in \Sigma(X)$ for some p-powers q_1', q_3. Now (5-2) implies the contradiction

$$2q_1'q_1^{-1} = \langle q_1'\alpha_1 + q_3\alpha_3, q_1^{-1}\alpha_1^\vee\rangle = \langle q_1'\alpha_1 + q_3\alpha_3, q_2^{-1}\alpha_2^\vee\rangle = 0. \qquad \square$$

6. Localization at Σ

Localization at S allows one to pass from S to a subset. There is a second, older, kind of localization which does the same thing with $\Sigma(X)$. Geometrically, it simply corresponds to looking at an orbit in a carefully chosen toroidal embedding. The next result summarizes what was already known about localization at Σ:

Proposition 6.1. *Let X be a toroidal spherical variety and let $Y \subseteq X$ be an orbit. Put $V := \langle \mathscr{C}_Y \rangle^\perp \subseteq \Xi_\mathbb{Q}(X)$. Then*:

a) $\Xi_p(Y) = \Xi_p(X) \cap V$.

b) $\Sigma(Y) = \Sigma(X) \cap V$.

c) $S^{(p)}(Y) = S^{(p)}(X)$.

Proof. Part a) follows, for example, from [Knop 1991, Theorem 1.3]. Moreover, $\mathscr{V}(Y) = (\mathscr{V}(X)+V)/V$, where $V = \langle \mathscr{C} \rangle_\mathbb{Q}$ (this follows from [Knop 1993, Satz 7.4]), which implies b). Part c) follows, for example, from the fact that all closed orbits in any toroidal compactification of X are of the form G/Q with $Q^{\text{red}} = {}^-P$, where P is the parabolic corresponding to $S^{(p)}$. □

If $p \neq 2$, then the remark after Proposition 5.3 allows us now to determine the type of a simple root for Y.

$$S^{(p)}(Y) = S^{(p)}(X), \tag{6-1}$$

$$S^{(a)}(Y) = S^{(a)}(X) \cap V, \tag{6-2}$$

$$S^{(2a)}(Y) = S^{(2a)}(X) \cap V, \tag{6-3}$$

$$S^{(b)}(Y) = S^{(b)}(X) \cup \bigl(S^{(a)}(X) \setminus V\bigr) \cup \bigl(S^{(2a)}(X) \setminus V\bigr). \tag{6-4}$$

For $p = 2$, equations (6-2) and (6-3) have to be replaced by the weaker equality

$$S^{(a)}(Y) \cup S^{(2a)}(Y) = \bigl(S^{(a)}(X) \cap V\bigr) \cup \bigl(S^{(2a)}(X) \cap V\bigr). \tag{6-5}$$

The next lemma shows (in particular) that moreover

$$S^{(2a)}(Y) \subseteq S^{(2a)}(X) \cap V. \tag{6-6}$$

Lemma 6.2. *Let X and Y be as above and let $\alpha \in S \cap V$ be of type* (a) *for X. Then α is also of type* (a) *for Y. Moreover, let D be a color of X which is moved by α. Then $E = (D \cap Y)^{\text{red}}$ is a color of Y which is moved by α and there is a p-power q such that δ_E is the restriction of $q\delta_D$ to $\Xi_p(Y)$.*

Proof. We plan to use localization at S but face the problem that $\mathscr{C}_Y$ may not meet $N^-_\mathbb{Q}(X)$, the image of the antidominant Weyl chamber in $N_\mathbb{Q}(X)$. To bypass this problem, we go one dimension up: the group $\overline{G} := G \times \boldsymbol{G}_m$ acts on $\overline{X}^0 := X \times \boldsymbol{G}_m$. Then $N_\mathbb{Q}(\overline{X}^0) = N_\mathbb{Q}(X) \oplus \mathbb{Q}$ and $\mathscr{V}(\overline{X}^0) = \mathscr{V}(X) \times \mathbb{Q}$. Now choose any v_0 in the relative interior of $N^-(X) \cap \{\alpha = 0\} \subseteq \mathscr{V}(X)$ and put $\overline{\mathscr{C}} := (\mathscr{C} \times 0) + \mathbb{Q}_{\geq 0}(v_0, 1)$. Choose any fan $\mathscr{F}$ whose support is $\mathscr{V}(\overline{X}^0)$ and which contains $\overline{\mathscr{C}}$. This gives rise to a toroidal $\overline{G}$-variety $\overline{X}$. Moreover, $\overline{X}$ contains $Y \times \boldsymbol{G}_m$ since $\mathscr{C}_Y$ is a face of $\overline{\mathscr{C}}$. Its closure is denoted by $\overline{Y}$.

Now choose any $v \in \mathscr{C}^0$ small enough that $v + v_0$ is still in the relative interior of $N^-(X) \cap \{\alpha = 0\}$, and choose $a \in \mathbb{Z}_{>0}$ such that $v_1 := a(v + v_0, 1)$ is the image of a 1-parameter subgroup λ of $\overline{G}$. Then, by construction, $S^\lambda = \{\alpha\}$ and $\overline{G}^\lambda$ is of semisimple rank 1. Moreover, $\overline{X}^\lambda$ is contained in $\overline{Y}$ by Proposition 4.8. Thus, we get a diagram

$$\begin{tikzcd} \tilde{Y}_\lambda \arrow[r, two heads] \arrow[d, "\tilde{\pi}_\lambda"] & \overline{Y}_\lambda \arrow[r, hook] \arrow[d, "\pi_\lambda"] & \overline{X}_\lambda \arrow[d, "\pi_\lambda"] \\ \tilde{Y}^\lambda \arrow[r, two heads, "\nu"] & \overline{X}^\lambda \arrow[r, equal] & \overline{X}^\lambda \end{tikzcd} \tag{6-7}$$

where $\tilde{Y}$ is the normalization of $\overline{Y}$ and where the vertical arrows represent Białynicki-Birula contractions on the open cell. By Proposition 4.9, the type of α on $\overline{X}^\lambda$ is (a). This means that the open $\overline{G}^\lambda$-orbit in $\overline{X}^\lambda$ is of the form $\overline{G}^\lambda/H_0$, where H_0^{red} is diagonalizable.

Now we argue that ν is purely inseparable. Indeed, the open orbit in $\tilde{Y}^\lambda$ is $\overline{G}^\lambda/H_1$, with $H_1^{\mathrm{red}} \subseteq H_0^{\mathrm{red}} \subseteq T$. This already shows that α is of type (a) for $\tilde{Y}$, and hence for $\overline{Y}$ and Y. Since both H_1^{red} and H_0^{red} are linearly reductive abelian groups, we have

$$\big[H_0^{\mathrm{red}} : H_1^{\mathrm{red}}\big] = \big[\Xi_p(\tilde{Y}^\lambda) : \Xi_p(\overline{X}^\lambda)\big]. \tag{6-8}$$

On the other hand,

$$\Xi_p(\tilde{Y}^\lambda) = \Xi_p(\tilde{Y}) = \Xi_p(\overline{Y}) = \Xi_p(\overline{X}) \cap \overline{V} = \Xi_p(\overline{X}^\lambda), \tag{6-9}$$

where $\overline{V} := \langle \overline{\mathscr{C}} \rangle$. This shows that ν is generically injective and therefore purely inseparable.

The color D gives rise to the color $D \times \boldsymbol{G}_m$ of $\overline{X}$. Its image D_0 in $\overline{X}^\lambda$ is a color of $\overline{X}^\lambda$. Now $E_0 = \nu^{-1}(D_0)^{\mathrm{red}}$ is a color of $\tilde{Y}^\lambda$. Clearly $\nu^{-1}(D_0) = qE_0$ for some p-power q. Finally, the closure of $(\tilde{\pi}_\lambda)^{-1}(E_0)$ is a color of $\overline{Y}$ which is of the form $E \times \boldsymbol{G}_m$, where $E = (D \cap Y)^{\mathrm{red}}$.

Recall $V = \langle \mathscr{C} \rangle_{\mathbb{Q}}^{\perp}$. Then

$$\overline{V} = \{(\chi, -v_0(\chi)) : \chi \in V\} \cong V \tag{6-10}$$

and $\Xi_{\mathbb{Q}}(\overline{X}) = \Xi_{\mathbb{Q}}(X) \oplus \mathbb{Q}$, $\Xi_{\mathbb{Q}}(\overline{Y}) = V \oplus \mathbb{Q}$, and $\Xi_{\mathbb{Q}}(\overline{X}^\lambda)$. Thus, for $\chi \in \Xi_{\mathbb{Q}}(X)$,

$$\delta_D(\chi) = \delta_{D\times \boldsymbol{G}_m}(\chi, 0) = \delta_{D\times \boldsymbol{G}_m}(\chi, -v_0(\chi)) = \delta_{D_0}(\chi, -v_0(\chi)); \tag{6-11}$$

and similarly, $\delta_E(\chi) = \delta_{E_0}(\chi, -v_0(\chi))$ for all $\chi \in V$. Thus, for $\chi \in V$,

$$\delta_D(\chi) = \delta_{D_0}(\chi, -v_0(\chi)) = q\delta_{E_0}(\chi, -v_0(\chi)) = q\delta_E(\chi). \qquad \square$$

Remark. With these results it is possible to recover all colors of Y and which color is being moved by which root. In characteristic zero this is good enough to compute the entire spherical system of Y. In positive characteristic we are missing information on the degrees $q_{D,\alpha}$ of Y, though. We plan to return to this question in the future.

Localization at Σ is, a priori, not possible for all subsets of $\Sigma(X)$. Therefore, we define:

Definition 6.3. A subset Σ' of $\Sigma(X)$ is called a *set of neighbors* if there is $v \in \mathcal{V}(X)$ such that

$$\Sigma' = \{\sigma \in \Sigma(X) : v(\sigma) = 0\}. \tag{6-12}$$

Equivalently, Σ' is a set of neighbors if $\mathbb{Q}_{\geq 0}\Sigma'$ is a face of $\mathbb{Q}_{\geq 0}\Sigma(X)$. Two spherical roots α and β are called *neighbors* if they are distinct and if $\{\alpha, \beta\}$ is a set of neighbors.

Clearly, if $\Sigma(X)$ is linearly independent, then all subsets are sets of neighbors. This is always the case if $p \neq 2$ (see [Brion 1990] for $\operatorname{char} k = 0$ and [Knop 2013, Corollary 4.8] for the general case). For $p = 2$ and $X = \mathrm{SL}(4)/\mathrm{SO}(4)$, Schalke has shown (unpublished) that $\Sigma(X) = \{\alpha_1, \alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3\}$. Since $\alpha_1 + (\alpha_2 + \alpha_3) = (\alpha_1 + \alpha_2) + \alpha_3$, the pairs $(\alpha_1, \alpha_2 + \alpha_3)$ and $(\alpha_1 + \alpha_2, \alpha_3)$ are not neighbors. In fact, $\mathcal{V}(X)$ is the cone over a quadrangle and the given pairs correspond to opposite faces.

Lemma 6.4. *If $\alpha, \beta \in \Sigma(X)$ are multiples of simple roots, then they are neighbors.*

Proof. The set $\mathbb{Q}_{\geq 0}\alpha + \mathbb{Q}_{\geq 0}\beta$ is already a face of $\mathbb{Q}_{\geq 0}S$ and therefore also of the smaller cone $\mathbb{Q}_{\geq 0}\Sigma(X)$. □

The following statement can be used to exclude certain configurations of colors and roots (see [Knop 2013, Proof of Theorem 4.5]).

Proposition 6.5. *Let $D \in \Delta(X)^{(a)}$ be moved by $\alpha \in S^{(a)}$ and let $\sigma \in \Sigma(X)$ be a neighbor of α with $\delta_D(\sigma) > 0$. Then $\sigma \in S^{(a)}$ and D is moved by σ, as well.*

Proof. We first reduce to the case that X is of rank 2 with $\Sigma = \{\alpha, \sigma\}$. Because α and σ are neighbors, one can choose a pointed cone $\mathcal{C}$ inside $\mathcal{V}(X) \cap \{\alpha = \sigma = 0\}$ which is of codimension 2 in $N_{\mathbb{Q}}(X)$. Let $\overline{X}$ be the simple embedding corresponding to $\mathcal{C}$ and let Y be its closed orbit. Then Proposition 6.1 implies $\Sigma(Y) = \{\alpha, \sigma\}$. Moreover, using the remark after Lemma 5.2, there is a spherical variety Y' with $\Sigma(Y') = \{\alpha, \sigma\}$ and $\operatorname{rk} Y' = 2$. Moreover, Lemma 6.2 implies that Y' still has a color E moved by α with $\delta_E(\sigma) = q\delta_D(\sigma) > 0$. Let us assume that we can prove that $\sigma \in S^{(a)}(Y')$ and that E is moved by σ. Clearly, D is moved by σ in X as well. Moreover, since $\alpha \in \Sigma(X)$, either $\alpha \in S^{(a)}(X)$ or $p = 2$ and $\alpha \in S^{(2a)}(X)$. But the latter case cannot happen, since then D could not be moved by any other simple root (Lemma 2.5). So $\sigma \in S^{(a)}(X)$, which finishes the reduction step.

From now on we assume that $\operatorname{rk} X = 2$ and, without loss of generality, that $X = G/H$ where H is reduced. Since $\delta_D(\sigma) > 0$ by assumption and $\delta_D(\alpha) = q_{D,\alpha} > 0$, the cone generated by $\mathcal{V}(X)$ and δ_D is the entire space $N := N_{\mathbb{Q}}(X)$. From that we get a morphism $X = G/H \to G/P$ with $\operatorname{rank} G/P = \dim N/N = 0$ [Knop 1991, Corollary 4.6].[1] Hence P is a parabolic subgroup with an identification $\Delta(G/P) = \Delta(G/H) \setminus \{D\}$. We may choose P in such a way that it is *opposite* to B.

Every $\beta \in S \setminus S^{(p)}$ moves at least one color and α moves even two. Assume first that these colors are *not* all different. Then, according to Lemma 2.5 and Proposition 5.4, there are two possibilities:

a) $\sigma = \alpha_1 + q\alpha_2$ with $\alpha_1, \alpha_2 \in S$ orthogonal. But then α_1 and α_2 would move the same color in G/P, which is impossible.

b) $S^{(a)}$ contains another element β besides α. But then $\beta \in \Sigma(X)$ (Proposition 5.3), and hence $\sigma = \beta \in S^{(a)}$. Moreover, there is a color D' moved by both α and σ. We claim that $D' = D$. Suppose not. Then D and D' are the two colors moved by α, and Proposition 2.3 implies the contradiction

$$\delta_D^{(\alpha)}(\sigma) = \langle \sigma, \alpha^\vee \rangle - \delta_{D'}^{(\alpha)}(\sigma) = \langle \sigma, \alpha^\vee \rangle - \frac{q_{D',\alpha}}{q_{D',\sigma}} \delta_{D'}^{(\sigma)}(\sigma) = \langle \sigma, \alpha^\vee \rangle - \frac{q_{D',\alpha}}{q_{D',\sigma}} < 0. \tag{6-13}$$

Thus, we are exactly in the asserted situation, that is, $\sigma \in S^{(a)}$ moving D.

So, assume from now on that the colors moved by all the $\beta \in S \setminus S^{(p)}$ are different. Then

$$\#\Delta(G/P) \geq \#S \setminus S^{(p)}. \tag{6-14}$$

Consider a toroidal completion $\overline{X}$ of X. Then the morphism $X \to G/P$ extends to $\overline{X}$ [Knop 1991, Theorem 4.1]. Every closed orbit in $\overline{X}$ is isogenous to $G/{}^{-}P_X$, where P_X is the parabolic attached to $S^{(p)}$. Hence $P_X \subseteq P^{\mathrm{red}}$. Thus (6-14) implies $P = P_X$.

Let $Y = G/H_1 \subset \overline{X}$ be the rank-1 orbit corresponding to the half-line $\mathcal{V}(X) \cap \{\sigma = 0\}$. Then $\Sigma(Y) = \{\sigma\}$. Because of $P = P_X$, the fiber P_X/H_1^{red} is one-dimensional and therefore isomorphic to $\boldsymbol{P}^1$, $\boldsymbol{G}_m$, or $\boldsymbol{A}^1$. The first case is impossible since H_1 is not parabolic. The second case is excluded since H_1 is not horospherical. Thus $P_X/H_1^{\mathrm{red}} \cong \boldsymbol{A}^1$.

This means in particular that (a conjugate of) H_1 contains the maximal torus T of G and that H_1 contains all root subgroups U_β which are contained in P_X except for one, which is denoted by γ, and which lies in P_X^u. The U_β corresponding to $\beta \in S$ generate the maximal unipotent subgroup of G. This implies $\gamma \in S$. Moreover, U_γ is a 1-dimensional module for the Levi part of P_X. This shows that H_1^{red} is, in fact, induced from $\mathrm{PGL}(2)/\boldsymbol{G}_m$ (on induction in arbitrary characteristic, see [Knop 2013, §2, in particular Lemma 2.1]). Hence $\sigma \in S^{(a)}$. But in that case, (6-14) would be a strict inequality, which is not true because of $P = P_X$. □

[1] The idea for this construction is due to Guido Pezzini.

Proposition 6.5 can be used to give bounds for $\delta_D(\sigma)$:

Corollary 6.6. *With α and σ as above, assume that $\sigma \notin S$ or that $\sigma \in S$ but does not move either color moved by α. Then*

$$q_{D,\alpha}^{-1}\langle\sigma,\alpha^\vee\rangle \le \delta_D(\sigma) \le 0. \tag{6-15}$$

Proof. The right-hand inequality follows directly from Proposition 6.5. For the left-hand inequality, apply Proposition 6.5 to the other color D' moved by α and observe that $\delta_{D'}^{(\alpha)} = \alpha^r - \delta_D^{(\alpha)}$ (Proposition 2.3). □

In positive characteristic, these bounds are less valuable since we didn't derive a bound on the denominator of $\delta_D(\sigma)$. Such a bound exists (in terms of the $q_{D,\alpha}$'s) and will be included in a future paper. Then (6-15) leaves only finitely many possibilities for $\delta_D(\sigma)$.

7. The p-spherical system

We summarize what we have proved so far in terms of a combinatorial structure which generalizes Luna's spherical systems. But first we need some more terminology:

Definition 7.1. Let G be a connected reductive group.

a) An element $\sigma \in \Xi_{\mathbb{Q}}(T)$ is called a *spherical root for G* if there is a spherical G-variety X such that σ is a spherical root for X. The set of spherical roots for G is denoted by $\Sigma(G)$.

b) A spherical root $\sigma \in \Sigma(G)$ is *compatible* with a subset $S^{(p)} \subseteq S$ if there is a spherical G-variety X with $\sigma \in \Sigma(X)$ and $S^{(p)} = S^{(p)}(X)$.

Remarks. 1. Proposition 6.1 shows that in the definition above, one may assume, without loss of generality, $\operatorname{rk} X = 1$.

2. Spherical varieties of rank 1 have been classified by Akhiezer [1983] in characteristic zero and Knop [2013] in general. In particular, for every G, there is a complete description of $\Sigma(G)$ (see [Knop 2013, §2 and §7]).

3. One result of that classification is that $\Sigma(G)$ is infinite unless $\operatorname{char} k = 0$ or G is simple of rank ≤ 2.

For the following, recall that $\mathbb{Z}_p = \mathbb{Z}[1/p]$ and $\Xi_p := \Xi \otimes \mathbb{Z}_p$ for any abelian group Ξ.

Definition 7.2. Let $p \neq 2$. Then a p-spherical system for G consists of

- a subgroup $\Xi \subseteq \Xi(T)$,
- a subset $\Sigma \subseteq \Xi_p \cap \Sigma(G)$,
- a subset $S^{(p)} \subseteq S$,

- a finite set $\Delta^{(a)}$,
- a map $\delta : \Delta^{(a)} \to \mathrm{Hom}(\Xi, \mathbb{Z}) : D \mapsto \delta_D$, and
- a map $S \setminus (S^{(p)} \cup S^{(a)}) \to p^{\mathbb{N}} : \alpha \mapsto q_\alpha$, where $S^{(a)} := S \cap \Sigma$.

Of course, these data are subject to some conditions. Here, we list only those which are straightforward generalizations of Luna's axioms. It is safe to say that more axioms have to be imposed which deal specifically with issues of positive characteristic. We keep the notation that α^r denotes the restriction of $\alpha^\vee$ to Ξ.

A1 All $\sigma \in \Sigma$ are primitive vectors of $\mathbb{Z}S \cap \Xi_p$.

A2 $\alpha^r = 0$ for all $\alpha \in S^{(p)}$.

A3 Every $\sigma \in \Sigma$ is compatible with $S^{(p)}$.

A4 For all $D \in \Delta^{(a)}$ and $\sigma \in \Sigma \setminus S^{(a)}$, we have $\delta_D(\sigma) \le 0$.

A5 For every $\alpha \in S^{(a)}$, there are exactly two $D \in \Delta^{(a)}$ with $\delta_D(\alpha) > 0$. Conversely, for every $D \in \Delta^{(a)}$, there is at least one $\alpha \in S^{(a)}$ with $\delta_D(\alpha) > 0$.

A6 For $\alpha \in S^{(a)}$, let $D^+ \ne D^- \in \Delta^{(a)}$ with $\delta_{D^\pm}(\alpha) > 0$. Then $q_{\alpha,D^\pm} := \delta_{D^\pm}(\alpha)^{-1} \in p^{\mathbb{N}}$ and $q_{\alpha,D^+}\delta_{D^+} + q_{\alpha,D^-}\delta_{D^-} = \alpha^r$.

A7 Let $\alpha \in S$ with $2\alpha \in \Sigma$. Then $\alpha \notin \Xi_p$ and $\frac{1}{2}\alpha^r(\Xi_p) \subseteq \mathbb{Z}_p$. Moreover, $\alpha^r(\sigma) \le 0$ for all $\sigma \in \Sigma \setminus \{2\alpha\}$.

A8 Let $q\alpha_1 + \alpha_2 \in \Sigma$ with $\alpha_1 \perp \alpha_2$. Then $q^{-1}\alpha_1^r = \alpha_2^r$ and $q^{-1}q_{\alpha_1} = q_{\alpha_2}$.

The point is, of course, that for $p \ne 2$ every homogeneous spherical variety, X gives rise to a p-spherical system. More specifically, we put

$$\begin{aligned} \Xi &:= \Xi(X), & \Sigma &:= \Sigma(X), & S^{(p)} &:= S^{(p)}(X), \\ \Delta^{(a)} &:= \Delta^{(a)}(X), & \delta_D &:= \delta_D^X. \end{aligned} \tag{7-1}$$

The only new constituents are the p-powers. For $\alpha \in S \setminus (S^{(p)} \cup S^{(a)}) = S^{(b)} \cup S^{(2a)}$, we define q_α as $q_{\alpha,D}$ from Proposition 2.3, where D is the unique color moved by α.

Now we verify all axioms.

A1 Holds by definition of $\Sigma(X)$.

A2 See Proposition 2.3.

A3 Follows from the definition of "compatibility".

A4 This is Corollary 6.6 in conjunction with [Knop 2013, Corollary 4.8], which implies that for $p \ne 2$, any two spherical roots are neighbors.

A5 The first part follows also from Corollary 6.6 and Proposition 2.3. The second part holds by definition of $\Delta^{(a)}(X)$.

A6 This follows from Proposition 2.3.

A7 The first part is Proposition 5.3b) and Corollary 2.4. The second follows from [Knop 2013, Theorem 4.5].

A8 This is Propositions 5.4 and 2.3.

Remarks. 1. In characteristic 0, Luna [2001, 5.1] used Wasserman's tables [1996] of spherical rank-2 varieties to verify the axioms. So our approach is more conceptual in that it uses only the classification of rank-1 but not of rank-2 varieties.[2]

2. The case $p = 2$ requires some modifications. To distinguish simple roots of type (a) and $(2a)$, we redefine $S^{(a)}$ as

$$S^{(a)} := \{\alpha \in S \cap \Sigma : \delta_D(\alpha) > 0 \text{ for some } D \in \Delta^{(a)}\}. \tag{7-2}$$

This works indeed for spherical systems coming from spherical varieties: Suppose there are $\alpha \in S^{(2a)}(X)$ and $D \in \Delta^{(a)}$ with $\delta_D(\alpha) > 0$. Then D is moved by some $\beta \in S^{(a)}$. Since α and β are neighbors (Lemma 6.4), we get a contradiction to Proposition 6.5.

With this change, all axioms hold for $p = 2$ except for one: in A4, one has to require that σ and α are neighbors. Observe that A7 is vacuously satisfied.

3. It is a natural question whether spherical varieties are classified by their p-spherical system. In characteristic zero, the answer is "yes" according to work by Luna [2001], Losev [2009], Cupit-Foutou [2010], and Bravi and Pezzini [2011a; 2011b; 2011c]. For $p \neq 2$ or 3, it might be possible that the p-spherical system determines the variety uniquely. For example, all complete homogeneous varieties are classified by p-spherical systems with $\Xi = 0$ (see the example before Lemma 2.2). Furthermore, the author convinced himself that this also holds for spherical varieties of rank 1. If $p = 2$ or $p = 3$, then uniqueness does not even hold for complete homogeneous varieties (see [Wenzel 1994, Proposition 4]) due to exceptional isogenies. If $p = 2$, then uniqueness is wrong already for $G = \mathrm{SL}(2)$, as then G contains nonstandard horospherical subgroup schemes (see [Knop 1995a]).

4. The above list of axioms A1–A8 is definitely only preliminary. Even in the rank-1 case, they do not suffice. For example, there is no axiom bounding the lattice Ξ from below. We plan to return to this problem in the future.

Acknowledgment

I would like to thank Guido Pezzini for many discussions on the matter of this paper. In particular, the main idea in the proof of Proposition 6.5 is due to him.

---

[2] Note that the tables in [Wasserman 1996] are slightly incomplete: of the series $G = \mathrm{Sp}(2n) \times \mathrm{Sp}(2)$, $H = \mathrm{Sp}(2n-2) \times \mathrm{Sp}(2)$ with $n \geq 2$, only the first case, $n = 2$, is stated. I would like to thank Guido Pezzini for pointing that out to me.

References

[Akhiezer 1983] D. Akhiezer, "Equivariant completions of homogeneous algebraic varieties by homogeneous divisors", *Ann. Global Anal. Geom.* **1**:1 (1983), 49–78. MR 85j:32052 Zbl 0537.14033

[Białynicki-Birula 1976] A. Białynicki-Birula, "Some properties of the decompositions of algebraic varieties determined by actions of a torus", *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **24**:9 (1976), 667–674. MR 56 #12020 Zbl 0355.14015

[Białynicki-Birula and Święcicka 1982] A. Białynicki-Birula and J. Święcicka, "Complete quotients by algebraic torus actions", pp. 10–22 in *Group actions and vector fields* (Vancouver, B.C., 1981), edited by J. B. Carrell, Lecture Notes in Math. **956**, Springer, Berlin, 1982. MR 85d:14065 Zbl 0493.14027

[Bravi and Pezzini 2011a] P. Bravi and G. Pezzini, "A constructive approach to the classification of wonderful varieties", preprint, 2011. arXiv 1103.0380

[Bravi and Pezzini 2011b] P. Bravi and G. Pezzini, "Primitive wonderful varieties", preprint, 2011. arXiv 1106.3187

[Bravi and Pezzini 2011c] P. Bravi and G. Pezzini, "The spherical systems of the wonderful reductive subgroups", preprint, 2011. arXiv 1109.6777

[Brion 1990] M. Brion, "Vers une généralisation des espaces symétriques", *J. Algebra* **134**:1 (1990), 115–143. MR 91i:14039 Zbl 0729.14038

[Cupit-Foutou 2010] S. Cupit-Foutou, "Wonderful varieties: A geometrical realization", preprint, 2010. arXiv 0907.2852

[Grothendieck 1965] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* 24 (1965), 231. MR 33 #7330 Zbl 0135.39701

[Haboush and Lauritzen 1993] W. Haboush and N. Lauritzen, "Varieties of unseparated flags", pp. 35–57 in *Linear algebraic groups and their representations* (Los Angeles, CA, 1992), edited by R. S. Elman et al., Contemp. Math. **153**, Amer. Math. Soc., Providence, RI, 1993. MR 95a:14055 Zbl 0806.14037

[Jouanolou 1983] J.-P. Jouanolou, *Théorèmes de Bertini et applications*, Progress in Mathematics **42**, Birkhäuser, Boston, MA, 1983. MR 86b:13007 Zbl 0519.14002

[Knop 1991] F. Knop, "The Luna–Vust theory of spherical embeddings", pp. 225–249 in *Proceedings of the Hyderabad Conference on Algebraic Groups* (Hyderabad, 1989), edited by S. Ramanan et al., Manoj Prakashan, Madras, 1991. MR 92m:14065 Zbl 0812.20023

[Knop 1993] F. Knop, "Über Bewertungen, welche unter einer reduktiven Gruppe invariant sind", *Math. Ann.* **295**:2 (1993), 333–363. MR 94a:14050 Zbl 0789.14040

[Knop 1995a] F. Knop, "Homogeneous varieties for semisimple groups of rank one", *Compositio Math.* **98**:1 (1995), 77–89. MR 96i:14042 Zbl 0860.14040

[Knop 1995b] F. Knop, "On the set of orbits for a Borel subgroup", *Comment. Math. Helv.* **70**:2 (1995), 285–309. MR 96c:14039 Zbl 0828.22016

[Knop 2013] F. Knop, "Spherical roots of spherical varieties", preprint, 2013. arXiv 1303.2466

[Losev 2009] I. V. Losev, "Uniqueness property for spherical homogeneous spaces", *Duke Math. J.* **147**:2 (2009), 315–343. MR 2010c:14055 Zbl 1175.14035

[Luna 1997] D. Luna, "Grosses cellules pour les variétés sphériques", pp. 267–280 in *Algebraic groups and Lie groups*, edited by G. Lehrer et al., Austral. Math. Soc. Lect. Ser. **9**, Cambridge Univ. Press, 1997. MR 99g:14059 Zbl 0902.14037

[Luna 2001] D. Luna, “Variétés sphériques de type A”, *Publ. Math. Inst. Hautes Études Sci.* 94 (2001), 161–226. MR 2003f:14056 Zbl 1085.14039

[Luna and Vust 1983] D. Luna and T. Vust, “Plongements d'espaces homogènes”, *Comment. Math. Helv.* **58**:2 (1983), 186–245. MR 85a:14035 Zbl 0545.14010

[Schalke 2011] B. Schalke, *Sphärische Einbettungen in positiver Charakteristik*, Diplomarbeit, Universität Erlangen, 2011.

[Sumihiro 1974] H. Sumihiro, “Equivariant completion”, *J. Math. Kyoto Univ.* **14** (1974), 1–28. MR 49 #2732 Zbl 0277.14008

[Wasserman 1996] B. Wasserman, “Wonderful varieties of rank two”, *Transform. Groups* **1**:4 (1996), 375–403. MR 97k:14051 Zbl 0921.14031

[Wenzel 1993] C. Wenzel, “Classification of all parabolic subgroup-schemes of a reductive linear algebraic group over an algebraically closed field”, *Trans. Amer. Math. Soc.* **337**:1 (1993), 211–218. MR 93g:20090 Zbl 0785.20024

[Wenzel 1994] C. Wenzel, “On the structure of nonreduced parabolic subgroup-schemes”, pp. 291–297 in *Algebraic groups and their generalizations: Classical methods* (University Park, PA, 1991), edited by W. J. Haboush and B. J. Parshall, Proc. Sympos. Pure Math. **56**, Amer. Math. Soc., Providence, RI, 1994. MR 95h:20059 Zbl 0828.20031

Communicated by Raman Parimala
Received 2013-05-31 Revised 2013-10-25 Accepted 2013-11-22

friedrich.knop@fau.de *Department Mathematik, Emmy-Noether-Zentrum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.729

Lefschetz operator and local Langlands modulo ℓ: the limit case

Jean-François Dat

Let K be a finite extension of $\mathbb{Q}_p$ with residue field $\mathbb{F}_q$, and let ℓ be a prime such that $q \equiv 1 (\mathrm{mod}\ \ell)$. We investigate the cohomology of the Lubin–Tate towers of K with coefficients in $\overline{\mathbb{F}}_\ell$, and we show how it encodes Vignéras' Langlands correspondence for unipotent $\overline{\mathbb{F}}_\ell$-representations.

1. Main results

This paper is part of a project, outlined in [Dat 2012b], that aims at providing a geometric interpretation of the Vignéras correspondence for modulo-ℓ representations of p-adic linear groups.

1.1. *Motivation.*

1.1.1. *The problem.* Let K be a local p-adic field, ℓ a prime distinct from p, and $d \geqslant 1$ an integer. Vignéras [2001] established a bijection between (classes of) irreducible smooth $\overline{\mathbb{F}}_\ell$-representations of $\mathrm{GL}_d(K)$ and (classes of) d-dimensional Weil–Deligne $\overline{\mathbb{F}}_\ell$-representations for K. On the one hand we have fairly natural "automorphic objects", but on the other hand we get fairly *unnatural* "Galois objects". Indeed, the nilpotent part of a Weil–Deligne $\overline{\mathbb{F}}_\ell$-representation has no obvious Galois interpretation, in contrast with $\overline{\mathbb{Q}}_\ell$-representations, where it is related to the infinitesimal action of the tame inertia subgroup on some associated continuous $\overline{\mathbb{Q}}_\ell$ representation of the Weil group. Therefore in the $\overline{\mathbb{F}}_\ell$ case, this nilpotent part appears as an "extra datum", from the arithmetic point of view. In

MSC2010: primary 11S37; secondary 11F70, 14G35.
Keywords: local Langlands, Lefschetz operator, modulo ℓ.

fact, Vignéras' correspondence was obtained by purely representation-theoretic arguments (a classification theorem à la Zelevinsky), and our aim is to find a geometric interpretation for it.

1.1.2. *The project.* Since Carayol's formulation of "nonabelian Lubin–Tate theory", the cohomology of the Lubin–Tate tower $(\mathcal{M}_{\mathrm{LT},n})_{n\in\mathbb{N}}$ of height d is a natural place to look for a realization of any Langlands-type local correspondence. Let us fix a completed algebraic closure K^{ca} of K and denote base changes to K^{ca} by adding the exponent "ca". It has long been noticed by the author that in order to get something correct for nonsupercuspidal representations, one should use the full "cohomology complex" $R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)$ as an object in the derived category $D^b(\mathrm{Rep}^{\infty}_{\mathbb{Z}_\ell}(\mathrm{GL}_d(K)))$ endowed with an action of the Weil group W_K and of the unit group $D^\times$ of the division algebra with center K and invariant $1/d$. Then for an irreducible representation π (over $\overline{\mathbb{F}}_\ell$ or $\overline{\mathbb{Q}}_\ell$), one considers the "derived π-coisotypical component"

$$R^*_\pi := \mathcal{H}^*\big(R\,\mathrm{Hom}_{\mathbb{Z}_\ell\,\mathrm{GL}_d(K)}(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell), \pi)\big)[1-d],$$

which is a finite-dimensional graded smooth/continuous $D^\times \times W_K$-module supported in the range $[1-d, d-1]$. In [Dat 2012c], we proved the equality

$$[R^*_\pi] = \mathrm{LJ}(\pi) \otimes \sigma^{\mathrm{ss}}(\pi) \tag{1.1.3}$$

in the Grothendieck group of smooth/continuous $D^\times \times W_K$-modules. Here LJ stands for the so-called Langlands–Jacquet transfer of [Dat 2012d] and $\sigma^{\mathrm{ss}}(\pi)$ is the Weil part (that is, the semisimple part), of the Weil–Deligne representation $\sigma(\pi) = (\sigma^{\mathrm{ss}}(\pi), N_\pi)$ attached to π.

We want to enrich R^*_π with a nilpotent operator so that a similar formula holds in a suitable Grothendieck group of Weil–Deligne representations. The inspiration for this is Arthur's second SL_2 factor in the theory of automorphic forms. An hint that this may be useful for our purpose is the relation between switching the two SL_2-factors of a local A-parameter and the Zelevinsky involution. Indeed, Vignéras' correspondence is more a $\overline{\mathbb{F}}_\ell$-analog of the "Zelevinsky correspondence" than of the Langlands correspondence for $\overline{\mathbb{Q}}_\ell$-representations of $\mathrm{GL}_d(K)$.[1]

The origin of Arthur's SL_2 lies in the Lefschetz decomposition of the intersection cohomology of Shimura varieties. Building on this analogy, we defined in [Dat 2012b] a "Lefschetz operator"

$$L : R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell) \longrightarrow R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)[2](1)$$

[1] Recall that both correspondences are intertwined by the Zelevinsky involution.

as the cup product by the Chern class of a natural equivariant line bundle on the Lubin–Tate tower. In turn, this operator induces a graded equivariant map

$$L_\pi^* : R_\pi^* \longrightarrow R_\pi^*[2](1).$$

In its roughest formulation, our hope is that the pair (R_π^*, L_π^*) encodes the Weil–Deligne representation associated to π by Vignéras' correspondence. More precisely, one can define the class $[R_\pi^*, L_\pi^*] = [R_\pi^{\mathrm{even}}, L_\pi^*] - [R_\pi^{\mathrm{odd}}, L_\pi^*]$ of (R_π^*, L_π^*) in the Grothendieck group of Weil–Deligne $\overline{\mathbb{F}}_\ell D^\times$-representations, as in [Dat 2012a, 2.2.2]. The best we can hope for is then the equality

$$[R_\pi^*, L_\pi^*] = \mathrm{LJ}(\pi) \otimes [\sigma^{\mathrm{ss}}(\pi), N_\pi]. \tag{1.1.4}$$

1.1.5. *Previous results.* We proved such an equality in the following cases:

- When π is any irreducible $\overline{\mathbb{Q}}_\ell$-representation (in this case, by "Vignéras correspondence", we just mean the Zelevinsky correspondence) [Dat 2012b].
- When π is any *unipotent* irreducible $\overline{\mathbb{F}}_\ell$-representation and the order of q in $\mathbb{F}_\ell^\times$ is at least d [Dat 2012a].

In Vignéras' terminology, a representation is "unipotent" if it belongs to the principal block of the category $\mathrm{Rep}^\infty_{\overline{\mathbb{F}}_\ell}(G)$, that is, the unique block that contains the trivial representation. For an irreducible π, this is equivalent to π occurring as a subquotient of some $\mathrm{Ind}_B^G(\chi)$ for χ an unramified character of a Borel subgroup B. We note that, in the second case above, R_π^* can be computed in greater generality,[2] but the author is still unable to control L_π^* when π is not unipotent. Also, the arguments are much more difficult in the "Coxeter congruence" case (when the order of q in $\mathbb{F}_\ell^\times$ is d) than in the "banal" case (when this order is greater than d), due to some representation-theoretic complications. However, a common feature of these two cases is that one can still use enough of the theory of weights (on the Galois side) and exponents (on the GL_d side) so as to split the complex and compute explicitly Yoneda extensions. Moreover, in all the cases above, (1.1.3) and (1.1.4) involve no cancellation because R_π^* turns out to be either oddly or evenly graded. In fact, the "énoncé optimiste" from [Dat 2012b, 1.3.3] holds true in these cases.

1.2. ***This paper.*** Here we study the case when $q \equiv 1 (\mathrm{mod}\ \ell)$ and $\ell > d$. This is called the "limit case" in [Clozel et al. 2008, §5]. From the point of view of weights or exponents, this is the most degenerate case, due to the congruence on q. For example, in this situation, the unipotent summand of the complex $R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)$ can be shown to be indecomposable. Moreover, as we have noted in [Dat 2012c, 2.2.7],

[2] And, in fact, for any irreducible π, if one admits that the $\mathbb{Z}_\ell$-cohomology of the Lubin–Tate tower is torsion free, as recently announced by Boyer.

R^*_π is generally not evenly nor oddly graded. Despite this bad news, computations are still feasible because of the additional assumption that $\ell > d$, which simplifies significantly the representation theory, as indicated by Vignéras' appendix to [Clozel et al. 2008].

1.2.1. *The result of the computation.* We explain in Corollary 3.1.3 that for π *unipotent*, R^*_π vanishes unless π is a subquotient of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$, or equivalently, unless $\sigma^{\mathrm{ss}}(\pi)$ is the trivial representation of W_K of dimension d. Then $\sigma(\pi) = (\sigma^{ss}(\pi), N)$ is given by a nilpotent $d \times d$ matrix N, whose Jordan form has shape λ for some partition λ of d. Through Vignéras' correspondence we thus get a parametrization $\lambda \in \mathcal{P}(d) \mapsto \pi_\lambda$ of all π's that occur as a subquotient of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$. We explain how to construct π_λ explicitly in Section 2.2.5. We now describe algebraically (R^*_π, L^*_π) and provide geometric intuition for the result obtained.

For a finite-dimensional $\overline{\mathbb{F}}_\ell$-vector space V with dual $\check{V}$, consider the graded space $H^* := \bigwedge^* \check{V} \otimes \bigwedge^* V$ endowed with the operator L^* of degree 2 which on the (p, q) part is given by

$$L : \bigwedge^p \check{V} \otimes \bigwedge^q V \xrightarrow{\mathrm{Id} \otimes \mathrm{Id}_V \otimes \mathrm{Id}} \bigwedge^p \check{V} \otimes \check{V} \otimes V \otimes \bigwedge^q V \xrightarrow{\wedge \otimes \wedge} \bigwedge^{p+1} \check{V} \otimes \bigwedge^{q+1} V.$$

When the dimension of V is less than ℓ, this satisfies the hard Lefschetz property; see Section A.1.4. In particular, this holds for V the *standard representation* of $\mathfrak{S}_d$ (which has dimension $d-1$). In this case, H^* also carries an action of $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$ that commutes with L^*, so we may decompose it as $(H^*, L^*) = \sum_{\lambda \in \mathcal{P}(d)} (H^*_\lambda, L^*_\lambda)$ by applying central primitive idempotents associated to partitions.

Theorem. *For $\lambda \in \mathcal{P}(d)$, the action of $D^\times$ on $R^*_{\pi_\lambda}$ is trivial and that of W_K is unipotent. Moreover, there is an isomorphism $(R^*_{\pi_\lambda}, L^*_{\pi_\lambda})^{\oplus d_\lambda} \simeq (H^*_\lambda, L^*_\lambda)$ of graded $\overline{\mathbb{F}}_\ell$-vector spaces compatible with L operators.*

Here d_λ is the dimension of the simple $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-module associated to λ.

There is geometric intuition behind this result. Consider the diagonal torus T in PGL_d and the discrete cocompact subgroup ϖ^X of T^{an} obtained by evaluating cocharacters at a fixed uniformizer ϖ of $\mathcal{O}_K$, and let A be the abelian variety T^{an}/ϖ^X, which has an action by the Weyl group $\mathfrak{S}_d$. Its cohomology is equivariantly isomorphic to H^* (note that $V = X_*(T) \otimes \overline{\mathbb{F}}_\ell$ is the standard representation of $\mathfrak{S}_d$), and there's a natural choice of a $\mathfrak{S}_d$-equivariant ample invertible sheaf on A whose associated Chern class can be put in the form described above. Now the special fiber (analytic reduction) of Mumford's formal model of A turns out to be isomorphic to the quotient of the special fiber of Deligne's formal model of Drinfeld's symmetric space by the action of B. This suggests a relation between the cohomology of A and $R^*_{i_B}$ with $i_B = \mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell) = \bigoplus_\lambda \pi_\lambda^{\oplus d_\lambda}$. In general, however, there is no such a relation because some multiplicities appear when one tries to compare vanishing cycles on both sides, but somehow these multiplicities disappear when $q = 1$ in the coefficients.

By representation theory of $\mathfrak{S}_d$ the theorem implies that $R^*_{\pi_\lambda}$ vanishes unless the Young tableau of λ is a hook or a double-hook. We explain in Section 2.2.7 that λ is a hook if and only if π_λ is *elliptic*, that is, $\mathrm{LJ}(\pi_\lambda) \neq 0$. In this case we can make the theorem more explicit.

Corollary. *Assume $\lambda = (d - j, 1^{(j)})$ for some $j \in \{0, \dots, d-1\}$ and put $j' = d-1-j$. Then we have $(R^*_{\pi_\lambda}, L^*_{\pi_\lambda}) \simeq (\boldsymbol{P}_{j'})^{\oplus j+1} \oplus (\boldsymbol{P}_{j'-1})^{\oplus j}$ where $\boldsymbol{P}_k$ denotes the cohomology of a projective space of dimension k, shifted by $-k$, and with its tautological Lefschetz operator.*

In particular, the space $R^*_{\pi_\lambda}$ has total dimension $2jj' + d$. When $jj' \neq 0$, the pair $(R^*_{\pi_\lambda}, L^*_{\pi_\lambda})$ does not have the right dimension, and what is worse, it does not seem related to the Vignéras pair $(\sigma^{\mathrm{ss}}(\pi_\lambda), N_\lambda)$ in any reasonable Grothendieck group of Weil–Deligne representations. In other words, (1.1.4) fails in this case.

However, it is still true that it encodes the Vignéras pair, provided one uses extra structure.

1.2.2. *Main result.* Observe that R^*_π has the structure of a graded right module over the derived endomorphism algebra $\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell G}(\pi, \pi)$. Consider the subalgebra $\mathscr{E}^*_\pi$ generated by extensions that "come from the boundary", namely by the kernel of the map $\mathrm{Ext}^1_{\overline{\mathbb{F}}_\ell G}(\pi, \pi) \longrightarrow \mathrm{Ext}^1_{\overline{\mathbb{F}}_\ell \bar{G}}(\bar{\pi}, \bar{\pi})$, where $\bar{G} = \mathrm{GL}_d(\mathbb{F}_q)$ and $\bar{\pi} = \pi^{1+\varpi M_d(\mathbb{O})}$. For a unipotent π, we'll see that $\mathscr{E}^*_\pi$ is also the image of a natural map $\mathrm{Ext}^*_{\mathscr{H}}(\pi^I, \pi^I) \longrightarrow \mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell G}(\pi, \pi)$, where I is an Iwahori subgroup and $\mathscr{H}$ is the corresponding Hecke algebra. This is a local graded algebra and we denote by $\mathscr{E}^+_\pi$ its maximal ideal. In the cases when (1.1.4) has been established, one also observes that either $\mathscr{E}^+_\pi = 0$ or at least its action on R^*_π vanishes. In contrast, in the limit case under study here, this action is nonzero and is somehow responsible for R^*_π being "too big". So, define $R^{\mathrm{red}}_\pi := R^*_\pi / R^*_\pi \mathscr{E}^+_\pi$. This is still a graded $\overline{\mathbb{F}}_\ell$-representation of $W_K \times D^\times$, and L^*_π induces an operator $L^{\mathrm{red}}_\pi : R^{\mathrm{red}}_\pi \longrightarrow R^{\mathrm{red}}_\pi[2](1)$. Let us finally denote by $[R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi]$ the image of the pair $(R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi)$ in the Grothendieck group of Weil–Deligne representations.

Theorem. *Let π be an elliptic unipotent irreducible $\overline{\mathbb{F}}_\ell$-representation of $\mathrm{GL}_d(K)$. As above, let $\pi = \pi_\lambda$ for $\lambda = (d - j, 1^{(j)})$. Then we have a $\overline{\mathbb{F}}_\ell$-linear isomorphism $(R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi) \simeq \boldsymbol{P}_{j'} \oplus (\boldsymbol{P}_0)^{\oplus j}[j']$. Hence in the Grothendieck group of Weil–Deligne $\overline{\mathbb{F}}_\ell D^\times$-representations we get*

$$[R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi] = \mathrm{LJ}(\pi) \otimes [\sigma^{\mathrm{ss}}(\pi), N_\pi].$$

When π is not elliptic but has a nonzero R^*_π (that is, π is associated to a double hook) we expect that $[R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi]$ vanishes, but we don't prove this here.

1.2.3. *A sketch of the argument.* Representation theoretic considerations tell us that, under our assumption on ℓ, the graded space R^*_π for π unipotent is the abutment of a spectral sequence whose E_2^{pq} term is $\mathrm{Ext}^p_{\mathscr{H}}(H^q_c(\mathcal{M}^{\mathrm{ca}}_{\mathrm{LT},I}, \overline{\mathbb{F}}_\ell), \pi^I)$, where $\mathcal{M}_{\mathrm{LT},I}$ is the Lubin–Tate space at Iwahori level, and $\mathscr{H}$ is the Hecke–Iwahori algebra as above. The main point of the paper is to compute this E_2 term and show that the spectral sequence degenerates at E_2. There are three ingredients coming into this computation.

- We can compute the $\overline{\mathbb{F}}_\ell$-cohomology of $\mathcal{M}_{\mathrm{LT},I}$ because the $\overline{\mathbb{Q}}_\ell$-cohomology is known and the simple geometry of $\mathcal{M}_{\mathrm{LT},I}$ shows the $\mathbb{Z}_\ell$-cohomology has no torsion.
- Representation theory of p-adic groups, under our assumption on ℓ, reduces the computation of the E_2 term to that of known multiplicities in certain tensor product representations of a symmetric group.
- Some numerical coincidences appear, that force degeneration of the spectral sequence.

Then comes the computation of L^*_π. Here we have to understand cup-products between the extensions mentioned above, and this also boils down to a problem in representation theory of the symmetric group that we solve in the Appendix. Once cup-products are understood, we need a handle on the Lefschetz operator (after all, it could be trivial!). This is provided by the explicit form of L^*_π when π is the trivial representation, which itself comes from the very definition of the Lefschetz operator as the Chern class of a bundle that is lifted from the crystalline (or Gross–Hopkins) period space. It turns out that knowing this particular L^*_π is enough to compute L^*_π for all unipotent elliptic π. Eventually, our computation of cup-products also allows going from the pair (R^*_π, L^*_π) to $(R^{\mathrm{red}}_\pi, L^{\mathrm{red}}_\pi)$.

Remark. Part of the above arguments can be generalized to approach the computation of R^*_π for any irreducible representation π (under the same hypothesis $q \equiv 1 \pmod{\ell}$ and $\ell > d$). Indeed, Boyer's announcement that the $\mathbb{Z}_\ell$-cohomology of the Lubin–Tate tower is torsion free implies that the only π's for which R^*_π is nonzero contain a simple type (J, τ). Then, in the above spectral sequence, one should replace $\mathscr{H}$ by the Hecke ring of (J, τ), π^I by $\mathrm{Hom}_J(\tau, \pi)$, and $H^q_c(\mathcal{M}^{\mathrm{ca}}_{\mathrm{LT},I}, \overline{\mathbb{F}}_\ell)$ by $R^q\,\mathrm{Hom}_J(\tau, R\Gamma_c(\mathcal{M}^{\mathrm{ca}}_{\mathrm{LT}}, \overline{\mathbb{F}}_\ell))$. The main problem may then be to show that the latter is torsion free. Granted this, and since the Hecke ring of (J, τ) is known to be isomorphic to a Hecke–Iwahori algebra, all the combinatorics should be the same as in this paper. *However*, as in the regular case, at the moment we don't see how to get any handle on L^*_π when π is not unipotent.

1.2.4. *Organization of the paper.* Section 2 deals with most of the representation-theoretic prerequisites. We recall and expand on Vignéras' appendix to [Clozel et al.

2008] to describe the unipotent block and the elliptic unipotent representations in our context. We then compute extensions of Iwahori-invariants of these representations. This involves representation theory of the symmetric group, and in particular some knowledge of the decomposition of tensor products. We postpone to the Appendix a delicate computation of cup-products in this context, which we use in the study of R_π^{red}. Section 3 deals with the cohomological study. The main point is to show that some spectral sequences degenerate, which implies that the cohomology complex at Iwahori level is split. With this splitting property and our knowledge of extensions and cup-products from Section 2 in hand, the results listed above are quite easy computations. The theorem we gave in Section 1.2.1 is proved in Section 3.3 and our main theorem in Section 3.4.

2. Representation theory

2.1. ***The unipotent block.*** We put $G := \mathrm{GL}_d(K)$ and denote by $\mathrm{Rep}_R^\infty(G)$ the abelian category of smooth representations of G with coefficients in the commutative ring R. Let $\boldsymbol{b}$ be the unique primitive idempotent of the center $\mathfrak{Z}_{\mathbb{Z}_\ell}(G)$ of the category $\mathrm{Rep}_{\mathbb{Z}_\ell}^\infty(G)$ which is nonzero on the trivial representation. Denote by $\mathrm{Rep}_{\boldsymbol{b}}^\infty(G)$ the full subcategory of all objects on which $\boldsymbol{b}$ acts by the identity. This is a Serre subcategory, called the *unipotent block of* $\mathrm{Rep}_{\mathbb{Z}_\ell}^\infty(G)$.

Let I be the standard Iwahori subgroup of G and let I^ℓ be the maximal prime-to-ℓ subgroup of I. This is a distinguished open subgroup of I and the quotient I/I^ℓ is isomorphic to the ℓ-Sylow subgroup $\mathrm{Syl}_\ell(\mathbb{F}_q^\times)^d$ of $(\mathbb{F}_q^\times)^d$.

2.1.1. Proposition. *The unipotent block is generated by the projective representation* $\mathbb{Z}_\ell[G/I^\ell]$.

Proof. When $\mathbb{Z}_\ell$ is replaced by $\overline{\mathbb{F}}_\ell$, this is explained in [Clozel et al. 2008, Appendix 1], and we could probably reduce our claim to this reference. Here is another argument relying on our description of the unipotent block in [Dat 2012a, §3.1]. Indeed, by [Dat 2012a, Proposition 3.1.2], our claim reduces to a claim about the finite group $\overline{G} := \mathrm{GL}_d(\mathbb{F}_q)$. Namely, let $\overline{B}^\ell$ be the maximal prime-to-ℓ subgroup of the standard Borel subgroup $\overline{B}$ of $\overline{G}$, and let $\boldsymbol{b}_{\overline{G}}$ be the central idempotent in $\mathbb{Z}_\ell[\overline{G}]$ corresponding to the sum of all blocks that contain a unipotent rational series. Explicitly, $\boldsymbol{b}_{\overline{G}}$ is the sum in $\overline{\mathbb{Q}}_\ell[\overline{G}]$ of all central primitive idempotents e_π corresponding to irreducible $\overline{\mathbb{Q}}_\ell\overline{G}$-representations π whose semisimple part s_π in Lusztig's classification is an ℓ-element. Then the claim is that $\overline{P} := \mathbb{Z}_\ell[\overline{G}/\overline{B}^\ell]$ *is a projective generator of the category* $\boldsymbol{b}_{\overline{G}}\,\mathrm{Rep}_{\mathbb{Z}_\ell}(\overline{G})$.

It is indeed clear that $\overline{P}$ is projective. Moreover, the Jordan–Hölder constituents of $\overline{P} \otimes_{\mathbb{Z}_\ell} \overline{\mathbb{Q}}_\ell$ are all the irreducible representations whose semisimple part in Lusztig's parametrization is an ℓ-element of a *split* torus of the dual group $\overline{G}^*$. But because $q \equiv 1 (\mathrm{mod}\ \ell)$, every semisimple ℓ-element of $\overline{G}^*$ lies in a split torus. □

If H is an open compact subgroup of G, we denote by $\mathcal{H}_R(G, H)$ the Hecke algebra of left and right H-invariant measures on G with coefficients in R.

2.1.2. Corollary. *The functor $V \mapsto V^{I^\ell}$ induces an equivalence of categories*

$$\mathrm{Rep}^\infty_{\boldsymbol{b}}(G) \xrightarrow{\sim} \mathrm{Mod}(\mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)),$$

a quasi-inverse of which is the functor $M \mapsto \mathbb{Z}_\ell[G/I^\ell] \otimes_{\mathcal{H}_{\mathbb{Z}_\ell}(G,I^\ell)} M$.

The "intersection" $\mathrm{Rep}^\infty_{\overline{\mathbb{F}}_\ell}(G) \cap \mathrm{Rep}^\infty_{\boldsymbol{b}}(G)$ is Vignéras' unipotent block. In particular, the irreducible $\overline{\mathbb{F}}_\ell$-objects in $\mathrm{Rep}^\infty_{\boldsymbol{b}}(G)$ are the irreducible $\overline{\mathbb{F}}_\ell$-representations which appear as subquotients of some representation $\mathrm{Ind}_B^G(\chi)$, induced from an unramified $\overline{\mathbb{F}}_\ell$-character χ of a Borel subgroup B; see [Dat 2012a, Proposition 3.1.3]. Via the functor of the above corollary, these irreducible objects are in bijection with simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$-modules.

Let $\mathfrak{m}_I$ be the maximal ideal of the local subalgebra $\mathbb{Z}_\ell[I/I^\ell]$ of $\mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)$.

2.1.3. Proposition. *The ideal $\mathfrak{m} := \mathfrak{m}_I \mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)$ is two sided and is equal to $\mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)\mathfrak{m}_I$. The map $[I^\ell g I^\ell] \mapsto [IgI]$ induces an isomorphism of $\mathbb{F}_\ell$-algebras*

$$\mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)/\mathfrak{m} \xrightarrow{\sim} \mathcal{H}_{\mathbb{F}_\ell}(G, I).$$

Before proving the proposition, we introduce some more notation. Let T be the diagonal torus in G and let $N := N_G(T)$ be its normalizer. We denote by T^0 the maximal compact subgroup of T, and by T^ℓ the maximal prime-to-ℓ subgroup of T^0. Both are normal subgroups of N.

Proof. Since $T^\ell \subset I^\ell$, any element $w \in N/T^\ell$ gives rise to a well-defined Hecke operator $[I^\ell w I^\ell]$. By the Iwahori decomposition, the Hecke operators $[I^\ell w I^\ell]$, with w running on N/T^ℓ, form a $\mathbb{Z}_\ell$-basis of $\mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)$. Among them, the operators $[I^\ell t I^\ell]$ with $t \in T^0/T^\ell$ form a basis of the subalgebra $\mathbb{Z}_\ell[I/I^\ell]$. Then, the formula

$$[I^\ell w I^\ell] * [I^\ell t I^\ell] = [I^\ell w t w^{-1} I^\ell] * [I^\ell w I^\ell]$$

shows that $\mathfrak{m}_I \mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell) = \mathcal{H}_{\mathbb{Z}_\ell}(G, I^\ell)\mathfrak{m}_I$ is a two-sided ideal, since $\mathfrak{m}_I$ is generated by elements $1 - [I^\ell t I^\ell]$, $t \in T/T^\ell$. The same formula shows that the map of the proposition is an isomorphism of $\mathbb{F}_\ell$-vector spaces. That it is a morphism of algebras follows from the definition of convolution products on both sides, and the fact that for an element $w \in N$ the obvious map

$$(I^\ell \cap w I^\ell w^{-1})\backslash I^\ell \longrightarrow (I \cap w I w^{-1})\backslash I$$

is a bijection. □

We note that if M is an $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$-module, then $M/\mathfrak{m}M$ identifies with the I-coinvariants M_I, where I acts through I/I^ℓ.

2.1.4. Corollary. *Any simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$-module is killed by* $\mathfrak{m}$, *and the map of the previous proposition induces a bijection between simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$-modules and simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$-modules. Equivalently, for any irreducible $\overline{\mathbb{F}}_\ell$-representations V of G, we have $V^{I^\ell} = V^I \xrightarrow{\sim} V_I$ and the functor $V \mapsto V^I$ induces a bijection between irreducible $\overline{\mathbb{F}}_\ell$-representations in* $\mathrm{Rep}^\infty_{\boldsymbol{b}}(G)$ *and simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$-modules.*

2.2. *Elliptic unipotent representations.* We first recall the structure of the Iwahori–Hecke algebra, taking into account the fact that $q = 1$ in $\overline{\mathbb{F}}_\ell$.

2.2.1. Fact. *The map $w \mapsto [IwI]$ is an isomorphism of $\overline{\mathbb{F}}_\ell$-algebras*

$$\overline{\mathbb{F}}_\ell[N/T^0] \xrightarrow{\sim} \mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I).$$

Proof. In general, this map induces an isomorphism of algebras $T_w \mapsto [IwI]$ from the Iwahori–Hecke algebra $\mathcal{H}_q(\tilde{W}) \otimes \overline{\mathbb{F}}_\ell$ with parameter q of the extended Weyl group $\tilde{W} = N/T^0$ of (G, T) to $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$. The multiplication in $\mathcal{H}_q(\tilde{W})$ is determined by the braid relation and the formula $T_s^2 = (q-1)T_s + q$ for each simple reflection s associated to I. Specializing at $q = 1$ therefore gives $\mathcal{H}_q(\tilde{W}) \otimes \overline{\mathbb{F}}_\ell = \overline{\mathbb{F}}_\ell[\tilde{W}]$. $\square$

We now revisit the classical relation between parabolic induction and compact induction in the context where $q = 1$ in $\overline{\mathbb{F}}_\ell$. Let $\chi : T/T^0 \longrightarrow \overline{\mathbb{F}}_\ell^\times$ be an unramified character of T and consider the (unnormalized) induction $\mathrm{Ind}_B^G(\chi)$. Let us also identify the symmetric group $\mathfrak{S}_d$ with the subgroup of permutation matrices of G in the usual way. Because of the double cosets decomposition $G = \bigsqcup_{w \in \mathfrak{S}_d} IwB = \bigsqcup_{w \in \mathfrak{S}_d} I^\ell wB$, we see that

$$\mathrm{Ind}_B^G(\chi)^{I^\ell} = \mathrm{Ind}_B^G(\chi)^I. \tag{2.2.2}$$

In particular, the action of $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$ on $\mathrm{Ind}_B^G(\chi)^{I^\ell}$ factors through $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$. Through the previous isomorphism $\overline{\mathbb{F}}_\ell[N/T^0] \xrightarrow{\sim} \mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$, this action is given as follows.

2.2.3. Fact. *For $w \in \mathfrak{S}_d$, let $[IwB]_\chi$ be the unique element of $\mathrm{Ind}_B^G(\chi)^I$ that is supported on IwB and takes value 1 on w. Similarly, let $[wT]_\chi$ be the unique element of $\mathrm{ind}_T^N(\chi)$ that is supported on wT and takes value 1 on w. Then the map $[wT]_\chi \mapsto [IwB]_\chi$ is an isomorphism of $\mathbb{F}_\ell[N/T^0]$-modules*

$$\mathrm{ind}_T^N(\chi) \xrightarrow{\sim} \mathrm{Ind}_B^G(\chi)^I.$$

Proof. The mixed Bruhat decomposition shows that $([IwB]_\chi)_{w \in \mathfrak{S}_d}$ is a basis of $\mathrm{ind}_B^G(\chi)^I$ over $\overline{\mathbb{F}}_\ell$, therefore the map is an isomorphism of $\overline{\mathbb{F}}_\ell$-vector spaces. It is elementary to check that $[IwI] * [IB]_\chi = [IwB]_\chi$ for all $w \in \mathfrak{S}_d$, showing that the map is $\mathfrak{S}_d$-equivariant. Moreover if $t \in T$ dilates the unipotent radical of B, we see that $[ItI] * [IB]_\chi = \chi^{-1}(t)[IB]_\chi$. Since the semigroup T^+ of all elements that dilate the radical of B generates the group T, this equality is true for all $t \in T$. The T-equivariance of the map follows. $\square$

As a particular case, we get $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)^{I^\ell} \simeq \overline{\mathbb{F}}_\ell[N/T] = \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$. Because of our assumption that $\ell > d$, the right-hand side is a semisimple $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G, I^\ell)$-module. We summarize this as follows.

2.2.4. Corollary. $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ *is a semisimple representation of* G *and the functor* $V \mapsto V^I$ *induces an isomorphism between the poset of subrepresentations of* $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ *and that of subrepresentations of the regular* $\overline{\mathbb{F}}_\ell$*-representation of the symmetric group* $\mathfrak{S}_d$.

2.2.5. *More notation.* We put $S := \{1, \dots, d-1\}$ and we think of S as the set of simple roots of T in the upper triangular matrices, numbered by rows. To each subset $J \subset S$ is associated a unique standard parabolic subgroup P_J which contains B and such that J is the set of simple roots of T in the upper triangular matrices of the Levi component L_J of P_J. Denote the Weyl group of L_J by $\mathfrak{S}_J$, a parabolic subgroup of the Weyl group $\mathfrak{S}_S = \mathfrak{S}_d$ of G. Then we have an isomorphism of $\overline{\mathbb{F}}_\ell[N/T^0]$-modules

$$\mathrm{Ind}_{P_J}^G(\overline{\mathbb{F}}_\ell)^{I^\ell} = \mathrm{Ind}_{P_J}^G(\overline{\mathbb{F}}_\ell)^I \simeq \overline{\mathbb{F}}_\ell[\mathfrak{S}_d/\mathfrak{S}_J].$$

In fact, the image of the submodule $\mathrm{Ind}_{P_J}^G(\overline{\mathbb{F}}_\ell)^I$ of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)^I$ by the map of Corollary 2.2.4 is the submodule $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d/\mathfrak{S}_J]$ of $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$. As usual in this context, for any ring R we put

$$i_J(R) := \mathrm{Ind}_{P_J}^G(R) \quad \text{and} \quad v_J(R) := i_J(R) \Big/ \sum_{K \supset J} i_K(R).$$

Recall that $\mathrm{Ind}_B^G(\overline{\mathbb{Q}}_\ell)$ is multiplicity free, with pairwise distinct irreducible subquotients all $v_J(\overline{\mathbb{Q}}_\ell)$, $J \subset S$. In contrast, $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ is not multiplicity free, and $v_J(\overline{\mathbb{F}}_\ell)$ need not be irreducible.

2.2.6. *Notation again.* We denote by $P(d)$ the set of partitions $\lambda = (\lambda_1 \geqslant \lambda_2 \geqslant \cdots)$ of d. To such a partition are associated the parabolic subgroup $\mathfrak{S}_\lambda := \mathfrak{S}_{\lambda_1} \times \mathfrak{S}_{\lambda_2} \times \cdots$ of $\mathfrak{S}_d$, the permutation module $M_\lambda := \overline{\mathbb{F}}_\ell[\mathfrak{S}_d/\mathfrak{S}_\lambda]$, and the simple $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-module S_λ. The latter appears with multiplicity one in M_λ and may be inductively characterized by equalities $M_\lambda = S_\lambda + \sum_{\mu > \lambda} m_{\lambda,\mu} S_\mu$ in the Grothendieck group of $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-modules. We will denote by π_λ the unique irreducible $\overline{\mathbb{F}}_\ell$-representation of G such that $(\pi_\lambda)^I \simeq S_\lambda$.

To a subset $J \subset S$ we associate the unique partition λ_J such that $\mathfrak{S}_{\lambda_J}$ is conjugate to $\mathfrak{S}_J$. We then have $i_J(\overline{\mathbb{F}}_\ell)^I \simeq M_{\lambda_J}$, so that

$$i_J(\overline{\mathbb{F}}_\ell) = \pi_{\lambda_J} + \sum_{\mu > \lambda_J} m_{\lambda_J,\mu} \pi_\mu$$

in the Grothendieck group of finite-length $\overline{\mathbb{F}}_\ell$-representations of G. We can also write

$$v_J(\overline{\mathbb{F}}_\ell) = \pi_{\lambda_J} + \sum_{\mu > \lambda_J} m'_{\lambda_J,\mu}\pi_\mu,$$

but in general $m'_{\lambda_J,\mu}$ need not vanish.

2.2.7. *Elliptic unipotent representations.* An irreducible representation of G is called *elliptic* if it is not a virtual sum of parabolically induced representations. We know from [Dat 2012d, lemme 3.2.1] that up to unramified twist, an elliptic unipotent representation occurs as a subquotient of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$. However, in contrast with the regular case, not all such subquotients are elliptic.

Proposition. *The representation π_λ is elliptic if and only if λ is hook-shaped, that is, if $\lambda = (i, 1^{(d-i)})$ for some $i \in \{1, \dots, d\}$.*

Proof. The set $[M_\mu]_{\mu\in P(d)}$ is a basis of the Grothendieck group of $\overline{\mathbb{F}}_\ell$-representations of $\mathfrak{S}_d$ (recall that $\ell > d$). Write $[S_\lambda] = \sum_{\mu\geqslant\lambda} a_{\lambda,\mu}[M_\mu]$. By the foregoing, π_λ is elliptic if and only if $a_{\lambda,(d)} \neq 0$. It is proved in [James and Kerber 1981, 2.3.17] that this is equivalent to λ being a hook. □

Therefore, there are only d elliptic constituents in $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$, in high contrast with the ℓ-adic or banal case (2^{d-1} of them) or the regular nonbanal case ($2^d - 1$ of them).

There is a convenient realization of the modules $S_{(i,1^{(n-i)})}$. Denote by Std the standard $(d-1)$-dimensional $\overline{\mathbb{F}}_\ell$-representation of $\mathfrak{S}_d$. This is the subrepresentation of the permutation representation on $\overline{\mathbb{F}}_\ell^d$ on the subspace of vectors whose sum of coordinates vanish.

Fact. *For $i = 1, \dots d$, we have $S_{(i,1^{(d-i)})} = \bigwedge^{d-i}\mathrm{Std}$. In particular, $S_{(d)}$ is the trivial representation, and $S_{1^{(d)}}$ is the sign representation.*

The next fact will be an important technical tool in the study of the unipotent part of the cohomology complex of the Lubin–Tate tower.

2.2.8. Proposition. *For $i \in \{1, \dots, d\}$, we have*

$$v_{\{1,\dots,i-1\}}(\overline{\mathbb{F}}_\ell) \simeq v_{\{d-i+1,\dots,d-1\}}(\overline{\mathbb{F}}_\ell) \simeq \pi_{(i,1^{(d-i)})}.$$

Proof. Because $\pi_{(i,1^{(d-i)})}$ is a Jordan–Hölder factor of both $v_{\{1,\dots,i-1\}}(\overline{\mathbb{F}}_\ell)$ and $v_{\{d-i+1,\dots,d-1\}}(\overline{\mathbb{F}}_\ell)$, it suffices to prove the following equalities of dimensions:

$$\dim_{\overline{\mathbb{F}}_\ell} v_{\{1,\dots,i-1\}}(\overline{\mathbb{F}}_\ell)^I = \dim_{\overline{\mathbb{F}}_\ell} v_{\{d-i+1,\dots,d-1\}}(\overline{\mathbb{F}}_\ell)^I = \dim_{\overline{\mathbb{F}}_\ell} S_{(i,1^{(d-i)})}.$$

From the previous fact or from the hook-length formula, we have

$$\dim_{\overline{\mathbb{F}}_\ell} S_{(i,1^{(d-i)})} = \binom{d-1}{d-i}.$$

On the other hand, for any subset $J \subset S = \{1, \dots, d-1\}$, we have by definition $v_J(\overline{\mathbb{F}}_\ell) = v_J(\mathbb{Z}_\ell) \otimes_{\mathbb{Z}_\ell} \overline{\mathbb{F}}_\ell$ and we know from [Schneider and Stuhler 1991, Corollary 4.5] that $v_J(\mathbb{Z}_\ell)$ is free over $\mathbb{Z}_\ell$. Therefore we have

$$\dim_{\overline{\mathbb{F}}_\ell} v_J(\overline{\mathbb{F}}_\ell)^I = \dim_{\overline{\mathbb{F}}_\ell} v_J(\overline{\mathbb{F}}_\ell)^{I^\ell} = \dim_{\overline{\mathbb{Q}}_\ell} v_J(\overline{\mathbb{Q}}_\ell)^{I^\ell} = \dim_{\overline{\mathbb{Q}}_\ell} v_J(\overline{\mathbb{Q}}_\ell)^I.$$

Denote by r_B the normalized Jacquet functor along B. Then Borel's theorem on principal series representations tells us that for any $J \subset S$, we have

$$\dim_{\overline{\mathbb{Q}}_\ell} v_J(\overline{\mathbb{Q}}_\ell)^I = \dim_{\overline{\mathbb{Q}}_\ell} r_B(v_J(\overline{\mathbb{Q}}_\ell)).$$

Now, for $w \in \mathfrak{S}_d$, put $J(w) := \{j \in \{1, \dots, d-1\}, w(j) < w(j+1)\}$. By [Dat 2012a, Fact 2.1.1 and subsequent paragraph], we have

$$\dim_{\overline{\mathbb{Q}}_\ell} r_B(v_J(\overline{\mathbb{Q}}_\ell)) = \#\{w \in \mathfrak{S}_d,\ J(w) = J\}.$$

Observe that in the cases where $J = \{1, \dots, i-1\}$ or $J = \{d-i, \dots, d-1\}$, the map $w \mapsto w_{|J}$ induces a bijection from $\{w \in \mathfrak{S}_d,\ J(w) = J\}$ to the set of nondecreasing maps $J \to \{1, \dots, d-1\}$. Therefore, in the same cases, the map $w \mapsto w(J)$ induces a bijection from $\{w \in \mathfrak{S}_d,\ J(w) = J\}$ to the set of subsets of size $|J| = i-1$ in S, whence the desired equalities

$$\dim_{\overline{\mathbb{F}}_\ell} v_{\{1,\dots,i-1\}}(\overline{\mathbb{F}}_\ell)^I = \dim_{\overline{\mathbb{F}}_\ell} v_{\{d-i+1,\dots,d-1\}}(\overline{\mathbb{F}}_\ell)^I = \binom{d-1}{i-1}. \qquad \square$$

2.3. ***Extensions between some simple $\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G, I)$-modules.***

2.3.1. Let G^0 be the subgroup of G generated by compact elements. We have $G^0 = \ker(|\det|_K : G \longrightarrow \mathbb{R}_+^\times)$. The isomorphism of Fact 2.2.1 restricts to an isomorphism

$$\overline{\mathbb{F}}_\ell[(N \cap G^0)/T^0] \xrightarrow{\sim} \mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0, I).$$

The group $(N \cap G^0)/T^0$ is an extension

$$(T \cap G^0)/T^0 \hookrightarrow (N \cap G^0)/T^0 \twoheadrightarrow \mathfrak{S}_d,$$

where $(T \cap G^0)/T^0$ is a free abelian group of rank $d-1$ on which the conjugation action of $N \cap G^0$ factors through $\mathfrak{S}_d$ and is the standard representation, namely

$$\overline{\mathbb{F}}_\ell \otimes_{\mathbb{Z}} ((T \cap G^0)/T^0) \simeq_{\mathfrak{S}_d} \mathrm{Std}.$$

2.3.2. Proposition. *Let A and B be two $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-modules, that we may see as $\overline{\mathbb{F}}_\ell[N/T^0]$-modules via the projection $N/T^0 \twoheadrightarrow \mathfrak{S}_d$.*

(i) *There is a natural isomorphism*

$$\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, B) \xrightarrow{\sim} \mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}(A, B \otimes {\textstyle\bigwedge^*} \mathrm{Std})$$

functorial in A and B.

(ii) *If C is another $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-module, cup-products are given by the following compositions*:

$$\begin{aligned}
&\mathrm{Ext}^k_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, B)\otimes \mathrm{Ext}^l_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(B, C)\\
\xrightarrow{\sim}\ &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}(A, B\otimes {\textstyle\bigwedge^k}\mathrm{Std})\otimes \mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}(B, C\otimes {\textstyle\bigwedge^l}\mathrm{Std})\\
\longrightarrow\ &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}(A, C\otimes {\textstyle\bigwedge^l}\mathrm{Std}\otimes {\textstyle\bigwedge^k}\mathrm{Std})\\
\longrightarrow\ &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}(A, C\otimes {\textstyle\bigwedge^{k+l}}\mathrm{Std})\\
\xrightarrow{\sim}\ &\mathrm{Ext}^{k+l}_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, C),
\end{aligned}$$

where the second map is composition and the third is induced by the exterior product.

Proof. As with any free abelian group of finite rank, there is a natural isomorphism of graded algebras

$$\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(T\cap G^0)/T^0]}(\overline{\mathbb{F}}_\ell, \overline{\mathbb{F}}_\ell)\simeq {\textstyle\bigwedge^*}\big(\overline{\mathbb{F}}_\ell\otimes_{\mathbb{Z}}((T\cap G^0)/T^0)\big).$$

This isomorphism is compatible with automorphisms of the group $(T\cap G^0)/T^0$, and in particular with the action of $\mathfrak{S}_d$. As already noted above the proposition, the right-hand side with its $\mathfrak{S}_d$ action is $\bigwedge^*\mathrm{Std}$. With A and B as in the proposition, we thus get

$$\begin{aligned}
\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, B) &\simeq \mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(T\cap G^0)/T^0]}(A, B)^{\mathfrak{S}_d}\\
&\simeq \big(\mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(A, B)\otimes {\textstyle\bigwedge^*}\mathrm{Std}\big)^{\mathfrak{S}_d}\\
&\simeq \mathrm{Hom}_{\overline{\mathbb{F}}_\ell\mathfrak{S}_d}(A, B\otimes {\textstyle\bigwedge^*}\mathrm{Std}).
\end{aligned}$$

Here in the first line we have used that $\ell > d$, so that $\mathfrak{S}_d$ has no higher cohomology on $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-modules. This also shows (ii) on the cup-products, since the algebra structure on $\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(T\cap G^0)/T^0]}(\overline{\mathbb{F}}_\ell, \overline{\mathbb{F}}_\ell)$ is given by the exterior product. □

We see in particular that the dimension of $\mathrm{Ext}^k_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(S_\lambda, S_\mu)$ equals the multiplicity of S_λ as a constituent of $S_\mu\otimes S_{(n-k,1^{(k)})}$. Computing such multiplicities is a notoriously difficult problem and remains largely open. Fortunately, enough is known for our purposes in this paper.

2.3.3. Corollary. *For $i, j, k\in\{0,\dots,d-1\}$, the dimension over $\overline{\mathbb{F}}_\ell$ of the extension space $\mathrm{Ext}^k_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}\big(\bigwedge^i\mathrm{Std}, \bigwedge^j\mathrm{Std}\big)$ is either 0 or 1. It is 1 if and only if the following inequalities hold*:

$$i+j\geqslant k,\quad j+k\geqslant i,\quad k+i\geqslant j,\quad i+j+k\leqslant 2d-2.$$

Remark. The symmetry of the above conditions should not be surprising since the dimension we are interested in is that of $\big(\bigwedge^i \mathrm{Std} \otimes \bigwedge^j \mathrm{Std} \otimes \bigwedge^k \mathrm{Std}\big)^{\mathfrak{S}_d}$ by the last proposition and the self-duality of irreducible representations of $\mathfrak{S}_d$.

The above conditions are also invariant under the transformation

$$(i, j) \mapsto (i', j') := (d-1-i, d-1-j).$$

This corresponds to the fact that $\bigwedge^{i'} \mathrm{Std} \simeq \bigwedge^i \mathrm{Std} \otimes \bigwedge^{d-1} \mathrm{Std}$.

A less symmetric formulation of the inequalities of the corollary, which is sometimes more convenient, is:

$$|j-i| \leqslant k \leqslant \min(i+j, i'+j'). \tag{2.3.4}$$

Proof of the Corollary. This is Theorem 2.1 of [Remmel 1989], formulated in a more symmetric way, and corrected. More precisely, assume, as we may from the above remark, that $i \leqslant j$ and $i+j \leqslant d-1$. Then our claim is that we have nonvanishing (and multiplicity one) if and only if $j-i \leqslant k \leqslant i+j$. To match the notation of [Remmel 1989], we put $n := d$, $r := n-i$, and $s := n-j$. Then Remmel asserts that nonvanishing (and multiplicity one) holds if and only if $j-i \leqslant k \leqslant i+j+1$, which seems incompatible with our claim, and which is obviously false when $i = j = 0$. However, there is a slight mistake in the proof, located in the third line of [Remmel 1989, p. 113], where it is asserted that "there are two possibilities for the positions of the remaining green cells [...]". Indeed, in the case that there are no green cells at all (that is, when $x = n-i-j$) there is only *one* possibility. Once corrected, we get our claim. $\square$

In our study of the cohomology complex of the Lubin–Tate tower we will need some cup-products between some nonvanishing Ext spaces of the above corollary. In order to simplify the notation a bit, we will abbreviate $\Lambda^i := \bigwedge^i \mathrm{Std}$ and write $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$ instead of $\mathrm{Ext}^k_{\overline{\mathbb{F}}_\ell[(G^0\cap T)/T^0]}\big(\bigwedge^i \mathrm{Std}, \bigwedge^j \mathrm{Std}\big)$.

2.3.5. Theorem. *For $i = 1, \dots, d-2$, we fix a generator $\zeta^1_{i,i}$ of $\mathrm{Ext}^1(\Lambda^i, \Lambda^i)$. Let i, j, k be integers such that $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$ is nonzero and $\mathrm{Ext}^{k+1}(\Lambda^i, \Lambda^j)$ is also nonzero. Then, both the cup-product maps from $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$ to $\mathrm{Ext}^{k+1}(\Lambda^i, \Lambda^j)$,*

$$\xi^1_{i,i} \cup - \quad \textit{and} \quad - \cup \xi^1_{j,j},$$

are isomorphisms.

We postpone the proof of this theorem to the Appendix, in order to lighten the exposition a bit. Let us mention the following corollary, in which we use the notation $i' = d-1-i$ introduced above.

2.3.6. Corollary. *For $i \in \{1, \dots, d-1\}$, the self-extension algebra $\mathrm{Ext}^*(\Lambda^i, \Lambda^i)$ is (graded) isomorphic to $\overline{\mathbb{F}}_\ell[X]/(X^{2\min(i,i')})$ via $X \mapsto \xi^1_{i,i}$. Moreover, for any other*

$j \in \{1, \dots, d-1\}$, *the graded space* $\mathrm{Ext}^*(\Lambda^i, \Lambda^j)$ *is a* (*shifted*) *cyclic module over* $\mathrm{Ext}^*(\Lambda^i, \Lambda^i)$.

The proof of the next result is also postponed to the Appendix, Section A.1.4.

2.3.7. Proposition. *Let* $i \leqslant j \in \{0, \dots, d-1\}$ *and let* $\xi_{j,i}^{j-i}$ *be a fixed generator of* $\mathrm{Ext}^{j-i}(\Lambda^j, \Lambda^i)$. *For any simple module* S *over* $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$, *the cup-product*

$$\xi_{j,i}^{j-i} \cup - : \mathrm{Ext}^{j'}(\Lambda^i, S) \longrightarrow \mathrm{Ext}^{i'}(\Lambda^j, S)$$

is an isomorphism.

3. Cohomology and the Lefschetz operator

As in Section 1, we denote by $R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)$ the cohomology complex of the Lubin–Tate tower of height d of the field K. This is an object of the derived category $D^b(\mathrm{Rep}_{\mathbb{Z}_\ell}^\infty(G))$ with an action of the Weil group W_K and of the unit group $D^\times$ of the division algebra with invariant $1/d$ over K. We refer to [Dat 2007, §3.2] for a precise definition of this object.

We want to compute the graded $\overline{\mathbb{F}}_\ell[D^\times \times W_K]$-module

$$R_\pi^* := \mathscr{H}^*(R\,\mathrm{Hom}_{\mathbb{Z}_\ell G}(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell), \pi))[1-d]$$

for π a unipotent irreducible $\overline{\mathbb{F}}_\ell$-representation of G. Notice the shift by $1-d$, which is here for convenience. Indeed, by [Dat 2012c, Proposition 2.1.3], the graded space R_π^* is supported in the range $[1-d, d-1]$.

3.1. ***The unipotent part of the cohomology complex.*** Thanks to the equivalence of categories of Corollary 2.1.2 we have

$$R_\pi^* \simeq R^*\,\mathrm{Hom}_{\mathscr{H}_{\mathbb{Z}_\ell}(G, I^\ell)}(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)^{I^\ell}, \pi^{I^\ell})[1-d].$$

Moreover, since the action of I on π^{I^ℓ} is trivial, Proposition 2.1.3 implies

$$\begin{aligned} R_\pi^* &\simeq \mathscr{H}^*\big(R\,\mathrm{Hom}_{\mathscr{H}_{\mathbb{F}_\ell}(G,I)}(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)^{I^\ell} \otimes^L_{\mathscr{H}_{\mathbb{Z}_\ell}(G,I^\ell)} \mathscr{H}_{\mathbb{F}_\ell}(G, I), \pi^I)\big)[1-d] \\ &\simeq \mathscr{H}^*\big(R\,\mathrm{Hom}_{\mathscr{H}_{\mathbb{Z}_\ell}(G,I)}(L_I(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)), \pi^I)\big)[1-d], \end{aligned}$$

where L_I denotes the left-derived functor of the I-coinvariant functor. Let $\mathcal{M}_{\mathrm{LT},I}$ denote the Iwahori level of the Lubin–Tate tower (a quotient of the tame level $\mathcal{M}_{\mathrm{LT},1}$). We have $L_I(R\Gamma_c(\mathcal{M}_{\mathrm{LT}}^{\mathrm{ca}}, \mathbb{Z}_\ell)) \simeq R\Gamma_c(\mathcal{M}_{\mathrm{LT},I}^{\mathrm{ca}}, \mathbb{Z}_\ell)$. Now recall that the G-tower $\mathcal{M}_{\mathrm{LT}}$ is induced from a G^0-tower $\mathcal{M}_{\mathrm{LT}}^{(0)}$ (the notation we use is that of [Dat 2007, 3.5.1]). The analytic space $\mathcal{M}_{\mathrm{LT},I}^{(0)}$ is the deformation space with Iwahori level structure of a formal $\mathcal{O}_K$-module of height d over $\bar{k}$. Finally we have obtained

$$R_\pi^* \simeq \mathscr{H}^*\big(R\,\mathrm{Hom}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(R\Gamma_c(\mathcal{M}_{\mathrm{LT},I}^{\mathrm{ca},(0)}, \overline{\mathbb{F}}_\ell), \pi^I)\big)[1-d]. \tag{3.1.1}$$

Let us abbreviate

$$\mathscr{C}_I := R\Gamma_c(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I}, \overline{\mathbb{F}}_\ell)[d-1] \in D^b(\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)).$$

Its cohomology is quite easy to describe, although the author does not know any elementary proof. Recall that any $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-module inflates to a $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)$-module via the isomorphism $\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0] \xrightarrow{\sim} \mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)$ and the projection

$$(N\cap G^0)/T^0 \twoheadrightarrow \mathfrak{S}_d.$$

3.1.2. Proposition. *The cohomology of $\mathscr{C}_I$ vanishes outside the range $\{0,\dots,d-1\}$. For $i\in\{0,\dots,d-1\}$, we have $\mathscr{H}^i(\mathscr{C}_I)\simeq \bigwedge^{i'}\mathrm{Std}$, where $i' := d-1-i$.*

Proof. The deformation ring $\mathscr{R}_I$ with Iwahori level structure of a formal $\mathbb{O}_K$-module of height d is known to be isomorphic to $\breve{\mathbb{O}}[\![X_1,\dots,X_d]\!]/(X_1X_2\cdots X_d - \varpi)$, where $\breve{\mathbb{O}}$ is the completed maximal unramified extension of $\mathbb{O}_K$ and ϖ is a uniformizer of $\mathbb{O}_K$. A reference is [Taylor and Yoshida 2007, top of p. 483]. It follows that the vanishing cycles $\Psi^i(\mathscr{R}_I,\mathbb{Z}_\ell)$ are isomorphic, as $\mathbb{Z}_\ell$-modules, to $\bigwedge^i(\mathbb{Z}_\ell^{d-1})$. Since $\mathscr{H}^i(\mathscr{C}_I) = H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\overline{\mathbb{F}}_\ell) = \Psi^{d-1-i}(\mathscr{R}_I,\overline{\mathbb{F}}_\ell)^\vee$, we get at least the expected dimension for these cohomology spaces.

Unfortunately, computing the action of the Hecke algebra is not so easy. However, here we observe that $H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\mathbb{Z}_\ell)$ is torsion free, so that

$$H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\overline{\mathbb{F}}_\ell) = H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\mathbb{Z}_\ell)\otimes\overline{\mathbb{F}}_\ell,$$

and we may hope to deduce $H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\overline{\mathbb{F}}_\ell)$ by reduction modulo ℓ of

$$H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\overline{\mathbb{Q}}_\ell).$$

Indeed, by Proposition 2.2.8, it suffices to know that

$$H_c^{d-1+i}(\mathcal{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I},\overline{\mathbb{Q}}_\ell)\simeq v_{\{1,\dots,i\}}(\overline{\mathbb{Q}}_\ell)^I.$$

There are two ways to infer such an isomorphism. It follows from Boyer's local theorem in [Boyer 2009], which uses global arguments, but remains "on the Lubin–Tate side". It also follows by purely local arguments, from the Faltings–Fargues theorem [Fargues 2008] that the cohomology of the Lubin–Tate tower coincides with that of the Drinfeld tower, and the Schneider–Stuhler computation [Schneider and Stuhler 1991] of the cohomology of the Drinfeld symmetric space. □

3.1.3. Corollary. *We have $R^*_\pi = 0$ unless π is a subquotient of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ and $\pi^I = \pi_\lambda$ with λ a hook or double-hook partition. Moreover, the action of $D^\times$ is trivial and that of W_K is unipotent.*

Proof. By the proposition we have a spectral sequence

$$E_2^{p,q} := \mathrm{Ext}^p_{\mathscr{H}(G^0,I)}\big(\textstyle\bigwedge^{q'}\mathrm{Std},\pi^I\big)\Rightarrow R_\pi^{p-q}.$$

So R^*_π vanishes unless $\mathrm{Ext}^p_{\mathcal{H}(G^0,I)}\left(\bigwedge^{q'}\mathrm{Std}, \pi^I\right) \neq 0$ for some p and q. In this case, π^I has to be trivial on $(T\cap G^0)/T^0 \subset \mathcal{H}(G^0, I)$, so that π is a subquotient of $\mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ and π^I comes from a simple $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-module. Then, by Proposition 2.3.2 this simple module occurs in $\bigwedge^{q'}\mathrm{Std}\otimes\bigwedge^p\mathrm{Std}$. It follows from Remmel's theorem [Remmel 1989] that this simple module is associated to a double-hook or hook partition.

Let us turn to the actions of W_K and $D^\times$. We know that W_K acts trivially on the cohomology of $\mathscr{C}_I$ (because of $q=1$ in $\overline{\mathbb{F}}_\ell$), therefore W_K acts unipotently on $\mathscr{C}_I$ hence also on R^*_π. For the same reason, the action of $D^\times$ on $\mathscr{C}_I$ has to be unipotent. However, the center $F^\times$ of $D^\times$ acts on R^*_π by the same character as $F^\times$ acts on π, that is, the trivial character. Since ℓ does not divide the pro-order of $D^\times/F^\times$, we deduce that $D^\times$ acts trivially. □

One consequence of the next section will be the following theorem.

3.1.4. Theorem. *The complex $\mathscr{C}_I$ is split in $D^b(\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0, I))$. Namely, we have (noncanonically) $\mathscr{C}_I \simeq \bigoplus_{i=0}^{n-1}\bigwedge^{i'}\mathrm{Std}[-i]$ in $D^b(\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0, I))$.*

Proof. In the proof of Theorem 3.2.1 below, we get the following property on $\mathscr{C}_I$. For all $i=0,\dots,d-1$, the spectral sequence

$$E_2^{pq} = \mathrm{Ext}^p_{\mathcal{H}(G^0,I)}(H^q(\mathscr{C}_I), H^i(\mathscr{C}_I)) \Rightarrow \mathrm{Hom}_{D^b(\mathcal{H}(G^0,I))}(\mathscr{C}_I[q-p], H^i(\mathscr{C}_I))$$

degenerates at E_2. But then, it follows from the proof of the implication (i) ⇒ (ii) of [Deligne 1968, Proposition (1.2)] (or rather a dual version of it, as in [ibid., Remark (1.4)]) that the complex $\mathscr{C}_I$ is split. □

Remark. In contrast, the complex $\boldsymbol{b}_G R\Gamma_c(\mathcal{M}_{\mathrm{LT}}, \overline{\mathbb{F}}_\ell)$ is certainly *not split* in $D^b_{\overline{\mathbb{F}}_\ell}(G)$. Equivalently, $\mathscr{C} := R\Gamma_c(\mathcal{M}_{\mathrm{LT}}, \overline{\mathbb{F}}_\ell)^{I^\ell}$ is not split in $D^b(\mathcal{H}(G, I^\ell))$. Indeed, it is a perfect complex of $\overline{\mathbb{F}}_\ell[I/I^\ell]$-modules whose cohomology spaces are not of finite projective dimension since I acts trivially on them.

3.2. *The graded dimension of R^*_π when π is elliptic unipotent.* For $j=0,\dots,d-1$, we put $\pi_j := \pi_{(d-j,1^{(j)})}$, so that $(\pi_j)^I \simeq \bigwedge^j\mathrm{Std}$. As in Proposition 3.1.2, we put

$$j' := d-1-j.$$

3.2.1. Theorem. *The graded vector space $R^*_{\pi_j}$ is supported in the range $[-j', j']$. For $k\in[-j', j']$ we have*

$$\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j}) = \begin{cases} j+1, & \text{if } k-j' \text{ is even},\\ j, & \text{if } k-j' \text{ is odd}.\end{cases}$$

Proof. We prove equality of dimensions by proving inequalities in both directions.

In order to bound above $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j})$, we use the spectral sequence

$$E_2^{p,q} := \mathrm{Ext}^p_{\mathcal{H}(G^0,I)}(\mathcal{H}^q(\mathcal{C}_I), \pi_j^I) \Rightarrow R^{p-q}_{\pi_j}. \tag{3.2.2}$$

Proposition 3.1.2 tells us that $E_2^{p,q} = \mathrm{Ext}^p_{\mathcal{H}(G^0,I)}\big(\bigwedge^{q'}\mathrm{Std}, \bigwedge^j \mathrm{Std}\big)$ and Corollary 2.3.3 then ensures that $\dim_{\overline{\mathbb{F}}_\ell}(E_2^{p,q}) \leqslant 1$ for all p and q and

$$\begin{aligned}\dim_{\overline{\mathbb{F}}_\ell}(E_2^{p,q}) = 1 &\iff (-j' \leqslant p-q \leqslant j' \text{ and } -j \leqslant p+q-(d-1) \leqslant j)\\ &\iff (p,q) \text{ lies in the rectangle } (0,j'), (j',0), (d-1,j), (j,d-1).\end{aligned}$$

This rectangle is contained in the square $[0,d-1]\times[0,d-1]$ and its faces have slopes ± 1. Since this spectral sequence has finite support, it converges and we have

$$\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j}) = \sum_{i=0}^{d-1}\dim(E_\infty^{k+i,i}) \leqslant \sum_{i=0}^{d-1}\dim(E_2^{k+i,i}).$$

In particular, we see that $R^k_{\pi_j}$ vanishes unless $-j' \leqslant k \leqslant j'$, in which case we get

$$\begin{aligned}\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j}) &\leqslant \#\{i \in \{0,\dots,d-1\}, -j \leqslant k+2i-(d-1) \leqslant j\}\\ &= \#\{i \in \{0,\dots,d-1\}, -j+k' \leqslant 2i \leqslant j-k'\}\\ &= \#\{\text{even integers in the range } [-j+k', j+k']\}.\end{aligned}$$

For the last equality, we use that $[-j+k', j+k'] \subset [0, 2d-2]$, which is indeed equivalent to $-j' \leqslant k \leqslant j'$. Now the last expression in the right-hand side above is $j+1$ if $-j+k' = j'-k$ is even, and is j otherwise.

We now look for lower bounds on $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j})$. We will use the fact that, by Proposition 2.2.8, we have

$$\pi_j \simeq v_{\{1,\dots,j'\}}(\overline{\mathbb{F}}_\ell) \simeq v_{\{j+1,\dots,d-1\}}(\overline{\mathbb{F}}_\ell).$$

Denote by $\breve{\mathbb{Z}}_\ell$ the Witt vectors of $\overline{\mathbb{F}}_\ell$. We put

$$\omega_j^+ := v_{\{1,\dots,j'\}}(\breve{\mathbb{Z}}_\ell) \quad \text{and} \quad \omega_j^- := v_{\{j+1,\dots,d-1\}}(\breve{\mathbb{Z}}_\ell).$$

As recalled in the proof of Proposition 2.2.8, these are liftings of π_j over $\breve{\mathbb{Z}}_\ell$, that is, admissible free $\breve{\mathbb{Z}}_\ell$-representations of G such that

$$\omega_j^\pm \otimes_{\breve{\mathbb{Z}}_\ell} \overline{\mathbb{F}}_\ell \simeq \pi_j.$$

Therefore we have universal coefficients exact sequences

$$R^k_{\omega_j^\pm} \otimes_{\breve{\mathbb{Z}}_\ell} \overline{\mathbb{F}}_\ell \hookrightarrow R^k_{\pi_j} \twoheadrightarrow R^{k+1}_{\omega_j^\pm}[\ell],$$

for all $k \in \mathbb{Z}$, and where the $[\ell]$ denotes ℓ-torsion (kernel of multiplication by ℓ).

Since the $R^k_{\omega_j^\pm}$ are finitely generated $\breve{\mathbb{Z}}_\ell$-modules, we have equalities

$$\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^\pm}\otimes_{\breve{\mathbb{Z}}_\ell}\overline{\mathbb{F}}_\ell)=\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^\pm}[\ell])+\dim_{\breve{\mathbb{Q}}_\ell}(R^k_{\omega_j^\pm}[1/\ell]),$$

where $\breve{\mathbb{Q}}_\ell=\breve{\mathbb{Z}}_\ell[1/\ell]$ is the fraction field of $\breve{\mathbb{Z}}_\ell$. Therefore we get

$$\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j})=\dim_{\overline{\mathbb{F}}_\ell}(R^{k+1}_{\omega_j^\pm}[\ell])+\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^\pm}[\ell])+\dim_{\breve{\mathbb{Q}}_\ell}(R^k_{\omega_j^\pm}[1/\ell]). \tag{3.2.3}$$

Since $\omega_j^+[1/\ell]=v_{\{1,\dots,j'\}}(\breve{\mathbb{Q}}_\ell)$ and $\omega_j^-[1/\ell]=v_{\{j+1,\dots,n-1\}}(\breve{\mathbb{Q}}_\ell)$, we have already computed the last summand of the right-hand side in [Dat 2006] (see, more precisely, the display below [Dat 2006, Lemma 4.4.1]). This gives

$$\dim_{\breve{\mathbb{Q}}_\ell}(R^k_{\omega_j^\pm}[1/\ell])=1 \quad \text{if } -j'<k<j' \text{ and } j'-k \text{ is even,} \tag{3.2.5}$$

$$\dim_{\breve{\mathbb{Q}}_\ell}(R^{-j'}_{\omega_j^+}[1/\ell])=j+1, \quad \dim_{\breve{\mathbb{Q}}_\ell}(R^{j'}_{\omega_j^+}[1/\ell])=1, \tag{3.2.6}$$

$$\dim_{\breve{\mathbb{Q}}_\ell}(R^{-j'}_{\omega_j^-}[1/\ell])=1, \quad \dim_{\breve{\mathbb{Q}}_\ell}(R^{j'}_{\omega_j^-}[1/\ell])=j+1, \tag{3.2.7}$$

$$\dim_{\breve{\mathbb{Q}}_\ell}(R^k_{\omega_j^\pm}[1/\ell])=0 \quad \text{in all other cases.} \tag{3.2.8}$$

Case $k=-j'$. In this case, the equality $\dim_{\overline{\mathbb{F}}_\ell}(R^{-j'}_{\pi_j})=j+1$ follows from (3.2.3) applied to ω_j^+ in degree $-j'$, (3.2.6), and our previously obtained upper bound.

Case $k=j'$. Similarly, the equality $\dim_{\overline{\mathbb{F}}_\ell}(R^{j'}_{\pi_j})=j+1$ follows from (3.2.3) applied to ω_j^- in degree j', (3.2.7), and our previously obtained upper bound.

Case $-j'<k<j'$. For k in this range, we are going to prove that

$$\begin{aligned}&\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^+}[\ell])=j \text{ and } \dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^-}[\ell])=0 \quad \text{if } j'-k \text{ is even,}\\ &\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^+}[\ell])=0 \text{ and } \dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^-}[\ell])=j \quad \text{if } j'-k \text{ is odd.}\end{aligned} \tag{3.2.8}$$

Because of (3.2.3) and (3.2.5), this implies our desired equalities:

$$\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\pi_j})=\begin{cases} j+1 & \text{if } j'-k \text{ is even,}\\ j & \text{if } j'-k \text{ is odd.}\end{cases}$$

We will prove (3.2.8) by induction on k. The first case is $k=-j'+1$. When (3.2.3) is applied to ω_j^- in degree $-j'$ it reads

$$j+1=\dim_{\overline{\mathbb{F}}_\ell}(R^{-j'+1}_{\omega_j^-}[\ell])+\dim_{\overline{\mathbb{F}}_\ell}(R^{-j'}_{\omega_j^-}[\ell])+1.$$

The same equation in degree $-j'-1$ tells us that $R^{-j'}_{\omega_j^-}[\ell]=0$, whence the desired equality $\dim_{\overline{\mathbb{F}}_\ell}(R^{-j'+1}_{\omega_j^-}[\ell])=j$. On the other hand, (3.2.3) applied to ω_j^+ in degree $-j'$ immediately implies that $\dim_{\overline{\mathbb{F}}_\ell}(R^{-j'+1}_{\omega_j^+}[\ell])=0$.

We now assume that (3.2.8) has been proved up to $k-1$ and we want to prove it for k. We distinguish two cases.

Suppose first that $j'-k$ is even. Then our induction hypothesis tells us that $\dim_{\overline{\mathbb{F}}_\ell}(R^{k-1}_{\omega_j^-}[\ell])=j$ so that the upper bound already obtained and (3.2.3) for ω_j^- in degree $k-1$ imply that $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^-}[\ell]) = 0$ and also that $\dim_{\overline{\mathbb{F}}_\ell}(R^{k-1}_{\pi_j}) = j$. Then, (3.2.3) for ω_j^+ in degree $k-1$ together with the vanishing of $R^{k-1}_{\omega_j^+}[\ell]$ (induction hypothesis) and (3.2.8) tell us that $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^+}[\ell]) = j$, as desired.

Next, suppose that $j'-k$ is odd, and apply (3.2.3) to ω_j^+ in degree $k-1$. By the induction hypothesis, the upper bound, and (3.2.5), we get that $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^+}[\ell]) = 0$ and also that $\dim_{\overline{\mathbb{F}}_\ell}(R^{k-1}_{\pi_j}) = j+1$. Apply then (3.2.3) to ω_j^- in degree $k-1$. Again the induction hypothesis and (3.2.5) tell us that $\dim_{\overline{\mathbb{F}}_\ell}(R^k_{\omega_j^-}[\ell]) = j$, as desired. $\square$

Corollary (of the proof). *The spectral sequence* (3.2.2) *degenerates at* E_2.

In particular, Theorem 3.1.4 is now proved. We may use it to recast the foregoing result in the following way.

3.2.9. Corollary. *Any splitting* $\bigoplus_{q=0}^{d-1} \bigwedge^{q'} \mathrm{Std}[-q] \xrightarrow{\sim} \mathscr{C}_I$ *as in Theorem 3.1.4 induces a graded isomorphism*

$$R^*_{\pi_j} \xrightarrow{\sim} \bigoplus_{\substack{-j' \leqslant p-q \leqslant j' \\ -j \leqslant p+q+1-d \leqslant j}} \mathrm{Ext}^p_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}\big(\textstyle\bigwedge^{q'} \mathrm{Std}, \bigwedge^j \mathrm{Std}\big)[q-p].$$

Moreover, *each term of the above sum has dimension* 1.

3.3. ***The description of the pair* (R^*_π, L^*_π).** In this section we prove the theorem from Section 1.2.1 and its corollary.

3.3.1. Let us write $i_B := \mathrm{Ind}_B^G(\overline{\mathbb{F}}_\ell)$ and consider the graded $\overline{\mathbb{F}}_\ell$-vector space $R^*_{i_B}$. By (3.1.1) we have

$$R^*_{i_B} \simeq \mathscr{H}^*\big(R\,\mathrm{Hom}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(R\Gamma_c(\mathscr{M}^{\mathrm{ca},(0)}_{\mathrm{LT},I}, \overline{\mathbb{F}}_\ell), (i_B)^I)\big)[1-d].$$

On the other hand we have on the left $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)$-module $(i_B)^I = \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$ a right module structure over $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$ which induces a left module structure on $R^*_{i_B}$. Now let $\lambda \in \mathscr{P}(d)$ and denote by ε_λ the central idempotent corresponding to the simple module S_λ, as well as $d_\lambda := \dim_{\overline{\mathbb{F}}_\ell} S_\lambda$. We then may recover $R^*_{\pi_\lambda}$ by applying ε_λ:

$$(R^*_{\pi_\lambda})^{\oplus d_\lambda} = \varepsilon_\lambda R^*_{i_B}.$$

As in Corollary 3.2.9, the splitting property of $\mathscr{C}_I$ shows that for $k = 0, \dots, 2d-2$ we have

$$R^{k+1-d}_{i_B} \simeq \bigoplus_{p-q=k+1-d} \mathrm{Ext}^p_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}\big(\textstyle\bigwedge^{q'} \mathrm{Std}, \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]\big).$$

Inserting Proposition 2.3.2 we get

$$R^{k+1-d}_{i_B} \simeq \bigoplus_{p+q'=k} \mathrm{Hom}_{\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]}\big(\textstyle\bigwedge^{q'} \mathrm{Std}, \bigwedge^p \mathrm{Std} \otimes \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]\big).$$

By Frobenius reciprocity and self-duality we finally get a $\mathfrak{S}_d$-equivariant isomorphism

$$R^{k+1-d}_{i_B} \simeq \bigoplus_{p+q'=k} \textstyle\bigwedge^{q'} \mathrm{Std} \otimes \bigwedge^{p} \mathrm{Std},$$

which shows that, as a graded vector space $R^*_{i_B}[d-1]$ is $\mathfrak{S}_d$-equivariantly isomorphic to the graded space H^* considered in the theorem given in Section 1.2.1. So we have obtained half of this theorem and we now have to study compatibility with Lefschetz operators.

3.3.2. We refer to [Dat 2012b] for the precise definition of the Lefschetz operator

$$L : R\Gamma_c(\mathcal{M}^{\mathrm{ca}}_{\mathrm{LT}}, \overline{\mathbb{F}}_\ell) \longrightarrow R\Gamma_c(\mathcal{M}^{\mathrm{ca}}_{\mathrm{LT}}, \overline{\mathbb{F}}_\ell)[2](1)$$

on the cohomology complex of the Lubin–Tate tower. By functoriality, the latter induces a graded equivariant map $L^*_\pi : R^*_\pi \longrightarrow R^*_\pi[2](1)$ for any smooth $\overline{\mathbb{F}}_\ell$-representation of G. For our purposes here, the most useful feature of L is that it is lifted from the Chern class of the tautological bundle on the crystalline period space $\mathbb{P}^{d-1}$ of the Lubin–Tate space. This explains the following description of the pair $(R^*_{\pi_0}, L^*_{\pi_0})$ where π_0 denotes the unit representation of G over $\overline{\mathbb{F}}_\ell$ (see the first paragraph of the proof of [Dat 2012a, Theorem 4.2.2]).

Fact. *We have an isomorphism* $R^*_{\pi_0} \simeq H^*(\mathbb{P}^{d-1,\mathrm{ca}}, \overline{\mathbb{F}}_\ell)[d-1] = \bigoplus_{i=0}^{d-1} \overline{\mathbb{F}}_\ell[d-1-2i]$ *with* $L^*_{\pi_0}$ *corresponding to the Chern class of the tautological sheaf on* $\mathbb{P}^{d-1}$.

Because of our assumption that $q \equiv 1[\ell]$, we may and will forget all Tate twists in the sequel. We will also denote by

$$L_I \in \mathrm{Hom}_{D^b(\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0,I))}(\mathcal{C}_I, \mathcal{C}_I[2])$$

the morphism induced by L on the complex at Iwahori level. When π is in the unipotent block and $\pi^I = \pi^{I^\ell}$, the morphism L^*_π is induced by L_I through the identification (3.1.1). We also denote by $L^{(k)} := L[2k-2] \circ \cdots \circ L : R\Gamma_c \longrightarrow R\Gamma_c[2k]$ the k-th iterate of L, and similarly for $(L_I)^{(k)}$ or $(L^*_\pi)^{(k)}$.

3.3.3. Theorem. *Let* π *be any unipotent irreducible* $\overline{\mathbb{F}}_\ell$*-representation of* G. *Then* $(L^*_\pi)^{(k)}$ *induces an isomorphism* $R^{-k}_\pi \xrightarrow{\sim} R^k_\pi$ *for any* $k \geqslant 0$.

Proof. We know from Theorem 3.1.4 that $\mathcal{C}_I$ is a split complex. Let us choose a splitting $\mathcal{C}_I \xrightarrow{\sim} \bigoplus_{i=0}^{d-1} \mathcal{H}^i(\mathcal{C}_I)[-i]$. As in Corollary 3.2.9, this induces an isomorphism from R^*_π to the graded space associated to the bigraded space

$$(p, q) \mapsto E^{p,q}_\pi := \mathrm{Ext}^p_{\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\Lambda^{q'}, \pi^I).$$

This also induces an isomorphism

$$\mathrm{Hom}_{D^b(\mathcal{H}_{\overline{\mathbb{F}}_\ell}(G^0,I))}(\mathcal{C}_I, \mathcal{C}_I[2]) \simeq \bigoplus_{i,j=0}^{d-1} \mathrm{Ext}^{i-j+2}(\mathcal{H}^i(\mathcal{C}_I), \mathcal{H}^j(\mathcal{C}_I))$$

according to which we have a decomposition $L_I = \sum_{i,j=0}^{d-1} L_I^{i,j}$. By Proposition 3.1.2 we have

$$\mathrm{Ext}^{i-j+2}(\mathcal{H}^i(\mathcal{C}_I), \mathcal{H}^j(\mathcal{C}_I)) \simeq \mathrm{Ext}^{i-j+2}(\Lambda^{i'}, \Lambda^{j'})$$

and by Corollary 2.3.3 (see (2.3.4)), the latter has dimension 1 if $|i-j| \leqslant i-j+2 \leqslant \min(i+j, i'+j')$ and vanishes otherwise. In particular, when it does not vanish, we have $i-j+2 \geqslant 1$ with equality if and only if $j = i+1$. Now, each $L_I^{i,j}$ acts on the bigraded space E_π by a map of degree $(i-j+2, i-j)$. Since $(i-j+2)+(i-j) \geqslant 0$, it follows that L_I preserves the decreasing filtration on E_π defined by $\mathrm{Fil}_r\, E_\pi := \bigoplus_{p+q\geqslant r} E_\pi^{p,q}$, hence that L_π^* preserves the filtration induced on R_π^*. In particular, we may check the expected property of $(L_\pi^*)^k$ on the associated graded space. Concretely, this means that it suffices to prove that for all $0 \leqslant p \leqslant q \leqslant d-1$, the map

$$E_\pi^{p,q} \hookrightarrow E_\pi \xrightarrow{(L_I)^{(q-p)}} E_\pi \twoheadrightarrow E_\pi^{q,p}$$

is an isomorphism. This map is the composition

$$E^{p,q} \xrightarrow{L_I^{p,p+1}} E^{p+1,q-1} \xrightarrow{L_I^{p+1,p+2}} \cdots \xrightarrow{L_I^{q-1,q}} E^{q,p}$$

and is given by cup-product:

$$(L_I^{p,p+1} \cup \cdots \cup L_I^{q-1,q}) \cup - : \mathrm{Ext}^p(\Lambda^{q'}, \pi^I) \longrightarrow \mathrm{Ext}^q(\Lambda^{p'}, \pi^I).$$

But Proposition 2.3.7 and the next lemma imply that the latter map is an isomorphism. □

Lemma. *The element* $L_I^{p,p+1} \cup \cdots \cup L_I^{q-1,q}$ *is nonzero in* $\mathrm{Ext}^{q-p}(\Lambda^{p'}, \Lambda^{q'})$.

Proof. Clearly, it suffices to prove this for $p = 0$ and $q = d-1$. Let us consider the space $\mathrm{Hom}_{D^b(\mathcal{H}_{\bar{\mathbb{F}}_\ell}(G^0,I))}(\mathcal{C}_I, \mathcal{C}_I[2d-2])$. It is isomorphic to

$$\bigoplus_{i,j=0}^{d-1} \mathrm{Ext}^{i-j+2d-2}(\mathcal{H}^i(\mathcal{C}_I), \mathcal{H}^j(\mathcal{C}_I))$$

via our splitting of $\mathcal{C}_I$. But for $i, j \in \{0, \ldots, d-1\}$ the space $\mathrm{Ext}^{i-j+2d-2}(\Lambda^{i'}, \Lambda^{j'})$ vanishes unless $i-j+2d-2 \leqslant d-1$, which happens only when $i = 0$ and $j = d-1$. In other words, we have

$$\mathrm{Hom}_{D^b(\mathcal{H}_{\bar{\mathbb{F}}_\ell}(G^0,I))}(\mathcal{C}_I, \mathcal{C}_I[2d-2]) \simeq \mathrm{Ext}^{d-1}(\mathcal{H}^0(\mathcal{C}_I), \mathcal{H}^{d-1}(\mathcal{C}_I)) = \mathrm{Ext}^{d-1}(\Lambda^{d-1}, \Lambda^0).$$

Moreover, through this identification we have

$$(L_I)^{(d-1)} = L_I^{0,1} \cup L_I^{1,2} \cup \cdots \cup L_I^{d-2,d-1}.$$

So we are left to show that $(L_I)^{(d-1)}$ is nonzero.

Now, in the case that $\pi = \pi_0$ is the unit representation of G over $\overline{\mathbb{F}}_\ell$, the explicit description recalled in the fact given in Section 3.3.2 shows that $(L^*_{\pi_0})^{(d-1)}$ is nonzero, and hence so is $(L_I)^{(d-1)}$. □

3.3.4. Theorem 3.3.3 tells us that $(R^*_{i_B}, L^*_{i_B})$ satisfies the "hard Lefschetz theorem". Therefore, by the discussion in Section 3.3.1 it is $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$-equivariantly isomorphic to the pair (H^*, L^*) shifted by $1-d$ of the theorem given in Section 1.2.1, which finishes the proof of the latter theorem.

3.3.5. For π elliptic, we will now describe precisely the isomorphism class of the pair (R^*_π, L^*_π) in the category of finite-dimensional graded vector spaces endowed with a degree-2 endomorphism. This category is abelian artinian and its indecomposable objects are isomorphic, up to shift, to some

$$\boldsymbol{P}_k := \left(\bigoplus_{i=0}^{k} \overline{\mathbb{F}}_\ell[k-2i], L_k\right) \quad \text{with } L_k \text{ the unique map of degree 2 and rank } k.$$

Note also that

$$\boldsymbol{P}_k \simeq (H^*(\mathbb{P}^k, \overline{\mathbb{F}}_\ell)[k], c(\mathcal{O}_{\mathbb{P}^k(1)})),$$

the shifted cohomology of a projective space of dimension k with its tautological Lefschetz operator.

Corollary. *For $j = 0, \ldots, d-1$, we have $(R^*_{\pi_j}, L^*_{\pi_j}) \simeq (\boldsymbol{P}_{j'})^{\oplus j+1} \oplus (\boldsymbol{P}_{j'-1})^{\oplus j}$.*

Proof. By Theorem 3.2.1, the graded pieces $R^{-j'}_{\pi_j}, R^{-j'+2}_{\pi_j}, \ldots, R^{j'}_{\pi_j}$ all have dimension $j+1$. By Theorem 3.3.3 we see that $L^*_{\pi_j}$ induces isomorphisms

$$R^{-j'}_{\pi_j} \xrightarrow{\sim} R^{-j'+2}_{\pi_j} \xrightarrow{\sim} \cdots \xrightarrow{\sim} R^{j'}_{\pi_j},$$

whence the summand $(\boldsymbol{P}_{j'})^{\oplus j+1}$. Similarly, $L^*_{\pi_j}$ induces isomorphisms

$$R^{-j'+1}_{\pi_j} \xrightarrow{\sim} R^{-j'+3}_{\pi_j} \xrightarrow{\sim} \cdots \xrightarrow{\sim} R^{j'-1}_{\pi_j}$$

between all these j-dimensional spaces, whence the summand $(\boldsymbol{P}_{j'-1})^{\oplus j}$. □

3.4. ***Computation of*** $(\boldsymbol{R}^{\mathrm{red},*}_{\boldsymbol{\pi}}, \boldsymbol{L}^{\mathrm{red}}_{\boldsymbol{\pi}})$**.** We now assume that the unipotent representation π is elliptic. So it is of the form $\pi = \pi_j$ for some $j \in \{0, \ldots, d-1\}$. The particular case $j = 0$ corresponds to the trivial representation $\pi_j \simeq \overline{\mathbb{F}}_\ell$. In this case we find that $R^*_{\pi_0} = \bigoplus_{i=0}^{d-1} \overline{\mathbb{F}}_\ell[-2i+d-1]$ so that the total dimension of $R^*_{\pi_0}$ is d. In the other extreme case, when $j = d-1$ so that π_j is the Steinberg representation, we get that $R^*_{\pi_{d-1}}$ is concentrated in degree 0 and has dimension d. These are, however, the only cases where the total dimension is d, in contrast with the other situations we have studied in previous papers (ℓ-adic, banal, and regular cases).

3.4.1. In order to recover a d-dimensional vector space, we consider the following "reduced" version of R^*_π. We put

$$\mathscr{E}^*_{\pi^I} := R^*\operatorname{End}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\pi^I) = \bigoplus_{k\geqslant 0}\operatorname{Ext}^k_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\pi^I,\pi^I),$$

the self-extension algebra of the $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)$-module π^I. This is a positively graded algebra and we denote by $\mathscr{E}^+_{\pi^I} := \bigoplus_{k>0}\operatorname{Ext}^k_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\pi^I,\pi^I)$ its augmentation ideal. Via the isomorphism (3.1.1), the graded $\overline{\mathbb{F}}_\ell$-vector space R^*_π carries a graded right-module structure over $\mathscr{E}^*_{\pi^I}$. We then put

$$R^{\mathrm{red},*}_\pi := R^*_\pi / R^*_\pi \mathscr{E}^+_{\pi^I}.$$

For π elliptic, we will show that this graded vector space has total dimension d, as desired.

3.4.2. Theorem. *Choose a splitting* $\bigoplus_{q=0}^{d-1}\bigwedge^{q'}\operatorname{Std}[-q] \xrightarrow{\sim} \mathscr{C}_I$ *as in Theorem 3.1.4. Then, through the isomorphism of Corollary 3.2.9, we have*

$$R^*_{\pi_j}\mathscr{E}^+_{\pi_j^I} \xrightarrow{\sim} \bigoplus_{\substack{-j'<p-q\leqslant j'\\ -j<p+q+1-d\leqslant j}} \operatorname{Ext}^p_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}\big(\textstyle\bigwedge^{q'}\operatorname{Std}, \bigwedge^j\operatorname{Std}\big)[q-p].$$

Therefore we also get an isomorphism

$$R^{\mathrm{red}}_{\pi_j} \simeq \bigoplus_{q=0}^{j'}\operatorname{Ext}^{j'-q}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\Lambda^{q'},\Lambda^j)[-j'+2q] \oplus \bigoplus_{q=j'+1}^{d-1}\operatorname{Ext}^{q-j'}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\Lambda^{q'},\Lambda^j)[j'].$$

In particular $R^{\mathrm{red}}_{\pi_j}$ *has total dimension d and its graded dimension is given by*

$$\dim_{\overline{\mathbb{F}}_\ell} R^{\mathrm{red},k}_{\pi_j} = \begin{cases} j+1 & \text{if } k=-j',\\ 1 & \text{if } k=-j'+2i \text{ with } 0<i\leqslant j',\\ 0 & \text{else.}\end{cases}$$

Proof. Let D be the square defined by the inequalities $-j'\leqslant p-q\leqslant j'$ and $-j\leqslant p+q-d+1\leqslant j$ and let D^+ be the complement in D of the left-side edges defined by $p-q=-j'$ and $p+q=-j+d-1$. Corollary 3.2.9 expresses $R^*_{\pi_j}$ as the graded vector space associated to the double graded vector space $(p,q)\mapsto E^{pq} := \operatorname{Ext}^p(\Lambda^{q'},\Lambda^j)$, whose support is D. For $l>0$, an element of $\mathscr{E}^l_{\pi^I} = \operatorname{Ext}^l(\Lambda^j,\Lambda^j)$ acts through a bigraded map of degree $(l,0)$. Therefore, $R^*_{\pi_j}\mathscr{E}^+_{\pi_j^I}$ is the graded space associated to a bigraded subspace of $E^{*\bullet}$ whose support is contained in D^+. In other words, the isomorphism of Corollary 3.2.9 takes $R_{\pi_j}\mathscr{E}^+_{\pi_j^I}$ in the right-hand side of our statement.

To get the other inclusion, we need to understand the action of a generator $\xi^1_{j,j}$ of $\mathscr{E}^1_{\pi_j} = \mathrm{Ext}^1(\Lambda^j, \Lambda^j)$, which is given by the cup-product

$$E^{p,q} = \mathrm{Ext}^p_{\mathscr{H}(G^0,I)}(\Lambda^{q'}, \Lambda^j) \xrightarrow{-\cup\xi^1_{j,j}} E^{p+1,q} = \mathrm{Ext}^{p+1}_{\mathscr{H}(G^0,I)}(\Lambda^{q'}, \Lambda^j).$$

But Theorem 2.3.5 tells us that this is an isomorphism, whenever both sides are nonzero. It follows that $E^{p+1,q} \subset R^*_{\pi_j}\mathscr{E}^+_{\pi_j^I}$, and finally that the isomorphism of Corollary 3.2.9 induces the claimed isomorphism. The claim on $R^{\mathrm{red}}_{\pi_j}$ is a direct a consequence of the latter. □

By definition the graded map L^*_π commutes with $\mathscr{E}^+_{\pi^I}$, so L^*_π induces a graded linear map $L^{\mathrm{red}}_\pi : R^{\mathrm{red}}_\pi \longrightarrow R^{\mathrm{red}}_\pi[2]$.

3.4.3. Corollary. *For $j = 0, \dots, d-1$, we have $(R^{\mathrm{red}}_{\pi_j}, L^{\mathrm{red}}_{\pi_j}) \simeq \boldsymbol{P}_{j'} \oplus (\boldsymbol{P}_0)^{\oplus j}[j']$. In particular we have the equality*

$$[R^{\mathrm{red}}_{\pi_j}, L^{\mathrm{red}}_{\pi_j}] = \mathrm{LJ}(\pi_j) \otimes (\sigma^{\mathrm{ss}}(\pi_j), N_{\pi_j})$$

in the Grothendieck group of Weil–Deligne $\overline{\mathbb{F}}_\ell D^\times$-representations.

Proof. By the corollary in Section 3.3.5 (and its proof) and the very definition of $R^{\mathrm{red}}_{\pi_j}$ as a quotient of $R^*_{\pi_j}$, we see that $L^{\mathrm{red}}_{\pi_j}$ induces a *surjective* map $R^{\mathrm{red},k}_{\pi_j} \twoheadrightarrow R^{\mathrm{red},k+2}_{\pi_j}$ for all $k \geqslant -j'$. So our first claim follows from the description of the graded dimension of $R^*_{\pi_j}$ in the above theorem. The second claim follows from the equalities

- $\mathrm{LJ}(\pi_j) = (-1)^j \,\mathrm{LJ}(\pi_0) = (-1)^{j'} \,\mathrm{LJ}(\pi_{d-1}) = (-1)^{j'}[1_{D^\times}]$ and
- $\sigma^{\mathrm{ss}}(\pi_j) = (1_{W_K})^{\oplus d}$ and N_{π_j} has Jordan type $\lambda = (d-j, 1^{(j)})$.

The latter equalities are seen, for example, by reducing modulo ℓ the corresponding equalities for $v_{\{1,\dots,j'\}}(\overline{\mathbb{Q}}_\ell)$, since we have $\pi_j = v_{\{1,\dots,j'\}}(\overline{\mathbb{F}}_\ell)$. □

3.4.4. Remark. There are dual versions of the theorem and corollary. Instead of considering $R^{\mathrm{red},*}_\pi$, one could consider

$$R^{\mathrm{cored},*}_\pi := R^*_\pi[\mathscr{E}^+_{\pi^I}]$$

the graded subvector space of R^*_π which is annihilated by the augmentation ideal $\mathscr{E}^+_{\pi^I}$. Then a very similar argument to the proof of the above theorem shows that the isomorphism of Corollary 3.2.9 induces the following one:

$$R^{\mathrm{cored}}_{\pi_j} \simeq \bigoplus_{q=0}^{j-1} \mathrm{Ext}^{j'+q}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\Lambda^{q'}, \Lambda^j)[j'] \oplus \bigoplus_{q=j}^{d-1} \mathrm{Ext}^{j+q'}_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\Lambda^{q'}, \Lambda^j)[j'-2q].$$

In particular $R^{\mathrm{cored},*}_{\pi_j}$ has total dimension d and graded dimension given by

$$\dim_{\overline{\mathbb{F}}_\ell} R^{\mathrm{red},k}_{\pi_j} = \begin{cases} j+1 & \text{if } k = j', \\ 1 & \text{if } k = j'-2i \text{ with } 0 < i \leqslant j', \\ 0 & \text{else.} \end{cases}$$

Similar arguments as in the last corollary show that L^*_π takes R^{cored}_π to itself, and that for $j=0,\dots,d-1$, we have $(R^{\mathrm{cored}}_{\pi_j}, L^{\mathrm{cored}}_{\pi_j}) \simeq \boldsymbol{P}_{j'} \oplus (\boldsymbol{P}_0)^{\oplus j}[-j']$, leading to the same conclusion that $[R^{\mathrm{cored}}_{\pi_j}, L^{\mathrm{cored}}_{\pi_j}] = \mathrm{LJ}(\pi_j) \otimes (\sigma^{\mathrm{ss}}(\pi_j), N_{\pi_j})$.

3.4.5. We now give the relation between the algebra $\mathscr{E}^*_{\pi^I}$ used here and that used in the statement of the main theorem in Section 1.2.2. The inflation functor from $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I)$-modules to $\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0, I^\ell)$-modules along the map of Proposition 2.1.3 yields a morphism of graded algebras

$$\mathscr{E}^*_{\pi^I} \longrightarrow \mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell G^0}(\pi,\pi)$$

and the right action of $\mathscr{E}^*_{\pi^I}$ on R^*_π that one deduces from this morphism coincides with the right-action used above, which was obtained from the expression (3.1.1). We claim that the image of the above map is the algebra $\mathscr{E}^*_\pi$ defined in Section 1.2.2. Since $\mathscr{E}^*_{\pi^I}$ is generated in degree 1, we only need to see that sequence

$$\mathrm{Ext}^1_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\pi^I,\pi^I) \longrightarrow \mathrm{Ext}^1_{\overline{\mathbb{F}}_\ell G^0}(\pi,\pi) \longrightarrow \mathrm{Ext}^1_{\overline{\mathbb{F}}_\ell \overline{G}}(\bar\pi,\bar\pi)$$

is exact. But through the equivalence of Corollary 2.1.2 this sequence reads

$$\mathrm{Ext}^1_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I)}(\pi^I,\pi^I) \longrightarrow \mathrm{Ext}^1_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(G^0,I^\ell)}(\pi^I,\pi^I) \longrightarrow \mathrm{Ext}^1_{\mathscr{H}_{\overline{\mathbb{F}}_\ell}(\mathrm{GL}_d(\mathbb{O}),I^\ell)}(\pi^I,\pi^I).$$

On the other hand, since $\ell > d$, the right-hand side coincides with $\mathrm{Ext}^1_{\overline{\mathbb{F}}_\ell[I^\ell/I]}(\pi^I,\pi^I)$. Therefore the exactness is equivalent to the fact that any extension E of π^I by π^I as $\mathscr{H}(G^0, I^\ell)$-modules is a $\mathscr{H}(G^0, I)$-module if and only if it splits as an extension of I^ℓ/I-modules. But this is clear.

Appendix: Proof of Theorem 2.3.5 and Proposition 2.3.7

In order to compute enough cup-products between the nonvanishing Ext spaces of Corollary 2.3.3, we need to exhibit explicit generators of some of these Ext spaces.

A.1. ***Construction of extensions.*** Recall that we have abbreviated $\Lambda^i := \bigwedge^i \mathrm{Std}$ and $\mathrm{Ext}^k_{\overline{\mathbb{F}}_\ell[(G^0\cap T)/T^0]}\big(\bigwedge^i \mathrm{Std}, \bigwedge^j \mathrm{Std}\big)$ by $\mathrm{Ext}^k(\Lambda^i,\Lambda^j)$.

A.1.1. *Further notation.* The standard representation $\Lambda^1 = \mathrm{Std}$ can be presented as the quotient of the permutation module $\overline{\mathbb{F}}_\ell^d$ by the line generated by $(1,1,\dots,1)$. Let $e_1,\dots,e_d$ be the image of the canonical basis of $\overline{\mathbb{F}}_\ell^d$ in Λ^1. This set of vectors enjoys the following properties:

- it is a generating set with "only" linear relation $\sum_{r=1}^d e_r = 0$ and
- the action of $\mathfrak{S}_d$ is given by $\sigma(e_r) = e_{\sigma(r)}$, where we see $\mathfrak{S}_d$ as the permutation group of $\{1,\dots,d\}$.

More generally, if I is a subset of $\{1, \dots, d\}$ of size $|I| = i$ we put

$$e_I := e_{r_1} \wedge \cdots \wedge e_{r_i} \in \Lambda^i,$$

where $I = \{r_1, r_2, \dots, r_i\}$ and $r_1 < r_2 < \cdots < r_i$. For a collection $I_1, \dots, I_n$ of subsets of $\{1, \dots, d\}$ we define $\varepsilon(I_1, \dots, I_n)$ by the equality

$$e_{I_1} \wedge \cdots \wedge e_{I_n} = \varepsilon(I_1, \dots, I_n) e_{I_1 \cup \cdots \cup I_n}.$$

Thus $\varepsilon(I_1, \dots, I_n)$ is 0 unless all I_t's are pairwise disjoint, in which case it is a sign. Now, if we fix i, the set of vectors $\{e_I, I \subset \{1, \dots, d\}, |I| = i\}$ enjoys the following properties:

- It is a generating set whose space of linear relations is generated by the following ones: $\sum_{r=1}^d \varepsilon(J, \{r\}) e_{J \cup \{r\}} = 0$ for each $J \subset \{1, \dots, d\}$ of size $|J| = i-1$. In particular, the subset of all e_I's for I contained in $\{1, \dots, d-1\}$ is a basis of Λ^i.
- The action of $\mathfrak{S}_d$ is given by $\sigma(e_I) = \mathrm{sgn}(\sigma_{|I}) e_{\sigma(I)}$, where $\sigma_{|I}$ is considered as a permutation of $\{1, \dots, i\}$ via the orderings on I and $\sigma(I)$ inherited from that of $\{1, \dots, d\}$.

When $i = i_1 + \cdots + i_n$ is a composition series of i, we have a canonical map (exterior product) $\Lambda^{i_1} \otimes \cdots \otimes \Lambda^{i_n} \xrightarrow{\text{prod}} \Lambda^i$. We will later need the quasisection $\Lambda^i \xrightarrow{\text{can}} \Lambda^{i_1} \otimes \cdots \otimes \Lambda^{i_n}$ of this map defined by

$$\begin{aligned}\mathrm{can}(v_1 \wedge \cdots \wedge v_i) \\ = \sum_{\tau \in [\mathfrak{S}_i / \mathfrak{S}_{i_1} \times \cdots \times \mathfrak{S}_{i_n}]} \mathrm{sgn}(\tau) (v_{\tau(1)} \wedge \cdots \wedge v_{\tau(i_1)}) \otimes \cdots \otimes (v_{\tau(i-i_n+1)} \wedge \cdots \wedge v_{\tau(i)}),\end{aligned}$$

where the index set stands for permutations τ that are increasing on each $\left[\sum_{s<t} i_s + 1, \sum_{s \leqslant t} i_s\right]$. Note that prod $\circ$ can is the multiplication by $i!/(i_1! \cdots i_n!)$, which is nonzero in $\overline{\mathbb{F}}_\ell$, due to our assumption that $\ell > d$. Note also that can is $\mathfrak{S}_d$-equivariant and the evaluation on e_I is given by

$$\mathrm{can}(e_I) = \sum_{I_1 \sqcup \cdots \sqcup I_n = I, |I_t| = i_t} \varepsilon(I_1, \dots, I_n) e_{I_1} \otimes \cdots \otimes e_{I_n}. \tag{A.1.2}$$

A.1.3. Definition. *For $d-1 \geqslant i \geqslant j \geqslant 0$, we denote by $\xi_{i,j}^{i-j}$ the generator of $\mathrm{Ext}^{i-j}(\Lambda^i, \Lambda^j)$ provided by the nonzero $\mathfrak{S}_d$-equivariant map $\Lambda^i \xrightarrow{\text{can}} \Lambda^j \otimes \Lambda^{i-j}$ via the isomorphism of Proposition 2.3.2(i).*

A.1.4. *Proof of Proposition 2.3.7.* Since the regular representation $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d]$ is semi-simple, it is sufficient to prove that for all $i \leqslant j$ the cup-product

$$\xi_{j,i}^{j-i} \cup - : \mathrm{Ext}^{j'}(\Lambda^i, \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]) \longrightarrow \mathrm{Ext}^{i'}(\Lambda^j, \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]) \tag{A.1.5}$$

is an isomorphism. Note that, as a $\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]$-module, we have $\overline{\mathbb{F}}_\ell[\mathfrak{S}_d] = \mathrm{ind}_{T\cap G^0}^{N\cap G^0}(\overline{\mathbb{F}}_\ell)$, so Frobenius reciprocity tells us that

$$\mathrm{Ext}^{j'}_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(\Lambda^i, \overline{\mathbb{F}}_\ell[\mathfrak{S}_d]) \simeq \mathrm{Ext}^{j'}_{\overline{\mathbb{F}}_\ell[(T\cap G^0)/T^0]}(\Lambda^i, \overline{\mathbb{F}}_\ell) \simeq \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^i, \Lambda^{j'}).$$

Now, by the formula for cup-products in Proposition 2.3.2(ii) and the expression of $\xi^{j-i}_{j,i}$ in Definition A.1.3, we see that the map (A.1.5) is isomorphic to the map

$$\text{(A.1.6)}\quad \begin{aligned} &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^i, \Lambda^{j'}) \to \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^j, \Lambda^{i'}) \\ &\qquad \varphi \mapsto \psi : (\Lambda^j \xrightarrow{\text{can}} \Lambda^i \otimes \Lambda^{j-i} \xrightarrow{\varphi\otimes\mathrm{Id}} \Lambda^{j'} \otimes \Lambda^{i'-j'} \xrightarrow{\wedge} \Lambda^{i'}). \end{aligned}$$

That the latter map is an isomorphism is an avatar of the Lefschetz decomposition of the exterior algebra of an hermitian space (see, for example, [Griffiths and Harris 1978, Chapter 0.7]). Since our field of coefficients $\overline{\mathbb{F}}_\ell$ has positive characteristic we briefly review the argument in order to ensure that it is still valid in our context.

Let us consider the graded space $H^* := \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^*, \Lambda^*) = \check{\Lambda}^* \otimes \Lambda^*$, where $\check{\Lambda}$ is the $\overline{\mathbb{F}}_\ell$-linear dual space of Λ. We endow it with an operator of degree 2 which on the (p, q) part is given by

$$\begin{aligned} L :\quad &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^p, \Lambda^q) \to \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^{p+1}, \Lambda^{q+1}) \\ &\qquad \varphi \mapsto L\varphi : (\Lambda^{p+1} \xrightarrow{\text{can}} \Lambda^p \otimes \Lambda \xrightarrow{\varphi\otimes\mathrm{Id}} \Lambda^q \otimes \Lambda \xrightarrow{\wedge} \Lambda^{q+1}), \end{aligned}$$

and an operator of degree -2 which on the (p, q)-part is given by

$$\begin{aligned} {}^tL :\quad &\mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^p, \Lambda^q) \to \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^{p-1}, \Lambda^{q-1}) \\ &\qquad \varphi \mapsto {}^tL\varphi : (\Lambda^{p-1} \xrightarrow{\wedge} \Lambda^p \otimes \check{\Lambda} \xrightarrow{\varphi\otimes\mathrm{Id}} \Lambda^q \otimes \check{\Lambda} \xrightarrow{\text{can}} \Lambda^{q-1}). \end{aligned}$$

Here the last map is the composition of can with the evaluation map $\Lambda \otimes \check{\Lambda} \longrightarrow \overline{\mathbb{F}}_\ell$. Explicitly, denoting by $e_1^*, \dots, e_{d-1}^*$ the dual basis of $e_1, \dots, e_{d-1}$, we have $L\varphi = \varphi \wedge \left(\sum_i e_i^* \otimes e_i\right)$, or, even more explicitly, we have for $I, J \subset \{1, \dots, d-1\}$ with $|I| = p$ and $|J| = q$, denoting by I^c the complementary subset,

$$L(e_I^* \otimes e_J) = \sum_{k\in I^c\cap J^c} \varepsilon(I, k)\varepsilon(J, k) e^*_{I\cup\{k\}} \otimes e_{J\cup\{k\}},$$

while

$${}^tL(e_I^* \otimes e_J) = \sum_{k\in I\cap J} \varepsilon(I \setminus \{k\}, k)\varepsilon(J \setminus \{k\}, k) e^*_{I\setminus\{k\}} \otimes e_{J\setminus\{k\}}.$$

Then a simple computation shows that commutator $[L, {}^tL]$ acts on $\mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^p, \Lambda^q)$ by multiplication by $p + q - d + 1$, hence it acts on H^k by multiplication by $k - d + 1$. It follows that the triple $(L, [L, {}^tL], {}^tL)$ is a $\mathfrak{sl}_2$-triple, that is, is the image of the canonical basis $(E_{12}, E_{11} - E_{22}, E_{21})$ of $\mathfrak{sl}_2(\overline{\mathbb{F}}_\ell)$ by a unique structure of $\mathfrak{sl}_2(\overline{\mathbb{F}}_\ell)$-modules on H^*. Moreover, the weights are $1-d, 2-d, \dots, d-1$ and the eigenspace of weight $1 - d + k$ is H^k. Now, we use our assumption that $\ell > d$.

This means that the simple $\mathfrak{sl}_2(\overline{\mathbb{F}}_\ell)$-modules with weights in the above range are constructed in the same way as for $\mathfrak{sl}_2(\mathbb{C})$. In particular, for each $\lambda \in \{0, \dots, d-1\}$ there is a unique simple module (V_λ, r_λ) of highest weight λ, its weights are $-\lambda, \lambda+2, \dots, \lambda$, and $r_\lambda(E_{12})$ induces an isomorphism between eigenspaces $V_\lambda^k \xrightarrow{\sim} V_\lambda^{k+2}$ for $-\lambda \leqslant k < \lambda$. Now, taking a filtration of H^* with simple subquotients, we see that L^k induces an isomorphism $H^{d-1-k} \xrightarrow{\sim} H^{d-1+k}$ for all $k = 0, \dots, d-1$, and consequently an isomorphism

$$L^k : \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^p, \Lambda^q) \xrightarrow{\sim} \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(\Lambda^{p+k}, \Lambda^{q+k})$$

whenever $p+q = d-1-k$. But we have $L^k\varphi = \varphi \wedge \big(k! \sum_{|K|=k} e_K^* \otimes e_K\big)$, so that L^k is also given by

$$L^k\varphi = k!(\Lambda^{p+k} \xrightarrow{\text{can}} \Lambda^p \otimes \Lambda^k \xrightarrow{\varphi\otimes\mathrm{Id}} \Lambda^q \otimes \Lambda^k \xrightarrow{\wedge} \Lambda^{q+k}).$$

Now, taking $p = i$, $k = j - i = i' - j'$, and $q = j'$ we get that (A.1.6) is an isomorphism, as desired. □

A.1.7. Lemma. *For $i = 1, \dots, d-1$, we define inductively an element $F_i \in \Lambda^i \otimes \Lambda^i$ by setting*

$$F_1 := \frac{1}{d}\sum_{r=1}^{d} e_r \otimes e_r, \quad \textit{and} \quad F_i = \frac{1}{d}\sum_{r=1}^{d} e_r \wedge F_{i-1} \wedge e_r, \ \textit{for } i > 1. \tag{A.1.8}$$

Then, F_i is a generator of $(\Lambda^i \otimes \Lambda^i)^{\mathfrak{S}_d}$ and we have the formula

$$F_i = (-1)^{i(i-1)/2}\frac{i!}{d^i} \sum_{I\subset\{1,\dots,d\},|I|=i} e_I \otimes e_I. \tag{A.1.9}$$

Proof. Formula (A.1.9) is easily checked by induction, using the fact that

$$e_r \wedge e_I \otimes e_I \wedge e_r = \varepsilon(\{r\}, I)\varepsilon(I, \{r\})e_{I\cup\{r\}} \otimes e_{I\cup\{r\}}$$

is 0 unless $r \in \{1, \dots, d\} \setminus I$, in which case we have $\varepsilon(\{r\}, I)\varepsilon(I, \{r\}) = (-1)^{|I|}$. Now, F_i is clearly $\mathfrak{S}_d$-invariant, so it only remains to check it is nonzero. Consider the element

$$E_i := e_{\{i+1,\dots,d-1\}} \wedge F_i \wedge e_{\{i+2,\dots,d\}} \in \Lambda^{d-1} \otimes \Lambda^{d-1} \simeq \overline{\mathbb{F}}_\ell.$$

In our formula for F_i, the only subset I that will contribute to E_i is $I = \{1, \dots, i\}$, so that we get

$$E_i = c_i \cdot e_{\{1,\dots,d\}\setminus\{d-1\}} \otimes e_{\{1,\dots,d\}\setminus\{i+1\}},$$

with $c_i = \pm i!/d^i$ nonzero. Hence $E_i \neq 0$ and therefore $F_i \neq 0$. □

The element F_i defines a nonzero morphism $\xi^i_{0,i} : \Lambda^0 = \overline{\mathbb{F}}_\ell \longrightarrow \Lambda^i \otimes \Lambda^i$, and therefore provides a generator $\xi^i_{0,i}$ for $\mathrm{Ext}^i(\Lambda^0, \Lambda^i)$. More generally, for $i \leqslant j$, the map $\Lambda^i \longrightarrow \Lambda^j \otimes \Lambda^{j-i}$, $v \mapsto v \wedge F_{j-i}$ is a $\mathfrak{S}_d$-equivariant map and the above proof shows that it takes nonzero value on $v = e_{\{d-i,\dots,d-1\}}$, for example.

A.1.10. Definition. *For $0 \leqslant i \leqslant j \leqslant d-1$, we denote by $\xi^{j-i}_{i,j}$ the generator of* $\mathrm{Ext}^{j-i}(\Lambda^i, \Lambda^j)$ *provided by the nonzero $\mathfrak{S}_d$-equivariant map $v \mapsto v \wedge F_{j-i}$, $\Lambda^i \xrightarrow{\text{can}} \Lambda^j \otimes \Lambda^{j-i}$ via the isomorphism of Proposition 2.3.2*(i).

A.1.11. Lemma. *For $0 < i < d-1$, the mapping $e_r \mapsto e_r \wedge F_{i-1} \wedge e_r - F_i$ extends uniquely to a $\mathfrak{S}_d$-equivariant $\overline{\mathbb{F}}_\ell$-linear nonzero map*

$$\xi^i_{1,i} : \Lambda^1 \longrightarrow \Lambda^i \otimes \Lambda^i.$$

Proof. Since the e_r's generate Λ^1, the extension is unique. For its existence we need to check compatibility with the linear relation $\sum_r e_r = 0$. But this follows from the recursive definition (A.1.8). Equivariance is clear, given the invariance of F_i and F_{i-1}. It remains to check nonvanishing. But we have

$$\begin{aligned} e_{\{i+1,\dots,d-1\}} \wedge \xi^i_{1,i}(e_d) \wedge e_{\{i+2,\dots,d\}} &= e_{\{i+1,\dots,d-1\}} \wedge (e_d \wedge F_{i-1} \wedge e_d - F_i) \wedge e_{\{i+2,\dots,d\}} \\ &= -\, e_{\{i+1,\dots,d-1\}} \wedge F_i \wedge e_{\{i+2,\dots,d\}}, \end{aligned}$$

which was shown to be nonzero in the previous proof. □

The homomorphism $\xi^i_{1,i}$ therefore provides a generator of $\mathrm{Ext}^i(\Lambda, \Lambda^i)$. We will now construct generators of some other nonzero Ext spaces.

A.1.12. *Explicit basis of* Ext *spaces.* In this section we fix integers $i, j, k \in \{0, \dots, d-1\}$ and we assume that $i \leqslant j$ and $i + j \leqslant d-1$. Under these assumptions, Corollary 2.3.3 tells us that $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$ has dimension 1 exactly when $j - i \leqslant k \leqslant j + i$.

Definition. *In this context, if k is of the form $k = j - i + 2l$, we define*

$$\begin{aligned} \xi^k_{i,j} : \quad \Lambda^i \xrightarrow{\text{can}} \Lambda^{i-l} \otimes \Lambda^l &\to \Lambda^j \otimes \Lambda^k \\ v \otimes w &\mapsto v \wedge F_{k-l} \wedge w. \end{aligned}$$

If k is of the form $k = j - i + 2l + 1$, we define

$$\begin{aligned} \xi^k_{i,j} : \quad \Lambda^i \xrightarrow{\text{can}} \Lambda^{i-l-1} \otimes \Lambda^1 \otimes \Lambda^l &\to \Lambda^j \otimes \Lambda^k \\ v \otimes u \otimes w &\mapsto v \wedge \xi^{k-l}_{1,k-l}(u) \wedge w. \end{aligned}$$

By construction, the map $\xi^k_{i,j}$ is $\mathfrak{S}_d$-equivariant, and we denote by the same symbol the corresponding element of $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$ *via Proposition 2.3.2*(i).

Remark. This definition is consistent with Definition A.1.10. Moreover, by Proposition 2.3.2(ii), we see that by construction we have a factorization $\xi_{i,j}^k = \xi_{i,i-l}^l \cup \xi_{i-l,j}^{k-l}$ with l as in the definition above.

The following proposition is the key to Theorem 2.3.5.

Proposition. *For all k such that $j-i \leqslant k \leqslant j+i$, the element $\xi_{i,j}^k$ is a generator of* $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$. *Moreover, if $k < j+i$, we have*

(i) $\xi_{i,i}^1 \cup \xi_{i,j}^k \in \overline{\mathbb{F}}_\ell^\times \cdot \xi_{i,j}^{k+1}$ *and*

(ii) $\xi_{i,j}^k \cup \xi_{j,j}^1 \in \overline{\mathbb{F}}_\ell^\times \cdot \xi_{i,j}^{k+1}$.

Proof. The proofs of (i) and (ii) are somewhat lengthy and complicated computations, and are postponed to Section A.2 in order to lighten a bit this section.

Let us deduce that $\xi_{i,j}^k$ is a generator of $\mathrm{Ext}^k(\Lambda^i, \Lambda^j)$. We know that this is equivalent to $\xi_{i,j}^k$ being nonzero. By either (i) and (ii), it is enough to prove that $\xi_{i,j}^{i+j}$ is nonzero. By definition, the homomorphism $\xi_{i,j}^{i+j} : \Lambda^i \longrightarrow \Lambda^j \otimes \Lambda^{i+j}$ is given by $v \mapsto F_j \wedge v$. But the proof of Lemma (A.1.9) shows that, for example, $F_j \wedge e_{\{d-i+1,\ldots,d\}}$ is nonzero (note that $j+2 \leqslant d-i+1$). $\square$

A.1.13. *Proof of Theorem 2.3.5.* The last proposition implies the claim in the case where $i \leqslant j$ and $i+j \leqslant d-1$. Using duality we get the case where $i \geqslant j$ and $i+j \leqslant d-1$. Here by duality we mean contragredient $A \mapsto A^* = \mathrm{Hom}_{\overline{\mathbb{F}}_\ell}(A, \overline{\mathbb{F}}_\ell)$ of $\overline{\mathbb{F}}_\ell$-representations of $\mathfrak{S}_d$. Indeed, for such representations we have functorial isomorphisms

$$\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, B) \xrightarrow{\sim} \mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(B^*, A^*)$$

that are compatible with cup-products in the obvious sense, and, on the other hand, we have $(\Lambda^i)^* \simeq \Lambda^i$ and $(\Lambda^j)^* \simeq \Lambda^j$. So the claim is now proved for $i+j \leqslant d-1$.

In order to get the case $i+j \geqslant d-1$ we use the endoequivalence of the category of $\overline{\mathbb{F}}_\ell$-representations of $\mathfrak{S}_d$ given by $A \mapsto A \otimes \Lambda^{d-1}$. Again we have functorial isomorphisms

$$\mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A, B) \xrightarrow{\sim} \mathrm{Ext}^*_{\overline{\mathbb{F}}_\ell[(N\cap G^0)/T^0]}(A \otimes \Lambda^{d-1}, B \otimes \Lambda^{d-1})$$

that are compatible with cup-products. On the other hand, we have

$$\Lambda^i \otimes \Lambda^{d-1} \simeq \Lambda^{d-1-i} \quad \text{and} \quad \Lambda^j \otimes \Lambda^{d-1} \simeq \Lambda^{d-1-j},$$

and $(d-1-i)+(d-1-j) \leqslant d-1$ whenever $i+j \geqslant d-1$.

A.2. ***Proof of the preceding proposition.*** We now prove items (i) and (ii) of the proposition appearing on this page. We will evaluate both sides of the claimed equalities on elements e_I for $I \subset \{1, \dots, d\}$ of size i. Inserting (A.1.2) in our definitions, we get

$$\begin{aligned}\xi^1_{i,i}(e_I) &= \sum_{r\in I} \varepsilon(I \setminus \{r\}, \{r\}) e_{I\setminus\{r\}} \wedge (e_r \otimes e_r - F_1)\\ &= \sum_{r\in I} e_I \otimes e_r - \sum_{r\in I} \varepsilon(I \setminus \{r\}, \{r\}) e_{I\setminus\{r\}} \wedge F_1.\end{aligned}$$

Similarly, when $k = j - i + 2l$ we get

$$\xi^k_{i,j}(e_I) = \sum_{I=I_{i-l}\sqcup I_l} \varepsilon(I_{i-l}, I_l) e_{I_{i-l}} \wedge F_{k-l} \wedge e_{I_l},$$

and when $k = j - i + 2l + 1$ we get

$$\xi^k_{i,j}(e_I) = \sum_{I=I_{i-l-1}\sqcup\{r\}\sqcup I_l} \varepsilon(I_{i-l-1}, \{r\}, I_l) e_{I_{i-l-1}} \wedge (e_r \wedge F_{k-l-1} \wedge e_r - F_{k-l}) \wedge e_{I_l}.$$

A.2.1. *Proof of* (i) *in the case* $k = j - i + 2l$. We assume that $k = j - i + 2l$. From the two expressions above and from the dictionary for cup-products in Proposition 2.3.2(ii), we get

$$(\xi^1_{i,i} \cup \xi^k_{i,j})(e_I) = A - B,$$

with

$$A = \sum_{r\in I} \xi^k_{i,j}(e_I) \wedge e_r = \sum_{r\in I} \sum_{I=I_{i-l}\sqcup I_l} \varepsilon(I_{i-l}, I_l) e_{I_{i-l}} \wedge F_{k-l} \wedge e_{I_l} \wedge e_r$$

and

$$B = \frac{1}{d} \sum_{r\in I} \varepsilon(I \setminus \{r\}, \{r\}) \sum_{s=1}^{d} \xi^k_{i,j}(e_{I\setminus\{r\}} \wedge e_s) \wedge e_s.$$

Now in the expression A, each summand is 0 unless $r \in I_{i-l}$, in which case we have

$$\begin{aligned}&\varepsilon(I_{i-l}, I_l) e_{I_{i-l}} \wedge F_{k-l} \wedge e_{I_l} \wedge e_r\\ &\qquad = (-1)^l \varepsilon(I_{i-l}, I_l)\varepsilon(I_{i-l} \setminus \{r\}, \{r\}) e_{I_{i-l}\setminus\{r\}} \wedge (e_r \wedge F_{k-l} \wedge e_r) \wedge e_{I_l}\\ &\qquad = (-1)^l \varepsilon(I_{i-l} \setminus \{r\}, \{r\}, I_l) e_{I_{i-l}\setminus\{r\}} \wedge (e_r \wedge F_{k-l} \wedge e_r) \wedge e_{I_l},\end{aligned}$$

so that eventually we may rewrite

$$A = (-1)^l \sum_{I=I_{i-l-1}\sqcup\{r\}\sqcup I_l} \varepsilon(I_{i-l-1}, \{r\}, I_l) e_{I_{i-l-1}} \wedge (e_r \wedge F_{k-l} \wedge e_r) \wedge e_{I_l}$$

and thus recognize one summand of the expression of $\xi^{k+1}_{i,j}(e_I)$. Moreover, in expression B each summand is 0 unless s is outside $I \setminus \{r\}$, in which case, putting

$I^{-r+s} := I \setminus \{r\} \sqcup \{s\}$, we may write

$$\begin{aligned}
&\xi_{i,j}^k(e_{I\setminus\{r\}} \wedge e_s) \wedge e_s \\
&\quad = \varepsilon(I \setminus \{r\}, \{s\})\xi_{i,j}^k(e_{I^{-r+s}}) \wedge e_s \\
&\quad = \varepsilon(I \setminus \{r\}, \{s\}) \sum_{I^{-r+s}=I_{i-l}^{-r+s}\sqcup I_l^{-r+s}} \varepsilon(I_{i-l}^{-r+s}, I_l^{-r+s})\, e_{I_{i-l}^{-r+s}} \wedge F_{k-l} \wedge e_{I_l^{-r+s}} \wedge e_s.
\end{aligned}$$

Now in the last expression, each summand vanishes unless $s \in I_{i-l}^{-r+s}$ and, as above, we have

$$\begin{aligned}
&\varepsilon(I_{i-l}^{-r+s}, I_l^{-r+s})\, e_{I_{i-l}^{-r+s}} \wedge F_{k-l} \wedge e_{I_l^{-r+s}} \wedge e_s \\
&\qquad\qquad = (-1)^l \varepsilon(I_{i-l}^{-r+s} \setminus \{s\}, \{s\}, I_l^{-r+s})\, e_{I_{i-l}^{-r+s}\setminus\{s\}} \wedge (e_s \wedge F_{k-l} \wedge e_s) \wedge e_{I_l^{-r+s}}.
\end{aligned}$$

Hence we get

$$\begin{aligned}
&\xi_{i,j}^k(e_{I\setminus\{r\}} \wedge e_s) \wedge e_s \\
&= (-1)^l \varepsilon(I \setminus \{r\}, \{s\}) \sum_{I\setminus\{r\}=I_{i-l-1}\sqcup I_l} \varepsilon(I_{i-l-1}, \{s\}, I_l)\, e_{I_{i-l-1}} \wedge (e_s \wedge F_{k-l} \wedge e_s) \wedge e_{I_l}.
\end{aligned}$$

Because of the equality

$$\varepsilon(I_{i-l-1}, \{s\}, I_l) = (-1)^l \varepsilon(I_{i-l-1}, I_l, \{s\}) = (-1)^l \varepsilon(I_{i-l-1}, I_l)\varepsilon(I \setminus \{r\}, \{s\}),$$

we have

$$\varepsilon(I \setminus \{r\}, \{s\})\varepsilon(I_{i-l-1}, \{s\}, I_l) = \varepsilon(I \setminus \{r\}, \{r\})\varepsilon(I_{i-l-1}, \{r\}, I_l),$$

so that eventually

$$\begin{aligned}
B &= \frac{(-1)^l}{d} \sum_{s=1}^{d} \sum_{I=I_{i-l-1}\sqcup\{r\}\sqcup I_l} \varepsilon(I_{i-l-1}, \{r\}, I_l)\, e_{I_{i-l-1}} \wedge (e_s \wedge F_{k-l} \wedge e_s) \wedge e_{I_l} \\
&= (-1)^l \sum_{I=I_{i-l-1}\sqcup\{r\}\sqcup I_l} \varepsilon(I_{i-l-1}, \{r\}, I_l)\, e_{I_{i-l-1}} \wedge F_{k-l+1} \wedge e_{I_l},
\end{aligned}$$

which gives the second summand in the expression of $\xi_{i,j}^{k+1}(e_I)$. Namely, we have shown that $\xi_{i,i}^1 \cup \xi_{i,j}^k = (-1)^l \xi_{i,j}^{k+1}$ in this case.

A.2.2. *Proof of* (i) *in the case* $k = j - i + 2l + 1$. We now assume that $k = j - i + 2l + 1$. Again we can write

$$(\xi_{i,i}^1 \cup \xi_{i,j}^k)(e_I) = A - B,$$

with

$$A = \sum_{r\in I} \xi_{i,j}^k(e_I) \wedge e_r$$

and

$$B = \frac{1}{d} \sum_{r \in I} \varepsilon(I \setminus \{r\}, \{r\}) \sum_{s=1}^{d} \xi_{i,j}^{k} (e_{I \setminus \{r\}} \wedge e_s) \wedge e_s.$$

Inserting the expression of $\xi_{i,j}^{k}$ in A we may decompose $A = A_1 + A_2$ with

$$A_1 = \sum_{r \in I} \sum_{I = I_{i-l-1} \sqcup \{t\} \sqcup I_l} \varepsilon(I_{i-l-1}, t, I_l) e_{I_{i-l-1}} \wedge e_t \wedge F_{k-l-1} \wedge e_t \wedge e_{I_l} \wedge e_r$$

and

$$A_2 = -\sum_{r \in I} \sum_{I = I_{i-l-1} \sqcup \{t\} \sqcup I_l} \varepsilon(I_{i-l-1}, t, I_l) e_{I_{i-l-1}} \wedge F_{k-l} \wedge e_{I_l} \wedge e_r.$$

Each summand of A_1 vanishes unless $r \in I_{i-l-1}$. This remark allows us to rewrite the sum in the following way:

$$A_1 = \sum_{I = I_{i-l} \sqcup I_l} \varepsilon(I_{i-l}, I_l) \sum_{r,t \in I_{i-l}} e_{I_{i-l}} \wedge F_{k-l-1} \wedge e_t \wedge e_{I_l} \wedge e_r = 0.$$

Indeed, the sum vanishes because the summands are antisymmetric in (r, t). As for A_2, we split it further as $A_{21} + A_{22}$, according to whether $r = t$ or $r \neq t$:

$$A_{21} = -\sum_{I = I_{i-l-1} \sqcup \{t\} \sqcup I_l} \varepsilon(I_{i-l-1}, t, I_l) e_{I_{i-l-1}} \wedge F_{k-l} \wedge e_{I_l} \wedge e_t,$$

$$A_{22} = -\sum_{I = I_{i-l-1} \sqcup \{t\} \sqcup I_l} \sum_{r \neq t} \varepsilon(I_{i-l-1}, t, I_l) e_{I_{i-l-1}} \wedge F_{k-l} \wedge e_{I_l} \wedge e_r.$$

We are then pleased to see that

$$\begin{aligned} A_{21} &= -(-1)^l \sum_{I = I_{i-l-1} \sqcup \{t\} \sqcup I_l} \varepsilon(I_{i-l-1}, I_l \cup \{t\}) e_{I_{i-l-1}} \wedge F_{k-l} \wedge e_{I_l \cup \{t\}} \\ &= (-1)^{l+1} (l+1) \sum_{I = I_{i-l-1} \sqcup I_{l+1}} \varepsilon(I_{i-l-1}, I_{l+1}) e_{I_{i-l-1}} \wedge F_{k-l} \wedge e_{I_{l+1}} \\ &= (-1)^{l+1} (l+1) \xi_{i,j}^{k+1} (e_I). \end{aligned}$$

The term A_{22} will be canceled by a term occurring in B, so we now turn to B, and begin to decompose it as $B_1 + B_2$, with

$$\begin{aligned} B_1 = \sum_{r \in I} \varepsilon(I \setminus r, r) \Bigg(&\frac{1}{d} \sum_{s=1}^{d} \varepsilon(I \setminus r, s) \\ &\times \sum_{I^s = I^s_{i-l-1} \sqcup \{t\} \sqcup I^s_l} \varepsilon(I^s_{i-l-1}, \{t\}, I^s_l) e_{I^s_{i-l-1}} \wedge (e_t \wedge F_{k-l-1} \wedge e_t) \wedge e_{I^s_l} \wedge e_s \Bigg) \end{aligned}$$

and

$$B_2 = -\sum_{r\in I}\varepsilon(I\setminus r, r)\Bigg(\frac{1}{d}\sum_{s=1}^{d}\varepsilon(I\setminus r, s) \times \sum_{I^s = I^s_{i-l-1}\sqcup\{t\}\sqcup I^s_l}\varepsilon(I^s_{i-l-1}, \{t\}, I^s_l)e_{I^s_{i-l-1}}\wedge F_{k-l}\wedge e_{I^s_l}\wedge e_s,\Bigg)$$

where we have written $I^s := I\setminus\{r\}\cup\{s\}$. We may split further, $B_2 = B_{21}+B_{22}$, according to whether $t = s$ or $t\neq s$. Then we have

$$\begin{aligned}
B_{21} &= -\sum_{r\in I}\varepsilon(I\setminus r, r)\Bigg(\frac{1}{d}\sum_{s=1}^{d}\varepsilon(I\setminus r, s) \times \sum_{I\setminus\{r\} = I_{i-l-1}\sqcup I_l}\varepsilon(I_{i-l-1}, \{s\}, I_l)e_{I_{i-l-1}}\wedge F_{k-l}\wedge e_{I_l}\wedge e_s\Bigg)\\
&= -\sum_{r\in I}\varepsilon(I\setminus r, r)\frac{1}{d}\sum_{s=1}^{d}\sum_{I\setminus\{r\} = I_{i-l-1}\sqcup I_l}(-1)^l\varepsilon(I_{i-l-1}, I_l)e_{I_{i-l-1}}\wedge F_{k-l}\wedge e_{I_l}\wedge e_s\\
&= -\sum_{r\in I}\varepsilon(I\setminus r, r)\sum_{I\setminus\{r\} = I_{i-l-1}\sqcup I_l}(-1)^l\varepsilon(I_{i-l-1}, I_l)e_{I_{i-l-1}}\wedge F_{k-l}\wedge e_{I_l}\wedge\Bigg(\frac{1}{d}\sum_{s=1}^{d}e_s\Bigg)\\
&= 0 \quad \Bigg(\text{since } \sum_{s=1}^{d}e_s = 0\Bigg).
\end{aligned}$$

As for B_{22}, we note that each summand vanishes unless $s\in I^s_{i-l-1}$. In this case, we may write $e_{I^s_{i-l-1}} = \varepsilon(I_{i-l-2}, s)e_{I_{i-l-2}}\wedge e_s$ for I_{i-l-2} a subset of $I\setminus\{r, t\}$, and we note that I^s_l is also a subset of $I\setminus\{r, s\}$, so we simply denote it by I_l. Therefore we may rearrange the sum in the following way:

$$B_{22} = -\sum_{r,t\in I, r\neq t}\ \sum_{I\setminus\{r,t\} = I_{i-l-2}\sqcup I_l}\frac{1}{d}\sum_{s=1}^{d}\mathrm{sign}(r, s, t, I_{i-l-2}, I_l)e_{I_{i-l-2}}\wedge e_s\wedge F_{k-l}\wedge e_{I_l}\wedge e_s,$$

where

$$\begin{aligned}
\mathrm{sign}(r, s, t, I_{i-l-2}, I_l) &= \varepsilon(I\setminus r, r)\varepsilon(I\setminus r, s)\varepsilon(I_{i-l-2}\cup\{s\}, \{t\}, I_l)\varepsilon(I_{i-l-2}, \{s\})\\
&= \varepsilon(I\setminus r, r)\varepsilon(I\setminus r, s)\varepsilon(I_{i-l-2}, \{s\}, \{t\}, I_l)\\
&= (-1)^{l+1}\varepsilon(I\setminus r, r)\varepsilon(I\setminus r, s)\varepsilon(I_{i-l-2}, \{t\}, I_l, \{s\})\\
&= (-1)^{l+1}\varepsilon(I\setminus r, r)\varepsilon(I_{i-l-2}, \{t\}, I_l)\\
&= (-1)^{l+1}\varepsilon(I_{i-l-2}, \{t\}, I_l, \{r\})\\
&= \varepsilon(I_{i-l-2}, \{r\}, \{t\}, I_l).
\end{aligned}$$

But since $\varepsilon(I_{i-l-2},\{r\},\{t\},I_l)=-\varepsilon(I_{i-l-2},\{t\},\{r\},I_l)$, we see that $B_{22}=0$. It remains to deal with B_1. As in the case of B_{22}, we see that each summand vanishes unless $s\in I^s_{i-l-1}$, so that we may rearrange the sum

$$\begin{aligned}
B_1 &= \sum_{r,t\in I, r\neq t}\ \sum_{I\setminus\{r,t\}=I_{i-l-2}\sqcup I_l} \frac{1}{d}\sum_{s=1}^{d}\varepsilon(I_{i-l-2},\{r\},\{t\},I_l) \\
&\qquad\qquad \times e_{I_{i-l-2}}\wedge e_s\wedge(e_t\wedge F_{k-l-1}\wedge e_t)\wedge e_{I_l}\wedge e_s \\
&= (-1)^l \sum_{r,t\in I, r\neq t}\ \sum_{I\setminus\{r,t\}=I_{i-l-2}\sqcup I_l} \varepsilon(I_{i-l-2},\{r\},\{t\},I_l)e_{I_{i-l-2}}\wedge(e_t\wedge F_{k-l}\wedge e_t)\wedge e_{I_l} \\
&= -\sum_{r,t\in I, r\neq t}\ \sum_{I\setminus\{r,t\}=I_{i-l-2}\sqcup I_l} \varepsilon(I_{i-l-2}\cup\{t\},\{r\},I_l)e_{I_{i-l-2}\cup\{t\}}\wedge F_{k-l}\wedge e_{I_l}\wedge e_t \\
&= -\sum_{I=I_{i-l-1}\sqcup\{r\}\sqcup I_l}\ \sum_{t\neq r}\varepsilon(I_{i-l-1},r,I_l)e_{I_{i-l-1}}\wedge F_{k-l}\wedge e_{I_l}\wedge e_t \\
&= A_{22}.
\end{aligned}$$

Finally we have proved that in the case $k=j-i+2l+1$,

$$(\xi^1_{i,i}\cup\xi^k_{i,j})(e_I)=A_1+A_{21}+A_{22}-B_1-B_{21}-B_{22}=A_{21}=(-1)^{l+1}(l+1)\xi^{k+1}_{i,j}(e_I).$$

A.2.3. *Proof of* (ii). It is certainly possible to do a direct computation as above, but it seems easier to use case (i) and prove (ii) by induction on k. Indeed, if we assume that $\xi^{k-1}_{i,j}\cup\xi^1_{j,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^k_{i,j}$, then we get, thanks to (i),

$$\xi^k_{i,j}\cup\xi^1_{j,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^1_{i,i}\cup\xi^{k-1}_{i,j}\cup\xi^1_{j,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^1_{i,i}\cup\xi^k_{i,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^{k+1}_{i,j}.$$

Therefore all we have to do is initialize the induction by proving that $\xi^{j-i}_{i,j}\cup\xi^1_{j,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^{j-i+1}_{i,j}$. But since we already know that $\xi^1_{j,j}\cup\xi^1_{j,j}\in\overline{\mathbb{F}}^\times_\ell\xi^2_{j,j}$, it will suffice to check that $\xi^{j-i}_{i,j}\cup\xi^2_{j,j}\in\overline{\mathbb{F}}^\times_\ell\cdot\xi^{j-i+2}_{i,j}$ (this involves less computation). Again we evaluate both sides on elements e_I. Let us write

$$\xi^{j-i}_{i,j}(e_I)=e_I\wedge F_{j-i}=C_{j-i}\sum_{|K|=j-i}e_I\wedge e_K\otimes e_K,$$

with $C_{j-i}=(-1)^{(j-i)(j-i-1)/2}(j-i)!/(d^{j-i})$. We then have

$$(\xi^{j-i}_{i,j}\cup\xi^2_{j,j})(e_I)=C_{j-i}\sum_{|K|=j-i}\varepsilon(I,K)\xi^2_{j,j}(e_{I\cup K})\wedge e_K.$$

Inserting the expression

$$\xi^2_{j,j}(e_{I\cup K})=\sum_{r\in I\cup K}\varepsilon(I\cup K\setminus r,r)e_{I\cup K\setminus r}\wedge F_1\wedge e_r,$$

which is valid for I and K disjoint, we get

$$(\xi_{i,j}^{j-i} \cup \xi_{j,j}^2)(e_I) = C_{j-i} \sum_{|K|=j-i} \sum_{r \in I \cup K} \varepsilon(I, K)\varepsilon(I \cup K \setminus r, r) e_{I \cup K \setminus r} \wedge F_1 \wedge e_r \wedge e_K.$$

We see that those summands where $r \in K$ vanish. Hence we may restrict the second sum to I:

$$\begin{aligned}(\xi_{i,j}^{j-i} \cup \xi_{j,j}^2)(e_I) &= C_{j-i} \sum_{|K|=j-i} \sum_{r \in I} \varepsilon(I, K)\varepsilon(I \setminus r \cup K, r) e_{I \setminus r \cup K} \wedge F_1 \wedge e_r \wedge e_K \\ &= C_{j-i} \sum_{|K|=j-i} \sum_{r \in I} \mu(I, K, r) e_{I \setminus r} \wedge e_K \wedge F_1 \wedge e_K \wedge e_r,\end{aligned}$$

with

$$\begin{aligned}\mu(I, K, r) &= \varepsilon(I, K)\varepsilon(I \setminus r \cup K, r)\varepsilon(I \setminus r, K)(-1)^{j-i} \\ &= \varepsilon(I, K)\varepsilon(I \setminus r, K, r)(-1)^{j-i} = \varepsilon(I, K)\varepsilon(I \setminus r, r, K) = \varepsilon(I \setminus r, r).\end{aligned}$$

Eventually we get

$$\begin{aligned}(\xi_{i,j}^{j-i} \cup \xi_{j,j}^2)(e_I) &= C_{j-i} \sum_{|K|=j-i} \sum_{r \in I} \varepsilon(I \setminus r, r) e_{I \setminus r} \wedge e_K \wedge F_1 \wedge e_K \wedge e_r \\ &= \sum_{r \in I} \varepsilon(I \setminus r, r) e_{I \setminus r} \wedge \left(C_{j-i} \sum_{|K|=j-i} e_K \wedge F_1 \wedge e_K \right) \wedge e_r \\ &= \sum_{r \in I} \varepsilon(I \setminus r, r) e_{I \setminus r} \wedge F_{j-i+1} \wedge e_r \\ &= \xi_{i,j}^{j-i+2}(e_I).\end{aligned}$$

References

[Boyer 2009] P. Boyer, "Monodromie du faisceau pervers des cycles évanescents de quelques variétés de Shimura simples", *Invent. Math.* **177**:2 (2009), 239–280. MR 2010i:14042 Zbl 1172.14016

[Clozel et al. 2008] 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 2010j:11082 Zbl 1169.11020

[Dat 2006] J. F. Dat, "Espaces symétriques de Drinfeld et correspondance de Langlands locale", *Ann. Sci. École Norm. Sup.* (4) **39**:1 (2006), 1–74. MR 2007j:22026 Zbl 1141.22004

[Dat 2007] J.-F. Dat, "Théorie de Lubin–Tate non-abélienne et représentations elliptiques", *Invent. Math.* **169**:1 (2007), 75–152. MR 2008g:22022 Zbl 1180.22020

[Dat 2012a] J.-F. Dat, "Lefschetz operator and local Langlands mod ℓ: The regular case", *Nagoya Math. J.* **208** (2012), 1–38. MR 3006695 Zbl 06127266

[Dat 2012b] J.-F. Dat, "Opérateur de Lefschetz sur les tours de Drinfeld et Lubin–Tate", *Compos. Math.* **148**:2 (2012), 507–530. MR 2904195 Zbl 1247.14025

[Dat 2012c] J.-F. Dat, "Théorie de Lubin–Tate non Abélienne ℓ-entière", *Duke Math. J.* **161**:6 (2012), 951–1010. MR 2913099 Zbl 1260.11070

[Dat 2012d] J.-F. Dat, "Un cas simple de correspondance de Jacquet–Langlands modulo ℓ", *Proc. Lond. Math. Soc.* (3) **104**:4 (2012), 690–727. MR 2908780 Zbl 1241.22020

[Deligne 1968] P. Deligne, "Théorème de Lefschetz et critères de dégénérescence de suites spectrales", *Inst. Hautes Études Sci. Publ. Math.* 35 (1968), 259–278. MR 39 #5582 Zbl 0159.22501

[Fargues 2008] L. Fargues, "L'isomorphisme entre les tours de Lubin–Tate et de Drinfeld et applications cohomologiques", pp. 1–325 in *L'isomorphisme entre les tours de Lubin–Tate et de Drinfeld*, Progr. Math. **262**, Birkhäuser, Basel, 2008. MR 2441312 Zbl 1136.14001

[Griffiths and Harris 1978] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley-Interscience, New York, 1978. Reprinted Wiley, New York, 1994. MR 80b:14001 Zbl 0408.14001

[James and Kerber 1981] G. James and A. Kerber, *The representation theory of the symmetric group*, Encyclopedia of Mathematics and its Applications **16**, Addison-Wesley, Reading, MA, 1981. MR 83k:20003 Zbl 0491.20010

[Remmel 1989] J. B. Remmel, "A formula for the Kronecker products of Schur functions of hook shapes", *J. Algebra* **120**:1 (1989), 100–118. MR 90h:20015 Zbl 0684.20012

[Schneider and Stuhler 1991] P. Schneider and U. Stuhler, "The cohomology of p-adic symmetric spaces", *Invent. Math.* **105**:1 (1991), 47–122. MR 92k:11057 Zbl 0751.14016

[Taylor and Yoshida 2007] R. Taylor and T. Yoshida, "Compatibility of local and global Langlands correspondences", *J. Amer. Math. Soc.* **20**:2 (2007), 467–493. MR 2007k:11193 Zbl 1210.11118

[Vignéras 2001] M.-F. Vignéras, "Correspondance de Langlands semi-simple pour $\mathrm{GL}(n, F)$ modulo $\ell \neq p$", *Invent. Math.* **144**:1 (2001), 177–223. MR 2003f:11182 Zbl 1031.11068

Communicated by Richard Taylor
Received 2013-06-18 Revised 2013-11-07 Accepted 2013-12-10

dat@math.jussieu.fr *Institut de Mathématiques de Jussieu, Université Pierre et Marie Curie (Paris 6), 4, place Jussieu, 75252 Paris, France*

ALGEBRA AND NUMBER THEORY 8:3 (2014)
dx.doi.org/10.2140/ant.2014.8.767

Splitting tower and degree of tt-rings

Paul Balmer

After constructing a splitting tower for separable commutative ring objects in tensor-triangulated categories, we define and study their degree.

Introduction

Let $\mathcal{K}$ be a tensor-triangulated category (tt-category, for short). Denote its tensor by $\otimes : \mathcal{K} \times \mathcal{K} \to \mathcal{K}$ and its $\otimes$-unit by $\mathbb{1}$. Let A be a ring object in $\mathcal{K}$, that is, an associative monoid $\mu : A \otimes A \to A$ with unit $\eta : \mathbb{1} \to A$. We want to study the *degree* of such a ring object under the assumption that A is what we call a *tt-ring*, that is, is commutative and separable. We focus on tt-rings because their Eilenberg–Moore category, A-$\mathrm{Mod}_{\mathcal{K}}$, of A-modules in $\mathcal{K}$ remains a tt-category and extension of scalars $F_A : \mathcal{K} \to A$-$\mathrm{Mod}_{\mathcal{K}}$ is a tt-functor (a fact which also explains the terminology: tt-rings preserve tt-categories). See Section 1.

In practice, tt-rings appear in commutative algebra as finite étale algebras and in representation theory of finite groups as the amusing algebras $A = \Bbbk(G/H)$ associated to subgroups $H < G$; see [Balmer 2012]. In the latter case, if $\mathcal{K} = \mathcal{K}(G)$ is the derived or the stable category of the group G over a field $\Bbbk$, then A-$\mathrm{Mod}_{\mathcal{K}}$ is nothing but the corresponding category $\mathcal{K}(H)$ for the subgroup. These two sources already provide an abundance of examples. Furthermore, the topological reader will find tt-rings among ring spectra, equivariant or not.

Let us contemplate the problem of defining a reasonable notion of degree, i.e., an integer $\deg(A)$ measuring the size of the tt-ring A in a general tt-category $\mathcal{K}$. When working over a field $\Bbbk$, it is tempting to use $\dim_{\Bbbk}(A)$. When A is a projective separable R-algebra over a commutative ring R, its rank must be finite [DeMeyer and Ingraham 1971] and provides a fine notion of degree for A viewed in the tt-category of perfect complexes $\mathrm{D}^{\mathrm{perf}}(R) = \mathrm{K}^{\mathrm{b}}(R\text{-proj})$. However, tt-geometry covers more than commutative algebra. Unorthodox separable algebras already emerge in representation theory, for instance, as the above $A = \Bbbk(G/H)$. In $\mathrm{D}^{\mathrm{b}}(\Bbbk G\text{-mod})$, one can still forget the G-action and take dimension over $\Bbbk$ as a possible degree —

Research supported by NSF grant DMS-1303073.
MSC2010: primary 18E30; secondary 13B40, 20J05, 55U35.
Keywords: separable, tensor triangulated category, degree.

which of course yields the index $[G:H]$ in that example — but one step further, in the stable category $\mathcal{K} = \mathrm{stab}(\Bbbk G)$, dimension over $\Bbbk$ becomes a slippery notion whereas the tt-ring $A = \Bbbk(G/H)$ remains equally important. Such questions become even harder in general stable homotopy categories (see [Hovey et al. 1997]) where there is simply no ground field $\Bbbk$ to deal with in the first place. So a good concept of degree in the broad generality of tt-geometry requires a new idea.

Our solution relies on the following splitting theorem, which echoes a classical property of usual separable rings (see [Chase et al. 1965]). Note that such a result is completely wrong for nonseparable rings, already with $B = A[X]$, for instance.

Theorem 2.1. *Let $f : A \to B$ and $g : B \to A$ be homomorphisms of tt-rings such that $g \circ f = \mathrm{id}_A$. Then there exists a tt-ring C and a ring isomorphism*

$$\begin{pmatrix} g \\ * \end{pmatrix} : B \xrightarrow{\sim} A \times C.$$

Using this theorem, we construct (Definition 3.1) a tower of tt-rings and homomorphisms

$$A =: A^{[1]} \to A^{[2]} \to \cdots \to A^{[n]} \to \cdots$$

such that after extending scalars to $A^{[n]}$ our A splits as the product of $\mathbb{1} \times \cdots \times \mathbb{1}$ (n times) with $A^{[n+1]}$. The *degree* of A is defined to be the last n such that $A^{[n]} \neq 0$ (Definition 3.4).

We prove a series of results which show that this concept of degree behaves according to intuition and provides a reasonable invariant. In basic examples, we recover expected values, like $[G:H]$ in the case of $\Bbbk(G/H)$ in $\mathrm{D^b}(\Bbbk G\text{-mod})$. In the stable category however, $\deg(\Bbbk(G/H))$ can be smaller than $[G:H]$. In the extreme case of $H < G$ strongly p-embedded, we even get $\deg(\Bbbk(G/H)) = 1$ in $\mathrm{stab}(\Bbbk G)$; see Example 4.6. We prove in Section 4 that the degree is finite for every tt-ring in the derived category of perfect complexes over a scheme, in the bounded derived category of a finite-dimensional cocommutative Hopf algebra and in the stable homotopy category of finite spectra.

It is an open question whether the degree must always be finite, at least locally.

Several aspects of this work extend to nontriangulated additive tensor categories. This is discussed in Remark 3.13.

In [Balmer 2013], our degree theory will be used to control the fibers of the map $\mathrm{Spc}(A\text{-Mod}_{\mathcal{K}}) \to \mathrm{Spc}(\mathcal{K})$. We shall notably reason by induction on the degree, thanks to this result:

Theorem. *Let A be a tt-ring of finite degree d in $\mathcal{K}$. Then in the tt-category of A-modules, we have an isomorphism of tt-rings $F_A(A) \simeq \mathbb{1} \times A^{[2]}$ where the tt-ring $A^{[2]}$ has degree $d-1$ in $A\text{-Mod}_{\mathcal{K}}$.*

Convention. All our tt-categories are essentially small and idempotent complete.

1. The tt-category of A-modules

We quickly list standard properties of the Eilenberg–Moore category $A\text{-Mod}_{\mathcal{K}}$ of A-modules in $\mathcal{K}$; see [Eilenberg and Moore 1965; Mac Lane 1998, Chapter VI; Balmer 2011].

As A is *separable* (that is, $\mu : A \otimes A \to A$ has a section σ as A, A-bimodules) the category $A\text{-Mod}_{\mathcal{K}}$ admits a unique triangulation such that both extension of scalars $F_A : \mathcal{K} \to A\text{-Mod}_{\mathcal{K}}$, $x \mapsto A \otimes x$, and its forgetful right adjoint $U_A : A\text{-Mod}_{\mathcal{K}} \to \mathcal{K}$ are exact; see [Balmer 2011]. Also, $A\text{-Mod}_{\mathcal{K}}$ is equivalent to the idempotent completion of the *Kleisli category* $A\text{-Free}_{\mathcal{K}}$ of free A-modules; see [Kleisli 1965]. Objects of $A\text{-Free}_{\mathcal{K}}$ are the same as those of $\mathcal{K}$, denoted $F_A(x)$ for every $x \in \mathcal{K}$, and morphisms $\mathrm{Hom}_A(F_A(x), F_A(y)) := \mathrm{Hom}_{\mathcal{K}}(x, A \otimes y)$. Denote by $\bar{f} : F_A(x) \to F_A(y)$ the morphism in $A\text{-Free}_{\mathcal{K}}$ corresponding to $f : x \to A \otimes y$ in $\mathcal{K}$.

As our tt-ring A is furthermore *commutative*, there is a tensor structure $- \otimes_A - :$ $A\text{-Mod}_{\mathcal{K}} \times A\text{-Mod}_{\mathcal{K}} \to A\text{-Mod}_{\mathcal{K}}$ making $F_A : \mathcal{K} \to A\text{-Mod}_{\mathcal{K}}$ a tt-functor. Indeed, one can define $\otimes_A$ on the Kleisli category by $F_A(x) \otimes_A F_A(y) := F_A(x \otimes y)$ and $\bar{f} \otimes_A \bar{g} = \overline{(\mu \otimes 1 \otimes 1)(23)(f \otimes g)}$ if $f : x \to A \otimes x'$ and $g : y \to A \otimes y'$, thus:

$$x \otimes y \xrightarrow{f \otimes g} A \otimes x' \otimes A \otimes y' \xrightarrow{(23)} A \otimes A \otimes x' \otimes y' \xrightarrow{\mu \otimes 1 \otimes 1} A \otimes x' \otimes y'.$$

Idempotent completion then yields $\otimes_A$ on $A\text{-Mod}_{\mathcal{K}}$. One can also describe $\otimes_A$ on all modules directly. First only assume that A is separable with a chosen A, A-bimodule section $\sigma : A \to A \otimes A$ of μ. Let (x_1, ϱ_1) and (x_2, ϱ_2) be right and left A-modules in $\mathcal{K}$, respectively. Then the endomorphism v in $\mathcal{K}$,

$$v : x_1 \otimes x_2 \xrightarrow{1 \otimes \eta \otimes 1} x_1 \otimes A \otimes x_2 \xrightarrow{1 \otimes \sigma \otimes 1} x_1 \otimes A \otimes A \otimes x_2 \xrightarrow{\varrho_1 \otimes \varrho_2} x_1 \otimes x_2, \tag{1.1}$$

is an idempotent: $v \circ v = v$. Hence one can define $x_1 \otimes_A x_2 := \mathrm{Im}(v)$ as the corresponding direct summand of $x_1 \otimes x_2$. We obtain a split coequalizer in $\mathcal{K}$:

$$x_1 \otimes A \otimes x_2 \underset{1 \otimes \varrho_2}{\overset{\varrho_1 \otimes 1}{\rightrightarrows}} x_1 \otimes x_2 \xrightarrow{v} \mathrm{Im}(v) = x_1 \otimes_A x_2,$$

as in the traditional definition of $\otimes_A$. When A is commutative, left and right A-modules coincide and one induces an A-action on $x_1 \otimes_A x_2$ from the usual formula on $x_1 \otimes x_2$. One verifies that this coincides with the tensor constructed above.

Proposition 1.2 (Projection formula). *Let A be a tt-ring in $\mathcal{K}$. For all $y \in \mathcal{K}$ and $x \in A\text{-Mod}_{\mathcal{K}}$, we have a natural isomorphism $U_A(x \otimes_A F_A(y)) \cong U_A(x) \otimes y$ in $\mathcal{K}$.*

Proof. By construction of $\otimes_A$, it suffices to prove the existence of such an isomorphism for $x \in A\text{-Free}_{\mathcal{K}}$, which is natural in x in that category (and in y too, but that is easy). So, let $x = F_A(z)$ for some $z \in \mathcal{K}$. Then $U_A(x \otimes F_A(y)) =$

$U_A(F_A(z) \otimes_A F_A(y)) = U_A(F_A(z \otimes y)) = A \otimes (z \otimes y) \cong (A \otimes z) \otimes y = U_A(x) \otimes y$. This looks trivial, but the point is that this isomorphism is natural with respect to morphisms $\bar{f} : x = F_A(z) \to F_A(z') = x'$ in A-$\mathrm{Free}_{\mathcal{K}}$ for $f : z \to A \otimes z'$ in $\mathcal{K}$ (not just natural in z). This is now an easy verification. □

Remark 1.3. For two ring objects A and B, the ring object $A \times B$ is $A \oplus B$ with componentwise structure. The ring object $A \otimes B$ has multiplication $(\mu_1 \otimes \mu_2)(23) : (A \otimes B)^{\otimes 2} \to A \otimes B$ and obvious unit. The *opposite* A^{op} is A with $\mu^{\mathrm{op}} = \mu(12)$. The *enveloping ring* A^{e} is $A \otimes A^{\mathrm{op}}$. Left A^{e}-modules are just A, A-bimodules.

If A and B are separable, then so are $A \times B$, $A \otimes B$, and A^{op}. Conversely, if $A \times B$ is separable then so are A and B (restrict the section "σ" to each factor).

Remark 1.4. Let $h : A \to B$ be a homomorphism of tt-rings in $\mathcal{K}$ (that is, h is compatible with multiplications and units). We also say that B is an *A-algebra* or a tt-ring *over A*. Then idempotent-complete the functor $F_h : A$-$\mathrm{Free}_{\mathcal{K}} \to B$-$\mathrm{Free}_{\mathcal{K}}$ defined on objects by $F_h(F_A(x)) = F_B(x)$ and on morphisms by $F_h(\bar{f}) = \overline{(h \otimes 1) \circ f}$. Alternatively, equip B with a right A-module structure via h and define for every A-module $x \in A$-$\mathrm{Mod}_{\mathcal{K}}$, its extension $F_h(x) = B \otimes_A x$ equipped with the left B-module structure on the B factor. Both define the same tt-functor $F_h : A$-$\mathrm{Mod}_{\mathcal{K}} \to B$-$\mathrm{Mod}_{\mathcal{K}}$ and the following diagram commutes up to isomorphism:

$$\begin{tikzcd} & \mathcal{K} \arrow[dl, "F_A"'] \arrow[dr, "F_B"] & \\ A\text{-}\mathrm{Mod}_{\mathcal{K}} \arrow[rr, "F_h \cong B \otimes_A -"] & & B\text{-}\mathrm{Mod}_{\mathcal{K}}. \end{tikzcd}$$

Furthermore, if $k : B \to C$ is another homomorphism then $F_{kh} \cong F_k \circ F_h$.

Remark 1.5. For A a tt-ring in $\mathcal{K}$, there is a one-to-one correspondence between

(i) A-algebras in $\mathcal{K}$, that is, homomorphism $h : A \to B$ of tt-rings in $\mathcal{K}$, and

(ii) tt-rings $\bar{B}$ in A-$\mathrm{Mod}_{\mathcal{K}}$.

The correspondence is the obvious one: To every tt-ring $\bar{B} = (\bar{B}, \bar{\mu}, \bar{\eta})$ in A-$\mathrm{Mod}_{\mathcal{K}}$, associate $B := U_A(\bar{B})$ and $h := U_A(\bar{\eta})$. The ring structure on B is given by $B \otimes B = U_A(\bar{B}) \otimes U_A(\bar{B}) \overset{v}{\twoheadrightarrow} U_A(\bar{B} \otimes_A \bar{B}) \overset{\bar{\mu}}{\to} U_A(\bar{B}) = B$ and $\eta_B : \mathbb{1} \overset{\eta_A}{\to} A \overset{h}{\to} B$. Conversely, if $h : A \to B$ is a homomorphism, then one can use h to equip $\bar{B} := B$ with an A-module structure and verify that $\mu : B \otimes B \to B$ respects the idempotent v of (1.1), and hence defines $\bar{\mu} : B \otimes_A B \to B$. Then B is separable in $\mathcal{K}$ (with section σ of μ) if and only if $\bar{B}$ is separable in A-$\mathrm{Mod}_{\mathcal{K}}$ (with section $v\sigma$ of $\bar{\mu}$).

We tacitly use this dictionary below. If we need to distinguish the A-algebra B in $\mathcal{K}$ from the tt-ring $\bar{B}$ in A-$\mathrm{Mod}_{\mathcal{K}}$, we shall write $U_A(\bar{B})$ for the former.

Under this correspondence, if $\mathscr{L} := A\text{-Mod}_{\mathscr{K}}$, there is an equivalence $B\text{-Mod}_{\mathscr{K}} \cong \bar{B}\text{-Mod}_{\mathscr{L}}$ such that the following diagram commutes up to isomorphism:

$$\begin{tikzcd} \mathscr{K} \arrow[r, "F_A"] \arrow[d, "F_B"'] & \mathscr{L} = A\text{-Mod}_{\mathscr{K}} \arrow[d, "F_{\bar{B}}"] \\ B\text{-Mod}_{\mathscr{K}} \arrow[r, dash, "\simeq"] & \bar{B}\text{-Mod}_{\mathscr{L}}. \end{tikzcd}$$

On Kleisli categories, it maps $F_B(x)$ to $F_{\bar{B}}(F_A(x))$ for every $x \in \mathscr{K}$ and follows the sequence of isomorphisms $\mathrm{Hom}_B(F_B(x), F_B(y)) \cong \mathrm{Hom}_{\mathscr{K}}(x, B \otimes y) \cong \mathrm{Hom}_A(F_A(x), B \otimes y) \cong \mathrm{Hom}_A(F_A(x), \bar{B} \otimes_A F_A(y)) \cong \mathrm{Hom}_{\bar{B}}(F_{\bar{B}} F_A(x), F_{\bar{B}} F_A(y))$ on morphisms. Idempotent completion does the rest.

Remark 1.6. Let $F : \mathscr{K} \to \mathscr{L}$ be a tt-functor. Let A be a tt-ring in $\mathscr{K}$ and let $B := F(A)$ its image in $\mathscr{L}$. Then B is also a tt-ring and there exists a tt-functor $\bar{F} : A\text{-Mod}_{\mathscr{K}} \to B\text{-Mod}_{\mathscr{L}}$:

$$\begin{tikzcd} \mathscr{K} \arrow[r, "F"] \arrow[d, shift right, "F_A"'] & \mathscr{L} \arrow[d, shift right, "F_B"'] \\ A\text{-Mod}_{\mathscr{K}} \arrow[r, "\bar{F}"] \arrow[u, shift right, "U_A"'] & B\text{-Mod}_{\mathscr{L}}, \arrow[u, shift right, "U_B"'] \end{tikzcd} \tag{1.7}$$

such that $\bar{F} F_A = F_B F$ and $U_B \bar{F} = F U_A$. Explicitly, for every A-module (x, ϱ), we have $\bar{F}(x, \varrho) = (F(x), F(\varrho))$, where

$$B \otimes F(x) = F(A) \otimes F(x) \cong F(A \otimes x) \xrightarrow{F(\varrho)} F(x).$$

On morphisms, $\bar{F}(f) = F(f)$. The "Kleislian" description of $\bar{F}$ is equally easy.

2. Splitting theorems

We will iteratively use the following splitting result:

Theorem 2.1. *Let $f : A \to B$ and $g : B \to A$ be homomorphisms of tt-rings in $\mathscr{K}$ such that $g \circ f = \mathrm{id}_A$. Then there exists a tt-ring C and a ring isomorphism $h : B \xrightarrow{\sim} A \times C$ such that $\mathrm{pr}_1\, h = g$. Consequently, C becomes an A-algebra, via $\mathrm{pr}_2\, hf$. Moreover, if C' is another A-algebra and $h' : B \xrightarrow{\sim} A \times C'$ is another A-algebra isomorphism such that $\mathrm{pr}_1\, h = g$, then there exists an isomorphism of A-algebras $\ell : C \xrightarrow{\sim} C'$ such that $h' = (1 \times \ell)h$.*

We start with a couple of additive lemmas.

Lemma 2.2. *Let B be a ring object, B_1 and B_2 two B, B-bimodules and*

$$h : B \xrightarrow{\sim} B_1 \oplus B_2$$

an isomorphism of B, B-bimodules. Then B_1 and B_2 admit unique structures of ring objects such that h becomes a ring isomorphism $B \xrightarrow{\sim} B_1 \times B_2$.

Proof. Write the given isomorphisms

$$h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} : B \xrightarrow{\sim} B_1 \oplus B_2 \quad \text{and} \quad h^{-1} = \begin{pmatrix} k_1 & k_2 \end{pmatrix} : B_1 \oplus B_2 \xrightarrow{\sim} B.$$

If h is to be a ring isomorphism, we must have for $i = 1, 2$ that the multiplication $\mu_i : B_i \otimes B_i \to B_i$ is given by $\mu_i = h_i \mu(k_i \otimes k_i)$ and the unit $\eta_i : \mathbb{1} \to B_i$ by $\eta_i = h_i \eta$. Hence we have uniqueness. Conversely, let us see that these formulas provide the wanted ring structures. Let $\rho : B \otimes B_2 \to B_2$ be the left B-action on B_2. By left B-linearity of $k_2 : B_2 \to B$, we have $\mu(1 \otimes k_2) = k_2 \rho : B \otimes B_2 \to B$. Note that $hh^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ implies $h_i k_j = 0$ when $i \neq j$. Therefore $h_1 \mu(1 \otimes k_2) = h_1 k_2 \rho = 0$. Similarly, B, B-linearity of k_1 and k_2 gives $h_i \mu(1 \otimes k_j) = 0$ and $h_i \mu(k_j \otimes 1) = 0$ when $i \neq j$. So the bottom square of the following diagram commutes, in which the top one commutes by definition:

$$\begin{tikzcd}
(B_1 \oplus B_2) \otimes (B_1 \oplus B_2) \arrow[rr, "\mu_{B_1 \times B_2}"] \arrow[d, "\cong"] \arrow[dd, bend right=60, "h^{-1} \otimes h^{-1}"', "\simeq"] & & B_1 \oplus B_2 \arrow[d, equal] \\
(B_1 \otimes B_1) \oplus (B_1 \otimes B_2) \oplus (B_2 \otimes B_1) \oplus (B_2 \otimes B_2) \arrow[rr, "\left(\begin{smallmatrix} \mu_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_2 \end{smallmatrix}\right)"] \arrow[d, "(k_1 \otimes k_1 \ \ k_1 \otimes k_2 \ \ k_2 \otimes k_1 \ \ k_2 \otimes k_2)"] & & B_1 \oplus B_2 \\
B \otimes B \arrow[rr, "\mu"] & & B. \arrow[u, "h = \left(\begin{smallmatrix} h_1 \\ h_2 \end{smallmatrix}\right)", "\simeq"']
\end{tikzcd}$$

Hence $h : B \xrightarrow{\sim} B_1 \oplus B_2$ is an isomorphism of objects-equipped-with-multiplications. Since B is associative and unital, B_1 and B_2 must have the same properties. □

Lemma 2.3. *Let C and C' be ring objects and*

$$\begin{pmatrix} 1 & 0 \\ s & \ell \end{pmatrix} : \mathbb{1} \times C \xrightarrow{\sim} \mathbb{1} \times C'$$

a ring isomorphism. Then $s = 0$ and ℓ is a ring isomorphism.

Proof. Let us denote by (C, μ, η) and (C', μ', η') the structures. Clearly ℓ is an isomorphism of objects. From the fact that $\left(\begin{smallmatrix} 1 & 0 \\ s & \ell \end{smallmatrix}\right)$ preserves the structures it follows that $\eta' = s + \ell \eta$ and that

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu' \end{pmatrix} \left(\begin{pmatrix} 1 & 0 \\ s & \ell \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ s & \ell \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 \\ s & \ell \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix},$$

giving in particular $\mu'(s \otimes \ell) = 0$ and $\mu'(\ell \otimes \ell) = \ell \mu$. Composing the former with $(1 \otimes \ell^{-1} \eta') : \mathbb{1} \otimes \mathbb{1} \to \mathbb{1} \otimes C$, we get

$$0 = \mu'(s \otimes \ell)(1 \otimes \ell^{-1} \eta') = \mu'(s \otimes \eta') = \mu'(1_{C'} \otimes \eta') s = s$$

and therefore $\ell \eta = \eta'$. Hence ℓ preserves multiplication and unit. □

Proof of Theorem 2.1. Via the morphism $g : B \to A$, we can equip A with a structure of B, B-bimodule, so that g becomes B^{e}-linear. Since $B^{\mathrm{e}} = B \otimes B^{\mathrm{op}}$ is separable, the category $B^{\mathrm{e}}\text{-Mod}_{\mathcal{K}}$ of B, B-bimodules is triangulated in such a way that $U_{B^{\mathrm{e}}} : B^{\mathrm{e}}\text{-Mod}_{\mathcal{K}} \to \mathcal{K}$ is exact. Choose a distinguished triangle over g in $B^{\mathrm{e}}\text{-Mod}_{\mathcal{K}}$ say $C \to B \xrightarrow{g} A \xrightarrow{z} \Sigma C$. Forgetting the B^{e}-action, since g is split by f in $\mathcal{K}$, we see that $U_{B^{\mathrm{e}}}(z) = 0$. Since $U_{B^{\mathrm{e}}}$ is faithful, z is also zero in the triangulated category $B^{\mathrm{e}}\text{-Mod}_{\mathcal{K}}$, which in turn yields an isomorphism $h : B \xrightarrow{\sim} A \oplus C$ of B, B-bimodules such that $\mathrm{pr}_1\, h = g$. By Lemma 2.2, A and C can be equipped with ring structures so that h is a ring isomorphism. We are left to verify that this new ring structure on A is indeed the original one. This follows from the fact that $g : B \to A$ is a split epimorphism which is a homomorphism from B to A with both structures (the original one by hypothesis and the new one because $h = \left(\begin{smallmatrix} g \\ * \end{smallmatrix}\right)$ is a homomorphism). Note that C is separable by Remark 1.3. Finally, for uniqueness of C as A-algebra, with the notation of the statement, we obtain an isomorphism $k := h' \circ h^{-1} : \mathbb{1} \times C \xrightarrow{\sim} \mathbb{1} \times C'$ in $A\text{-Mod}_{\mathcal{K}}$ such that $\mathrm{pr}_1\, k = \mathrm{pr}_1$ which means that k has the form $\left(\begin{smallmatrix} 1 & 0 \\ s & \ell \end{smallmatrix}\right)$, and we conclude by Lemma 2.3. □

Theorem 2.4. *Let A be a tt-ring in $\mathcal{K}$. Then there exists a ring isomorphism $h : A \otimes A \xrightarrow{\sim} A \times A'$ for some tt-ring A' in such a way that* $\mathrm{pr}_1\, h = \mu$. *Moreover, the A-algebra A' is unique up to isomorphism with this property, where $A \otimes A$ is considered as an A-algebra on the left* (*via the homomorphism* $1 \otimes \eta : A \to A \otimes A$).

Proof. Apply Theorem 2.1 to the tt-ring $B = A \otimes A$ with $g = \mu : A \otimes A \to A$ and $f = 1_A \otimes \eta = F_A(\eta) : A \to A \otimes A$. □

Remark 2.5. From the isomorphism $\left(\begin{smallmatrix} \mu \\ * \end{smallmatrix}\right) : A \otimes A \xrightarrow{\sim} A \oplus A'$, we observe that $A' \simeq \Sigma^{-1}\,\mathrm{cone}(\mu) \simeq \mathrm{cone}(1_A \otimes \eta) \simeq A \otimes \mathrm{cone}(\eta)$ in $\mathcal{K}$. Furthermore,

$$\mathrm{supp}(A') \subseteq \mathrm{supp}(A).$$

3. Splitting tower and degree

Definition 3.1. We define *the splitting tower* of a tt-ring A,

$$A^{[0]} \to A^{[1]} \to A^{[2]} \to \cdots \to A^{[n]} \to A^{[n+1]} \to \cdots,$$

as follows: We start with $A^{[0]} = \mathbb{1}$, $A^{[1]} = A$, and $\eta : A^{[0]} \to A^{[1]}$. Then for $n \geq 1$ we define $A^{[n+1]} = (A^{[n]})'$ in the notation of Theorem 2.4 applied to the tt-ring $A^{[n]}$ in the tt-category $A^{[n-1]}\text{-Mod}_{\mathcal{K}}$ (see Remark 1.5). Equivalently, $A^{[n+1]}$ is characterized as an $A^{[n]}$-algebra by the existence of an isomorphism of $A^{[n]}$-algebras

$$h : A^{[n]} \otimes_{A^{[n-1]}} A^{[n]} \xrightarrow{\sim} A^{[n]} \times A^{[n+1]}, \tag{3.2}$$

such that $\mathrm{pr}_1\, h = \mu$, where $A^{[n]} \otimes_{A^{[n-1]}} A^{[n]}$ is an $A^{[n]}$-algebra via the left factor. This tower $\{A^{[n]}\}_{n \geq 0}$ is well-defined up to isomorphism.

Remark 3.3. By Remark 2.5, $\mathrm{supp}(A^{[n+1]}) \subseteq \mathrm{supp}(A^{[n]})$, and if $A^{[n]} = 0$ for some n then $A^{[m]} = 0$ for all $m \geq n$. Also, by construction, if we consider $A^{[n]}$ as a tt-ring in $A^{[n-1]}$-$\mathrm{Mod}_{\mathcal{K}}$, we have $(A^{[n]})^{[m]} \cong A^{[n+m-1]}$ for all $m \geq 1$.

Definition 3.4. We say that A *has finite degree* d if $A^{[d]} \neq 0$ and $A^{[d+1]} = 0$. In that case, we write $\deg(A) = d$ or $\deg_{\mathcal{K}}(A) = d$ if we need to stress the category. If $A^{[n]} \neq 0$ for all $n \geq 0$, we say that A has *infinite degree*.

Example 3.5. For $A = \mathbb{1} \times \mathbb{1}$, we have $A \otimes A \simeq A \times A$. Hence $A^{[2]} = A = A^{[1]}$. If one was to compute $A^{[2]} \otimes A^{[2]}$ one would get $A^{[2]} \times A$ again and misreading Definition 3.4 could lead to the false impression that $A^{[3]}$ is A again and that all $A^{[n]}$ are equal. This is not the way to compute $A^{[3]}$! One needs to compute $A^{[2]} \otimes_{A^{[1]}} A^{[2]} = A \otimes_A A = A = A^{[2]} \times 0$ and therefore $A^{[3]} = 0$. So, the tt-ring $\mathbb{1} \times \mathbb{1}$ has degree 2. In (3.2), it is important to perform the tensor ***over*** $\mathbf{A^{[n-1]}}$.

An immediate gain of having a numerical invariant like the degree is the possibility of making proofs by induction. This is applied in [Balmer 2013] using the splitting theorem (Theorem 2.4), in the following form:

Theorem 3.6. *Let A be a tt-ring of finite degree d in a tt-category $\mathcal{K}$. Then we have a ring isomorphism $F_A(A) \simeq \mathbb{1}_A \times A^{[2]}$ and $\deg(A^{[2]}) = d - 1$ in A-$\mathrm{Mod}_{\mathcal{K}}$.*

Proof. Since $A^{[2]} = A'$, this is simply Theorem 2.4 with A-algebras replaced by tt-rings in A-$\mathrm{Mod}_{\mathcal{K}}$ (see Remark 1.5 if necessary), together with the observation that $(A^{[2]})^{[n]} = A^{[n+1]}$ for all $n \geq 1$, which gives $\deg_{A\text{-}\mathrm{Mod}_{\mathcal{K}}}(A^{[2]}) = \deg_{\mathcal{K}}(A) - 1$. □

Before showing in Section 4 that many tt-rings have finite degree, let us build our understanding of this $\deg(A) \in \mathbb{N} \cup \{\infty\}$, starting with functorial properties.

Theorem 3.7. *Let A be a tt-ring in $\mathcal{K}$.*

(a) *Let $F : \mathcal{K} \to \mathcal{L}$ be a tt-functor. Then for every $n \geq 0$, we have $F(A)^{[n]} \simeq F(A^{[n]})$ as tt-rings. In particular*, $\deg(F(A)) \leq \deg(A)$.

(b) *Let $F : \mathcal{K} \to \mathcal{L}$ be a tt-functor. Suppose that F is "weakly conservative on* $\mathrm{supp}(A)$*", that is, for $x \in \mathcal{K}_{\mathrm{supp}(A)}$ if $F(x) = 0$ then $x^{\otimes m} = 0$ for $m \geq 0$; for instance, if F is just conservative. Then* $\deg(F(A)) = \deg(A)$.

(c) *Suppose that $B \in \mathcal{K}$ is a tt-ring such that* $\mathrm{supp}(B) \supseteq \mathrm{supp}(A)$; *for instance, if* $\mathrm{supp}(B) = \mathrm{Spc}(\mathcal{K})$. *Then $\deg(A)$ is equal to the degree of $F_B(A)$ in B-$\mathrm{Mod}_{\mathcal{K}}$.*

(d) *Suppose $\mathcal{K}$ is local and that $B \in \mathcal{K}$ is a nonzero tt-ring. Then the degree of A in $\mathcal{K}$ is equal to the degree of $F_B(A)$ in B-$\mathrm{Mod}_{\mathcal{K}}$.*

Proof. To prove (a) by induction on n, simply apply F to (3.2), which characterizes the splitting tower $A^{[*]}$. So, if $\deg(A) < \infty$ then $F(A)^{[\deg(A)+1]} \simeq F(A^{[\deg(A)+1]}) = 0$ and $\deg(F(A)) \leq \deg(A)$. For (b), recall that $A^{[n]} \in \mathcal{K}_{\mathrm{supp}(A)}$ for every $n \geq 1$ (Remark 3.3). As $\deg(F(A)) =: d < \infty$ implies $F(A^{[d+1]}) \simeq F(A)^{[d+1]} = 0$ we get by weak-conservativity of F that $A^{[d+1]}$ is $\otimes$-nilpotent, and hence zero (every ring object is a direct summand of its $\otimes$-powers via the unit). This $A^{[d+1]}=0$ means $\deg(A)\le d=\deg(F(A))$, which finishes (b). Then (c) follows since $\operatorname{supp}(B)\supseteq\operatorname{supp}(A)$ implies that $F_B:\mathcal{K}\to B\text{-Mod}_{\mathcal{K}}$ is weakly conservative on $\mathcal{K}_{\operatorname{supp}(A)}$. Indeed, if $\operatorname{supp}(x)\subseteq\operatorname{supp}(A)$ and $F_B(x)=0$ then $B\otimes x=U_BF_B(x)=0$ and $\varnothing=\operatorname{supp}(B\otimes x)=\operatorname{supp}(B)\cap\operatorname{supp}(x)=\operatorname{supp}(x)$, which implies that x is $\otimes$-nilpotent. For (d), recall that a tt-category is *local* if $x\otimes y=0$ implies that x or y is $\otimes$-nilpotent. Hence for the nonzero tt-ring B, the functor $F_B:\mathcal{K}\to B\text{-Mod}_{\mathcal{K}}$ is weakly conservative on the whole of $\mathcal{K}$ and we can apply (b). □

Let us now describe the local nature of the degree. Recall that for every prime $\mathcal{P}\in\operatorname{Spc}(\mathcal{K})$, the local category $\mathcal{K}_{\mathcal{P}}=(\mathcal{K}/\mathcal{P})^\natural$ at $\mathcal{P}$ is the idempotent completion of the Verdier quotient $\mathcal{K}/\mathcal{P}$, hence comes with a tt-functor $q_{\mathcal{P}}:\mathcal{K}\twoheadrightarrow\mathcal{K}/\mathcal{P}\hookrightarrow\mathcal{K}_{\mathcal{P}}$.

Theorem 3.8. *Let A be a tt-ring in $\mathcal{K}$. Suppose that $q_{\mathcal{P}}(A)$ has finite degree in the local tt-category $\mathcal{K}_{\mathcal{P}}$ for every point $\mathcal{P}\in\operatorname{Spc}(\mathcal{K})$. Then A has finite degree and*

$$\deg(A)=\max_{\mathcal{P}\in\operatorname{Spc}(\mathcal{K})}\deg(q_{\mathcal{P}}(A))=\max_{\mathcal{P}\in\operatorname{supp}(A)}\deg(q_{\mathcal{P}}(A)).$$

Proof. There exists, for every $\mathcal{P}\in\operatorname{Spc}(\mathcal{K})$, an integer $n_{\mathcal{P}}\ge1$ such that $q_{\mathcal{P}}(A^{[n_{\mathcal{P}}]})=(q_{\mathcal{P}}(A))^{[n_{\mathcal{P}}]}=0$. Hence $\mathcal{P}$ belongs to the open $\mathcal{U}(A^{[n_{\mathcal{P}}]}):=\operatorname{Spc}(\mathcal{K})-\operatorname{supp}(A^{[n_{\mathcal{P}}]})$. Putting all those open subsets together, we cover $\operatorname{Spc}(\mathcal{K})$. But the spectrum is always quasicompact and $\mathcal{U}(A^{[n]})\subseteq\mathcal{U}(A^{[n+1]})$, hence there exists $n\ge0$ such that $\mathcal{U}(A^{[n]})=\operatorname{Spc}(\mathcal{K})$. This means $A^{[n]}=0$, that is, $d:=\deg(A)<\infty$. By Theorem 3.7(a) we have $d=\deg(A)\ge\max_{\mathcal{P}\in\operatorname{Spc}(\mathcal{K})}\deg(q_{\mathcal{P}}(A))\ge\max_{\mathcal{P}\in\operatorname{supp}(A)}\deg(q_{\mathcal{P}}(A))$. Since $A^{[d]}\ne0$ there exists $\mathcal{P}\in\operatorname{supp}(A^{[d]})\subseteq\operatorname{supp}(A)$ with $0\ne q_{\mathcal{P}}(A^{[d]})\simeq(q_{\mathcal{P}}(A))^{[d]}$ and hence $\deg(q_{\mathcal{P}}(A))\ge d=\deg(A)$, wrapping up all the above inequalities into equalities. □

We now discuss the link between the degree and the trivial tt-ring $\mathbb{1}$.

Theorem 3.9. *Let A be a tt-ring in $\mathcal{K}$. Suppose $\mathcal{K}\ne0$.*

(a) *For every $n\ge1$, we have* $\deg(\mathbb{1}^{\times n})=n$.

(b) *For every $n\ge1$ we have $F_{A^{[n]}}(A)\simeq\mathbb{1}^{\times n}\times A^{[n+1]}$ as tt-rings in $A^{[n]}$-*$\text{Mod}_{\mathcal{K}}$.

(c) *If* $\deg(A)<\infty$ *then $B:=A^{[\deg(A)]}$ is nonzero and we have in B-*$\text{Mod}_{\mathcal{K}}$

$$F_B(A)\simeq\mathbb{1}^{\times\deg(A)}. \tag{3.10}$$

(d) *If a tt-functor $F:\mathcal{K}\to\mathcal{L}$ is weakly conservative on* $\mathcal{K}_{\operatorname{supp}(A)}$ (see Theorem 3.7(b) for this notion — for example, if F is conservative), *and if $F(A)\simeq\mathbb{1}^{\times d}$ in $\mathcal{L}$, then* $\deg(A)=d$.

(e) *Let B be a tt-ring such that $F_B(A)\simeq\mathbb{1}^{\times d}$ as tt-rings in B-*$\text{Mod}_{\mathcal{K}}$*. Suppose either that* $\operatorname{supp}(B)\supseteq\operatorname{supp}(A)$*, or that $\mathcal{K}$ is local and $B\ne0$. Then $d=$* $\deg(A)$.

We need another additive lemma, whose naive proof (with a permutation) fails.

Lemma 3.11. *Let $A = \mathbb{1}^{\times n}$. Then there exists an isomorphism $h : F_A(A) = A \otimes A \xrightarrow{\sim} A \times A^{\times(n-1)}$ of A-algebras such that* $\mathrm{pr}_1\, h = \mu$.

Proof. To keep track of the various copies of $\mathbb{1}$, write $A = \bigoplus_{i=1}^{n} \mathbb{1}_i$ and $A^{\times(n-1)} = \bigoplus_{\ell=1}^{n-1} \bigoplus_{i=1}^{n} \mathbb{1}_{i\ell}$ where $\mathbb{1}_i = \mathbb{1}_{i\ell} = \mathbb{1}$ for all i and ℓ. Then $A \otimes A = \bigoplus_{i,j} \mathbb{1}_i \otimes \mathbb{1}_j$. Define h by mapping the summand $\mathbb{1}_i \otimes \mathbb{1}_i = \mathbb{1}$ identically to $\mathbb{1}_i \hookrightarrow A$ and $\mathbb{1}_i \otimes \mathbb{1}_j = \mathbb{1}$ identically to $\mathbb{1}_{ij} \hookrightarrow A^{\times(n-1)}$ when $i \neq j$ and $j \leq n-1$, but mapping $\mathbb{1}_i \otimes \mathbb{1}_n = \mathbb{1}$ *diagonally* to $\bigoplus_{\ell=1,\,\ell\neq i}^{n-1} \mathbb{1}_{i\ell} \hookrightarrow A^{\times(n-1)}$ for all $i < n$. Verifications are now an exercise. □

Proof of Theorem 3.9. We prove (a) by induction on n. The result is clear for $n = 1$. If $A = \mathbb{1}^{\times n}$ for $n \geq 2$ then Lemma 3.11 gives $A^{[2]} \simeq A^{\times(n-1)} \simeq \mathbb{1}_A^{\times(n-1)}$ in A-$\mathrm{Mod}_{\mathcal{K}}$. By induction hypothesis applied to the tt-category A-$\mathrm{Mod}_{\mathcal{K}}$ we get $\deg(A^{[2]}) = n-1$ and hence the result by the definition of the degree. For (b), we need to prove that there are $A^{[n]}$-algebra isomorphisms $A^{[n]} \otimes A \simeq A^{[n]} \times \cdots \times A^{[n]} \times A^{[n+1]}$ (with n factors $A^{[n]}$). This is an easy induction on n, applying $A^{[n+1]} \otimes_{A^{[n]}} -$ and using (3.2) at each stage. Equation (3.10) follows since $A^{[\deg(A)+1]} = 0$. Parts (d) and (e) follow from (a) and Theorem 3.7(b)–(d). □

Corollary 3.12. *Suppose that $\mathcal{K}$ is local and that $A, B \in \mathcal{K}$ are two tt-rings of finite degree. Then $A \times B$ and $A \otimes B$ have finite degree with* $\deg(A \times B) = \deg(A) + \deg(B)$ *and* $\deg(A \otimes B) = \deg(A) \cdot \deg(B)$.

Proof. By Theorem 3.9(c), there exists two tt-rings $\bar{A} \neq 0$ and $\bar{B} \neq 0$ such that $F_{\bar{A}}(A) \simeq \mathbb{1}^{\times \deg(A)}$ and $F_{\bar{B}}(B) \simeq \mathbb{1}^{\times \deg(B)}$. Let then $\bar{C} = \bar{A} \otimes \bar{B}$. Extending scalars from $\bar{A}$ and from $\bar{B}$ to $\bar{C}$ gives $F_{\bar{C}}(A \times B) \simeq \mathbb{1}^{\times(\deg(A)+\deg(B))}$ and $F_{\bar{C}}(A \otimes B) \simeq \mathbb{1}^{\times(\deg(A)\cdot\deg(B))}$. Finally, $\bar{C} \neq 0$ since $\mathcal{K}$ is local; now apply Theorem 3.9(e). □

Remark 3.13. It will be clear to the interested reader that several arguments, mostly the early ones of Section 2, only depend on the property that split epimorphisms in $\mathcal{K}$ admit a kernel (a property which holds when $\mathcal{K}$ is triangulated, regardless of idempotent-completeness). The reader interested in using the degree in that generality will easily adapt our definition. However, all results which involve $\mathrm{Spc}(\mathcal{K})$, the support $\mathrm{supp}(A)$, or the local categories $\mathcal{K}/\mathcal{P}$, as well as the geometric applications in [Balmer 2013], only make sense when $\mathcal{K}$ is triangulated. It is nonetheless interesting to be able to speak of the degree in the generality of, say, the category of abelian groups, for instance.

4. Examples

We start by quickly discussing tt-rings of minimal degree (beyond $\deg(0) = 0$).

Proposition 4.1. *Let A be a tt-ring with* $\deg(A) = 1$, *that is*, *such that* $\mu : A \otimes A \to A$ *is an isomorphism. Then* $A \otimes - : \mathcal{K} \to \mathcal{K}$ *is a* (*very special*) *Bousfield localization with* $F_A : \mathcal{K} \to A\text{-}\mathrm{Mod}_{\mathcal{K}}$ *as* (*Verdier*) *localization. Also*, $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}})$ *is homeomorphic to the open and closed subset* $\mathrm{supp}(A)$ *of* $\mathrm{Spc}(\mathcal{K})$. *If* $\mathcal{K}$ *is rigid, this further implies a decomposition* $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2$ *under which* $A \cong (\mathbb{1}, 0)$.

Proof. Since μ is an isomorphism, so are its two right inverses $\eta \otimes 1$ and $1 \otimes \eta : A \to A^{\otimes 2}$, and these inverses coincide. So, $L := A \otimes -$ is a Bousfield localization ($\eta L = L\eta$ is an isomorphism). Let $C \to \mathbb{1} \xrightarrow{\eta} A \to \Sigma(C)$ be an exact triangle on η. Since $A \otimes \eta$ is an isomorphism, we have $A \otimes C = 0$. Therefore $\mathrm{Spc}(\mathcal{K}) = \mathrm{supp}(A) \sqcup \mathrm{supp}(C)$, hence $\mathrm{supp}(A)$ is open and closed. Since every object $x \in \mathcal{K}$ fits in an exact triangle $C \otimes x \to x \to A \otimes x \to \Sigma(C \otimes x)$, it is standard to show that the kernel of $A \otimes -$ is exactly the thick $\otimes$-ideal $\mathcal{J} := \langle C \rangle$ generated by C and that F_A induces an equivalence $\mathcal{K}/\mathcal{J} \xrightarrow{\sim} A\text{-}\mathrm{Mod}_{\mathcal{K}}$. Hence $\mathrm{Spc}(F_A)$ induces a homeomorphism $\mathrm{Spc}(A\text{-}\mathrm{Mod}_{\mathcal{K}}) \cong \{\mathcal{P} \in \mathrm{Spc}(\mathcal{K}) \mid \mathcal{J} \subseteq \mathcal{P}\} = \{\mathcal{P} \mid C \in \mathcal{P}\} = \mathcal{U}(C) = \mathrm{supp}(A)$. When $\mathcal{K}$ is rigid, $\mathrm{supp}(A) \cap \mathrm{supp}(C) = \varnothing$ forces, furthermore, $\mathrm{Hom}_{\mathcal{K}}(A, \Sigma C) = 0$, in which case the above triangle splits: $\mathbb{1} \simeq A \oplus C$. This gives the desired decomposition, where $\mathcal{K}_1 = A \otimes \mathcal{K}$ and $\mathcal{K}_2 = C \otimes \mathcal{K}$. □

We want to show that the degree is finite in examples. Our main tool is:

Theorem 4.2. *Suppose that* $\mathcal{K}$ *admits a conservative tt-functor* $F : \mathcal{K} \to \mathcal{L}$ *into a tt-category* $\mathcal{L}$ *such that every object of* $\mathcal{L}$ *is isomorphic to a sum of suspensions of* $\mathbb{1}_{\mathcal{L}}$.[1] *Then every tt-ring in* $\mathcal{K}$ *has finite degree. More precisely*, *if* $F(A) \simeq \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$ *for* $r_k, \ldots, r_\ell \in \mathbb{N}$ *then* $\deg(A) = \sum_{i=k}^{\ell} r_i$.

Proof. By Theorem 3.7(b), it suffices to prove that every tt-ring A in $\mathcal{L}$ has finite degree $d = \sum_{i=k}^{\ell} r_i$, where $A \simeq \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$ as objects in $\mathcal{L}$. First, let $B = A^{[d+1]}$. Then, by Theorem 3.9(b), we have $F_B(A) \simeq \mathbb{1}^{d+1} \oplus x$ in $B\text{-}\mathrm{Mod}_{\mathcal{L}}$. On the other hand, $F_B(A) \simeq \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$. Therefore there is a split monomorphism $\mathbb{1}^{d+1} \rightarrowtail \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$ in $B\text{-}\mathrm{Mod}_{\mathcal{L}}$ which can be described by a split injective $d \times (d+1)$ matrix with coefficients in the graded-commutative ring $S^\bullet = \mathrm{End}_B^\bullet(\mathbb{1}_B)$. This is impossible (by mapping to a graded residue field of $S^\bullet$) unless $S^\bullet = 0$, that is, $B = 0$ meaning $A^{[d+1]} = 0$. Hence $\deg(A) \le d$. Now, replace B by $A^{[\deg(A)]}$ and reason as above. We now have isomorphisms $F_B(A) \simeq \mathbb{1}^{\deg(A)}$ and $F_B(A) \simeq \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$ in $B\text{-}\mathrm{Mod}_{\mathcal{L}}$ with $B \neq 0$. The isomorphism $\mathbb{1}^{\deg(A)} \simeq \bigoplus_{i=k}^{\ell} \Sigma^i \mathbb{1}^{r_i}$ forces (periodicities $\Sigma^i \mathbb{1} \simeq \mathbb{1}$ in $B\text{-}\mathrm{Mod}_{\mathcal{K}}$ whenever $r_i \neq 0$ and) $\deg(A) = \sum_{i=k}^{\ell} r_i$. □

Corollary 4.3. *Let X be a quasicompact and quasiseparated scheme* (*for example*, *an affine or a noetherian scheme*). *Then every tt-ring in* $\mathrm{D}^{\mathrm{perf}}(X)$ *has finite degree.*

[1] Such an $\mathcal{L}$ is sometimes called a "field" but the author finds this definition too restrictive. Also note that the existence of such a functor F forces $\mathcal{K}$ to be local.

Proof. By Theorem 3.8, we can assume that $X = \mathrm{Spec}(R)$ with $(R, \mathfrak{m})$ local. Then, the functor $\mathrm{D^{perf}}(R) \to \mathrm{D^{perf}}(k)$ to the residue field $k = R/\mathfrak{m}$ is conservative. □

Example 4.4. Let A be a separable commutative R-algebra which is projective as an R-module (and finitely generated by [DeMeyer and Ingraham 1971, Proposition II.2.1]). Since A is R-flat, we can view it as the "same" tt-ring in $\mathrm{D^{perf}}(R)$. Then its degree can be computed in every residue field, hence $\deg(A)$ coincides with the rank of A as R-module.

Corollary 4.5. *Let $\mathbb{H}$ be a finite-dimensional cocommutative Hopf algebra over a field $\Bbbk$. Then every tt-ring in the bounded derived category* $\mathrm{D^b}(\mathbb{H}\text{-mod})$ *of finitely generated $\mathbb{H}$-modules* (*with $\otimes = \otimes_{\Bbbk}$*) *has finite degree.*

Proof. Apply Theorem 4.2 to the fiber functor $\mathrm{D^b}(\mathbb{H}\text{-mod}) \to \mathrm{D^b}(\Bbbk)$. □

Example 4.6. For any finite group G, all tt-rings in $\mathrm{D^b}(\Bbbk G\text{-mod})$ have finite degree. For every subgroup $H \le G$, the tt-ring $A = \Bbbk(G/H)$ has finite degree $\deg(A) = \dim_{\Bbbk}(A) = [G:H]$ in $\mathrm{D^b}(\Bbbk G\text{-mod})$. Hence A has also finite degree in

$$\mathrm{stab}(\Bbbk G) \cong \frac{\mathrm{D^b}(\Bbbk G\text{-mod})}{\mathrm{D^{perf}}(\Bbbk G)}$$

by Theorem 3.7(a). However, if $H < G$ is a strongly p-embedded subgroup then $F_A \cong \mathrm{Res}^G_H$ is an equivalence $\mathrm{stab}(\Bbbk G) \xrightarrow{\sim} \mathrm{stab}(\Bbbk H)$ and $\eta_A : \mathbb{1} \xrightarrow{\sim} A$ is an isomorphism, hence $\deg(A) = 1$ in $\mathrm{stab}(\Bbbk G)$. (Example: $p = 2$ and $C_2 < S_3$.)

Example 4.7. Let H_1 and H_2 be two nonconjugate cyclic subgroups of order p in G (for instance, two nonconjugate symmetries in D_8 for $p = 2$) and consider $A_i = \Bbbk(G/H_i)$ in $\mathcal{K} = \mathrm{stab}(\Bbbk G)$ as above. Then, by the Mackey formula, $A_1 \otimes A_2 = 0$. Consequently they have disjoint support and therefore both formulas of Corollary 3.12 fail in this case, showing the importance of our assumption that the category be local. Yet one can still deduce global formulas via Theorem 3.8.

Corollary 4.8. *In the stable homotopy category $\mathcal{K} = \mathrm{SH^{fin}}$ of finite* (*topological*) *spectra*, *every tt-ring has finite degree.*

Proof. First note that the result is true in the localizations $\mathrm{SH}^{\mathrm{fin}}_{\mathbb{Q}} \cong \mathrm{D^b}(\mathbb{Q}\text{-mod})$ and $\mathrm{SH}^{\mathrm{fin}}_p$, at zero and at each prime p. For the latter, it suffices to apply Theorem 4.2 to homology with coefficients in $\mathbb{Z}/p$, which is conservative on $\mathrm{SH}^{\mathrm{fin}}_p$ and takes values in $\mathrm{D^b}(\mathbb{Z}/p\text{-mod})$. Now, if A is a tt-ring in $\mathrm{SH^{fin}}$, then there exists $m \ge 1$ such that $A^{[m]}$ goes to zero in $\mathrm{SH}^{\mathrm{fin}}_{\mathbb{Q}}$ (since its degree is finite there). Replacing A by $A^{[m]}$, we can assume that A itself maps to zero in $\mathrm{SH}^{\mathrm{fin}}_{\mathbb{Q}}$, that is, A is torsion. But then A is nonzero in $\mathrm{SH}^{\mathrm{fin}}_p$ for only finitely many primes p. Therefore we can find n big enough that $A^{[n]} = 0$ everywhere. Hence $A^{[n]} = 0$. □

Acknowledgments

It is my pleasure to thank an anonymous referee for several useful suggestions, including the one discussed in Remark 3.13 and a simplification of the proof of Corollary 4.8.

References

[Balmer 2011] P. Balmer, "Separability and triangulated categories", *Adv. Math.* **226**:5 (2011), 4352–4372. MR 2012c:18021 Zbl 1236.18013

[Balmer 2012] P. Balmer, "Stacks of group representations", preprint, 2012, available at http://www.math.ucla.edu/~balmer/research/Pubfile/G-stacks.pdf. To appear in *J. Eur. Math. Soc.*

[Balmer 2013] P. Balmer, "Separable extensions in tt-geometry and generalized Quillen stratification", preprint, 2013. arXiv 1309.1808

[Chase et al. 1965] S. U. Chase, D. K. Harrison, and A. Rosenberg, "Galois theory and Galois cohomology of commutative rings", pp. 15–33 Mem. Amer. Math. Soc. No. **52**, Amer. Math. Soc., Providence, RI, 1965. MR 33 #4118 Zbl 0143.05902

[DeMeyer and Ingraham 1971] F. DeMeyer and E. Ingraham, *Separable algebras over commutative rings*, Lecture Notes in Mathematics **181**, Springer, Berlin, 1971. MR 43 #6199 Zbl 0215.36602

[Eilenberg and Moore 1965] S. Eilenberg and J. C. Moore, "Adjoint functors and triples", *Illinois J. Math.* **9** (1965), 381–398. MR 32 #2455 Zbl 0135.02103

[Hovey et al. 1997] M. Hovey, J. H. Palmieri, and N. P. Strickland, *Axiomatic stable homotopy theory*, Mem. Amer. Math. Soc. **610**, Amer. Math. Soc., Providence, RI, 1997. MR 98a:55017 Zbl 0881.55001

[Kleisli 1965] H. Kleisli, "Every standard construction is induced by a pair of adjoint functors", *Proc. Amer. Math. Soc.* **16** (1965), 544–546. MR 31 #1289 Zbl 0138.01704

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics **5**, Springer, New York, 1998. MR 2001j:18001 Zbl 0906.18001

Communicated by Yuri Manin
Received 2013-09-10 Revised 2013-12-02 Accepted 2014-01-08

balmer@math.ucla.edu *Mathematics Department, University of California, Los Angeles, Box 951555, Los Angeles, CA 90095-1555, 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 8 No. 3 2014

1937-0652(2014)8:3;1-8