Algebra & Number Theory

Volume 10
2016
No. 4

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2016 is US $290/year for the electronic version, and $485/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 MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.695

Moduli of morphisms of logarithmic schemes

Jonathan Wise

We show that there is a logarithmic algebraic space parameterizing logarithmic morphisms between fixed logarithmic schemes when those logarithmic schemes satisfy natural hypotheses. As a corollary, we obtain the representability of the stack of stable logarithmic maps from logarithmic curves to a fixed target without restriction on the logarithmic structure of the target.

An intermediate step requires a left adjoint to pullback of étale sheaves, whose construction appears to be new in the generality considered here, and which may be of independent interest.

1. Introduction

Let X and Y be logarithmic algebraic spaces over a logarithmic scheme S. Consider the functor $\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y)$[1] whose value on a logarithmic S-scheme S' is the set of logarithmic morphisms $X' \to Y'$, where $X' = X \times_S S'$ and $Y' = Y \times_S S'$. Under reasonable hypotheses on these data, we show that $\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y)$ is representable by a logarithmic algebraic space over S.

Supported by an NSA Young Investigator's Grant, Award #H98230-14-1-0107.
MSC2010: primary 14H10; secondary 14D23, 14A20.
Keywords: logarithmic geometry, moduli.

[1] See Section 1C for our conventions on notation.

Our strategy is to work relative to the space $\mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$ parameterizing morphisms of the underlying algebraic spaces $\underline{X}$ and $\underline{Y}$ of X and Y, respectively. More precisely, $\mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})(S')$ is the set of morphisms of schemes $\underline{X} \times_{\underline{S}} \underline{S}' \to \underline{Y} \times_{\underline{S}} \underline{S}'$. In order to guarantee that a morphism

$$\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y) \to \mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$$

exists, we need to assume that the morphism of logarithmic spaces $\pi : X \to S$ is integral, meaning $\pi^* M_S \to M_X$ is an integral morphism of sheaves of monoids.

Theorem 1.1. *Let $\pi : X \to S$ be a proper, flat, finite presentation, geometrically reduced, integral morphism of fine logarithmic algebraic spaces. Let Y be a logarithmic stack[2] over S. Then the morphism*

$$\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y) \to \mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$$

is representable by logarithmic algebraic spaces locally of finite presentation over S.

Combining the theorem with already known criteria for the algebraicity of $\mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$, such as [Hall and Rydh 2015, Theorem 1.2], we obtain:

Corollary 1.1.1. *In addition to the assumptions of Theorem 1.1, assume as well that Y is an algebraic stack over S that is locally of finite presentation with quasicompact and quasiseparated diagonal and affine stabilizers, and that $X \to S$ is of finite presentation. Then $\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y)$ is representable by a logarithmic algebraic stack locally of finite presentation over S.*

One application is to the construction of the stack of prestable logarithmic maps. Let $\mathfrak{M}$ denote the logarithmic stack of logarithmic curves. The algebraicity of $\mathfrak{M}$ may be verified in a variety of ways, e.g., [Gross and Siebert 2013, Proposition A.3]. For a logarithmic algebraic stack Y over S, we write $\mathfrak{M}(Y/S)$ for the logarithmic stack whose T-points are logarithmically commutative diagrams

$$\begin{array}{ccc} C & \longrightarrow & Y \\ \downarrow & & \downarrow \\ T & \longrightarrow & S \end{array}$$

in which C is a logarithmic curve over T.

Taking X to be the universal curve over $\mathfrak{M}$ in the previous corollary yields:

Corollary 1.1.2. *Suppose that $Y \to S$ is a morphism of logarithmic algebraic stacks with quasifinite and separated relative diagonal. Then $\mathfrak{M}(Y/S)$ is representable by a logarithmic algebraic stack locally of finite presentation over S.*

This improves on several previous results:

[2] It is not necessary for Y to be algebraic.

(1) [Chen 2014] required Y to have a rank-1 Deligne–Faltings logarithmic structure,

(2) [Abramovich and Chen 2014] required Y to have a generalized Deligne–Faltings logarithmic structure, and

(3) [Gross and Siebert 2013] required Y to have a Zariski logarithmic structure.

The evaluation space for stable logarithmic maps can also be constructed using Theorem 1.1. Recall that the standard logarithmic point P is defined by restricting the divisorial logarithmic structure of $\boldsymbol{A}^1$ to the origin. A family of standard logarithmic points in Y parameterized by a logarithmic scheme S is a morphism of logarithmic algebraic stacks $S \times P \to Y$. Following [Abramovich et al. 2010], we define $\wedge Y$ to be the fibered category of standard logarithmic points of Y.

Corollary 1.1.3 [Abramovich et al. 2010, Theorem 1.1.1]. *If Y is a logarithmic algebraic stack with quasifinite and quasiseparated diagonal then $\wedge Y$ is representable by a logarithmic algebraic stack locally of finite presentation over S.*

1A. ***Outline of the proof.*** Working relative to $\mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$, the question of the algebraicity of $\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y)$ is reduced to showing that, given a logarithmic algebraic space X and a logarithmic algebraic stack Y, both over S, as well as a commutative triangle

$$\begin{tikzcd} \underline{X} \arrow[rr, "f"] \arrow[dr, "\pi"'] & & \underline{Y} \arrow[dl] \\ & \underline{S} & \end{tikzcd}$$

of algebraic stacks, the lifts of f to an S-morphism of logarithmic algebraic stacks making the triangle commute are representable by a logarithmic algebraic stack.

This problem reduces immediately to the verification that morphisms of logarithmic structures $f^* M_Y \to M_X$ compatible with the maps from $\pi^* M_S$ are representable by a logarithmic algebraic space over S. We may therefore eliminate Y from our consideration by setting $M = f^* M_Y$ and restricting our attention to the functor $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ that parameterizes morphisms of logarithmic structures $M \to M_X$.

Stated precisely, the S'-points of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ are the morphisms of logarithmic structures $M' \to M'_X$, where M' and M'_X are the logarithmic structures deduced by base change on $\underline{X}' = \underline{X} \times_{\underline{S}} \underline{S}'$ that fit into a commutative triangle

$$\begin{tikzcd} & \pi'^* M_{S'} \arrow[dl] \arrow[dr] & \\ M' \arrow[rr] & & M'_X \end{tikzcd}$$

As is typical for logarithmic moduli problems, we now separate the question of the representability of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ by a logarithmic algebraic stack into a question about the representability of a larger stack

$$\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)) = \mathrm{Hom}_{\mathbf{Sch}/\mathbf{Log}(S)}(M, M_X)$$

over *schemes* (not logarithmic schemes), followed by the identification of an open substack of *minimal* objects within $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ that represents $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$.

When $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ is viewed as a category, its objects are the same as the objects of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$. Its fiber over a scheme T consists of tuples (M_T, f, α), where

(i) M_T is a logarithmic structure on T,

(ii) $f : (T, M_T) \to S$ is a morphism of logarithmic schemes, and

(iii) $\alpha : f^*M \to f^*M_X$ is a morphism of logarithmic structures on $\underline{X} \times_{\underline{S}} T$ that is compatible with the maps from M_T.

The categories $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ and $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ are slightly different: the morphisms in the former are required to be cartesian over the category of schemes, while in the latter they are only required to be cartesian over the category of logarithmic schemes. That is, $(T, M_T, f, \alpha) \to (T', M_{T'}, f', \alpha')$ in $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ lies in $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ only if the map $(T, M_T) \to (T', M_{T'})$ is strict.

We show in Section 2 that $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ is representable by an algebraic space relative to $\mathbf{Log}(S)$. As Olsson has proved that $\mathbf{Log}(S)$ is algebraic, the representability of $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ by an algebraic *stack* follows.

In Sections 3 and 4 we use Gillam's criterion (Appendix B) to prove that the fibered category $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ over logarithmic schemes is induced from a fibered category over schemes with a logarithmic structure. Section 3 treats the case where $X = S$ by adapting methods from homological algebra to commutative monoids. In Section 4, we transform the local minimal object of Section 3 to a global minimal object by means of a left adjoint to pullback for étale sheaves (constructed, under suitable hypotheses, in Section 4A), whose existence appears to be a new observation.

Gillam's criterion characterizes the fibered category over schemes, inducing the fibered category $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ over logarithmic schemes: it is the substack of *minimal objects* of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$. A slight augmentation of that criterion (described in Appendix B) implies that the substack of minimal objects is open in $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$. Combined with the algebraicity of $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ proved in Section 2, the verification of Gillam's criteria in Sections 3 and 4 implies that $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ is representable by a logarithmic algebraic *stack*. A direct analysis of the stabilizers of logarithmic maps in Section 5 then implies that $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ is representable by an algebraic *space*.

1B. *Remarks on hypotheses.* It is far from clear that all of the hypotheses of Theorem 1.1 are essential. We summarize how they are used in the proof: integrality of X over S is used to guarantee the existence of a morphism

$$\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y) \to \mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y});$$

it is also used in the construction of minimal objects. Properness and finite presentation are used to guarantee the representability of $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ by an algebraic stack. Flatness, finite presentation, and geometrically reduced fibers are used to guarantee the existence of a left adjoint to pullback of étale sheaves, used in the construction of global minimal objects from local ones.

1C. *Conventions.* We generally follow the notation of [Kato 1989] concerning logarithmic structures, except that we write $\underline{X}$ for the scheme (or fibered category) underlying a logarithmic scheme (or fibered category) X. The logarithmic structures that appear in this paper will all be fine, although we will usually point this out in context. If M is a logarithmic structure on X, we write $\exp : M \to \mathcal{O}_X$ for the structural morphism and $\log : \mathcal{O}_X^* \to M$ for the reverse inclusion.

It is occasionally convenient to pass only part of the way from a chart for a logarithmic structure to its associated logarithmic structure. We formalize this in the following definition:

Definition 1.2. A *quasilogarithmic structure* on a scheme X is an extension N of an étale sheaf of integral[3] monoids $\bar{N}$ by $\mathcal{O}_X^*$ and a morphism $N \to \mathcal{O}_X$ compatible with the inclusions of $\mathcal{O}_X^*$. We will say that a quasilogarithmic structure is *coherent* if its associated logarithmic structure is coherent. If $f : X' \to X$ is a morphism of schemes, the *pullback* f^*N of N to X' is obtained by pushout via $f^{-1}\mathcal{O}_X^* \to \mathcal{O}_{X'}$ from the pulled-back extension $f^{-1}N$:

$$\begin{tikzcd}
0 \arrow[r] & f^{-1}\mathcal{O}_X^* \arrow[r] \arrow[d] & f^{-1}N \arrow[r] \arrow[d] & f^*\bar{N} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{O}_{X'}^* \arrow[r] & f^*N \arrow[r] & f^*\bar{N} \arrow[r] & 0
\end{tikzcd}$$

We write $\mathrm{Hom}(A, B)$ for the set of morphisms between two objects of the same type. When A and B and the morphisms between them may reasonably be construed to vary with objects of a category $\mathscr{C}$, we write $\mathrm{Hom}_{\mathscr{C}}(A, B)$ for the

[3]The integrality assumption is not necessary in the definition. It is included to avoid qualifying every quasilogarithmic structure that appears below with the adjective "integral".

functor or fibered category of morphisms between A and B. Occasionally, we also employ a subscript on Hom to indicate restriction to homomorphisms preserving some additional structure. We rely on context to keep the two meanings of these decorations distinct.

2. Algebraicity relative to the category of schemes

We show that the morphism

$$\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)) \to \mathbf{Log}(S) \tag{2-1}$$

is representable by algebraic spaces. Combined with the algebraicity of $\mathbf{Log}(S)$ [Olsson 2003, Theorem 1.1], this implies that $\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X))$ is representable by an algebraic stack.

Proposition 2.1. *Let $S = (\underline{S}, M_S)$ be a logarithmic scheme and let X be a proper S-scheme with $\pi : X \to S$ denoting the projection. Assume given logarithmic structures M and M_X on X with morphisms of logarithmic structures $\pi^* M_S \to M$ and $\pi^* M_S \to M_X$. Assume as well that M_S, M_X, and M are all coherent. Then the morphism* (2-1) *is representable by algebraic spaces locally of finite presentation over S.*

We first treat the local problem in which $X = S$ and then pass to the general case.

Theorem 2.2 [Gross and Siebert 2013, Proposition 2.9]. *Suppose that P and Q are coherent logarithmic structures on a scheme X. Then* $\mathrm{Hom}_{\mathbf{Sch}/X}(P, Q)$ *is representable by an algebraic space locally of finite presentation over X.*

For isomorphisms, this is [Olsson 2003, Corollary 3.4].

Proof. The question of the algebraicity of $\mathrm{Hom}_{\mathbf{Sch}/X}(P, Q)$ may be separated into one about the algebraicity of $\mathrm{Hom}_{\mathbf{Sch}/X}(\bar{P}, \bar{Q})$ and another about the relative algebraicity of the map

$$\mathrm{Hom}_{\mathbf{Sch}/X}(P, Q) \to \mathrm{Hom}_{\mathbf{Sch}/X}(\bar{P}, \bar{Q}). \tag{2-2}$$

Lemma 2.2.1. *The functor* $\mathrm{Hom}_{\mathbf{Sch}/X}(\bar{P}, \bar{Q})$ *is representable by an étale algebraic space over X.*

Proof. Because $\bar{P}$ and $\bar{Q}$ are constructible, the natural map

$$f^* \mathrm{Hom}_{\text{ét}(X)}(\bar{P}, \bar{Q}) \to \mathrm{Hom}_{\text{ét}(X')}(f^* \bar{P}, f^* \bar{Q})$$

is an isomorphism for any morphism $f : X' \to X$. Therefore, we may represent $\mathrm{Hom}_{\mathbf{Sch}/X}(\bar{P}, \bar{Q})$ with the espace étalé of $\mathrm{Hom}_{\text{ét}(X)}(\bar{P}, \bar{Q})$. □

The relative algebraicity of (2-2) is equivalent to the following lemma:

Lemma 2.2.2 [Gross and Siebert 2013, Lemma 2.12]. *Let Q be a logarithmic structure on a scheme X and let P be a coherent quasilogarithmic structure on X. Fix a morphism $\bar{u} : \bar{P} \to \bar{Q}$. The lifts of $\bar{u}$ to a morphism $u : P \to Q$ of quasilogarithmic structures are parameterized by a relatively affine scheme of finite presentation over X.*

Our proof of this lemma only differs from that of [loc. cit.] superficially, but is nevertheless included for the sake of completeness. It is also possible to deduce Lemma 2.2.2 from Lemma 2.2.1 and [Olsson 2003, Corollary 3.4].

Proof. This is a local question in X, so we may freely pass to an étale cover. Furthermore, replacing P with a quasilogarithmic structure P_0 that has the same associated logarithmic structure does not change the morphisms to Q, by the universal property of the associated logarithmic structure. Since the logarithmic structure associated to P admits a chart étale locally, we can therefore select P_0 to be a quasilogarithmic structure whose sheaf of characteristic monoids $\bar{P}_0$ is constant. Replacing P with P_0, we can assume that the characteristic monoid of P is constant.

We wish to construct the space of completions of the diagram

$$\begin{tikzcd} P \arrow[r, dashed, "u"] \arrow[d] & Q \arrow[d] \\ \bar{P} \arrow[r, "\bar{u}"] & \bar{Q} \end{tikzcd}$$

(in which u is also required to be compatible with the maps $\exp : P \to \mathcal{O}_X$ and $\exp : Q \to \mathcal{O}_X$). Replacing Q with $\bar{u}^{-1}Q$, we can assume that $\bar{P} = \bar{Q}$ and $\bar{u} = \mathrm{id}_{\bar{P}}$.[4]

Let H be the moduli space of maps $u : P \to Q$ that are compatible with $\bar{u} = \mathrm{id}_{\bar{P}}$, ignoring the maps to $\mathcal{O}_X$. Locally such a map exists because P and Q are both extensions of $\bar{P}$ by $\boldsymbol{G}_m$ and $\bar{P}$ is generated by a finite collection of global sections. Indeed, this implies that P and Q are each determined by a finite collection of $\boldsymbol{G}_m$-torsors on X, all of which can be trivialized after passage to a suitable open cover of X. It follows that H is a torsor on X under $\mathrm{Hom}_{\mathbf{Sch}/X}(\bar{P}, \boldsymbol{G}_m)$ and in particular is representable by an affine scheme over X, which is of finite presentation since $\bar{P}$ is finitely generated, hence finitely presented [Rosales and García-Sánchez 1999, Theorem 5.12].

[4]At this point a morphism $P \to Q$ covering $\bar{u}$ must be an isomorphism, so we could complete the proof using [Olsson 2003, Corollary 3.4].

We may now work relative to H and assume that the map $u : P \to Q$ has already been specified. We argue that the locus where the diagram

$$\begin{tikzcd} P \arrow[d] \arrow[dr, "\exp"] & \\ Q \arrow[r, "\exp"'] & \mathcal{O}_X \end{tikzcd}$$

commutes is closed. In effect, we are looking at the locus where two $\boldsymbol{G}_m$-equivariant maps $P \to \boldsymbol{A}^1$ agree. But P is generated as a monoid with $\boldsymbol{G}_m$-action by a finite collection of sections, hence the agreement of the two maps $P \to \boldsymbol{A}^1$ corresponds to the agreement of a finite collection of pairs of sections of $\boldsymbol{A}^1$. But $\boldsymbol{A}^1$ is separated, so this is representable by a closed subscheme of finite presentation. □

This completes the proof of Theorem 2.2. □

We can now obtain a global variant:

Corollary 2.2.3. *Let X be a proper, flat, finite presentation algebraic space over S and let P, Q be logarithmic structures on X with P coherent. Then* $\mathrm{Hom}_{\mathbf{Sch}/S}(P, Q)$ *is representable by an algebraic space locally of finite presentation over S.*

Proof. Let $\pi : X \to S$ be the projection. Then we have

$$\mathrm{Hom}_{\mathbf{Sch}/S}(P, Q) = \pi_* \mathrm{Hom}_{\mathbf{Sch}/X}(P, Q).$$

We have already seen in Theorem 2.2 that $\mathrm{Hom}_{\mathbf{Sch}/X}(P, Q)$ is representable by an algebraic space over X that is quasicompact, quasiseparated, and locally of finite presentation. We may therefore apply [Hall and Rydh 2015, Theorem 1.3] (or any of a number of other representability results for schemes of morphisms) to deduce the algebraicity and local finite presentation of $\pi_* \mathrm{Hom}_{\mathbf{Sch}/X}(P, Q)$. □

Corollary 2.2.4. *Let X and S be as in the last corollary. Suppose that P, Q, and R are logarithmic structures on X with P and Q coherent and that morphisms $\alpha : P \to Q$ and $\beta : P \to R$ have been specified. Then there is an algebraic space, locally of finite presentation over S, parameterizing the commutative triangles*

$$\begin{tikzcd} P \arrow[d, "\alpha"'] \arrow[dr, "\beta"] & \\ Q \arrow[r] & R \end{tikzcd}$$

Proof. We recognize this functor as a fiber product,

$$\mathrm{Hom}_{\mathbf{Sch}/S}(Q, R) \times_{\mathrm{Hom}_{\mathbf{Sch}/S}(P,R)} \{\beta\}.$$ □

Proposition 2.1 is an immediate consequence of Corollary 2.2.4, applied with $P = \pi^* M_S$, $Q = M$, and $R = M_X$.

3. Local minimality

After Proposition 2.1, all that is left to demonstrate Theorem 1.1 is to verify Gillam's criteria for

$$\mathbf{Log}(\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)) = \mathrm{Hom}_{\mathbf{Sch}/\mathbf{Log}(S)}(M, M_X).$$

As in the proof of Proposition 2.1, we separate this problem into local and global variants, the local version being the case $\underline{S} = \underline{X}$. We treat the local problem in this section and deduce the solution to the global problem in the next one.

Let $\underline{X}$ be a scheme equipped with three fine logarithmic structures, denoted $\pi^* M_S$, M_X, and M in order to emphasize the application in the next section, and morphisms of logarithmic structures

$$\pi^* M_S \to M_X, \qquad \pi^* M_S \to M.$$

Let $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ be the set of commutative diagrams

$$\begin{tikzcd} \pi^* M_S \arrow[r] \arrow[d] \arrow[rr, bend left] & N \arrow[d] & M \arrow[dl, "\varphi"] \\ M_X \arrow[r] & N_X & \end{tikzcd}$$

in which N is a *quasilogarithmic structure*[5] (Definition 1.2) and the square on the left is cocartesian. These data are determined up to unique isomorphism by the quasilogarithmic structure N, the morphism $\pi^* M_S \to N$, and the morphism φ. We will refer to an object of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ with the pair (N, φ), with the morphism $\pi^* M_S \to N$ specified tacitly.[6]

Convention 3.1. As a matter of notation, whenever we have a morphism of monoids $\pi^* M_S \to N$ (resp. $\pi^* \bar{M}_S \to \bar{N}$), we write N_X (resp. $\bar{N}_X$) for the monoid obtained by pushout:

$$\begin{tikzcd} \pi^* M_S \arrow[r] \arrow[d] & M_X \arrow[d] \\ N \arrow[r] & N_X \end{tikzcd} \qquad \left(\text{resp.} \quad \begin{tikzcd} \pi^* \bar{M}_S \arrow[r] \arrow[d] & \bar{M}_X \arrow[d] \\ \bar{N} \arrow[r] & \bar{N}_X \end{tikzcd} \right)$$

[5]The use of quasilogarithmic structures here is entirely for convenience: it allows us to avoid repeated passage to associated logarithmic structures. The reader who would prefer not to worry about quasilogarithmic structures should feel free to assume N is a logarithmic structure and worry instead about remembering to take associated logarithmic structures at the right moments.

[6]Effectively, N is an object of the category of quasilogarithmic structures equipped with a morphism from $\pi^* M_S$.

When $N = N_S$ above, we simply write N_X rather than $(N_S)_X$.

The object of this section will be to prove the following two lemmas:

Lemma 3.2. *For any $\underline{X}$-scheme $\underline{Y}$, any object of* $\mathrm{GS}^{\mathrm{loc}}(\underline{Y})$ *admits a morphism from a minimal object.*

Lemma 3.3. *The pullback of a minimal object of* $\mathrm{GS}^{\mathrm{loc}}(\underline{Y})$ *via any morphism $\underline{Y}' \to \underline{Y}$ is also minimal.*

The construction of the minimal object appearing in Lemma 3.2 is done in Section 3A, while the proof of its minimality appears in Section 3B, along with the proof of Lemma 3.3.

3A. ***Construction of minimal objects.*** Fixing $(N, \varphi) \in \mathrm{GS}^{\mathrm{loc}}(\underline{X})$, we construct an object $(R, \rho) \in \mathrm{GS}^{\mathrm{loc}}(\underline{X})$ and a morphism $(R, \rho) \to (N, \varphi)$. In Section 3B we verify that (R, ρ) is minimal and that its construction is stable under pullback.

We assemble R in steps: first we build the associated group of its characteristic monoid, then we identify its characteristic monoid within this group, and finally we build the quasilogarithmic structure above the characteristic monoid.

Recall that the map $\varphi : M \to N_X$ induces a map $u : M \to N_X/\pi^* N_S \simeq M_X/\pi^* M_S$ known as the *type* of u. This generalizes [Gross and Siebert 2013, Definition 1.10]. For brevity, we write $\bar{M}_{X/S} = M_X/\pi^* M_S = \bar{M}_X/\pi^* \bar{M}_S$ below. Note that u is equivariant with respect to the action of $\pi^* M_S$ on M and the (trivial) action of $\pi^* M_S$ on the relative characteristic monoid $\bar{M}_{X/S}$. Therefore, u may equally well be considered a morphism $\bar{M}/\pi^* \bar{M}_S \to \bar{M}_{X/S}$.

Remark 3.4. The following construction is technical, so the reader may find it helpful to keep in mind that it is really an elaboration of an exercise in homological algebra:

If $0 \to A \to B \to C \to 0$ is an exact sequence of abelian groups, and $u : M \to C$ is a given homomorphism, there is a universal homomorphism $A \to A'$ such that u lifts to a homomorphism $M \to B'$, where $B' = A' \amalg_A B$:

$$\begin{tikzcd}
 & & & M \arrow[d] \arrow[ddl, dashed] & \\
0 \arrow[r] & A \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & C \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & A' \arrow[r] & B' \arrow[r] & C \arrow[r] & 0
\end{tikzcd}$$

Moreover, A' may be taken to be $M \times_C B$.

The associated group of the characteristic monoid of R. We set $\bar{R}_0^{\mathrm{gp}} = \bar{M}^{\mathrm{gp}} \times_{\bar{M}_{X/S}^{\mathrm{gp}}} \bar{M}_X^{\mathrm{gp}}$. This fits into a commutative diagram with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & \pi^* \bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d, equal] & \bar{R}_0^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}^{\mathrm{gp}} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \pi^* \bar{M}_S^{\mathrm{gp}} \arrow[r] & \bar{M}_X^{\mathrm{gp}} \arrow[r] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] & 0
\end{tikzcd}$$

Observe that $\bar{R}_0^{\mathrm{gp}}$ comes with *two* maps $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{R}_0^{\mathrm{gp}}$ corresponding to the two maps

$$\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}^{\mathrm{gp}}, \qquad \pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}_X^{\mathrm{gp}}.$$

We take $\epsilon : \bar{R}_0^{\mathrm{gp}} \to \bar{R}^{\mathrm{gp}}$ to be the quotient of $\bar{R}_0^{\mathrm{gp}}$ by the diagonal copy of $\pi^* \bar{M}_S^{\mathrm{gp}}$. We may then define $\bar{R}_X^{\mathrm{gp}}$ by pushout via $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}_X^{\mathrm{gp}}$ (Convention 3.1).

The homomorphism $\bar{\rho} : \bar{M}^{\mathrm{gp}} \to \bar{R}_X^{\mathrm{gp}}$. Let $\bar{M}_X^+$ be the pushout of $\bar{M}_X$ by the homomorphism *of monoids* $\pi^* \bar{M}_S \to \pi^* \bar{M}_S^{\mathrm{gp}}$. This is a submonoid of $\bar{M}_X^{\mathrm{gp}}$ and fits into the exact sequence in the middle row of the diagram below. Let $\bar{R}_0^+$ be the pullback of $\bar{M} \to \bar{M}_{X/S}$ to $\bar{M}_X^+$ (the upper right square of the diagram). The diagram is commutative except for the dashed arrows (which will be explained momentarily) and has exact rows:

$$\begin{tikzcd}
0 \arrow[r] & \pi^* \bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{R}_0^+ \arrow[r] \arrow[d, "\beta"] \arrow[ddl, dashed, "\epsilon"'] & \bar{M} \arrow[r] \arrow[d] \arrow[ddl, dashed, "\bar{\rho}"'] & 0 \\
0 \arrow[r] & \pi^* \bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}_X^+ \arrow[r] \arrow[d] & \bar{M}_{X/S} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \bar{R}^{\mathrm{gp}} \arrow[r, "\alpha"'] & (\bar{R}^{\mathrm{gp}})_X \arrow[r] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] & 0
\end{tikzcd}$$

Note that $(\bar{R}^{\mathrm{gp}})_X$ is the pushout of $\pi^* \bar{M}_S \to \bar{M}_X$ via $\pi^* \bar{M}_S \to \bar{R}^{\mathrm{gp}}$ *as a monoid*. Equivalently, it is the pushout of $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}_X^+$ via $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{R}^{\mathrm{gp}}$, again as a monoid. It is contained in but not necessarily equal to $\bar{R}_X^{\mathrm{gp}} = (\bar{R}_X)^{\mathrm{gp}}$.

The difference between the two compositions

$$\bar{R}_0^+ \xrightarrow{\beta} \bar{M}_X^+ \longrightarrow (\bar{R}^{\mathrm{gp}})_X, \qquad \bar{R}_0^+ \xrightarrow{\epsilon} \bar{R}^{\mathrm{gp}} \xrightarrow{\alpha} (\bar{R}^{\mathrm{gp}})_X$$

factors uniquely through a map $\bar{M}^{\mathrm{gp}} \to (\bar{R}^{\mathrm{gp}})_X \subset \bar{R}_X^{\mathrm{gp}}$. We take this as the definition of $\bar{\rho}$.

Remark 3.5. Observe that when the maps $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{R}^{\mathrm{gp}}$ and $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}^{\mathrm{gp}}$ are the canonical ones, the diagram on the left commutes but the diagram on the right does not:

$$
\begin{tikzcd}
\pi^*\bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}^{\mathrm{gp}} \arrow[d, "\bar{\rho}"] \\
\bar{R}^{\mathrm{gp}} \arrow[r, "\alpha"] & \bar{R}_X^{\mathrm{gp}}
\end{tikzcd}
\qquad
\begin{tikzcd}
\pi^*\bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}^{\mathrm{gp}} \arrow[d, "\bar{\rho}"] \\
\bar{R}_0^{\mathrm{gp}} \arrow[r] & (\bar{R}_0^{\mathrm{gp}})_X
\end{tikzcd}
$$

This is the reason we introduced the quotient ϵ earlier.

We view $(\bar{R}^{\mathrm{gp}}, \bar{\rho})$ as the initial object of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ *on the level of associated groups of characteristic monoids*. Justification for this attitude will be given in Section 3B (see the proof of Lemma 3.6).

The characteristic monoid $\bar{R}$. We identify the smallest sheaf of submonoids $\bar{R} \subset \bar{R}^{\mathrm{gp}}$ that contains the image of $\pi^*\bar{M}_S$ and whose pushout $\bar{R}_X$ contains the image of $\bar{\rho} : \bar{M} \to \bar{R}_X^{\mathrm{gp}}$. For each local section ξ of $\bar{M}$ we will identify a local section (or possibly a finite collection of local sections) of $\bar{R}^{\mathrm{gp}}$ for inclusion in $\bar{R}$; we will then take $\bar{R}$ to be the submonoid of $\bar{R}^{\mathrm{gp}}$ generated by these local sections. As M is assumed to be coherent, a finite number of these local sections suffice to generate $\bar{R}$, which guarantees that $\bar{R}$ is coherent.

Suppose that $\xi \in \Gamma(U, \bar{M})$ is a section over some quasicompact U that is étale over X. Recall that $\bar{\rho}(\xi)$ lies in $(\bar{R}^{\mathrm{gp}})_X$, which is the pushout of $\bar{R}^{\mathrm{gp}}$ via the integral homomorphism $\pi^*\bar{M}_S \to \bar{M}_X$. At least after passage to a finer quasicompact étale cover, we can represent $\bar{\rho}(\xi)$ as a pair (a, b), where $a \in \bar{R}^{\mathrm{gp}}$ and $b \in \bar{M}_X$ (see Appendix A).

Let $B \subset \Gamma(U, \bar{M}_X)$ be the collection of all $b \in \bar{M}_X$ such that $\bar{\rho}(\xi)$ can be represented as (a, b) for some $a \in \Gamma(U, \bar{R}^{\mathrm{gp}})$. As $\bar{R}^{\mathrm{gp}} \to \bar{R}_X^{\mathrm{gp}}$ is injective (it is integral), there is at most one a for any $b \in \Gamma(U, \bar{M}_X)$. Note that B carries an action of $\Gamma(U, \pi^*\bar{M}_S)$, for if $\bar{\rho}(\xi)$ is representable by (a, b) then it is also representable by $(a - c, b + c)$. The action of the sharp monoid $\pi^*\bar{M}_S$ gives B a partial order by setting $b \le b + c$ for all $c \in \Gamma(U, \pi^*\bar{M}_S)$. We will show that b has a least element with respect to this partial order.

Suppose b and b' are elements of B with (a, b) and (a', b') both representing $\bar{\rho}(\xi) \in \Gamma(U, \bar{R}_X^{\mathrm{gp}})$. As $\pi^*\bar{M}_S \to \bar{M}_X$ is integral, Lemma A.2 implies that there must be elements $d \in \bar{M}_X$ and $c, c' \in \pi^*\bar{M}_S$ with $a + c = a' + c'$ and $b = d + c$ and $b' = d + c'$. But then $\bar{\rho}(\xi)$ is also representable by $(a + c, d) = (a' + c', d)$. Therefore, for any pair $b, b' \in B$ there is a $d \in B$ with $d \le b$ and $d \le b'$.

It will now follow that B has a least element if we can show that every infinite decreasing chain of elements of B stabilizes. But B is a subset of $\Gamma(U, \bar{M}_X)$, and, at least provided U has been chosen small enough, this is a strict submonoid of a finitely generated abelian group. A strictly decreasing chain of elements must have strictly decreasing distance from the origin in $\Gamma(U, \bar{M}_X) \otimes \mathbb{R}$, and there can be only a finite number of elements of B whose image in $\Gamma(U, \bar{M}_X) \otimes \mathbb{R}$ is within a fixed distance of the origin. Therefore, the chain must stabilize and B has a least element.

Writing b for the least element of B and a for the corresponding element of $\Gamma(U, \bar{R}^{\mathrm{gp}})$ such that (a, b) represents $\bar{\rho}(\xi)$, we include a as an element of $\bar{R}$. As $\bar{M}$ is coherent, we can repeat this construction for each element in a finite collection of sections that generate $\bar{M}$ over U (provided that U has been chosen small enough).

By construction, $\bar{R}$ is locally of finite type and integral. Moreover, the following lemma says that $(\bar{R}, \bar{\rho})$ is the initial object of type u in $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ *on the level of characteristic monoids*. We defer its proof to Section 3B in order not to interrupt the construction of (R, ρ).

Lemma 3.6. *For any $(N, \varphi) \in \mathrm{GS}^{\mathrm{loc}}(\underline{X})$ of type u, there is a unique morphism $(\bar{R}, \bar{\rho}) \to (\bar{N}, \bar{\varphi})$.*

The quasilogarithmic structure R and the map ρ. This construction will require an object $(N, \varphi) \in \mathrm{GS}^{\mathrm{loc}}(\underline{X})$ and not just a type. Suppose (N, φ) has type u and $(\bar{R}, \bar{\rho})$ has been constructed as in the steps above. Then Lemma 3.6 implies that there is a canonical map $\bar{R} \to \bar{N}$ compatible with the tacit maps from $\pi^* \bar{M}_S$. By pulling back N from $\bar{N}$, we obtain a quasilogarithmic structure R with characteristic monoid $\bar{R}$.

As we have a factorization (again by Lemma 3.6)

$$\bar{M} \xrightarrow{\bar{\rho}} \bar{R}_X \longrightarrow \bar{N}_X$$

and R_X is pulled back from N_X over $\bar{N}_X$, the universal property of the fiber product yields an induced map $\rho : M \to R$.

3B. *Verification of Gillam's criteria.*

Proof of Lemma 3.6. To prove the lemma, we must show that there is a unique morphism $\bar{\mu} : \bar{R} \to \bar{N}$ such that the induced diagram

$$\begin{tikzcd} \bar{M} \arrow[r, "\bar{\rho}"] \arrow[dr, "\bar{\varphi}"'] & \bar{R}_X \arrow[d, "\bar{\mu}_X"] \\ & \bar{N}_X \end{tikzcd} \tag{3-1}$$

is commutative.

Given (N, φ), we have a diagram with exact rows that commutes except for some of the parts involving the dashed arrow:

$$\begin{tikzcd}
0 \arrow[r] & \pi^*\bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d, equal] & \bar{R}_0^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}^{\mathrm{gp}} \arrow[r] \arrow[d, "u"] \arrow[ddl, dashed, "\bar{\varphi}"] & 0 \\
0 \arrow[r] & \pi^*\bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}_X^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \bar{N}^{\mathrm{gp}} \arrow[r] & \bar{N}_X^{\mathrm{gp}} \arrow[r] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] & 0
\end{tikzcd}$$

The lower triangle involving $\bar{\varphi}$ commutes (because (N, φ) has type u) but the upper one may not. The difference of the two compositions

$$\bar{R}_0^{\mathrm{gp}} \longrightarrow \bar{M}^{\mathrm{gp}} \xrightarrow{\bar{\varphi}} \bar{N}_X^{\mathrm{gp}}, \qquad \bar{R}_0^{\mathrm{gp}} \longrightarrow \bar{M}_X^{\mathrm{gp}} \longrightarrow \bar{N}_X^{\mathrm{gp}}$$

factors uniquely through a map $\bar{\mu} : \bar{R}_0^{\mathrm{gp}} \to \bar{N}^{\mathrm{gp}}$. Moreover, $\bar{\mu}$ vanishes on the diagonal copy of $\pi^*\bar{M}_S^{\mathrm{gp}}$ inside $\bar{R}_0^{\mathrm{gp}}$. This gives a factorization

$$\pi^*\bar{M}_S^{\mathrm{gp}} \longrightarrow \bar{R}^{\mathrm{gp}} \xrightarrow{\bar{\mu}} \bar{N}^{\mathrm{gp}}.$$

Moreover, the construction of $\bar{\mu}$ is easily reversed to give a bijective correspondence between maps $\bar{\mu} : \bar{R}^{\mathrm{gp}} \to \bar{N}^{\mathrm{gp}}$ compatible with the tacit maps from $\pi^*\bar{M}_S^{\mathrm{gp}}$ and morphisms $\bar{\varphi} : \bar{M}^{\mathrm{gp}} \to \bar{N}_X^{\mathrm{gp}}$ compatible with the type (see Remark 3.4).

We must verify that the image of $\bar{\mu} : \bar{R} \to \bar{N}^{\mathrm{gp}}$ lies in $\bar{N}$. By definition, $\bar{N}_X \subset \bar{N}_X^{\mathrm{gp}}$ contains the image of $\bar{\varphi} : \bar{M} \to \bar{N}_X^{\mathrm{gp}}$. Write $\bar{R}' \subset \bar{R}^{\mathrm{gp}}$ for the preimage of $\bar{N}$ via the map $\bar{R}^{\mathrm{gp}} \to \bar{N}^{\mathrm{gp}}$; thus $\bar{R}' \to \bar{N}$ is an exact morphism of monoids [Kato 1989, Definition 4.6(1)]. Then $\bar{R}'_X \to \bar{N}_X$ is also exact [loc. cit.], so $\bar{R}'_X$ coincides with the preimage of $\bar{N}_X \subset \bar{N}_X^{\mathrm{gp}}$. Furthermore, $\bar{R}'_X$ contains the image of $\bar{\rho} : \bar{M} \to \bar{R}_X^{\mathrm{gp}}$ because $\bar{N}_X$ contains the image of $\bar{\varphi} : \bar{M} \to \bar{N}_X^{\mathrm{gp}}$. On the other hand, $\bar{R}$ was constructed as the smallest submonoid of $\bar{R}^{\mathrm{gp}}$ such that $\bar{R}_X$ contains the image of $\bar{\rho} : \bar{M} \to \bar{R}_X^{\mathrm{gp}}$. Therefore, $\bar{R} \subset \bar{R}'$ and the image of $\bar{R} \to \bar{N}_X$ is contained in $\bar{N}$.

Finally, to get the commutativity of (3-1), it is sufficient to work on the level of associated groups. Assemble the diagram

$$\begin{tikzcd}
0 \arrow[r] & \pi^*\bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{R}_0^{\mathrm{gp}} \arrow[r] \arrow[d, "\beta"] \arrow[dl, dashed, "\epsilon"'] \arrow[ddl, dashed, "\bar{\mu}\epsilon"] & \bar{M}^{\mathrm{gp}} \arrow[r] \arrow[d] \arrow[dl, dashed, "\bar{\rho}"'] \arrow[ddl, dashed, "\bar{\varphi}"] & 0 \\
0 \arrow[r] & \bar{R}^{\mathrm{gp}} \arrow[r] \arrow[d, "\bar{\mu}"'] & \bar{R}_X^{\mathrm{gp}} \arrow[r] \arrow[d, "\bar{\mu}_X"'] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \bar{N}^{\mathrm{gp}} \arrow[r] & \bar{N}_X^{\mathrm{gp}} \arrow[r] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] & 0
\end{tikzcd}$$

which is commutative except for some of the parts involving the dashed arrows. We write d for any of the horizontal arrows. To show that $\bar{\varphi} = \bar{\mu}_X \circ \bar{\rho}$, it is sufficient to show that $\bar{\varphi} \circ d = \bar{\mu}_X \circ \bar{\rho} \circ d$. Recall that $\bar{\rho}$ was constructed such that $\bar{\rho} \circ d = \beta - d \circ \epsilon$ and $\bar{\mu}$ was constructed such that $d \circ \bar{\mu} \circ \epsilon = \bar{\mu}_X \circ \beta - \bar{\varphi} \circ d$. Therefore,

$$\bar{\mu}_X \circ \bar{\rho} \circ d = \bar{\mu}_X \circ \beta - \bar{\mu}_X \circ d \circ \epsilon = \bar{\mu}_X \circ \beta - d \circ \bar{\mu} \circ \epsilon = \bar{\varphi} \circ d,$$

as required. □

Proof of Lemma 3.2. To minimize excess notation, we assume (without loss of generality) that $\underline{Y} = \underline{X}$.

Consider a morphism $(N', \varphi') \to (N, \varphi)$ of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$. We verify that the map $(R, \rho) \to (N, \varphi)$ factors in a unique way through (N', φ'). In Lemma 3.6, we have already seen that there is a unique map $\bar{R} \to \bar{N}'$, compatible with the maps from $\pi^* \bar{M}_S$, and a unique factorization of $\bar{\varphi} : \bar{M} \to \bar{N}'_X$ through $\bar{\rho} : \bar{M} \to \bar{R}_X$. In particular, the diagram

$$\begin{tikzcd} \bar{R} \arrow[r, "\bar{\varphi}'"] \arrow[dr, "\bar{\varphi}"'] & \bar{N}' \arrow[d] \\ & \bar{N} \end{tikzcd}$$

commutes. Since N' is pulled back from N by the vertical arrow in the diagram above, this gives a uniquely determined arrow $R \to N'$. Likewise, the diagrams of solid arrows below are commutative (the diagonal arrow on the left coming from Lemma 3.6):

$$\begin{tikzcd} \bar{M} \arrow[r] \arrow[d] & \bar{R}_X \arrow[d] \arrow[dl] \\ \bar{N}'_X \arrow[r] & \bar{N}_X \end{tikzcd} \qquad \begin{tikzcd} M \arrow[r] \arrow[d] & R_X \arrow[d] \arrow[dl, dashed] \\ N'_X \arrow[r] & N_X \end{tikzcd}$$

As N'_X is pulled back from N_X via the map $\bar{N}'_X \to \bar{N}_X$, there is a unique induced map $R_X \to N'_X$ completing the diagram on the right. This proves the minimality of (R, ρ). □

Proof of Lemma 3.3. We show that the tools used in the construction of R and ρ are all compatible with pullback of quasilogarithmic structures. Pullback of quasilogarithmic structures along a morphism $g : Y' \to Y$ involves two steps: pullback of étale sheaves along g followed by pushout of extensions along $g^{-1}\mathcal{O}_Y^* \to \mathcal{O}_{Y'}^*$. We verify that the construction of R and ρ commutes with these operations:

(1) $\bar{R}^{\mathrm{gp}}$ was constructed as a quotient of a fiber product, and both fiber products and quotients are preserved by pullback of étale sheaves;

(2) $\bar{\rho}$ was induced by the universal property of $\bar{M}$ as a quotient, and quotients are preserved by pullback of étale sheaves;

(3) $\bar{R}$ was built as the sheaf of submonoids of $\bar{R}^{\mathrm{gp}}$ generated by a collection of local sections, and this construction is compatible with pullback of étale sheaves;

(4) R was constructed as a pullback of an extension by $\mathcal{O}_Y^*$, and such pullbacks are preserved by pullback of étale sheaves and by pushout along $g^{-1}\mathcal{O}_Y^* \to \mathcal{O}_{Y'}^*$;

(5) ρ was induced by the universal property of a base change of extensions, and, as remarked above, base change of extensions is preserved by pullback of étale sheaves and pushout along the kernels. □

4. Global minimality

In this section, $X = (\underline{X}, M_X)$ and $S = (\underline{S}, M_S)$ will be fine logarithmic schemes. We assume that the projection $\pi : X \to S$ is proper and flat *with reduced geometric fibers* and that the morphism of logarithmic structures $\pi^* M_S \to M_X$ is integral. We also assume a second coherent logarithmic structure M on X has been specified, along with a morphism $\pi^* M_S \to M$. We will verify Gillam's minimality criterion (Proposition B.1) for the fibered category $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ over **LogSch**.

We continue to use Convention 3.1 to notate pushouts, as well as to work with quasilogarithmic structures instead of logarithmic structures. Define $\mathrm{GS}(S)$ to be the category of pairs (N_S, φ), where N_S is a quasilogarithmic structure on S equipped with a tacitly specified morphism $M_S \to N_S$ and φ fits into a commutative diagram

$$\begin{tikzcd} \pi^* M_S \arrow[r] \arrow[d] \arrow[rr, bend left] & \pi^* N_S \arrow[d] & M \arrow[dl, "\varphi"] \\ M_X \arrow[r] & N_X & \end{tikzcd}$$

whose square is cocartesian. By pullback of quasilogarithmic structures, we may assemble this definition into a fibered category over $\mathbf{LogSch}/S$. When T is strict over S, we write $\mathrm{GS}(\underline{T})$ instead of $\mathrm{GS}(T)$.

We may recognize $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)(T)$[7] inside of $\mathrm{GS}(T)$ as the category of pairs (N_T, φ), where N_T is a logarithmic, as opposed to merely quasilogarithmic, structure. We are free to work with $\mathrm{GS}(T)$ in place of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)(T)$, as minimal objects of the latter may be induced from minimal objects of the former by passage to the associated logarithmic structure.

[7] The notation $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ refers to the fibered category over $\mathbf{LogSch}/S$ whose value on a logarithmic scheme T over S is the set of morphisms of logarithmic structures $M|_T \to M_X|_T$ on $\underline{X} \times_{\underline{S}} \underline{T}$, compatible with the tacit morphisms from $\pi^* M_T$.

Lemma 4.1. *For any logarithmic scheme T over S, every object of* $\mathrm{GS}(T)$ *admits a morphism from a minimal object.*

Lemma 4.2. *If $f : T' \to T$ is a morphism of logarithmic schemes over S and (Q_T, ψ) is minimal in* $\mathrm{GS}(T)$, *then $f^*(Q_T, \psi)$ is minimal in* $\mathrm{GS}(T')$.

Note that passage to the associated logarithmic structure commutes with pullback of prelogarithmic structures [Kato 1989, (1.4.2)], so Lemma 4.2 implies that minimal objects of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)(\underline{T})$ pull back to minimal objects of $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)(\underline{T}')$.

The strategy of proof for Lemmas 4.1 and 4.2 will be to bootstrap from the minimal objects of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ constructed in Section 3. The essential tool in this construction is a left adjoint to pullback for étale sheaves, for whose construction in Section 4A we must assume X has reduced geometric fibers over S. Having dispensed with generalities in Section 4A, we take up the construction of minimal objects of $\mathrm{GS}(\underline{T})$ in Section 4B.

Zariski logarithmic structures. The following proposition will only be used in Appendix C. It shows that when M_S and M are Zariski logarithmic structures, one can replace M_X by its best approximation by a Zariski logarithmic structure for the purpose of constructing the category GS. At least in many situations, this means that one can work in the Zariski topology rather than the étale topology for the purpose of constructing a minimal logarithmic structure. See Appendix C for more details.

Lemma 4.3. *Let X be a scheme. Denote by $\tau : \mathrm{\acute{e}t}(X) \to \mathrm{zar}(X)$ the morphism of sites. Then $\mathrm{Hom}(\tau^* F, \tau^* G) = \mathrm{Hom}(F, G)$ for any sheaves F and G on $\mathrm{zar}(X)$. In particular, $\tau_* \tau^* \simeq \mathrm{id}$.*

Proof. By adjunction, we have $\mathrm{Hom}(\tau^* F, \tau^* G) = \mathrm{Hom}(F, \tau_* \tau^* G)$. But we can calculate that $\tau_* \tau^* G(U) = \tau^* G(U) = G(U)$ for any open $U \subset X$. The first equality is the definition; the second equality holds because, for example, the espace étalé of G (in the Zariski topology) is a scheme, hence satisfies étale descent. □

Proposition 4.4. *Suppose that M and M_S are Zariski logarithmic structures and that $\overline{M}_S^{\mathrm{gp}}$ is torsion-free. Let τ denote the canonical morphism from the étale site to the Zariski site. Define GS' to be the category obtained by imitating the definition of* GS, *with M_X replaced by $\tau^* \tau_* M_X$. Then $\mathrm{GS} \simeq \mathrm{GS}'$.*

Proof. By assumption, we have $M_S = \tau^* M'_S$ and $M = \tau^* M'$ for some logarithmic structures M'_S on S and M' on $\mathrm{zar}(X)$. Here τ^* denotes pullback of logarithmic structures.

We will begin by constructing a functor $\mathrm{GS}' \to \mathrm{GS}$. Observe that an object of GS' is a commutative diagram

$$\begin{tikzcd}
\pi^* M_S \arrow[r] \arrow[d] \arrow[rr, bend left] & \tau^*\tau_* M_X \arrow[d] & M \arrow[dl] \\
\pi^* N_S \arrow[r] & N'_X &
\end{tikzcd}$$

in which the square is cocartesian. Then composition with $\tau^*\tau_* M_X \to M_X$ induces

$$\begin{tikzcd}
\pi^* M_S \arrow[r] \arrow[d] \arrow[rrr, bend left] & \tau^*\tau_* M_X \arrow[r] \arrow[d] & M_X \arrow[d] & M \arrow[dl] \\
\pi^* N_S \arrow[r] & N'_X \arrow[r] & N_X &
\end{tikzcd}$$

and omitting $\tau^*\tau_* M_X$ and N'_X yields an object of GS.

Now we construct the functor $\mathrm{GS} \to \mathrm{GS}'$. Suppose that we have an object of GS,

$$\begin{tikzcd}
\pi^* M_S \arrow[r] \arrow[d] \arrow[rr, bend left] & M_X \arrow[d] & \tau^* M \arrow[dl] \\
\pi^* N_S \arrow[r] & N_X &
\end{tikzcd}$$

Applying τ_*, this gives the diagram

$$\begin{tikzcd}
\pi^*_{\mathrm{zar}} M'_S \arrow[r] \arrow[d] \arrow[rr, bend left] & \tau_* M_X \arrow[d] & M' \arrow[dl] \\
\tau_*\pi^* N_S \arrow[r] & \tau_* N_X &
\end{tikzcd} \tag{4-1}$$

Note that we have used $\tau_*\pi^* M_S = \tau_*\tau^*\pi^*_{\mathrm{zar}} M'_S = \pi^*_{\mathrm{zar}} M'_S$ and $\tau_* M = \tau_*\tau^* M' = M'$ by Lemma 4.3. The square in the diagram above is cocartesian. Indeed, first construct an exact sequence

$$0 \to \pi^* M_S^{\mathrm{gp}} \to (\pi^* N_S \times M_X)^{\sim} \to N_X \to 0,$$

where $(\pi^* N_S \times M_S)^{\sim}$ is the smallest submonoid of $\pi^* N_S^{\mathrm{gp}} \times M_X^{\mathrm{gp}}$ that contains $\pi^* N_S \times M_X$ and the image of $\pi^* M_S^{\mathrm{gp}}$. Applying τ_* to this gives an exact sequence

$$0 \to \tau_*\pi^* M_S^{\mathrm{gp}} \to \tau_*(\pi^* N_S \times M_X)^{\sim} \to \tau_* N_X \to R^1\tau_*\pi^* M_X^{\mathrm{gp}}. \tag{4-2}$$

The notation in the middle term is unambiguous because

$$\big(\tau_*(\pi^* N_S \times M_X)\big)^{\sim} = \tau_*\big((\pi^* N_S \times M_X)^{\sim}\big)$$

via the natural map. To see this, note first that it is sufficient to verify this at the level of characteristic monoids, since both sides are torsors under $\mathcal{O}_X^*$ over their characteristic monoids (the pushforward of an $\mathcal{O}_X^*$-torsor being an $\mathcal{O}_X^*$-torsor by Hilbert's Theorem 90). It is also sufficient to check this on stalks, so we may assume that X is the spectrum of a field, and in particular that $\bar{M}'_S$ is constant. An element of $\tau_*\big((\pi^*\bar{N}_S \times \bar{M}_X)^{\sim}\big)$ is then a section of $\pi^*\bar{N}_S^{\mathrm{gp}} \times \bar{M}_X^{\mathrm{gp}}$ that can be expressed as $x - y$ for some $x \in \pi^*\bar{N}_S \times \bar{M}_X$ and $y \in \bar{M}_S$ and is invariant under the action of the Galois group. An element of $\big(\tau_*(\pi^*\bar{N}_S \times \bar{M}_X)\big)^{\sim}$ is of the form $x - y$ where x is a Galois-invariant section of $\pi^*\bar{N}_S \times \bar{M}_X$ and y is a section of $\bar{M}'_S$. But the Galois action on $\bar{M}_S$ is trivial since $\bar{M}_S = \tau^*\bar{M}'_S$, so $x - y$ is Galois-invariant if and only if x is.

Now we show that $R^1\tau_*\pi^* M_S^{\mathrm{gp}} = 0$. We can verify this by passing to stalks and assume that X is the spectrum of a field. Note that $\pi^* M_S^{\mathrm{gp}}$ is an extension of a torsion-free abelian group by $\mathcal{O}_X^*$. We know that $R^1\tau_*\mathcal{O}_X^* = 0$ by Hilbert's Theorem 90, and $R^1\tau_*\pi^*\bar{M}_S^{\mathrm{gp}} = 0$ because we can identify it with homomorphisms from the Galois group, which is profinite, into the discrete, torsion-free abelian group $\bar{M}_S^{\mathrm{gp}}$.

Now the exact sequence (4-2) implies that the square in diagram (4-1) is cocartesian. Applying τ^* to diagram (4-1) we get an object of GS',

$$\begin{tikzcd}
\pi^* M_S \arrow[r] \arrow[d] \arrow[rr, bend left] & \tau^*\tau_* M_X \arrow[d] & M \arrow[dl] \arrow[dl, phantom] \arrow[ddl, bend left] \\
\tau^*\tau_*\pi^* N_S \arrow[r] \arrow[d] & \tau^*\tau_* N_X \arrow[d] & \\
\pi^* N_S \arrow[r] & N'_X &
\end{tikzcd}$$

Here N'_X is defined to make the bottom square cartesian. But pullback preserves cartesian diagrams, so both squares are cartesian and the outer part of the diagram is the desired object of GS'. We leave it to the reader to verify that these constructions are inverse to one another. □

4A. ***Left adjoint to pullback.*** In this section we prove that pullback for étale sheaves has a left adjoint under two natural hypotheses (flatness and local finite presentation) and one apparently unnatural one (reduced geometric fibers). When $f : X \to S$ is étale, the left adjoint to pullback exists for obvious reasons and is well-known: simply compose with the structure morphism of the espace étalé with f. The construction in the present generality appears to be new.

Our construction is based on the following observation: if F is an étale sheaf on X, write $F^{\text{ét}}$ for its espace étalé. When S is the spectrum of a separably closed field, $f_!F$ has no choice but to be the set of connected components of $F^{\text{ét}}$. In general, if the definition of $f_!$ is to be compatible with base change in S, this forces $f_!F$ to coincide with $\pi_0(F^{\text{ét}}/S)$, as defined by Laumon and Moret-Bailly [2000, Section 6.8] or Romagny [2011].

The results of [Laumon and Moret-Bailly 2000] and [Romagny 2011] guarantee the existence of $\pi_0(F^{\text{ét}}/S)$ as long as $F^{\text{ét}}$ is flat, is of finite presentation, and possesses reduced geometric fibers over S.[8] This suffices to treat a large enough class of étale sheaves to generate all others under colimits when X is merely locally of finite presentation over S. As $f_!$ must respect colimits where defined, we can extend the definition to any étale sheaf F by applying $f_!$ to a diagram of étale sheaves over X with colimit F and then taking the colimit of the resulting sheaves over S.

Flatness and local finite presentation appear to be natural hypotheses for the existence of $f_!$, in the sense that removing either leads immediately to counterexamples (the inclusion of a closed point or the spectrum of a local ring, respectively). It is less clear how essential it is to require reduced geometric fibers, as our construction makes use of that hypothesis only to ensure the existence of $\pi_0(F^{\text{ét}}/S)$ as an étale sheaf.

Gabber has argued (personal communication, 2014) that the natural condition on f under which f^* possesses a left adjoint is, in addition to suitable finiteness conditions, that the morphism $\widetilde{X} \to \widetilde{S}$ possess connected geometric fibers whenever $\widetilde{X}$ is the strict henselization of X at a geometric point x and $\widetilde{S}$ is the strict henselization of S at $f(x)$.

Theorem 4.5. *Suppose that $f : X \to S$ is a flat, local finite presentation morphism of algebraic spaces with reduced geometric fibers. Then the functor $f^* : \text{ét}(S) \to \text{ét}(X)$ on étale sheaves has a left adjoint, $f_!$.*

Proof. In this proof we will move freely between étale sheaves and their espaces étalés, which are algebraic spaces that are étale over the base. The étale site ét$(-)$ will be taken to mean the category of all étale *algebraic spaces* over the base, so that it coincides with the category of étale sheaves.

For any $U \in \text{ét}(X)$, we may define a functor

$$F_U : \text{ét}(S) \to \mathbf{Sets} : V \mapsto \mathrm{Hom}_S(U, V) = \mathrm{Hom}_X(U, f^*V).$$

[8] In fact, [Laumon and Moret-Bailly 2000] assumes that X is smooth over S, but, as we will see below, only flatness is necessary in the construction.

Consider the collection $\mathscr{C}$ of all $U \in \text{ét}(X)$ for which F_U is representable by an étale sheaf on S. The existence of $f_!$ is equivalent to the assertion that $\mathscr{C} = \text{ét}(X)$.

Step 1. We observe first that $\mathscr{C}$ is closed under colimits: suppose that $U = \varinjlim U_i$ and that F_{U_i} is representable by $f_! U_i$ for all i. Then we may take $f_! U = \varinjlim f_! U_i$:

$$\begin{aligned} F_U(V) &= \mathrm{Hom}_{\text{ét}(X)}\big(\varinjlim U_i, f^*V\big) \\ &= \varprojlim \mathrm{Hom}_{\text{ét}(X)}(U_i, f^*V) \\ &= \varprojlim \mathrm{Hom}_{\text{ét}(S)}(f_! U_i, V) \\ &= \mathrm{Hom}_{\text{ét}(S)}\big(\varinjlim f_! U_i, V\big). \end{aligned}$$

Every étale algebraic space over X is a colimit of étale algebraic spaces of finite presentation over S. For example, every étale algebraic space over X is a colimit of étale algebraic spaces that are affine over affine open subsets of S. Therefore, Step 1 implies that the construction of $f_! F$ for arbitrary étale sheaves F on X reduces to the construction for those representable by algebraic spaces of finite presentation over S.

Step 2. Following the construction of $\pi_0(X/S)$ from [Laumon and Moret-Bailly 2000, Section 6.8], we argue next that $\mathscr{C}$ contains all étale U over X that are of finite presentation over S. One could also use the construction of $\pi_0(X/S)$ from [Romagny 2011, Théorème 2.5.2(i)].

Suppose that U is flat, of finite presentation, and representable by schemes over S. By [EGA IV$_3$ 1966, Corollaire 15.6.5], there is an open subscheme $W \subset U \times_S U$ such that, for each point x of U, the open set $W \cap (\{x\} \times U) \subset U$ is the connected component of x in U. A field-valued point of $U \times_S U$ lies in W if and only if its two projections to U lie in the same connected component. Thus $W \subset U \times_S U$ is a flat equivalence relation on U, hence has a quotient $\pi_0(U/S) = U/W$ that is an algebraic space over S.

We verify that U/W is étale over S. It is certainly flat and locally of finite presentation since U is. It is therefore enough to verify that it is formally unramified. This condition can be verified after base change to the geometric points of S. As the definition of W commutes with base change, so does the quotient U/W. We can therefore assume S is the spectrum of a separably closed field, and then $U/W = \pi_0(U/S) = \pi_0(U)$ is simply the set of connected components of U, which is certainly unramified over S.

Now we show that $\pi_0(U/S)$ represents F_U. If $g : U \to V$ is a morphism from U to an étale S-scheme V then the preimages of points of V are open and closed in their fibers over S (since V has discrete fibers over S). Therefore, $U \times_V U$ contains W (viewing both as open subschemes of $U \times_S U$), so f factors through U/W. □

Corollary 4.5.1. *Let $f : X \to S$ be as in the statement of the theorem and let $f' : X' \to S'$ be deduced by base change via a morphism $g : S' \to S$. Then the natural morphism $f'_! g^* \to g^* f_!$ is an isomorphism.*

Proof. Since a morphism of étale sheaves is an isomorphism if and only if it is an isomorphism on stalks, it is sufficient to verify the assertion upon base change to all geometric points of S' and therefore to assume that S' is itself a geometric point. Since every étale sheaf is a colimit of representable étale sheaves that are of finite presentation over S (as in the proof of Theorem 4.5), it is sufficient to show that

$$f'_! g^* F \to g^* f_! F$$

when F is representable by a scheme that is of finite presentation over S. In that case, $f_! F = \pi_0(F^{\text{ét}}/S)$ and $f'_! g^* F = \pi_0(F^{\text{ét}} \times_S S'/S')$. But the fiber of $\pi_0(F^{\text{ét}}/S)$ over S' is $\pi_0(F^{\text{ét}} \times_S S'/S')$ by definition! □

The following proposition is well-known and included only for completeness.

Proposition 4.6. *Let X be a site. The inclusion of sheaves of abelian groups* (*resp. sheaves of commutative monoids*) *on X in sheaves of sets on X admits a left adjoint $F \mapsto \mathbb{Z}F$* (*resp. $F \mapsto \mathbb{N}F$*).

Proof. The proofs for abelian groups and for commutative monoids are identical, so we only write the proof explicitly for abelian groups.

It is equivalent to demonstrate that, for any sheaf of sets F on X there is an initial pair (G, φ), where G is a sheaf of abelian groups and $\varphi : F \to G$ is a morphism of sheaves of sets. Denote the category of pairs (G, φ) by $\mathscr{C}$. By the adjoint functor theorem, $\mathscr{C}$ has an initial object if it is closed under small limits and has an essentially small coinitial subcategory [Mac Lane 1998, Theorem X.2.1]. Closure under small limits is immediate.

For the essentially small coinitial subcategory, take the collection $\mathscr{C}_0$ of all (G, φ) such that $\varphi(F)$ generates G as a sheaf of abelian groups (i.e., the smallest subsheaf of abelian groups $G' \subset G$ that contains $\varphi(F)$ is G itself). The cardinalities of $G'(U)$ for all U in a set of topological generators may be bounded in terms of the cardinalities of the $F(U)$. It follows that $\mathscr{C}_0$ is essentially small and by the adjoint functor theorem that the inclusion of sheaves of abelian groups in sheaves of sets has a left adjoint. □

Proposition 4.7. *Let $f : X \to S$ be flat and locally of finite presentation with reduced geometric fibers. The functor f^* on sheaves of abelian groups* (*resp. sheaves of commutative monoids*) *has a left adjoint*, *denoted $f_!$.*

Proof. The proof is the same as the proof of the theorem above. We recognize that the class of sheaves of abelian groups (resp. sheaves of commutative monoids) F

for which $f_! F$ exists is closed under colimits. It contains $\mathbb{Z}U$ (resp. $\mathbb{N}U$) for all étale U over X since we may take

$$f_!(\mathbb{Z}U) = \mathbb{Z} f_!(U) \qquad (\text{resp. } f_!(\mathbb{N}U) = \mathbb{N} f_!(U)).$$

Finally, all sheaves of abelian groups are colimits of diagrams of $\mathbb{Z}U$ (resp. $\mathbb{N}U$) as above, so $f_!$ is defined for all sheaves of abelian groups on X. □

Zariski sheaves. We include a statement about the left adjoint to pullback on sheaves in the Zariski topology in a restricted situation when it agrees with the left adjoint on étale sheaves. In practice, this can be used to compute the left adjoint on étale sheaves by working in the Zariski topology, as in Appendix C.

Proposition 4.8. *Let S be the spectrum of an algebraically closed field and let $f : X \to S$ be a reduced, finite-type scheme over S. Then $f^* : \mathrm{zar}(S) \to \mathrm{zar}(X)$ has a left adjoint, given by $f_! \tau^*$, where $f_!$ denotes the left adjoint on étale sheaves.*

Proof. As in Lemma 4.3, we write τ for the morphism from the étale site to the Zariski site. We have

$$\mathrm{Hom}(f_! \tau^* F, G) = \mathrm{Hom}(\tau^* F, f^*_{\text{ét}} G) = \mathrm{Hom}(\tau^* F, \tau^* f^*_{\mathrm{zar}} G) = \mathrm{Hom}(F, f^*_{\mathrm{zar}} G),$$

as required. □

4B. ***Reduction to the local problem.*** When (N_S, φ) is an object of $\mathrm{GS}(\underline{S})$, the pair $(\pi^* N_S, \varphi)$ is an object of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$. This determines a functor $\mathrm{GS}(\underline{S}) \to \mathrm{GS}^{\mathrm{loc}}(\underline{X})$, and when $\pi : \underline{X} \to \underline{S}$ is an isomorphism this functor induces an equivalence between $\mathrm{GS}(\underline{S})$ and $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$.

Before giving the proof of Lemma 4.1, we explain the construction. In order to minimize notation, we construct minimal objects of $\mathrm{GS}(\underline{S})$; the same construction applies to $\mathrm{GS}(\underline{T})$ for any $\underline{S}$-scheme $\underline{T}$ after pulling back the relevant data.

We suppose that (N_S, φ) is an object of $\mathrm{GS}(\underline{S})$ and we construct a pair (Q_S, ψ) in $\mathrm{GS}(\underline{S})$ and a morphism $(Q_S, \psi) \to (N_S, \varphi)$. Then we show that (Q_S, ψ) is minimal in $\mathrm{GS}(\underline{S})$ and that the construction of (Q_S, ψ) is stable under pullback.

Let $(R, \rho) \to (\pi^* N_S, \varphi)$ be a morphism from a minimal object of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$, as guaranteed by Lemma 3.2.

Construction of Q_S. Define $\bar{Q}_S$ to be the pushout of the left half of the diagram

$$\begin{tikzcd}
\pi_! \pi^* \bar{M}_S \arrow[r] \arrow[d] & \pi_! \bar{R} \arrow[r] \arrow[d] & \pi_! \pi^* \bar{N}_S \arrow[d] \\
\bar{M}_S \arrow[r] \arrow[rr, bend right] & \bar{Q}_S \arrow[r, dashed] & \bar{N}_S
\end{tikzcd} \tag{4-3}$$

of étale sheaves of monoids. The commutativity of the diagram of solid lines and the universal property of pushout induce a morphism $\bar{Q}_S \to \bar{N}_S$, as shown above. Pulling back N_S via this map gives a quasilogarithmic structure Q_S with characteristic monoid $\bar{Q}_S$.

Remark 4.9. The construction of $\bar{Q}_S$ can be simplified when $\pi : X \to S$ has connected fibers. Under that hypothesis, the counit $\pi_! \pi^* \bar{M}_S \to \bar{M}_S$ is an isomorphism and $\bar{Q}_S = \pi_! \bar{R}$.

Construction of ψ. Pulling diagram (4-3) back to X, we get a commutative diagram

$$\begin{tikzcd}
\pi^* \bar{M}_S \arrow[r] \arrow[d] \arrow[dd, bend right, "\mathrm{id}"'] & \bar{R} \arrow[r] \arrow[d] & \pi^* \bar{N}_S \arrow[d] \arrow[dd, bend left, "\mathrm{id}"] \\
\pi^* \pi_! \pi^* \bar{M}_S \arrow[r] \arrow[d] & \pi^* \pi_! \bar{R}_S \arrow[r] \arrow[d] & \pi^* \pi_! \pi^* \bar{N}_S \arrow[d] \\
\pi^* \bar{M}_S \arrow[r] & \pi^* \bar{Q}_S \arrow[r] & \pi^* \bar{N}_S
\end{tikzcd} \tag{4-4}$$

of étale sheaves of monoids. The vertical arrows on the left and right sides of the diagram compose to identities because $\pi_!$ and π^* are adjoint functors. We isolate the commutative diagram

$$\begin{tikzcd}
\pi^* \bar{M}_S \arrow[r] \arrow[dr] & \bar{R} \arrow[r] \arrow[d] & \pi^* \bar{N}_S \\
& \pi^* \bar{Q}_S \arrow[ur] &
\end{tikzcd}$$

and push it out via $\pi^* \bar{M}_S \to \bar{M}_X$ to obtain the lower half of the diagram

$$\begin{tikzcd}
& \bar{M} \arrow[d, "\bar{\rho}"'] \arrow[dr, "\bar{\varphi}"] & \\
\bar{M}_X \arrow[r] \arrow[dr] & \bar{R}_X \arrow[r] \arrow[d] & \bar{N}_X \\
& \bar{Q}_X \arrow[ur] &
\end{tikzcd} \tag{4-5}$$

The upper half of the diagram is provided by the morphism $(R, \rho) \to (\pi^* N_S, \psi)$ of $\mathrm{GS}^{\mathrm{loc}}(\bar{X})$. By composing the vertical arrows in the center of the diagram, we obtain the definition of $\bar{\psi}$:

$$\bar{\psi} : \bar{M} \xrightarrow{\bar{\rho}} \bar{R}_X \longrightarrow \bar{Q}_X.$$

Now observe that Q_X is the pullback of N_X via the map $\bar{Q}_X \to \bar{N}_X$. The commutative triangle

$$\begin{tikzcd} \bar{M} \arrow[d, "\bar{\psi}"'] \arrow[dr, "\bar{\varphi}"] & \\ \bar{Q}_X \arrow[r] & \bar{N}_X \end{tikzcd}$$

yields a commutative triangle

$$\begin{tikzcd} M \arrow[d, dashed, "\psi"'] \arrow[dr, "\varphi"] & \\ Q_X \arrow[r] & N_X \end{tikzcd}$$

by the universal property of the fiber product.

Proof of Lemma 4.1. We check that the object (Q_S, ψ) constructed above satisfies the universal property of a minimal object of $\mathrm{GS}(\underline{S})$.

Consider a morphism $(N'_S, \varphi') \to (N_S, \varphi)$ of $\mathrm{GS}(\underline{S})$. We must show that there is a unique map $(Q_S, \psi) \to (N'_S, \varphi')$ rendering the triangle

$$\begin{tikzcd} (Q_S, \psi) \arrow[r, dashed] \arrow[dr] & (N'_S, \varphi') \arrow[d] \\ & (N_S, \varphi) \end{tikzcd} \tag{4-6}$$

commutative. To specify the dashed arrow above we must give a factorization of $Q_S \to N_S$ through N'_S and show that the induced triangle

$$\begin{tikzcd} & M \arrow[dl, "\psi"'] \arrow[d, "\varphi'"] \\ Q_X \arrow[r] & N'_X \end{tikzcd} \tag{4-7}$$

is commutative.

The dashed arrow in diagram (4-6) and the commutativity of diagram (4-7) are both determined at the level of characteristic monoids. That is, to give a factorization of $Q_S \to N_S$ through N'_S is the same as to give a factorization of $\bar{Q}_S \to \bar{N}_S$ through $\bar{N}'_S$ since the logarithmic structure N'_S is pulled back from N_S via $\bar{N}'_S \to \bar{N}_S$. Similarly, to verify the commutativity of diagram (4-7) it is sufficient to show that the induced diagram of characteristic monoids commutes.

By the definition of $\bar{Q}_S$ as a pushout in diagram (4-3), to give $\bar{Q}_S \to \bar{N}'_S$ compatible with the tacit maps from $\bar{M}_S$ is the same as to give $\pi_! \bar{R} \to \bar{N}'_S$ compatible with the maps from $\pi_! \pi^* \bar{M}_S$. The latter is equivalent, by adjunction, to giving

$\bar{R} \to \pi^* \bar{N}'_S$ compatible with the maps from $\pi^* \bar{M}_S$. But by the minimality of (R, ρ) in $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$, there is a unique factorization of $R \to \pi^* N_S$ through $\pi^* N'_S$ such that the induced triangle depicted in the upper half of the diagram

$$\begin{tikzcd}
M \arrow[d, "\rho"] \arrow[dr, "\varphi'"] \arrow[dd, bend right, "\psi"'] & \\
R_X \arrow[r] \arrow[d] & N'_X \\
Q_X \arrow[ur] &
\end{tikzcd}$$

is commutative. The lower triangle in the diagram is also commutative: as already noted, the map $R \to \pi^* N'_S$ gives a factorization of $\pi_! \bar{R} \to \bar{N}'_S$ through $\bar{Q}_S$, so by adjunction we get a factorization of $\bar{R} \to \pi^* \bar{N}'_S$ through $\pi^* \bar{Q}_S$, and therefore a factorization $R \to \pi^* \bar{Q}_S \to \bar{N}'_S$. The lower triangle is the pushout of this sequence via $R \to R_X$. The outer triangle, which coincides with diagram (4-7), is therefore commutative, as desired. □

Proof of Lemma 4.2. It is sufficient to treat the case where $T = S$. Suppose, then, that $f : \underline{S}' \to \underline{S}$ is a morphism of schemes and set $\underline{X}' = \underline{X} \times_{\underline{S}} \underline{S}'$. We show that $(f^* Q_S, f^* \psi)$ is a minimal object of $\mathrm{GS}(S')$. We verify that all of the data that go into the construction of Q_S are compatible with pullback:

(1) the minimal object (R, ρ) of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ pulls back to a minimal object of $\mathrm{GS}^{\mathrm{loc}}(X')$ by Lemma 3.3;

(2) the formation of $\pi_! \bar{R}$ is compatible with pullback by Corollary 4.5.1;

(3) the pushout $\bar{Q}_S$ is compatible with pullback by the preservation of colimits under pullback of sheaves;

(4) the quasilogarithmic structure Q_S is formed as a fiber product of sheaves and these are preserved by pullback.

This shows that the construction of Q_S is compatible with pullback. We make a similar verification for ψ:

(5) the compatibility of $\pi_!$ with f^* (Corollary 4.5.1) guarantees that diagram (4-4) pulls back to its analogue on X';

(6) compatibility of pullback of sheaves with colimits guarantees the compatibility of the lower half of diagram (4-5) with pullback;

(7) Lemma 3.3 and the assumption that (R, ρ) be minimal in $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$ guarantee that the pullback of the upper half of diagram (4-5) coincides with its analogue constructed in $\mathrm{GS}^{\mathrm{loc}}(X')$.

This shows that the construction of ψ is compatible with pullback and completes the proof. $\square$

5. Automorphisms of minimal logarithmic structures

Lemma 5.1. *Suppose that (R, ρ) is a minimal object of $\mathrm{GS}^{\mathrm{loc}}(\underline{X})$. Then the automorphism group of (R, ρ) is trivial.*

Proof. It is equivalent to show that only the identity automorphism of R^{gp} is compatible with both the inclusion of $\pi^* M_S^{\mathrm{gp}}$ and the map $\rho : M^{\mathrm{gp}} \to R_X^{\mathrm{gp}}$. Suppose that α is an automorphism of (R, ρ). Then the induced morphism $\bar{\alpha} : \bar{R} \to \bar{R}$ is the identity, by Lemma 3.6.

We use this to conclude that $\alpha = \mathrm{id}$. As $\bar{\alpha} = \mathrm{id}$, the automorphism α is determined by a homomorphism $\lambda : \bar{R}^{\mathrm{gp}} \to \mathcal{O}_X^*$ with

$$\alpha(x) = x + \log \lambda(\bar{x})$$

and $\bar{x}$ denoting the image of x under the projection $R \to \bar{R}$. By assumption, α restricts to the identity on $\pi^* \bar{M}_S^{\mathrm{gp}} \subset \bar{R}^{\mathrm{gp}}$ so λ restricts to the trivial homomorphism on $\pi^* \bar{M}_S^{\mathrm{gp}}$. Therefore, it must factor through the quotient $\bar{M}^{\mathrm{gp}} / \pi^* \bar{M}_S^{\mathrm{gp}}$ of $\bar{R}^{\mathrm{gp}}$ by $\pi^* \bar{M}_S^{\mathrm{gp}}$.

Let $\alpha_X : R_X \to R_X$ denote the automorphism induced by pushout of α. We investigate the condition that α_X commute with ρ in terms of λ and show that this implies $\lambda = 1$. Define λ_X by the formula $\log \lambda_X = \alpha_X - \mathrm{id}$ and note that $\lambda_X : \bar{R}_X^{\mathrm{gp}} \to \mathcal{O}_X^*$ is induced by pushout from the pair of morphisms $\lambda : \bar{R}^{\mathrm{gp}} \to \mathcal{O}_X^*$ and $1 : \bar{M}_X^{\mathrm{gp}} \to \mathcal{O}_X^*$. This implies $\lambda_X \circ q = 1$ and $\lambda_X \circ i = \lambda$ in the notation of the diagram

$$\begin{tikzcd}
0 \arrow[r] & \pi^* \bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bar{R}_0^{\mathrm{gp}} \arrow[r, "r"] \arrow[d, "q"] \arrow[dl, dashed, "\epsilon"'] & \bar{M}^{\mathrm{gp}} \arrow[r] \arrow[d] \arrow[dl, dashed, "\bar{\rho}"'] & 0 \\
0 \arrow[r] & \bar{R}^{\mathrm{gp}} \arrow[r, "i"] & \bar{R}_X^{\mathrm{gp}} \arrow[r] & \bar{M}_{X/S}^{\mathrm{gp}} \arrow[r] & 0
\end{tikzcd}$$

That $\lambda_X \circ q = 1$ follows from the fact that q factors through $\bar{M}_X^{\mathrm{gp}} \to \bar{R}_X^{\mathrm{gp}}$.

For any $y \in R_0^{\mathrm{gp}}$, let $r(y)$ be its image in M^{gp}. Then we have

$$\begin{aligned}
0 = \alpha_X(\rho r(y)) - \rho r(y) &= \log \lambda_X \bar{\rho} \bar{r}(\bar{y}) \\
&= \log \lambda_X (q(\bar{y}) - i\epsilon(\bar{y})) \\
&= -\log \lambda_X (i\epsilon(\bar{y})) \\
&= -\log \lambda(\epsilon \bar{y}).
\end{aligned}$$

As $\epsilon : R_0^{\mathrm{gp}} \to R^{\mathrm{gp}}$ is surjective, we conclude that $\alpha - \mathrm{id} = \log \lambda = 0$, so $\alpha = \mathrm{id}$. $\square$

Corollary 5.1.1. *The automorphism group of a minimal object of* $\mathrm{GS}(\underline{S})$ *is trivial.*

Proof. Suppose that α is an automorphism of a minimal object (Q_S, ψ). Then $\pi^*\alpha$ is an automorphism of the object $(\pi^* Q_S, \psi) \in \mathrm{GS}^{\mathrm{loc}}(\underline{X})$. Choose a map $\iota : (R, \rho) \to (\pi^* Q_S, \psi)$ with (R, ρ) minimal. Then, by the definition of minimality, there is a commutative diagram

$$\begin{tikzcd} (R, \rho) \arrow[r, dashed] \arrow[d, "\iota"'] & (R, \rho) \arrow[d, "\iota"] \\ (\pi^* Q_S, \psi) \arrow[r, "\pi^*\alpha"] & (\pi^* Q_S, \psi) \end{tikzcd}$$

The dashed arrow is an automorphism of (R, ρ), hence must be the identity by the lemma. Therefore, by adjunction, the diagram

$$\begin{tikzcd} & \pi_! \bar{R}^{\mathrm{gp}} \arrow[dl] \arrow[dr] & \\ \bar{Q}_S^{\mathrm{gp}} \arrow[rr, "\bar{\alpha}"] & & \bar{Q}_S^{\mathrm{gp}} \end{tikzcd}$$

must commute. But, by definition, $\bar{Q}_S^{\mathrm{gp}}$ is generated by $\pi_! \bar{R}^{\mathrm{gp}}$ and $\bar{M}_S^{\mathrm{gp}}$ (diagram (4-3) and the subsequent discussion). By assumption, α commutes with the map $\bar{M}_S^{\mathrm{gp}} \to Q_S^{\mathrm{gp}}$ and we have just shown it commutes with the map $\pi_! \bar{R}^{\mathrm{gp}}$. It follows that $\bar{\alpha}$ is the identity map.

This implies that α must be induced from a map $\delta : \bar{Q}_S^{\mathrm{gp}} \to \mathcal{O}_S^*$, which we would like to show is trivial. This map is induced from a map $\bar{Q}_S^{\mathrm{gp}} / \bar{M}_S^{\mathrm{gp}} \to \mathcal{O}_S^*$ since δ is trivial on $\bar{M}_S^{\mathrm{gp}}$. Since $\pi_!$ preserves colimits (it is a left adjoint), we use the pushout in the left half of diagram (4-3) to make identifications:

$$\bar{Q}_S^{\mathrm{gp}} / \bar{M}_S^{\mathrm{gp}} \simeq \pi_! \bar{R}^{\mathrm{gp}} / \pi_! \pi^* \bar{M}_S^{\mathrm{gp}} \simeq \pi_! \big(\bar{R}^{\mathrm{gp}} / \pi^* \bar{M}_S^{\mathrm{gp}} \big).$$

By adjunction, δ is therefore induced from a map

$$\tilde{\delta} : \bar{R}^{\mathrm{gp}} / \pi^* \bar{M}_S^{\mathrm{gp}} \to \pi^* \mathcal{O}_S^*. \tag{5-1}$$

But composing this with the map $\pi^* \mathcal{O}_S^* \to \mathcal{O}_X^*$ gives an automorphism of (R, ρ), which must be trivial, by the lemma. On the other hand, $\pi^* \mathcal{O}_S^*$ injects into $\mathcal{O}_X$ since X is flat over S, so the map $\tilde{\delta}$ — and by adjunction also δ — must be trivial. □

Corollary 5.1.2. *The functor* $\mathrm{Hom}_{\mathbf{LogSch}/S}(M, M_X)$ *is representable by a logarithmic algebraic space.*

Corollary 5.1.3. *The morphism* $\mathrm{Hom}_{\mathbf{LogSch}/S}(X, Y) \to \mathrm{Hom}_{\mathbf{LogSch}/S}(\underline{X}, \underline{Y})$ *is representable by logarithmic algebraic spaces.*

Appendix A: Integral morphisms of monoids

Recall that a morphism of monoids $f : P \to Q$ (written additively) is said to be *integral* if, for any $a, a' \in P$ and $b, b' \in Q$ such that

$$f(a) + b = f(a') + b',$$

there are elements $c, c' \in P$ and $d \in Q$ such that $a + c = a' + c'$ and $b = d + f(c)$ and $b' = d + f(c')$.

Continue to assume that $f : P \to Q$ is integral and let $P \to P'$ be another morphism of monoids. Consider the collection of pairs (a, b) where $a \in P'$ and $b \in Q$, modulo the relation $(a, b) \sim (a', b')$ if there are elements $c, c' \in P$ and $d \in Q$ such that $a + c = a' + c'$ and $b = d + f(c)$ and $b' = d + f(c')$.

The proofs of the following two lemmas are omitted as they are straightforward and likely well-known.

Lemma A.1. *If $f : P \to Q$ is integral then the relation defined above is an equivalence relation.*

Lemma A.2. *The monoid structure on $P' \times Q$ descends to the equivalence classes of the relation defined above and realizes the pushout of $P \to Q$ via $P \to P'$.*

Appendix B: Minimality

Gillam's criteria. This section is included only for convenience. All of the results here may be found in greater detail in [Gillam 2012]. Our Proposition B.1, below, is Descent Lemma 1.3 of op. cit.

Since our only application of this section is to the fibered category of logarithmic schemes, **LogSch**, over the category of schemes, **Sch**, we have not made any attempt to state the results below in their natural generality. The reader who is interested in that level of generality may easily verify that all of the arguments below are valid for an arbitrary fibered category.

Let **Sch** denote the category of schemes and let **LogSch** denote the category of logarithmic schemes. Note that **LogSch** is an étale stack (not fibered in groupoids) over **Sch**. Recall that a logarithmic structure on a fibered category F over **Sch** is a cartesian section of **LogSch** over F.

Suppose that F is a fibered category over **Sch** with a logarithmic structure $M : F \to$ **LogSch**. There is an induced fibered category $\mathfrak{L}(F, M)$ over **LogSch**: the objects of $\mathfrak{L}(F, M)$ are pairs (η, f), where $\eta \in F$ and where $f : Y \to M(\eta)$ is a morphism in **LogSch** such that $\underline{Y} = \underline{M(\eta)}$ and f lies above the identity morphism.

When F is representable by a logarithmic scheme, this is the familiar construction that associates to F the functor it represents on logarithmic schemes.

We give a characterization of the fibered categories G on **LogSch** that are equivalent to $\mathfrak{L}(F, M)$ for a fibered category F with logarithmic structure M.

An object ξ of G is called *minimal* if every diagram

$$\begin{tikzcd} \eta \arrow[rr] \arrow[dr] & & \xi \\ & \omega \arrow[ur, dashed] & \end{tikzcd}$$

of solid lines in $G(\underline{\xi})$ lying above $\mathrm{id}_{\underline{\xi}}$ admits a unique completion by a dashed arrow.

Proposition B.1. *A fibered category G over* **LogSch** *is equivalent to $\mathfrak{L}(F, M)$ if and only if the following two conditions hold*:

(1) *for every $\eta \in G$ there is a minimal object $\xi \in G(\underline{\eta})$ and a morphism $\eta \to \xi$ lying above the identity of $\underline{\eta}$, and*

(2) *the pullback of a minimal object of G via a morphism of* **Sch** *is minimal.*

Proof. Certainly $\mathfrak{L}(F, M)$ has this property. The minimal object associated to (α, f) is $(\alpha, \mathrm{id}_{\underline{\alpha}})$.

Conversely, let F be the full subcategory of minimal objects of G. By assumption, this is a fibered category over **Sch** with a map $M : F \to$ **Log Sch** by composition of the inclusion $F \subset G$ with the projection $G \to$ **LogSch**. This is cartesian over **Sch** because F is cartesian in G over **Sch** and $G \to$ **LogSch** is cartesian over **LogSch**, hence over **Sch**.

We have a functor $\mathfrak{L}(F, M) \to G$ sending (α, f) to $f^*\alpha$. We verify that this is an equivalence. As this is a cartesian functor between fibered categories, the verification can be done fiberwise over **Sch**. That is, it is enough to show that the functors $\mathfrak{L}(F, M)(S) \to G(S)$ are equivalences for all schemes S.

But $G(S)$ may be identified with

$$\coprod_{\eta \in F(S)} G(S)/\eta \simeq \coprod_{\eta \in F(S)} \mathbf{LogSch}/M(\eta) \simeq \mathfrak{L}(F, M)(S). \qquad \square$$

Openness of minimality. Unlike the previous section, this section is specific to logarithmic structures.

The proof of Proposition B.1 shows that G is $\mathfrak{L}(F, M)$, where $F \subset G$ is the substack of minimal objects. The next proposition shows that when F and G are fibered over *coherent* logarithmic schemes (in other words, when minimal objects are coherent), F is an *open* substack of G.

For any $\xi \in F$, the image of ξ in **LogSch** is a logarithmic scheme S. We refer to the logarithmic structure on S also as the logarithmic structure on ξ.

Theorem B.2. *Suppose that the logarithmic structure on each $\xi \in F$ is coherent. Then $F \subset$* **Log**(G) *is open.*

Proof. We must show that, for any $\eta \in \mathbf{Log}(G)$ lying above a scheme S, the locus in S where η is minimal is open in S. Let ξ be the minimal object admitting a morphism from η and let $M_\eta \to M_\xi$ be the associated morphism of logarithmic structures. The locus in S where η is minimal is the same as the locus where $\eta \to \xi$ restricts to an isomorphism. Since G is fibered in groupoids over **LogSch**, the map $\eta \to \xi$ restricts to an isomorphism if and only if $M_\xi \to M_\eta$ does. The following lemma therefore completes the proof. □

Lemma B.3. *Let $\alpha : M \to M'$ be a morphism of coherent logarithmic structures on a scheme S. The locus in S where α is an isomorphism is an open subset of S.*

Proof. It is sufficient to show that a morphism that is an isomorphism at a geometric point is an isomorphism in an étale neighborhood of that point. Choose charts $\bar{P} \to \bar{M}$ and $\bar{P}' \to \bar{M}'$ near a geometric point ξ,[9] fitting into a commutative diagram

$$\begin{tikzcd} \bar{P} \arrow[r] \arrow[d] & \bar{M} \arrow[r] \arrow[d] & \bar{M}_\xi \arrow[d] \\ \bar{P}' \arrow[r] & \bar{M}' \arrow[r] & \bar{M}'_\xi \end{tikzcd}$$

As the monoids $\bar{M}_\xi$ and $\bar{M}'_\xi$ are of finite presentation (because they are finitely generated [Rosales and García-Sánchez 1999, Theorem 5.12]), we can choose $\bar{P}$ and $\bar{P}'$ so that the maps $\bar{P} \to \bar{M}_\xi$ and $\bar{P}' \to \bar{M}'_\xi$ are isomorphisms, at least after shrinking the étale neighborhood of ξ. But then $\bar{P} \to \bar{P}'$ is an isomorphism and M and M' are therefore the logarithmic structures associated to the same quasilogarithmic structure. In particular, they are isomorphic. □

Appendix C: Explicit formulas, by Sam Molcho

The purpose of this appendix is to show that under certain reasonable simplifying assumptions, the minimal logarithmic structure constructed in the paper admits a rather concrete, simple description. Specifically, we will study minimal logarithmic structures in the case where we have a family of maps

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & V \arrow[dl] \\ S & \end{tikzcd} \tag{C-1}$$

[9] By a chart $\bar{P} \to \bar{M}$ we mean a homomorphism from a constant sheaf of monoids $\bar{P}$ to $\bar{M}$ such that, if P is defined to be the extension of $\bar{P}$ by $\mathcal{O}_S^*$ obtained by pulling back $M \to \bar{M}$, the associated logarithmic structure of $P \to M \to \mathcal{O}_S$ is M.

over a geometric point $S = \operatorname{Spec} k$, with π being log smooth and flat and having reduced fibers, and where the logarithmic structure on V is a Zariski log structure. To our knowledge, these assumptions hold in all previous work where minimal logarithmic structures have been studied, e.g., [Gross and Siebert 2013; Abramovich and Chen 2014; Chen 2014; Olsson 2008; Ascher and Molcho 2015]. In fact, in these papers the authors always begin with flat, proper families of schemes with reduced fibers, and construct a minimal logarithmic structure over each geometric point by writing down an explicit formula. They then prove that a logarithmic structure is minimal over a general family if and only if it restricts to a minimal logarithmic structure over each point. In the cited papers, what are actually considered are "absolute" families

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & V \\ S & \end{tikzcd}$$

We will show that in this absolute situation the notion of minimality defined in the present paper and the notion given in [Gross and Siebert 2013] (or, in fact, an evident extension of this notion) coincide.[10]

Structure of logarithmically smooth morphisms. Take a logarithmically smooth morphism $\pi : (X, M_X) \to (S, M_S)$ which is flat, proper, and has reduced fibers, with $S = \operatorname{Spec} k$ a geometric point. The geometry of (X, M_X) may be rather complicated; however, it is simple to understand at the loci where the relative characteristic $\bar{M}_{X/S}$ has rank 0 or 1. Specifically, we have:

Theorem C.1. *Suppose $\pi : (X, M_X) \to (S, M_S)$ is flat, proper, and has reduced fibers, and that $S = \operatorname{Spec} k$ is a geometric point. Then:*

(1) *Each irreducible component of X is smooth near its generic point η, and $\bar{M}_{X,\eta} = \bar{M}_S$.*

(2) *Whenever two irreducible components intersect, they intersect in a divisor of each, which we will call a node. A node then, generically, is the intersection of* ***precisely two*** *irreducible components, and, if q denotes the generic point of an irreducible component of the node, we have $M_{X,q} = M_S \oplus_{\mathbb{N}} \mathbb{N}^2$, where $\mathbb{N} \to \mathbb{N}^2$ is the diagonal map and $\rho_q : \mathbb{N} \to M_S$ is some homomorphism determined by π.*

(3) *There are certain divisors on the smooth locus of X such that, at the generic point p of such a divisor, we have $M_{X,p} = M_S \oplus \mathbb{N}$.*

[10] The formula of [Gross and Siebert 2013] was only presented for families of nodal curves, but it works for more general X and the agreement we prove here holds in that generality.

The divisors in (3) are the higher-dimensional analogue of marked points of logarithmic curves. In other words, the structure of a logarithmic morphism on the generic points of codimension-0 and codimension-1 strata is precisely the same as the structure of a logarithmic curve, as discussed by F. Kato. To see why these claims are true, we apply the chart criterion for logarithmic smoothness [Kato 1989] to obtain étale locally a commutative diagram

$$\begin{array}{ccccc} X & \longrightarrow & S \times_{\operatorname{Spec}\mathbb{Z}[Q]} \operatorname{Spec}\mathbb{Z}[P] & \longrightarrow & \operatorname{Spec}\mathbb{Z}[P] \\ & & \downarrow & & \downarrow \\ & & S & \longrightarrow & \operatorname{Spec}\mathbb{Z}[Q] \end{array}$$

in which the square is cartesian and the morphism from (X, M_X) to the fiber product with its induced logarithmic structure is smooth on the level of underlying schemes and strict. Since smooth morphisms preserve the information of how irreducible components intersect, the problem reduces to proving these claims for the fiber of a toric morphism of toric varieties $X(F, N) \to X(\kappa, Q)$ over the torus fixed point of $X(\kappa, Q)$. Here we use the notation $X(F, N)$ for the toric variety determined by a fan F in a lattice N, and similarly for $X(\kappa, Q)$, where we may assume κ is a single cone. A generic component of the fiber over the fixed point of $X(\kappa)$ then corresponds to a cone $\tau \in F$ such that τ maps isomorphically to κ, and a divisor in the fiber corresponds to a cone σ that maps onto κ with relative dimension 1. Condition (1) then is equivalent to saying that $\tau \cap N$ is isomorphic to $\kappa \cap Q$, since the duals of these monoids are charts for the logarithmic structures $M_{X,\eta}$ and M_S. Conditions (2) and (3) are equivalent to saying that every cone σ that maps onto κ with relative dimension 1 can have either one or two faces isomorphic to κ; and if σ has two such faces, then $N \cap \sigma \cong (Q \cap \kappa) \times_{\mathbb{N}} \mathbb{N}^2$ for some homomorphism $e_q : Q \cap \kappa \to \mathbb{N}$, while if σ has precisely one such face, then $\sigma \cap N \cong (Q \cap \kappa) \times \mathbb{N}$. These statements, and the construction of the homomorphism e_q are precisely the content of Lemmas 3.11 and 3.12 of [Ascher and Molcho 2015], or equivalently Lemmas 8 and 9 in [Gillam and Molcho 2013] from which the former lemmas are derived.

The description given above holds étale locally on X, for the étale sheaf M_X. In what follows, we also need to understand the induced Zariski sheaf $\tau_* M_X$ on X, where τ denotes the morphism of sites $\text{ét}(X) \to \text{zar}(X)$ — see Lemma 4.3 and Proposition 4.4. We claim that the description of $\tau_* M_X$ over the generic points η of irreducible components of X and over generic points q of irreducible components of nodes is only slightly more involved. Specifically, we have:

Corollary C.1.1. *The logarithmic structure $\tau_* M_X$ satisfies $\tau_* M_{X,\eta} = M_S$ on the generic point of an irreducible component η. On the generic point of an irreducible component of a node*, *we have either*

(1) $\tau_* M_{X,q} = M_S \oplus_{\mathbb{N}} \mathbb{N}^2$, *when the node is the intersection of two distinct irreducible components of X in the Zariski topology*, *or*

(2) $\tau_* M_{X,q} = M_S$, *when the node is the self-intersection of a single irreducible component.*

Proof. We prove the statement about nodes first. The question is local on X, so we may assume for simplicity that X is the spectrum of $\mathcal{O}_{X,q}$. According to Theorem C.1 above, we can choose an étale cover U of X, and a lift q' of q, with the property that $M_{U,q'} \cong M_S \oplus_{\mathbb{N}} \mathbb{N}^2$; here the two generators e_1, e_2 of $\mathbb{N}^2$ correspond to the two branches of U around q', and map to the two functions in $\mathcal{O}_{U,q}$ which define these two branches. For every étale cover V of U over X, and cover q'' of q', we have $M_{V,q''} \cong M_S \oplus_{\mathbb{N}} \mathbb{N}^2$ as well, as $M_{U,q'}$ and $M_{V,q''}$ are both pulled back from $M_{X,q}$. Thus, every étale cover V of U over X induces an automorphism of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$. On the other hand, the sheaf $\tau_* M_X$ is determined from M_U by descent, hence $\tau_* M_{X,q}$ is the submonoid of invariants of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$ under all automorphisms of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$ obtained by étale covers $V \to U$ over X. Every such automorphism further lives over S, hence must fix M_S; and since $e_1 + e_2 \in \mathbb{N}^2$ is in M_S, such an automorphism must fix $e_1 + e_2$. We are thus looking at automorphisms of $\mathbb{N}^2$ which fix $(1, 1)$, that is, matrices with coefficients in $\mathbb{N}$ and determinant 1, and that fix $(1, 1)$. The only two such matrices are the identity and the matrix $e_1 \to e_2$, $e_2 \to e_1$. Thus, the invariants of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$ are either all of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$ or the diagonal $M_S \oplus_{\mathbb{N}} \mathbb{N}(e_1 + e_2) \cong M_S$. To prove the corollary it remains to analyze when each case happens. Suppose first that q is the intersection of two distinct irreducible components of X. Since $\mathcal{O}_X$ is determined from $\mathcal{O}_U$ by descent as well, we see that the images of e_1, e_2 in U map to distinct functions in $\mathcal{O}_X$, which define q in each of the irreducible components. Thus, there can be no étale cover V of U over X which interchanges e_1, e_2 in the automorphism $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$, and we are in the situation where the invariants are all of $M_S \oplus_{\mathbb{N}} \mathbb{N}^2$. On the other hand, if q is the self-intersection of a single component, the only linear combinations of e_1 and e_2 that descend are the $ke_1 + ke_2$, which lie in M_S.

To see the statement about the logarithmic structure over the generic points of irreducible components η, we simply observe that the invariants of M_S over M_S are always M_S, hence $\tau_* M_{X,\eta} = M_S$ as well. $\square$

The minimal log structure over a geometric point. We consider a family of logarithmic schemes

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & V \arrow[dl] \\ S & \end{tikzcd} \tag{C-2}$$

with S a geometric point, and π a flat, logarithmically smooth morphism with reduced fibers. The morphism f is not assumed to be a logarithmic morphism. We write M_S for the logarithmic structure on S, M_X for the logarithmic structure on X, and M_V for the logarithmic structure on V. We further write $M = f^* M_V$ for the pullback of the logarithmic structure on V to X.

Out of X, we can create a category $\mathcal{C}$. The objects of $\mathcal{C}$ are the generic points of the strata in the minimal stratification on which the relative characteristic M_X is locally constant. Note that $\tau_* M_{X/S}$ is constant on these strata. We have a morphism $x \to y$ in $\mathcal{C}$ whenever $\{x\} \in \overline{\{y\}}$.

In the special case when X is a nodal curve, the category $\mathcal{C}$ is rather simple, with one object for each irreducible component η of X and an object for each node or marked point q, and a morphism $q \to \eta$ whenever q belongs to the component η. In fact, the same characterization holds for general X in depths 0 and 1: depth-0 objects correspond to irreducible components of X, and depth-1 objects correspond to either marked divisors in X or nodes where irreducible components of X intersect, according to Theorem C.1.

Lemma C.2. *Let F be a sheaf in the étale topology on X that is pulled back from a sheaf in the Zariski topology, which is constructible with respect to the category $\mathcal{C}$. Then*

$$\pi_!(F) = \varinjlim_{x \in \mathcal{C}} F_x.$$

Proof. By Proposition 4.8, $\pi_!$ is the left adjoint to π^*_{zar}, the pullback functor *on Zariski sheaves*, so we only need to verify that, for any sheaf G on S, we have $\operatorname{Hom}\big(\varinjlim_{x \in \mathcal{C}} F_x, G\big) = \operatorname{Hom}(F, \pi^* G)$. A sheaf G on S is simply a monoid, so $\pi^* G$ is the constant sheaf on X associated to G. Thus, to give a homomorphism $F \to \pi^* G$ is equivalent to giving homomorphisms $F_x \to G$ for each $x \in \mathcal{C}$ which are compatible with generization; but this is precisely the data of a homomorphism $\varinjlim_{x \in \mathcal{C}} F_x \to G$. □

Observe furthermore that the colimit over a finite indexing category only depends on the objects of depths 0 and 1, i.e., in this case, over the irreducible components η of X, the generic points q of the nodes of X, and the generic points of the marked divisors of X. Note that the marked divisors do not contribute to the colimit. The reason is that, for each $p \in \mathcal{C}$ corresponding to a marked divisor, there is a unique morphism $p \to \eta$, where η corresponds to the irreducible component containing the marked divisor. So the points corresponding to marked divisors can be omitted from the diagram without affecting the colimit. The same is true for nodes of X which are the self-intersection of a single irreducible component.

From here on, η is always going to denote the generic point of an irreducible component of X, and q is always going to denote the generic point of an irreducible component of the intersection of two *distinct* irreducible components.

In what follows, we will replace M_X with the sheaf $\tau^*\tau_* M_X$. Note that while X equipped with $\tau^*\tau_* M_X$ is no longer log smooth over (S, M_S), the minimal log structure obtained from $\tau^*\tau_* M_X$ and the minimal log structure obtained from M_X coincide, according to Proposition 4.4. The reason we do this replacement is because this allows us to use $\mathcal{C}$ in the calculation of the minimal log structure, according to Lemma C.2, rather than the far larger category of all étale specializations. Furthermore, according to the preceding remark, only irreducible components η and nodes q need to be included in the calculation, and there we have $\tau^*\tau_* M_{X,\eta} = M_{X,\eta}$, $\tau^*\tau_* M_{X,q} = M_{X,q}$, according to Corollary C.1.1. So this is, in fact, not a serious abuse of notation.

This allows us to work out an explicit presentation for the minimal logarithmic structure. We first determine its associated group. Let us recall the notation. Diagram (C-1) gives us a diagram

$$\begin{tikzcd} M_{X/S}^{\mathrm{gp}} & M_X^{\mathrm{gp}} \arrow[l] & M^{\mathrm{gp}} \arrow[l] \arrow[ll, bend right, "u"'] \\ & \pi^* M_S^{\mathrm{gp}} \arrow[u] \arrow[ur, "i"'] & \end{tikzcd}$$

of sheaves of monoids in which we have set $M = f^* M_V$ and in which u is the type of the morphism. With u fixed, the problem of finding a minimal logarithmic structure is the same as finding a minimal object N_S such that u factors through a homomorphism $M^{\mathrm{gp}} \to N_S^{\mathrm{gp}} \oplus_{\pi^* M_S^{\mathrm{gp}}} M_X^{\mathrm{gp}}$. To find the minimal logarithmic structure, we set

$$\bar{R}_0^{\mathrm{gp}} = \bar{M}_X^{\mathrm{gp}} \times_{\bar{M}_{X/S}^{\mathrm{gp}}} \bar{M}^{\mathrm{gp}} \quad \text{and} \quad \bar{R}^{\mathrm{gp}} = \bar{R}_0^{\mathrm{gp}} / \Delta(\pi^* \bar{M}_S^{\mathrm{gp}}),$$

where Δ is the diagonal map $\pi^* \bar{M}_S^{\mathrm{gp}} \to \bar{M}_X^{\mathrm{gp}} \times_{\bar{M}_{X/S}^{\mathrm{gp}}} \bar{M}^{\mathrm{gp}}$.

Applying Lemma C.2, we see that the associated group of the minimal logarithmic structure $\pi_! \bar{R}^{\mathrm{gp}}$ on S is given as the coequalizer

$$\varinjlim_{x \in C} \bar{R}_x^{\mathrm{gp}} = \varinjlim \biggl(\bigoplus_q \bar{R}_q^{\mathrm{gp}} \overset{\phi_1}{\underset{\phi_2}{\rightrightarrows}} \bigoplus_\eta \bar{R}_\eta^{\mathrm{gp}} \biggr),$$

where ϕ_1, ϕ_2 denote the two generization maps, induced by the generization morphisms $M_{X,q} \to M_{X,\eta}$ and $M_q \to M_\eta$ whenever q is contained in η. If we choose an ordering of the two components η_1, η_2 containing a node q, the coequalizer becomes the quotient

$$\bigoplus_q \bar{R}_q^{\mathrm{gp}} \xrightarrow{\phi_2 - \phi_1} \bigoplus_\eta \bar{R}_\eta^{\mathrm{gp}}.$$

On the other hand, even though R may be difficult to understand, its stalks at generic points and nodes are straightforward. We have

$$\bar{R}^{\mathrm{gp}}_{0,\eta} = \bar{M}^{\mathrm{gp}}_{X,\eta} \times \bar{M}^{\mathrm{gp}}_{\eta} \cong \bar{M}^{\mathrm{gp}}_{S} \times \bar{M}^{\mathrm{gp}}_{\eta}$$

and so

$$\bar{R}^{\mathrm{gp}}_{\eta} = \bar{M}^{\mathrm{gp}}_{\eta}.$$

Similarly, on a node q we have

$$\bar{R}^{\mathrm{gp}}_{0,q} = \bar{M}^{\mathrm{gp}}_{X,q} \times_{\mathbb{Z}} \bar{M}^{\mathrm{gp}}_{q}.$$

Now recall that $\bar{M}_{X,q} = \bar{M}_S \oplus_{\mathbb{N}} \mathbb{N}^2 = \{(a,b) : b - a = \rho_q d\} \subset \bar{M}_S \times \bar{M}_S$, where $\rho_q : \mathbb{N} \to \bar{M}_S$ is a homomorphism determined by π, as discussed on page 726. We abusively also write ρ_q for the image $\rho_q(1) \in \bar{M}_S$ of ρ_q. The morphism $\bar{M}^{\mathrm{gp}}_{X,q} \to \mathbb{Z}$ is the canonical projection $\bar{M}^{\mathrm{gp}}_{X,q} \to \bar{M}^{\mathrm{gp}}_{X/S}$, which, explicitly, is the homomorphism that sends a pair (a,b) such that $b - a = \rho_q d$ to d. Therefore,

$$\bar{R}^{\mathrm{gp}}_{0,q} = \big\{(a,b,m) : b - a = u_q(m)\rho_q\big\}.$$

The morphism $\pi^*\bar{M}^{\mathrm{gp}}_S \to \bar{R}^{\mathrm{gp}}_{0,q}$ is the diagonal $s \mapsto (s,s,i(s))$, and hence, passing to the quotient, we obtain an isomorphism

$$\bar{R}^{\mathrm{gp}}_{q} \cong \bar{M}^{\mathrm{gp}}_{q}$$

by sending $[(a,b,m)] \mapsto m - i(a)$, with inverse $m \mapsto [0, u_q(m)\rho_q, m]$. The two generization morphisms $\bar{R}^{\mathrm{gp}}_{0,q} \to \bar{R}^{\mathrm{gp}}_{0,\eta} \to \bar{R}^{\mathrm{gp}}_{\eta}$ are the two natural maps from $\bar{M}^{\mathrm{gp}}_{X,q} \times_{\mathbb{Z}} \bar{M}^{\mathrm{gp}}_{q}$ to $\bar{M}^{\mathrm{gp}}_{\eta}$, sending (a,b,m) to $(i(a) + \phi_1(m))$ or to $(i(b) + \phi_2(m))$, respectively, depending on whether η is the first or second irreducible component containing q, under our ordering. Thus, under the isomorphism above, the morphism $\bar{R}^{\mathrm{gp}}_{q} \to \bar{R}^{\mathrm{gp}}_{\eta_1} \times \bar{R}^{\mathrm{gp}}_{\eta_2}$ becomes the map $\bar{M}^{\mathrm{gp}}_{q} \to \bar{M}^{\mathrm{gp}}_{\eta_1} \times \bar{M}^{\mathrm{gp}}_{\eta_2}$ which sends m to $\big(-\phi_1(m), \phi_2(m) + i(u_q(m)\rho_q)\big)$. The associated group of the minimal logarithmic structure thus has the very simple quotient presentation

$$\bigoplus_q \bar{M}^{\mathrm{gp}}_{q} \xrightarrow{(-\phi_1, \phi_2 + i u_q \rho_q)} \bigoplus_\eta \bar{M}^{\mathrm{gp}}_{\eta}.$$

The characteristic monoid of the actual logarithmic structure $\pi_!\bar{R}$ is then the image of $\bigoplus \bar{M}_\eta$ in $\pi_!\bar{R}^{\mathrm{gp}}$.

Gross–Siebert minimality. We now specialize to the case when the family of morphisms $f : X \to V$ is absolute, i.e., of the form

$$\begin{array}{ccc} X & \xrightarrow{f} & V \\ {\scriptstyle\pi}\big\downarrow & & \\ S & & \end{array} \tag{C-3}$$

This point of view can be reconciled with that of the previous paragraph by simply redefining the logarithmic structure of V to be $M'_V = M_V \oplus_{\mathcal{O}_V^*} \pi^* M_S$. At any rate, to avoid confusion and keep in line with existing literature, we will denote the sheaf of monoids $f^*\bar{M}_V$ by P, and $f^* M'_V$ by M, as above. For a node q, specializing to two irreducible components η_1, η_2, let $\chi_i : P_q \to P_{\eta_i}$ denote the two generization maps.

Definition C.3. Let S be a geometric point. A logarithmic structure N_S on S is called a *Gross–Siebert minimal logarithmic structure* for f over M_S if its characteristic monoid $Q = \bar{N}_S$ has associated group isomorphic to the cokernel of

$$\Phi : \bigoplus_q P_q^{\mathrm{gp}} \xrightarrow{(\rho_q u_q, -\chi_1, \chi_2)} \bar{M}_S^{\mathrm{gp}} \oplus \bigoplus_\eta P_\eta^{\mathrm{gp}}$$

and Q is isomorphic to the image of $\bar{M}_S^{\mathrm{gp}} \oplus \bigoplus_\eta P_\eta^{\mathrm{gp}}$ in the associated group.

Remark C.4. In [Gross and Siebert 2013], the minimal logarithmic structure is in fact saturated by definition, i.e., Q is the saturation of the monoid in the definition above. Since our results go through when Q is merely integral, the definition is stated in this slightly more general form.

Remark C.5. In [Gross and Siebert 2013], the object of study is stable logarithmic maps, in which case $X \to S$ is a logarithmic curve. In this case, there are canonical logarithmic structures on X and S which make $X \to S$ logarithmically smooth. Over a geometric point, the characteristic monoid of the canonical logarithmic structure on S is $\mathbb{N}^m$, where m is the number of nodes of X. The definition of minimality given in [Gross and Siebert 2013] has this canonical logarithmic structure as M_S throughout. The definition presented here is the evident modification that allows the same flexibility as the present paper.

Theorem C.6. *A Gross–Siebert minimal logarithmic structure is minimal.*

Proof. Let $(X, N_X) \to (S, N_S)$ be any log morphism pulled back from $(X, M_X) \to (S, M_S)$ which admits a log morphism f to V, as in diagram (C-3). The morphism f induces morphisms $P_\eta^{\mathrm{gp}} \to \bar{N}_{X,\eta}^{\mathrm{gp}} = \bar{N}_S^{\mathrm{gp}}$ for each η. We thus have a unique extension to a summation morphism $\Sigma : \bar{M}_S^{\mathrm{gp}} \oplus \bigoplus_\eta P_\eta^{\mathrm{gp}} \to \bar{N}_S^{\mathrm{gp}}$. On the other hand, for each node q, the diagram

$$\begin{tikzcd}
& P_{\eta_1}^{\mathrm{gp}} \arrow[r] & \bar{N}_S^{\mathrm{gp}} & \\
P_q^{\mathrm{gp}} \arrow[ur] \arrow[dr] \arrow[rrr, dashed] & & & \bar{N}_{X,q}^{\mathrm{gp}} = \{(a, b) : b - a = \rho_q d\} \arrow[ul, bend right] \arrow[dl, bend left] \\
& P_{\eta_2}^{\mathrm{gp}} \arrow[r] & \bar{N}_S^{\mathrm{gp}} &
\end{tikzcd}$$

must commute; thus the map Σ must descend uniquely to the (unsharpened) quotient of Φ. As N_S^{gp} is assumed torsion-free, the morphism must descend to the sharpened quotient as well, i.e., the associated group of the minimal logarithmic structure. Since $\bar{M}_S \oplus \bigoplus_\eta P_\eta \to \bar{N}_S^{\mathrm{gp}}$ factors through $\bar{N}_S$, its image Q in Q^{gp} maps to $\bar{N}_S$, as desired. □

We are now ready for the comparison.

Theorem C.7. *Suppose $(X, M_X) \to (S = \operatorname{Spec}\mathbb{C}, M_S)$ is a logarithmic smooth morphism which is flat and has reduced fibers, and $f : X \to (V, M_V)$ is a morphism to a logarithmic scheme. The minimal logarithmic structure defined in the paper coincides with the Gross–Siebert minimal logarithmic structure.*

Proof. This is immediate, as both logarithmic structures satisfy the same universal property. □

It is actually rather interesting to give a direct proof of this fact, as the calculation is illuminating. We consider

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & V \\ S & \end{tikzcd}$$

and extend it to

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\pi"'] & V \arrow[dl] \\ S & \end{tikzcd}$$

as above by putting $M'_V = M_V \oplus_{\mathcal{O}_V^*} \pi^* M_S$, $M = f^* M'_V$. Then, $\bar{M}_q = P_q \oplus \bar{M}_S$ and $\bar{M}_\eta = P_\eta \oplus \bar{M}_S$. The morphism $m \mapsto \big(-\phi_1(m), \phi_2(m) + i(u_q(m)\rho_q)\big)$ is identified with $(m, s) \mapsto \big(-\chi_1(m), -s, \chi_2(m), u_q(m)\rho_q + s\big)$. Therefore, by the results of Section 1, we have that the associated group of the minimal logarithmic structure for the family is given as the quotient

$$\bigoplus_q \big(P_q^{\mathrm{gp}} \oplus \bar{M}_S^{\mathrm{gp}}\big) \xrightarrow{(-\chi_1, -\operatorname{id}, \chi_2, \operatorname{id} + u_q\rho_q)} \bigoplus_\eta \big(P_\eta^{\mathrm{gp}} \oplus M_S^{\mathrm{gp}}\big).$$

If Σ denotes the summation map $\bigoplus \bar{M}_S^{\mathrm{gp}} \to \bar{M}_S^{\mathrm{gp}}$, we obtain a commutative diagram

$$\begin{tikzcd} \bigoplus_q \big(P_q^{\mathrm{gp}} \oplus \bar{M}_S^{\mathrm{gp}}\big) \arrow[r, "(-\chi_1, -\operatorname{id}, \chi_2, \operatorname{id} + u_q\rho_q)"] \arrow[d] & \bigoplus_\eta \big(\bar{M}_S^{\mathrm{gp}} \oplus P_\eta\big) \arrow[d, "(\Sigma, \operatorname{id})"] \\ \bigoplus_q P_q^{\mathrm{gp}} \arrow[r, "(u_q\rho_q, -\chi_1, \chi_2)"'] & \bar{M}_S^{\mathrm{gp}} \oplus \bigoplus_\eta P_\eta \end{tikzcd}$$

Thus, the morphism (Σ, id) descends to a morphism of the quotients $\pi_! \bar{R}^{\mathrm{gp}} \to Q_S^{\mathrm{gp}}$, where Q is the characteristic monoid of the Gross–Siebert minimal logarithmic structure

$$\begin{tikzcd}
\bigoplus_q P_q^{\mathrm{gp}} \oplus \bar{M}_S^{\mathrm{gp}} \arrow[r] \arrow[d] & \bigoplus_\eta \big(\bar{M}_S^{\mathrm{gp}} \oplus P_\eta\big) \arrow[r] \arrow[d, "(\Sigma,\mathrm{id})"] & \pi_! \bar{R}^{\mathrm{gp}} \arrow[d] \\
\bigoplus_q P_q^{\mathrm{gp}} \arrow[r] & \bar{M}_S^{\mathrm{gp}} \oplus \bigoplus_\eta P_\eta \arrow[r] & Q_S^{\mathrm{gp}}
\end{tikzcd}$$

Observe that the kernel of the map $P_q^{\mathrm{gp}} \oplus \bar{M}_S^{\mathrm{gp}} \to P_q^{\mathrm{gp}}$ is simply $\bar{M}_S^{\mathrm{gp}}$. The kernel of the map on the right, on the other hand, is the kernel of the summation map $\bigoplus_\eta \bar{M}_S^{\mathrm{gp}} \to \bar{M}_S^{\mathrm{gp}}$. The induced map of kernels then fits into the sequence

$$\bigoplus_q \bar{M}_S^{\mathrm{gp}} \to \bigoplus_\eta \bar{M}_S^{\mathrm{gp}} \to \bar{M}_S^{\mathrm{gp}}.$$

But this is precisely the part of the complex that computes the homology of the geometric realization $B\mathcal{C}$ of the category $\mathcal{C}$ with coefficients in the group $\bar{M}_S^{\mathrm{gp}}$, where the morphism on the right is the evaluation map. Since X is connected, so is $\mathcal{C}$, and thus

$$H_0\big(B\mathcal{C}, \bar{M}_S^{\mathrm{gp}}\big) = \bar{M}_S^{\mathrm{gp}},$$

and thus $\bigoplus_q \bar{M}_S^{\mathrm{gp}}$ surjects onto the kernel of the evaluation map Σ. Thus, the map $\pi_! \bar{R}^{\mathrm{gp}} \to Q$ is injective. On the other hand, the morphism (Σ, id) is certainly surjective, so $\pi_! \bar{R}^{\mathrm{gp}} \to Q^{\mathrm{gp}}$ is also surjective. We thus get the isomorphism on the level of associated groups.

To extend the isomorphism on the level of actual monoids, note that since the summation morphism is also surjective on the level of monoids, the image of $\bigoplus_\eta \big(\bar{M}_S^{\mathrm{gp}} \oplus P_\eta\big)$ surjects onto the image of $\bar{M}_S \oplus \big(\bigoplus_\eta P_\eta\big)$, and hence $\pi_! \bar{R}$ surjects onto Q. Since $\pi_! \bar{R}$ and Q are integral and the morphism of associated groups is injective, we obtain the desired isomorphism.

Acknowledgements

The central construction of this paper was inspired by the construction of basic logarithmic maps in [Gross and Siebert 2013]. I am very grateful to Qile Chen and Steffen Marcus for their comments on an early draft, and to Dan Abramovich for his many detailed comments and his encouragement that I present these results in their natural generality. Samouil Molcho discovered an error in my original presentation of the main construction. I am very grateful to him for that correction, as well as for Appendix C.

References

[Abramovich and Chen 2014] D. Abramovich and Q. Chen, "Stable logarithmic maps to Deligne–Faltings pairs II", *Asian J. Math.* **18**:3 (2014), 465–488. MR 3257836 Zbl 1321.14025

[Abramovich et al. 2010] D. Abramovich, Q. Chen, W. D. Gillam, and S. Marcus, "The evaluation space of logarithmic stable maps", preprint, 2010. arXiv 1012.5416

[Ascher and Molcho 2015] K. Ascher and S. Molcho, "Logarithmic stable toric varieties and their moduli", preprint, 2015. arXiv 1412.3766

[Chen 2014] Q. Chen, "Stable logarithmic maps to Deligne–Faltings pairs I", *Ann. of Math.* (2) **180**:2 (2014), 455–521. MR 3224717 Zbl 1311.14028

[EGA IV_3 1966] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR 0217086 Zbl 0144.19904

[Gillam 2012] W. D. Gillam, "Logarithmic stacks and minimality", *Internat. J. Math.* **23**:7 (2012), 1250069, 38. MR 2945649 Zbl 1248.18008

[Gillam and Molcho 2013] W. D. Gillam and S. Molcho, "Stable log maps as moduli of flow lines", preprint, 2013, available at http://math.colorado.edu/∼samo2465/flows3.pdf.

[Gross and Siebert 2013] M. Gross and B. Siebert, "Logarithmic Gromov–Witten invariants", *J. Amer. Math. Soc.* **26**:2 (2013), 451–510. MR 3011419 Zbl 1281.14044

[Hall and Rydh 2015] J. Hall and D. Rydh, "Coherent Tannaka duality and algebraicity of Hom-stacks", preprint, 2015. arXiv 1405.7680

[Kato 1989] K. Kato, "Logarithmic structures of Fontaine–Illusie", pp. 191–224 in *Algebraic analysis, geometry, and number theory* (Baltimore, MD, 1988), edited by J.-I. Igusa, Johns Hopkins Univ. Press, Baltimore, MD, 1989. MR 1463703 Zbl 0776.14004

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics **39**, Springer, Berlin, 2000. MR 1771927 Zbl 0945.14005

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics **5**, Springer, New York, 1998. MR 1712872 Zbl 0906.18001

[Olsson 2003] M. C. Olsson, "Logarithmic geometry and algebraic stacks", *Ann. Sci. École Norm. Sup.* (4) **36**:5 (2003), 747–791. MR 2032986 Zbl 1069.14022

[Olsson 2008] M. C. Olsson, *Compactifying moduli spaces for abelian varieties*, Lecture Notes in Mathematics **1958**, Springer, Berlin, 2008. MR 2446415 Zbl 1165.14004

[Romagny 2011] M. Romagny, "Composantes connexes et irréductibles en familles", *Manuscripta Math.* **136**:1-2 (2011), 1–32. MR 2820394 Zbl 1266.14010

[Rosales and García-Sánchez 1999] J. C. Rosales and P. A. García-Sánchez, *Finitely generated commutative monoids*, Nova Science Publishers, Commack, NY, 1999. MR 1694173 Zbl 0966.20028

Communicated by Ravi Vakil
Received 2014-09-14 Revised 2015-10-28 Accepted 2016-03-22

jonathan.wise@math.colorado.edu
Department of Mathematics, University of Colorado, Campus Box 395, Boulder, CO 80309-0395, United States

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.737

Residual intersections and the annihilator of Koszul homologies

Seyed Hamid Hassanzadeh and Jose Naéliton

We study Cohen–Macaulayness, unmixedness, the structure of the canonical module and the stability of the Hilbert function of algebraic residual intersections. We establish some conjectures about these properties for large classes of residual intersections without restricting the local number of generators of the ideals involved. To determine the above properties, we construct a family of approximation complexes for residual intersections. Moreover, we determine some general properties of the symmetric powers of quotient ideals which were not known even for special ideals with a small number of generators. Finally, we show acyclicity of a prime case of these complexes to be equivalent to finding a common annihilator for higher Koszul homologies, which unveils a tight relation between residual intersections and the uniform annihilator of positive Koszul homologies, shedding some light on their structure.

1. Introduction

Understanding the residual of the intersection of two algebraic varieties is an intriguing concept in algebraic geometry. Similar to many other notions in intersection theory, the concept of residual intersections needs the proper interpretation to encompass the desired algebraic and geometric properties. In the sense of [Artin and Nagata 1972], which is our point of view, the notion is a vast generalization of linkage (or liaison), which is more ubiquitous, but also harder to understand. Let X and Y be irreducible closed subschemes of a Noetherian scheme Z with $\operatorname{codim}_Z(X) \leq \operatorname{codim}_Z(Y) = s$ and $Y \not\subset X$. Then Y is called a residual intersection of X if the number of equations needed to define $X \cup Y$ as a subscheme of Z is the smallest possible, that is, s. Precisely, if R is a Noetherian ring, I an ideal of height g and $s \geq g$ an integer, then:

Hassanzadeh was partially supported by a CNPq grant (300586/2012-4) and Naéliton worked under a CNPq Doctoral scholarship.
MSC2010: primary 13C40; secondary 13H10, 13D02, 14C17, 14M06.
Keywords: residual intersection, canonical module, type, sliding depth, approximation complex, Koszul annihilator.

- An (*algebraic*) *s-residual intersection* of I is a proper ideal J of R such that $\operatorname{ht}(J) \geq s$ and $J = (\mathfrak{a} :_R I)$ for some ideal $\mathfrak{a} \subset I$ generated by s elements.
- A *geometric s-residual intersection* of I is an algebraic s-residual intersection J of I such that $\operatorname{ht}(I + J) \geq s + 1$.

Based on a construction of Laksov for residual intersection, Fulton [1998, Definition 9.2.2] presents a formulation for residual intersection that, locally, can be expressed as follows: suppose that $X = \operatorname{Spec}(R)$ and that Y and S are closed subschemes of X defined by the ideals $\mathfrak{a}$ and I, respectively. Let $\tilde{X} = \operatorname{Proj}(\mathcal{R}_R(I))$ be the blow-up of X along S. Consider the natural map $\pi : \tilde{X} \to X$. Let $\tilde{Y} = \pi^{-1}(Y)$ and $\tilde{S} = \pi^{-1}(S)$. Then $\tilde{Y} = \operatorname{Proj}(\mathcal{R}_R(I)/\mathfrak{a}\mathcal{R}_R(I))$ and $\tilde{S} = \operatorname{Proj}(\mathcal{R}_R(I)/I\mathcal{R}_R(I))$ are closed subschemes of $\tilde{X}$ with ideal sheaves $\mathcal{I}_{\tilde{Y}}$ and $\mathcal{I}_{\tilde{S}}$, where $\mathcal{I}_{\tilde{S}}$ is an invertible sheaf. Let $Z' \subseteq \tilde{X}$ be the closed subscheme defined by the ideal sheaf $\mathcal{I}_{Z'} = \mathcal{I}_{\tilde{Y}} \cdot \mathcal{I}_{\tilde{S}}^{-1}$. Then Z' is called the *residual scheme to $\tilde{S}$ in $\tilde{Y}$*. Precisely $Z' = \operatorname{Proj}(\mathcal{R}_R(I)/\gamma\mathcal{R}_R(I))$ in which $\gamma = \mathfrak{a} \subseteq (\mathcal{R}_R(I))_{[1]}$. Finally, the residual intersection to S in Y [Fulton 1998, Definition 9.2.2] is the direct image of $\mathcal{O}_{Z'}$, i.e., $\pi_*(\mathcal{O}_{Z'})$. Since $\mathcal{O}_{Z'}$ is a coherent sheaf, by [Hartshorne 1977, Chapter III, Proposition 8.5],

$$\pi_*(\mathcal{O}_{Z'}) = \big(H^0(\tilde{X}, \mathcal{O}_{Z'})\big)^{\tilde{}} = \big(\Gamma(\tilde{X}, \mathcal{O}_{Z'})\big)^{\tilde{}}.$$

Note that $\Gamma(\tilde{X}, \mathcal{O}_{Z'})$ equals $\Gamma_*(\tilde{X}, \mathcal{O}_{Z'}(0))$, which is equal to the ideal transform $D_{\mathcal{R}_R(I)_+}(\mathcal{R}_R(I)/\gamma\mathcal{R}_R(I))_{[0]}$, by an application of the Čech complex. The latter is closely related, and in many cases determined, by the ideal

$$H^0_{\mathcal{R}_R(I)_+}(\mathcal{R}_R(I)/\gamma\mathcal{R}_R(I))_{[0]} = \bigcup(\mathfrak{a}I^i :_R I^{i+1}).$$

In [Hassanzadeh 2012] and in the current work, we consider the symmetric algebra $\operatorname{Sym}_R(I)$ instead of the Rees algebra $\mathcal{R}_R(I)$. This generalization has already proved its usefulness in studying multiple-point formulas by Kleiman; see [Fulton 1998, Example 17.6.2]. We then find a kind of *arithmetical* residual intersection to be $\bigcup\big(\gamma \operatorname{Sym}^i_R(I) :_R \operatorname{Sym}^{i+1}_R(I)\big)$, where $\gamma = \mathfrak{a} \subseteq \operatorname{Sym}^i_R(I)_{[1]}$. Comparing the above three definitions for residual intersection, we have

$$J = (\mathfrak{a} : I) \subseteq \bigcup\big(\gamma \operatorname{Sym}^i_R(I) :_R \operatorname{Sym}^{i+1}_R(I)\big) \subseteq \bigcup(\mathfrak{a}I^i :_R I^{i+1}).$$

Interestingly, these ideals coincide if the algebraic residual intersection J does not share any associated primes with I, e.g., if J is unmixed and the residual is geometric.

Determining the cases where the first inclusion above is an equality leads us to define a third variation of algebraic residual intersection.

Definition 1.1. An *arithmetic s-residual intersection* $J = (\mathfrak{a} : I)$ is an algebraic s-residual intersection such that $\mu_{R_\mathfrak{p}}((I/\mathfrak{a})_\mathfrak{p}) \leq 1$ for all prime ideals $\mathfrak{p} \supseteq (I + J)$ with $\operatorname{ht}(\mathfrak{p}) \leq s$. (Here μ denotes the minimum number of generators.)

Clearly any geometric s-residual intersection is arithmetic. Moreover, for any algebraic s-residual intersection $J = (\mathfrak{a} : (f_1, \dots, f_r))$ which is not geometric, all of the colon ideals $(\mathfrak{a} : f_i)$ are arithmetic s-residual intersections and at least one of them is not geometric.

In this paper we introduce a family of complexes, denoted by $\{{}_i\mathcal{Z}_\bullet^+\}_{i=0}^{\infty}$, to approximate the i-th symmetric power of $I/\mathfrak{a}$, which is denoted by $\mathrm{Sym}_R^i(I/\mathfrak{a})$ for $i > 1$. The idea to define this family is inspired by [Hassanzadeh 2012] in which the single complex ${}_0\mathcal{Z}_\bullet^+$ is treated. $H_0({}_0\mathcal{Z}_\bullet^+)$ is a cyclic module of the form R/K, where K is called the *disguised s-residual intersection of I with respect to* $\mathfrak{a}$; see Definition 2.1. The study of the other members of the above family of complexes sheds some more light on the structure of residual intersections. A flavor of our main results in Section 2 is the following:

Main results. *Let $(R, \mathfrak{m})$ be a Cohen–Macaulay (CM) local ring of dimension d and let I be an ideal with $\mathrm{ht}(I) = g > 0$. Let $s \geq g$ and $1 \leq k \leq s - g + 2$ and let $J = (\mathfrak{a} : I)$ be any (algebraic) s-residual intersection. Suppose that I is strongly Cohen–Macaulay (SCM). Then:*

1. *(Theorem 2.14) The canonical module of R/J is $\mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})$, provided the residual is arithmetic and R is Gorenstein.*
2. *(Corollary 2.8)* $\mathrm{depth}(R/\mathfrak{a}) = d - s$.
3. *(Corollary 2.8) J is unmixed of codimension s.*
4. *(Theorem 2.6 and Proposition 3.1) ${}_k\mathcal{Z}_\bullet^+$ is acyclic, $H_0({}_k\mathcal{Z}_\bullet^+) = \mathrm{Sym}_R^k(I/\mathfrak{a})$ and the latter is CM of dimension $d-s$ (the acyclicity of ${}_k\mathcal{Z}_\bullet^+$ implies conjecture (5) below in the arithmetic case).*

These results address the following (implicit) conjectures made during the development of the theory of algebraic residual intersections.

Conjectures. *Let R be a CM local ring and let I be SCM, or even just satisfying sliding depth (SD). Then:*

1. *R/J is Cohen–Macaulay.*
2. *The canonical module of R/J is the $(s - g + 1)$-st symmetric power of $I/\mathfrak{a}$, if R is Gorenstein.*
3. *$\mathfrak{a}$ is minimally generated by s elements.*
4. *J is unmixed.*
5. *The Hilbert series of R/J depends only on I and the degrees of the generators of $\mathfrak{a}$.*

The first conjecture essentially goes back to [Artin and Nagata 1972], and it has been asked as an open question in [Huneke and Ulrich 1988]. At the Sundance conference in 1990, Ulrich [1992] mentioned conjectures (1)–(4) above as desirable facts to be proved. The property of the Hilbert function is rather recent and has been analyzed by Chardin, Eisenbud and Ulrich [Chardin et al. 2001; Chardin et al. 2015].

It should be mentioned that these conjectures are proved if one supposes in addition that the ideal I has locally few generators, a condition which is called G_s, or if it has a deformation with the G_s property. Over time the G_s condition has become a "standard" assumption in the theory of residual intersections which is not avoidable in some cases. However, the desire is to prove the above assertions without restricting the local number of generators of I.

In comparison, obtaining the structure of the canonical module in the absence of the G_s condition is more challenging. To achieve this, we show that under the above hypotheses, $\mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})$ is a faithful maximal Cohen–Macaulay R/J-module of type 1. Concerning the type of modules, we prove even more. We show in Theorem 2.12 that the inequality

$$r_R\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big) \leq \binom{r+s-g-k}{r-1} r_R(R)$$

holds for any $1 \leq k \leq s-g+1$, where, for a finitely generated R-module M, $r_R(M) := \dim_{R/\mathfrak{m}} \mathrm{Ext}_R^{\mathrm{depth}(M)}(R/\mathfrak{m}, M)$ is the Cohen–Macaulay type of M.

In Section 3, we present several applications of the theorems and constructions so far. We state how much the Hilbert functions of R/J and $R/\mathfrak{a}$ depend on the generators and/or degrees of I and $\mathfrak{a}$. If I is SCM, then the Hilbert function of the disguised residual intersection, and that of $\mathrm{Sym}_R^k(I/\mathfrak{a})$, if $1 \leq k \leq s-g+2$, depends only on the *degrees* of the generators of $\mathfrak{a}$ and the Koszul homologies of I. In particular, $k=1$ implies that the Hilbert function of $R/\mathfrak{a}$ is constant on the open set of ideals $\mathfrak{a}$ generated by s forms of the given degrees such that $\mathrm{ht}(\mathfrak{a}:I) \geq s$. This is comparable with results in [Chardin et al. 2001], where the same assertion is concluded under some G_s hypotheses.

The graded structure of ${}_k\mathcal{Z}_\bullet^+$ shows that if I satisfies the SD_1 condition, then

$$\mathrm{reg}\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big) \leq \mathrm{reg}(R) + \dim(R_0) + \sigma(\mathfrak{a}) - (s-g+1-k)\,\mathrm{indeg}(I/\mathfrak{a}) - s$$

for $k \geq 1$. These applications were not known even when ideal I satisfies the G_s condition. Finally, in Proposition 3.4, by a combination of older and newer facts, we show that for any algebraic s-residual intersection $J = \mathfrak{a} : I$, if I is SCM and evenly linked to a G_s ideal (or has a deformation with these properties) then the disguised residual intersection and the algebraic residual intersection coincide. Based on this fact, we conjecture that (Conjecture 5.9) *in the presence of the sliding depth condition, the disguised residual intersection is the same as the algebraic residual intersection.*

In Section 4, we try to understand better the structure of the complex ${}_0\mathcal{Z}^+_\bullet$ in the case where $I/\mathfrak{a}$ is principal, e.g., $I = (\mathfrak{a}, b)$. We find in Theorem 4.4 that $H_i({}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})) \simeq bH_i(a_1, \dots, a_s)$ for all $i \geq 1$ and $H_0({}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})) \simeq R/(\mathfrak{a} : b)$. This fact shows how much the homologies of $\mathcal{Z}^+_\bullet$ complexes may depend on the generating sets, as well as showing a tight relation between the uniform annihilator of Koszul homologies and acyclicity of $\mathcal{Z}^+_\bullet$. As a byproduct, Corollary 4.6 suggests that to study the properties of colon ideals, instead of assuming $\operatorname{ht}((a_1, \dots, a_s) : I) \geq s$, one may only need to suppose that $IH_i(a_1, \dots, a_s) = 0$ locally at codimension $s - 1$.

Motivated by the facts in Section 4, we investigate the uniform annihilator of positive Koszul homologies in Section 5. Not much is known about the annihilator of Koszul homologies. In Corollary 5.6, we show that for a residual intersection $J = \mathfrak{a} : I$, where I satisfies SD and $\operatorname{depth}(R/I) \geq d - s$,

$$I \subseteq \bigcap_{j \geq 1} \operatorname{Ann}(H_j(\mathfrak{a})).$$

Surprisingly, this result contradicts one of the unpublished but well-known results of G. Levin [Vasconcelos 2005, Theorem 5.26], which yielded in [Corso et al. 2006] that $\operatorname{Supp}(H_1(\mathfrak{a})) = \operatorname{Supp}(H_0(\mathfrak{a}))$. Simple examples of residual intersection disprove this last claim. Moreover, we show in Theorem 5.4 that for an s-residual intersection $J = (\mathfrak{a} : I)$, if I satisfies SD and $\operatorname{depth}(R/I) \geq d - s$, so does $\mathfrak{a}$. This is an interesting result since for a long time it was known that the residual intersections of the ideal $\mathfrak{a}$ are Cohen–Macaulay although no one was aware of the SD property of $\mathfrak{a}$.

2. Residual approximation complexes

In this section we introduce a family of complexes which approximate the residual intersection and some of its related symmetric powers. We denote this family by $\{{}_i\mathcal{Z}^+_\bullet\}_{i=0}^\infty$. The complex ${}_0\mathcal{Z}^+_\bullet$ was already defined in [Hassanzadeh 2012] and used to prove the CM-ness of arithmetic residual intersections of ideals with sliding depth.

Throughout this section, R is a Noetherian ring of dimension d, and $I = (\boldsymbol{f}) = (f_1, \dots, f_r)$ is an ideal of grade $g \geq 1$. Although by adding one variable we could also treat the case $g = 0$, for simplicity we keep the assumption $g \geq 1$. Let $\mathfrak{a} = (a_1, \dots, a_s)$ be an ideal contained in I, with $s \geq g$, let $J = \mathfrak{a} :_R I$, and let $S = R[T_1, \dots, T_r]$ be a polynomial extension of R with indeterminates T_i. We denote the symmetric algebra of I over R by $\mathcal{S}_I$ or, in general, the symmetric algebra of an R-module M by $\operatorname{Sym}_R(M)$ and the k-th symmetric power of M by $\operatorname{Sym}^k_R(M)$. We consider $\mathcal{S}_I$ as an S-algebra via the ring homomorphism $S \to \mathcal{S}_I$ sending T_i to f_i as an element of $(\mathcal{S}_I)_1 = I$; then $\mathcal{S}_I = S/\mathcal{L}$. Let $a_i = \sum_{j=1}^r c_{ji} f_j$, $\gamma_i = \sum_{j=1}^r c_{ji} T_j$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_s)$ and $\mathfrak{g} := (T_1, \dots, T_r)$.

For a sequence of elements $\mathfrak{x}$ in a commutative ring A and an A-module M, we denote the Koszul complex by $K_\bullet(\mathfrak{x}; M)$, its cycles by $Z_i(\mathfrak{x}; M)$ and its homologies by $H_i(\mathfrak{x}; M)$. For a graded module M, $\operatorname{indeg}(M) := \inf\{i : M_i \neq 0\}$ and $\operatorname{end}(M) := \sup\{i : M_i \neq 0\}$. Setting $\deg(T_i) = 1$ for all i, S is a standard graded ring over $S_0 = R$.

To set one more convention, when we draw the picture of a double complex obtained from a tensor product of two finite complexes (in the sense of [Weibel 1994, 2.7.1]), say $\mathcal{A} \otimes \mathcal{B}$, we always put $\mathcal{A}$ in the vertical direction and $\mathcal{B}$ in the horizontal one. We also label the module which is in the upper-right corner as (0, 0) and consider the labels for the rest as the points in the third quadrant.

${}_k\mathcal{Z}^+_\bullet$ *complexes.* The first object in the construction of the family of ${}_k\mathcal{Z}^+_\bullet$ complexes is one of the approximation complexes — the $\mathcal{Z}$-complex [Herzog et al. 1983]. We consider the approximation complex $\mathcal{Z}_\bullet(\boldsymbol{f})$,

$$0 \to Z_{r-1} \otimes_R S(1-r) \to \cdots \to Z_1 \otimes_R S(-1) \to Z_0 \otimes_R S \to 0,$$

where $Z_i = Z_i(\boldsymbol{f})$ is the i-th cycle of the Koszul complex $K_\bullet(\boldsymbol{f}, R)$.

The second object is the Koszul complex $K_\bullet(\boldsymbol{\gamma}, S)$,

$$0 \to K_s(\gamma_1, \ldots, \gamma_s)(-s) \to \cdots \to K_1(\gamma_1, \ldots, \gamma_s)(-1) \to K_0(\gamma_1, \ldots, \gamma_s) \to 0.$$

Let $\mathcal{D}_\bullet = \operatorname{Tot}(K_\bullet(\boldsymbol{\gamma}, S) \otimes_S \mathcal{Z}_\bullet(\boldsymbol{f}))$. Then

$$\mathcal{D}_i = \bigoplus_{j=i-s}^{\min\{i,r-1\}} [Z_j \otimes_R S]^{\binom{s}{i-j}}(-i). \tag{2-1}$$

For a graded S-module M, the k-th graded component of M is denoted by $M_{[k]}$. Let $(D_\bullet)_{[k]}$ for $k \in \mathbb{Z}$ be the k-th graded strand of $D_\bullet$. We have $(D_i)_{[k]} = 0$ for all $k < i$; in particular,

$$H_k((D_\bullet)_{[k]}) = \operatorname{Ker}(D_k \to D_{k-1})_{[k]}. \tag{2-2}$$

Now, let $C^\bullet_{\mathfrak{g}} = C^\bullet_{\mathfrak{g}}(S)$ be the Čech complex of S with respect to the sequence $\mathfrak{g} = (T_1, \ldots, T_r)$,

$$C^\bullet_{\mathfrak{g}} : 0 \to C^0_{\mathfrak{g}}(= S) \to C^1_{\mathfrak{g}} \to \cdots \to C^r_{\mathfrak{g}} \to 0.$$

We then consider the bicomplex $C^\bullet_{\mathfrak{g}} \otimes_S D_\bullet$ with $C^0_{\mathfrak{g}} \otimes_S D_0$ in the corner. This bicomplex gives rise to two spectral sequences for which the second terms of the horizontal spectral are

$${}^2\mathrm{E}_{\mathrm{hor}}^{-i,-j} = H^j_{\mathfrak{g}}(H_i(D_\bullet)), \tag{2-3}$$

and the first terms of the vertical spectral are

$$\mathrm{E}_{\mathrm{ver}}^{-i,-j} = \begin{cases} 0 \to H^r_{\mathfrak{g}}(D_{r+s-1}) \to \cdots \to H^r_{\mathfrak{g}}(D_1) \to H^r_{\mathfrak{g}}(D_0) \to 0 & \text{if } j = r, \\ 0 & \text{otherwise.} \end{cases} \tag{2-4}$$

Since we have $H^r_{\mathfrak{g}}(D_i) = H^r_{\mathfrak{g}}\big(\bigoplus_j [Z_j \otimes S](-i)\big) = \bigoplus_j Z_j \otimes H^r_{\mathfrak{g}}(S)(-i)$ and $\operatorname{end}(H^r_{\mathfrak{g}}(S)) = -r$, it follows that $\operatorname{end}(H^r_{\mathfrak{g}}(D_i)) = i - r$, thus $H^r_{\mathfrak{g}}(D_i)_{[i-r+j]} = 0$, for all $j \geq 1$. We then define the following sequence of complexes indexed by $k \geq 0$:

$$0 \to H^r_{\mathfrak{g}}(D_{r+s-1})_{[k]} \to \cdots \to H^r_{\mathfrak{g}}(D_{r+k+1})_{[k]} \xrightarrow{\phi_k} H^r_{\mathfrak{g}}(D_{r+k})_{[k]} \to 0. \tag{2-5}$$

Since the vertical spectral (2-4) collapses at the second step, the horizontal spectral converges to the homologies of $H^r_{\mathfrak{g}}(D_\bullet)$. Since all of the homomorphisms are homogeneous of degree 0, the convergence inherits to any graded component. Therefore, for any $k \geq 0$, there exists a filtration $\cdots \subseteq \mathcal{F}_{2k} \subseteq \mathcal{F}_{1k} \subseteq \operatorname{Coker}(\phi_k)$ such that

$$\frac{\operatorname{Coker}(\phi_k)}{\mathcal{F}_{1k}} \simeq \big({}^{\infty}\mathrm{E}^{-k,0}_{\mathrm{hor}}\big)_{[k]}. \tag{2-6}$$

Observing that $H^{-t}_{\mathfrak{g}}(H_h(D_\bullet)) = 0$ for all $(t, h) \in \mathbb{N} \times \mathbb{N}_0$, one has ${}^{l}\mathrm{E}^{-k,0}_{\mathrm{hor}} \subseteq H^0_{\mathfrak{g}}(H_k(D_\bullet))$ for all $l \geq 2$. Hence we have the following chain of maps for which, except the isomorphism in the middle, all maps are canonical and the map on the right is given by (2-2):

$$\begin{tikzcd}
H^r_{\mathfrak{g}}(D_{r+k})_{[k]} \arrow[r] & \operatorname{Coker}(\phi_k) \arrow[r] & \frac{\operatorname{Coker}(\phi_k)}{\mathcal{F}_{1k}} \arrow[dll, "\simeq"'] & \\
{}^{\infty}\big(\mathrm{E}^{-k,0}_{\mathrm{hor}}\big)_{[k]} \arrow[r, "1-1"] & H^0_{\mathfrak{g}}(H_k(D))_{[k]} \arrow[r, "1-1"] & H_k(D)_{[k]} \arrow[r, "1-1"] & (D_k)_{[k]}
\end{tikzcd} \tag{2-7}$$

We denote the composition of the above chain of R-homomorphisms by τ_k.

Finally, we define the promised family of complexes as follows. For any integer $k \geq 0$, ${}_k\mathcal{Z}^+_\bullet$ is a complex of length s consisting of two parts: the right part is $(D_\bullet)_{[k]}$ and the left part is $({}^{1}\mathrm{E}_{\mathrm{ver}})_{[k]}$. These parts are joined via τ_k. More precisely,

$${}_k\mathcal{Z}^+_\bullet : 0 \to {}_k\mathcal{Z}^+_s \to \cdots \xrightarrow{\phi_k} {}_k\mathcal{Z}^+_{k+1} \xrightarrow{\tau_k} {}_k\mathcal{Z}^+_k \to \cdots \to {}_k\mathcal{Z}^+_0 \to 0, \tag{2-8}$$

wherein

$${}_k\mathcal{Z}^+_i = \begin{cases} (D_i)_{[k]}, & i \leq \min\{k, s\}, \\ H^r_{\mathfrak{g}}(D_{i+r-1})_{[k]}, & i > k. \end{cases} \tag{2-9}$$

The structure of ${}_k\mathcal{Z}^+_\bullet$ depends in two ways on the generating sets. Namely, it depends on the generating set of I, which is $\boldsymbol{f}$, and the expression of the generators of $\mathfrak{a}$ in terms of the generators of I, which are given by c_{ij}. However, for $k \geq 1$, we have

$$H_0({}_k\mathcal{Z}^+_\bullet) = H_0(D_\bullet)_{[k]} = (\mathcal{S}_I/(\gamma)\mathcal{S}_I)_{[k]} = \operatorname{Sym}^k_R(I/\mathfrak{a}).$$

The case where $k = 0$ is also very interesting. However, the structure of $H_0({}_0\mathcal{Z}^+_\bullet)$ is not as clear as the cases where $k > 0$.

Definition 2.1. Let R be a Noetherian ring and let $\mathfrak{a} \subseteq I$ be two ideals of R. The *disguised s-residual intersection of I with respect to* $\mathfrak{a}$ is the unique ideal K such that $H_0({}_0\mathcal{Z}_\bullet^+) = R/K$. The reasons for choosing the attribute *disguised* are as follows: K is contained in $J = (\mathfrak{a} : I)$ and it has the same radical as J has; if R is CM, J is an algebraic residual intersection and I satisfies some sliding depth condition, then K is Cohen–Macaulay; moreover, K coincides with J in the case where the residual is arithmetic, by [Hassanzadeh 2012, Theorem 2.11]. We conjecture (Conjecture 5.9) that under the above assumptions K is always the same as J despite that it does not appear so.

Acyclicity and Cohen–Macaulayness. One more occasion where the properties of ${}_k\mathcal{Z}_\bullet^+$ are independent of the generating sets of I and $\mathfrak{a}$ is the following lemma, which is crucial to proving the acyclicity of ${}_k\mathcal{Z}_\bullet^+$.

Lemma 2.2. *Let R be a Noetherian ring and let $\mathfrak{a} \subseteq I$ be two ideals of R. If $I = \mathfrak{a}$, then every complex ${}_k\mathcal{Z}_\bullet^+$ defined in* (2-8) *is exact.*

Proof. Since the approximation complex $\mathcal{Z}_\bullet(\boldsymbol{f})$ is a differential graded algebra, $H_i(\mathcal{Z}_\bullet)$ is an $\mathcal{S}_I = S/\mathcal{L}$-module for all i. Likewise, since the approximation complex $K_\bullet(\boldsymbol{\gamma})$ is a differential graded algebra, $H_i(K_\bullet(\boldsymbol{\gamma}))$ is an $S/(\boldsymbol{\gamma})$-module for all i. The bicomplex $K_\bullet(\boldsymbol{\gamma}) \otimes_S \mathcal{Z}_\bullet$ gives rise to the horizontal spectral sequence with second terms ${}^2\mathrm{E}_{\mathrm{hor}}^{-i,-j} = H_j(K_\bullet(\boldsymbol{\gamma};\, H_i(\mathcal{Z}_\bullet)))$. It follows that $(\mathcal{L} + (\boldsymbol{\gamma}))$ annihilates ${}^2\mathrm{E}_{\mathrm{hor}}^{-i,-j}$ and consequently annihilates ${}^\infty\mathrm{E}_{\mathrm{hor}}^{-i,-j}$, which is a subquotient of ${}^2\mathrm{E}_{\mathrm{hor}}^{-i,-j}$. By the convergence of the spectral sequence to the homologies of the total complex $H_\bullet(D_\bullet)$, it is straightforward to deduce that $H_i(D_\bullet)$ is an $(\mathcal{L} + (\boldsymbol{\gamma}))$-torsion module for all i, i.e., $(\mathcal{L} + (\boldsymbol{\gamma}))^N H_{i+j}(D_\bullet) = 0$ for all i, j and for some N. Considering the equation $I = \mathfrak{a}$ in $(\mathcal{S}_I)_1$, we have $\mathcal{L} + (\boldsymbol{\gamma}) = \mathcal{L} + \mathfrak{g}$. Therefore,

$$H^j_{\mathfrak{g}}(H_i(D_\bullet)) = H^j_{\mathfrak{g}+\mathcal{L}}(H_i(D_\bullet)) = H^j_{(\gamma)+\mathcal{L}}(H_i(D_\bullet)) = \begin{cases} H_i(D_\bullet), & j = 0, \\ 0, & j > 0. \end{cases} \tag{2-10}$$

Once more we study the spectral sequences arising from $C^\bullet_{\mathfrak{g}} \otimes D_\bullet$, which were already mentioned in (2-8). Applying (2-10), we draw both horizontal and vertical spectral sequences simultaneously as follows:

$$\begin{tikzcd}
{}^{\infty}\mathrm{E}_{\mathrm{hor}}: & 0 & \cdots & 0 & H_k(D_\bullet)_{[k]} & \cdots & H_0(D_\bullet)_{[k]} \\
{}^{1}\mathrm{E}_{\mathrm{ver}}: & 0 & \cdots & & 0 & \cdots & 0 \\
H^r_{\mathfrak{g}}(D_{r+s-1})_{[k]} \arrow[r] & \cdots \arrow[r, "\phi_k"] & H^r_{\mathfrak{g}}(D_{r+k})_{[k]} \arrow[r] \arrow[uurr, dashed, dash, "\tau_k"] & 0 & 0 & \cdots & 0
\end{tikzcd}$$

Since $H_j(D_\bullet)$ is a subquotient of D_j, we have $H_j(D_\bullet)_{[k]} = 0$ for all $j > k$. Also, by (2-10) there is only one nonzero row in the horizontal spectral. On the other hand, $\mathfrak{g}$ is a regular sequence on the D_j which in turn shows that the vertical spectral is just one line.

Consequently, $H_j({}_k\mathcal{Z}_\bullet^+) := H_{j+r-1}(H_\mathfrak{g}^r(D_\bullet))_{[k]} = H_{j-1}(D_\bullet)_{[k]} = 0$ whenever $j > k+1$. On the other hand, if $j < k$, then $H_j({}_k\mathcal{Z}_\bullet^+) := H_j(D_\bullet)_{[k]} = H_{j+r}(H_\mathfrak{g}^r(D_\bullet)_{[k]})$. However, the latter is zero since $\operatorname{end}(H_\mathfrak{g}^r(D_{r+j})) \leq j$. This shows that ${}_k\mathcal{Z}_\bullet^+$ is exact on the left and also on the right hand of τ_k.

It remains to prove the exactness in the joint points k and $k+1$. For this, notice that $\big({}^\infty\mathrm{E}_{\mathrm{hor}}^{-i,-j}\big)_{[k]} = 0$ for $j \geq 1$, hence $\mathcal{F}_{jk}/\mathcal{F}_{(j+1)k} = \big({}^\infty\mathrm{E}_{\mathrm{hor}}^{-k-j,-j}\big)_{[k]} = 0$ for $j \geq 1$. Consequently, $\mathcal{F}_{1k} = 0$ in (2-6). Therefore, the map τ_k defined in (2-7) is exactly the canonical map $H_\mathfrak{g}^r(D_{r+k})_{[k]} \xrightarrow{\text{Can.}} \operatorname{Coker}(\phi_k)_{[k]}$ as required. □

As already mentioned in the Introduction, some sliding depth conditions are needed to prove the acyclicity of the ${}_k\mathcal{Z}_\bullet^+$ complexes.

Definition 2.3. Let $(R, \mathfrak{m})$ be a Noetherian local ring of dimension d and let $I = (f_1, \dots, f_r)$ be an ideal of grade $g \geq 1$. Let k and t be two integers. We say that the ideal I satisfies SD_k at level t if $\operatorname{depth}(H_i(\boldsymbol{f}; R)) \geq \min\{d-g, d-r+i+k\}$ for all $i \geq r-g-t$. Whenever $t = r-g$, we simply say that I satisfies SD_k and we let SD stand for SD_0.

I is strongly Cohen–Macaulay, or SCM, if $H_i(\boldsymbol{f}; R)$ is CM for all i. Clearly SCM is equivalent to SD_{r-g}.

Similarly, we say that I satisfies the sliding depth condition on cycles, or SDC_k, at level t if $\operatorname{depth}(Z_i(\boldsymbol{f}, R)) \geq \min\{d-r+i+k, d-g+2, d\}$ for all $i \geq r-g-t$. Again if $t = r-g$, we simply say that I satisfies SDC_k and we use SDC instead of SDC_0.

Some of the basic properties and relations between conditions SD_k and SDC_k are explained in the following proposition.

Proposition 2.4. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d and let $I = (f_1, \dots, f_r)$ be an ideal of grade $g \geq 1$. Let k and t be two integers. Then*:

(1) *The properties* SDC_k *and* SD_k *at level t localize*; *they only depend on I and not the generating set*, *if $t = r-g$.*

(2) SD_k *implies* SDC_{k+1}.

(3) SD_k *at level $t < r-g$ implies* SDC_a *at level t for $a \leq g+t+2-d$.*

(4) SDC_{k+1} *at level t implies* SD_k *at level t for any t, if $g \geq 2$. This implication is also the case if $g = 1$ and $k = 0$.*

(5) SD_0 *at level $t \geq 1$ implies* SDC_0 *at level t.*

Proof. Part (1) is essentially proved in [Vasconcelos 1994]. Part (2) was proved in [Hassanzadeh 2012, Proposition 2.5]. Part (3) follows after analyzing the spectral sequences derived from the tensor product of the truncated Koszul complex

$$0 \to Z_i(f) \to K_i \to K_{i-1} \to \cdots \to K_0 \to 0$$

and the Čech complex $\mathcal{C}^{\bullet}_{\mathfrak{m}}(R)$. To prove (4), we consider the depth inequalities derived from the short exact sequences $0 \to Z_{i+1}(f) \to K_{i+1} \to B_i(f) \to 0$ and $0 \to B_i(f) \to Z_i(f) \to H_i(f) \to 0$. To show (5), we use the fact that SDC_k holds for any k at level -1, then a recursive induction applying $\operatorname{depth}(Z_i) \geq \min\{\operatorname{depth}(H_i), d, \operatorname{depth}(Z_{i+1}) - 1\}$ proves the assertion. □

In the next theorem we present sufficient conditions for the acyclicity of ${}_k\mathcal{Z}^+_{\bullet}$. The strategy taken here to prove the acyclicity is the same as the one applied in [Hassanzadeh 2012] (*however, the reader should notice that the complex* $\mathcal{C}_{\bullet}$ *defined in* [Hassanzadeh 2012] *is a little bit different from* ${}_0\mathcal{Z}^+_{\bullet}$ *here; in the former, the tail is substituted by a free complex but still* $\mathcal{C}_{\bullet}$ *remains quasi-isomorphic to* ${}_0\mathcal{Z}^+_{\bullet}$). Since the complex is finite, we avail ourselves of "lemme d'aciclicité" of Peskine and Szpiro: we assume some sliding depth conditions and prove the acyclicity in height $s-1$ wherein $I = \mathfrak{a}$. By Lemma 2.2, ${}_k\mathcal{Z}^+_{\bullet}$ is exact, if $I = \mathfrak{a}$. Then an induction will show the acyclicity globally.

Although the proof here is more involved than [Hassanzadeh 2012, Proposition 2.8], we prefer to omit it to go faster to newer theorems. The Cohen–Macaulay hypothesis in this theorem is needed to show that if for an R-module M we have $\operatorname{depth}(M) \geq d-t$, then for any prime $\mathfrak{p}$ we have $\operatorname{depth}(M_{\mathfrak{p}}) \geq \operatorname{ht}(\mathfrak{p})-t$ [Vasconcelos 1994, Section 3.3].

Proposition 2.5. *Let* $(R, \mathfrak{m})$ *be a CM local ring of dimension* d *and let* $J = (\mathfrak{a} : I)$ *be an* s*-residual intersection. Assume that* $I = (f_1, \ldots, f_r)$ *and* $\operatorname{ht}(I) = g \geq 1$. *Fix* $0 \leq k \leq \max\{s, s-g+2\}$. *Then the complex* ${}_k\mathcal{Z}^+_{\bullet}$ *is acyclic, if any of the following hypotheses holds*:

(1) $1 \leq s \leq 2$ *and* $s = k$, *or*

(2) $r+k \geq s+1,\ k \leq 2$, *and* I *satisfies* SDC_1 *at level* $s-g-k$, *or*

(3) $r+k \leq s$ *and* I *satisfies SD*, *or*

(4) $r+k \geq s+1,\ k \geq 3$, I *satisfies* SDC_1 *at level* $s-g-k$, *and* $\operatorname{depth}(Z_i(f)) \geq d-s+k$ *for* $0 \leq i \leq k$.

Consequently, having a finite acyclic complex whose components have sufficient depth, one can estimate the depth of the zeroth homology.

Theorem 2.6. *Let* $(R, \mathfrak{m})$ *be a CM local ring of dimension* d *and let* $I = (f_1, \ldots, f_r)$ *be an ideal with* $\operatorname{ht}(I) = g \geq 1$. *Let* $s \geq g$ *and fix* $0 \leq k \leq \min\{s, s-g+2\}$. *Suppose that one the following hypotheses holds*:

(i) $r+k \leq s$ *and* I *satisfies* SD, *or*

(ii) $r+k \geq s+1$, *I satisfies* SDC_1 *at level* $s-g-k$, *and* $\operatorname{depth}(Z_i(\boldsymbol{f})) \geq d-s+k$ *for* $0 \leq i \leq k$, *or*

(iii) $k \leq s-r+2$ *and I satisfies* SD, *or*

(iv) $\operatorname{depth}(H_i(\boldsymbol{f})) \geq \min\{d-s+k-2, d-g\}$ *for* $0 \leq i \leq k-1$ *and I satisfies* SD, *or*

(v) *I is strongly Cohen–Macaulay.*

Then for any s-residual intersection $J = (\mathfrak{a} : I)$, *the complex* ${}_k\mathcal{Z}_\bullet^+$ *is acyclic. Furthermore*, $\operatorname{Sym}_R^k(I/\mathfrak{a})$ *if* $1 \leq k \leq s-g+2$, *or the disguised residual intersection if* $k = 0$, *is CM of codimension s.*

In the above theorem, the cases where $k = 0, 1$ and $s - g + 1$ are the most important ones. For, $k = 0$ is related to the disguised residual intersection, $k = 1$ is connected to $R/\mathfrak{a}$, and $k = s - g + 1$ is related to the canonical module of R/J.

Remark 2.7. Any of the conditions (iii)–(v) are stronger than (i) and (ii). Condition (i) in Theorem 2.6 implies that $\operatorname{depth}(R/I) \geq d - s + k$ and condition (v) implies that $\operatorname{depth}(R/I) \geq d - g \geq d - s$. The other conditions, (ii), (iii) and (iv), yield $\operatorname{depth}(R/I) \geq d - s + k - 2$. Therefore, to have a better consequence of the theorem, it may not be too much to require $\operatorname{depth}(R/I) \geq d - s$.

The first two items in the next corollary settle conjectures (3) and (4) mentioned in the Introduction.

Corollary 2.8. *Let* $(R, \mathfrak{m})$ *be a CM local ring of dimension d and let* $I = (f_1, \ldots, f_r)$ *be an ideal with* $\operatorname{ht}(I) = g \geq 1$. *Suppose that* $\operatorname{depth}(R/I) \geq d-s$ *and that I satisfies any of the conditions in Theorem 2.6 for* $k = 1$, *for instance, if I satisfies SD. Then, for any algebraic s-residual intersection* $J = (\mathfrak{a} : I)$:

(1) $\operatorname{depth}(R/\mathfrak{a}) = d - s$ *and thus* $\mathfrak{a}$ *is minimally generated by s elements.*

(2) *J is unmixed of height s.*

(3) $\operatorname{Ass}(R/\mathfrak{a}) \subseteq \operatorname{Ass}(R/I) \cup \operatorname{Ass}(R/J)$. *Furthermore, if I is unmixed then the equality holds.*

(4) *If the residual is arithmetic, then* $\operatorname{Sym}_R^i(I/\mathfrak{a})$ *is a faithful R/J-module for all i.*

Proof. (1) Applying Theorem 2.6 for $k = 1$, we have $\operatorname{depth}(I/\mathfrak{a}) = d-s$. Therefore, the result follows from the exact sequence $0 \to I/\mathfrak{a} \to R/\mathfrak{a} \to R/I \to 0$ and the fact that $\operatorname{depth}(R/I) \geq d - s$ by hypothesis.

(2) J has codimension at least s, thus for any $\mathfrak{p} \in \operatorname{Ass}(R/J)$ we have $\operatorname{ht}(\mathfrak{p}) \geq s$. On the other hand, $\operatorname{Ass}(R/J) \subseteq \operatorname{Ass}(R/\mathfrak{a})$ and any prime ideal in the latter has codimension less than s by (1) (see [Bruns and Herzog 1998, Proposition 1.2.13]). This yields the unmixedness of J.

(3) Let $\mathfrak{p} \in \operatorname{Ass}(R/\mathfrak{a})$. Then $\operatorname{ht}(\mathfrak{p}) \leq s$, by (1). If $\operatorname{ht}(\mathfrak{p}) < s$ or $\mathfrak{p} \not\supseteq J$, then $\mathfrak{a}_\mathfrak{p} = I_\mathfrak{p}$ and thus $\mathfrak{p} \in \operatorname{Ass}(R/I)$; otherwise $\operatorname{ht}(\mathfrak{p}) = s$ and $\mathfrak{p} \supseteq J$, which clearly means $\mathfrak{p} \in \operatorname{Ass}(R/J)$. Verification of the last statement is straightforward.

(4) To see this, we refer to the proof of [Hassanzadeh 2012, Theorem 2.11] wherein it is shown that, in the spectral sequence defined in (2-3), $K := \big({}^{\infty}\mathrm{E}^{0,0}_{\mathrm{hor}}\big)_{[0]} \subseteq J \subseteq \big({}^{2}\mathrm{E}^{0,0}_{\mathrm{hor}}\big)_{[0]}$. Moreover, in the case where the residual is arithmetic, this spectral sequence, locally in height s, collapses in the second page. Hence all of these inclusions are equality. We just notice that

$$\big({}^{2}\mathrm{E}^{0,0}_{\mathrm{hor}}\big)_{[0]} = H^0_{\mathfrak{g}}(H_0(D_\bullet))_{[0]} = \bigcup_{i=0}^{\infty}(\boldsymbol{\gamma} \cdot \operatorname{Sym}^i_R(I) :_R \operatorname{Sym}^{i+1}_R(I)).$$

Thus once equality in the above line happens, $J = \boldsymbol{\gamma} \cdot \operatorname{Sym}^i_R(I) :_R \operatorname{Sym}^{i+1}_R(I)$ for all i. Therefore, $\operatorname{Sym}^i_R(I/\mathfrak{a}) = \operatorname{Sym}^{i+1}_R(I)/\boldsymbol{\gamma} \cdot \operatorname{Sym}^i_R(I)$ is a faithful R/J-module. □

The following examples show the accuracy of the conditions in Theorem 2.6 and Corollary 2.8(4).

Example 2.9. Let

$$R = \mathbb{Q}[x_1, \dots, x_7], \qquad M = \begin{pmatrix} 0 & x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 & x_7 \end{pmatrix}, \qquad \text{and} \qquad I = I_2(M).$$

Set

$$\begin{aligned} f_0 &= -x_4x_1, & f_3 &= -x_3x_4, \\ f_1 &= -x_2x_4, & f_4 &= x_1x_7 - x_3x_5, \\ f_2 &= x_1x_6 - x_2x_5, & f_5 &= x_2x_7 - x_3x_6. \end{aligned}$$

Macaulay2 computations show that $\operatorname{depth}(Z_1(I)) = 6, \operatorname{depth}(Z_2(I)) = 2$ and $\operatorname{depth}(Z_3(I)) = 6$, hence I satisfies SDC_1 at level 0 but not at level 1. However, Z_1 has enough depth, hence condition (ii) in Theorem 2.6 is fulfilled whenever $2 \leq s \leq 4$. The nontrivial case will be $s = 4$. Also, $\operatorname{depth}(R/I) = 4 > d - s = 3$. It then follows from Corollary 2.8 that for any 4-residual intersection $J = \mathfrak{a} : I$ we have $\operatorname{ht}(J) = 4 = \mu(\mathfrak{a})$. It is easy to see that I satisfies G_∞; hence [Eisenbud et al. 2004, Proposition 3.7bis] implies that $\ell(I) \geq 5$. On the other hand, the Plüker relations show that $f_2f_3 - f_1f_4 + f_0f_5 = 0$. Hence I is not generated by analytically independent elements; that is, $\ell(I) = 5$. Therefore, the same proposition in [loc. cit.] proves that there must exist a 5-residual intersection $J = \mathfrak{a} : I$ such that $\operatorname{ht}(J) > 5$. By the latter, Corollary 2.8 guarantees that, for that ideal $\mathfrak{a}$, the module $I/\mathfrak{a}$ cannot be CM and thus Theorem 2.6 does not work because SDC_1 at level 1 is not satisfied.

Another note about this example is that $\operatorname{depth}(R/I^2) = 2$, according to *Macaulay2*; therefore, $\operatorname{Ext}^5(R/I^2, R) \neq 0$. The vanishing of this Ext module would be sufficient to prove that I is 4-residually S_2 in [Chardin et al. 2001, Theorem 4.1], whereas we saw in the above corollary that I is 4-residually unmixed.

Example 2.10. The arithmetic hypothesis in Corollary 2.8(4) cannot be dropped. Let $R = \mathbb{Q}[x, y]$, $I = (x, y)$ and $\mathfrak{a} = (x^2, y^2)$. Then we have $J = (\mathfrak{a}, xy)$ and $\operatorname{Ann}(\operatorname{Sym}^2_R(I/\mathfrak{a})) = (\mathfrak{a}I : I^2) = (x, y)$. The latter is not J, since the residual (linkage) is not arithmetic.

In the case where the residual intersection is geometric, we have stronger corollaries.

Corollary 2.11. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d and let $I = (f_1, \dots, f_r)$ be an ideal with* $\operatorname{ht}(I) = g \geq 1$. *Suppose that* $\operatorname{depth}(R/I) \geq d - s$ *and that I satisfies any of the conditions in Theorem 2.6 for some $1 \leq k \leq s - g + 2$. Then, for any geometric s-residual intersection $J = (\mathfrak{a} : I)$ and for that k:*

(1) *$I^k/\mathfrak{a}I^{k-1}$ is isomorphic to* $\operatorname{Sym}^k_R(I/\mathfrak{a})$ *and it is a faithful maximal Cohen–Macaulay R/J-module.*

(2) *$\mathfrak{a}I^{k-1} = I^k \cap J$.*

(3) *$I^k + J$ is a CM ideal of height $s + 1$ and* $\operatorname{ht}(I^k + J/J) = 1$.

Proof.

(1) Consider the natural map
$$\varphi : \frac{\operatorname{Sym}^k_R(I)}{\boldsymbol{\gamma} \operatorname{Sym}^{k-1}_R(I)} \to \frac{I^k}{\mathfrak{a}I^{k-1}}$$
that sends T_i to f_i. The map φ is onto. Let $K = \operatorname{Ker}(\varphi)$. We show that $\operatorname{Ass}(K) = \varnothing$, hence $K = 0$. We have
$$\operatorname{Ass}(K) \subseteq \operatorname{Ass}\left(\frac{\operatorname{Sym}^k_R(I)}{\boldsymbol{\gamma} \operatorname{Sym}^{k-1}_R(I)}\right).$$
By Theorem 2.6,
$$\operatorname{depth}\left(\frac{\operatorname{Sym}^k_R(I)}{\boldsymbol{\gamma} \operatorname{Sym}^{k-1}_R(I)}\right) = d - s,$$
hence for any associated prime $\mathfrak{p}$ in the latter we have $\operatorname{ht}(\mathfrak{p}) \leq s$. However, since J is a geometric residual, for any $\mathfrak{p}$ of codimension $\leq s$, either $I_\mathfrak{p} = \mathfrak{a}_\mathfrak{p}$ or $I_\mathfrak{p} = (1)$. In the former case,
$$\frac{\operatorname{Sym}^k_R(I_\mathfrak{p})}{\boldsymbol{\gamma}_\mathfrak{p} \operatorname{Sym}^{k-1}_R(I_\mathfrak{p})} = 0$$
for $k \geq 1$, which implies that $\mathfrak{p} \notin \operatorname{Ass}(K)$. In the latter case, $\operatorname{Sym}^k_R(I_\mathfrak{p}) = R_\mathfrak{p}$, hence $\varphi_\mathfrak{p}$ is an isomorphism, which means $K_\mathfrak{p} = 0$; in particular, $\mathfrak{p} \notin \operatorname{Ass}(K)$. Therefore, $\operatorname{Ass}(K) = \varnothing$, hence $K = 0$. The CM-ness of $\operatorname{Sym}^k_R(I/\mathfrak{a})$ was already shown in

Theorem 2.6. To see that $I^k/\mathfrak{a}I^{k-1}$ is faithful, one may use Corollary 2.8 and part (1), or just notice that $J \subseteq \mathfrak{a}I^{k-1} : I^k$ and that the equality holds locally at all associated primes of J which are of height s.

(2) Consider the canonical map $\psi : I^k/(\mathfrak{a}I^{k-1}) \to (I^k + J)/J$. By (2), $I^k/\mathfrak{a}I^{k-1}$ is CM of dimension $d - s$, which in conjunction with the fact that the residual is geometric shows that ψ is an isomorphism for all $0 \le k \le s - g + 2$. Notice that $(I^k + J)/J \simeq I^k/(J \cap I^k)$. Thus the canonical map implies $J \cap I^k = \mathfrak{a}I^{k-1}$.

(3) By (1) and (2), $(I^k + J)/J$ is CM of dimension $d - s$. Also, by Theorem 2.6, R/J is CM of dimension $d - s$. Thus the exact sequence $0 \to (I^k + J)/J \to R/J \to R/(I^k + J) \to 0$ implies $\operatorname{depth}(R/I^k + J) \ge d - s - 1$. On the other hand, since the residual is geometric, $\operatorname{ht}(I^k + J) \ge s + 1$, which in turn completes the proof. To see that $\operatorname{ht}((I^k + J)/J) = 1$, it is enough to notice that R/J is CM. □

We now proceed to determine the structure of the canonical module of R/J. This achievement is via a study of the type of an appropriate candidate for the canonical module. Recall that in a Noetherian local ring $(R, \mathfrak{m})$ the type of a finitely generated module M is the dimension of the $R/\mathfrak{m}$-vector space $\operatorname{Ext}_R^{\operatorname{depth}(M)}(R/\mathfrak{m}, M)$, and it is denoted by $r_R(M)$ or just $r(M)$.

Theorem 2.12. *Let $(R,\mathfrak{m})$ be a CM local ring of dimension d and let $I = (f_1, \dots, f_r)$ be an ideal with $\operatorname{ht}(I) = g \ge 2$. Let $J = (\mathfrak{a} : I)$ be an s-residual intersection of I and let $1 \le k \le s - g + 1$. Suppose that I is SCM. Then*

$$r_R\big(\operatorname{Sym}_R^k(I/\mathfrak{a})\big) \le \binom{r + s - g - k}{r - 1} r_R(R).$$

Proof. First we deal with the case where $g = 2$. Consider the complex ${}_k\mathcal{Z}_\bullet^+$,

$$0 \to H_{\mathfrak{g}}^r(D_{r+s-1})_{[k]} \to \cdots \xrightarrow{\phi_k} H_{\mathfrak{g}}^r(D_{r+k})_{[k]} \xrightarrow{\tau_k} (D_k)_{[k]} \to \cdots \to (D_0)_{[k]} \to 0. \quad (2\text{-}11)$$

We study the depth of the components of this complex. Since $Z_{r-1} = R$, we have $\operatorname{depth}({}_k\mathcal{Z}_{r+s-1}^+) = d$. For $k \le j \le s - 2$, the first cycle which appears in D_{r+j} is Z_{r-s+j} and, by the SCM condition, it has depth equal to $d = d - g + 2$, which is at least $d - s + (j + 1) + 1$. As for the right part of the complex, we have $\operatorname{depth}(Z_i) = d - g + 2 \ge d - s + k + 1$ for all $0 \le i \le k$, since by assumption $k \le s - g + 1$.

Summing up, we have

$$\operatorname{depth}({}_k\mathcal{Z}_i^+) \ge d - s + i + 1 \quad \text{for all } i = 0, \dots, s - 1. \quad (2\text{-}12)$$

Now, let $F_\bullet \to R/\mathfrak{m}$ be a free resolution of $R/\mathfrak{m}$. We put the double complex $\operatorname{Hom}_R(F_\bullet, {}_k\mathcal{Z}_\bullet^+)$ in the third quadrant with $\operatorname{Hom}_R(F_0, {}_k\mathcal{Z}_0^+)$ in the center. The two arisen spectral sequences are

$$^{\infty}\mathrm{E}_{\mathrm{hor}}^{-i,-j}=\begin{cases}\mathrm{Ext}_R^j\big(R/\mathfrak{m},\mathrm{Sym}_R^k(I/\mathfrak{a})\big), & i=0,\\ 0, & i\neq 0,\end{cases}\quad\text{and}\quad {}^{1}\mathrm{E}_{\mathrm{ver}}^{-i,-j}=\mathrm{Ext}_R^j(R/\mathfrak{m},{}_k\mathcal{Z}_i^+).$$

The latter vanishes for all (i, j) such that $j-i\le d-s$ and $i\neq s$, according to (2-12). The convergence of both complexes to the homology module of the total complex implies that $\mathrm{Ext}_R^{d-s}\big(R/\mathfrak{m},\mathrm{Sym}_R^k(I/\mathfrak{a})\big)\simeq {}^{\infty}\mathrm{E}_{\mathrm{ver}}^{-s,-d}$. The latter is a submodule of $\mathrm{Ext}_R^d(R/\mathfrak{m},{}_k\mathcal{Z}_s^+)$. Therefore, $r_R\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big)\le\dim_{R/\mathfrak{m}}\big(\mathrm{Ext}_R^d(R/\mathfrak{m},{}_k\mathcal{Z}_s^+)\big)$.

Following from the construction of $\mathcal{D}_\bullet$, we have $D_{r+s-1}=S(-r+1-s)$. Hence

$${}_k\mathcal{Z}_s^+=H_{\mathfrak{g}}^r(S)_{[-r+1-s+k]}.$$

To calculate the dimension of the above Ext module, we just need to know how many copies of R appear in the inverse polynomial structure of the above local cohomology module. The answer is simply the number of positive solutions of the numerical equation $\alpha_1+\cdots+\alpha_r=r-1+s-k$, which is $\binom{r-k+s-2}{r-1}$. Therefore, $r_R\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big)\le\binom{r+s-2-k}{r-1}r(R)$ in the case where $g=2$.

Now, let $\mathrm{ht}(I)=g\ge 2$. We choose a regular sequence $\boldsymbol{a}$ of length $g-2$ inside $\mathfrak{a}$ which is a part of a minimal generating set of $\mathfrak{a}$. $I/\boldsymbol{a}$ is still SCM by [Huneke 1983, Corollary 1.5]; moreover, $J/\boldsymbol{a}=\mathfrak{a}/\boldsymbol{a}:I/\boldsymbol{a}$. Recall that $\mu(\mathfrak{a})=s$ by Corollary 2.8; therefore, $J/\boldsymbol{a}$ is an $(s-g+2)$-residual intersection of $I/\boldsymbol{a}$ which is of height 2. Hence we return to the case $g=2$ to obtain

$$\begin{aligned}r_R\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big)&=r_{R/\boldsymbol{a}}\big(\mathrm{Sym}_{R/\boldsymbol{a}}^k(I/\mathfrak{a})\big)\\&\le\binom{r+(s-g+2)-2-k}{r-1}r_{R/\boldsymbol{a}}(R/\boldsymbol{a})=\binom{r+s-g-k}{r-1}r_R(R)\end{aligned}$$

(see [Bruns and Herzog 1998, Exercise 1.2.26]). □

Remark 2.13. Clearly the above proof works for $k=0$ as well; hence if I is SCM of $\mathrm{ht}(I)\ge 2$ and K is a disguised s-residual intersection of I, then

$$r_R(R/K)\le\binom{r+s-g}{r-1}r_R(R).$$

Everything is now ready for us to explain the structure of the canonical module.

Theorem 2.14. *Suppose that $(R,\mathfrak{m})$ is a Gorenstein local ring of dimension d. Let I be an SCM ideal with $\mathrm{ht}(I)=g$ and let $J=(\mathfrak{a}:I)$ be an arithmetic s-residual intersection of I. Then $\omega_{R/J}$ is isomorphic to $\mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})$. It is isomorphic to $(I^{s-g+1}+J)/J$ in the geometric case. Furthermore, R/J is generically Gorenstein and R/J is Gorenstein if and only if $I/\mathfrak{a}$ is principal.*

Proof. Without loss of generality, we may suppose that $g\ge 2$; otherwise we may add a new variable x to R and consider ideals $\mathfrak{a}+(x)$ and $I+(x)$ in the local ring $R[x]_{\mathfrak{m}+(x)}$ wherein we have $J+(x)=(\mathfrak{a}+(x)):(I+(x))$.

R/J is CM by Theorem 2.6. Respectively, by Theorem 2.6, Corollary 2.8 and Theorem 2.12, $\mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})$ is a maximal CM, faithful R/J-module and of type 1. Therefore, it is the canonical module of R/J by [Bruns and Herzog 1998, Proposition 3.3.13]. In the geometric case, $\mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a}) \simeq (I^{s-g+1}+J)/J$ by Corollary 2.11.

To see that R/J is generically Gorenstein, notice that for any prime ideal $\mathfrak{p} \supseteq J$ of height s we have $r((R/J)_\mathfrak{p}) = \mu((\omega_{R/J})_\mathfrak{p}) = \mu\big(\mathrm{Sym}_{R_\mathfrak{p}}^{s-g+1}((I/\mathfrak{a})_\mathfrak{p})\big)$, which is 1 since the residual is arithmetic by the assumption. Finally, let $t = \mu(I/\mathfrak{a})$; then $r(R/J) = \mu(\omega_{R/J}) = \binom{s-g+t}{t-1}$ which is 1 if and only if $t = 1$. □

3. Hilbert function, Castelnuovo–Mumford regularity

Although the family $\{{}_k\mathcal{Z}_\bullet^+\}$ does not consist of free complexes, it can approximate R/J and $\mathrm{Sym}_R^k(I/\mathfrak{a})$ very closely. Several numerical invariants or functions such as Castelnuovo–Mumford regularity, projective dimension and the Hilbert function can be estimated via this family. Chardin, Eisenbud and Ulrich [Chardin et al. 2001] restated an old question [Stanley 1980] asking for which open sets of ideals $\mathfrak{a}$ the Hilbert function of $R/\mathfrak{a}$ depends only on the degrees of the generators of $\mathfrak{a}$. In [Chardin et al. 2001], the authors consider the following two conditions:

(A1) if the Hilbert function of $R/\mathfrak{a}$ is constant on the open set of ideals $\mathfrak{a}$ generated by s forms of the given degrees such that $\mathrm{codim}(\mathfrak{a} : I) \geq s$; and

(A2) if the Hilbert function of $R/(\mathfrak{a} : I)$ is constant on this set.

It is shown in [Chardin et al. 2001, Theorem 2.1] that ideals with some slight depth conditions in conjunction with G_{s-1} or G_s satisfy these two conditions. In this direction we have the following result which, in addition to showing the validity of conditions (A1) and (A2) under some sliding depth condition, provides a method for computing the desired Hilbert functions.

Proposition 3.1. *Let R be a CM graded ring over an Artinian local ring R_0, and let I and $\mathfrak{a}$ be two homogeneous ideals of R. Let $J = (\mathfrak{a} : I)$ be an s-residual intersection. Let $0 \leq k \leq s-g+2$ and suppose that I satisfies any of the following hypotheses*:

(1) $1 \leq s \leq 2$ *and* $s = k$, *or*

(2) $r+k \geq s+1$, $k \leq 2$, *and I satisfies* SDC_1 *at level* $s-g-k$, *or*

(3) $r+k \leq s$ *and I satisfies SD, or*

(4) $r+k \geq s+1$, $k \geq 3$, *I satisfies* SDC_1 *at level* $s-g-k$, *and* $\mathrm{depth}(Z_i(\boldsymbol{f})) \geq d-s+k$ *for* $0 \leq i \leq k$.

Then the Hilbert function of the disguised s-residual intersection, if $k=0$, and that of $\mathrm{Sym}_R^k(I/\mathfrak{a})$, *if* $1 \leq k \leq s-g+2$, *depends only on the* ***degrees*** *of the generators* *of* $\mathfrak{a}$ *and the Koszul homologies of I. In particular, if I satisfies any of the above conditions for* $k = 1$ *then the Hilbert function of* $R/\mathfrak{a}$ *satisfies* (A1).

Proof. By Proposition 2.5, the standing assumptions on I imply the acyclicity of ${}_k\mathcal{Z}_\bullet^+$. Hence the Hilbert function of $H_0({}_k\mathcal{Z}_\bullet^+)$ is derived from the Hilbert functions of the components of ${}_k\mathcal{Z}_\bullet^+$ which, according to (2-1) and (2-9), are just some direct sums of the Koszul cycles of I shifted by the twists appearing in the Koszul complex $K_\bullet(\boldsymbol{\gamma}, S)$. Since the Hilbert functions of Koszul cycles are inductively calculated in terms of those of the Koszul homologies, the Hilbert function of $H_0({}_k\mathcal{Z}_\bullet^+)$ depends on the Koszul homologies of I and just the degrees of the generators of $\mathfrak{a}$. In the case where $k = 1$, we know the Hilbert function of $I/\mathfrak{a} = \mathrm{Sym}_R^1(I/\mathfrak{a})$ and therefore we know the Hilbert function of $R/\mathfrak{a}$ by $0 \to I/\mathfrak{a} \to R/\mathfrak{a} \to R/I \to 0$. □

The next important numerical invariant associated to an algebraic or geometric object is Castelnuovo–Mumford regularity. Here we present an upper bound for the regularity of the disguised s-residual intersection K and the regularity of $\mathrm{Sym}_R^k(I/\mathfrak{a})$ (if $1 \le k \le s - g + 1$).

Assume that $R = \bigoplus_{n \ge 0} R_n$ is a positively graded *local Noetherian ring of dimension d over a Noetherian local ring $(R_0, \mathfrak{m}_0)$, and set $\mathfrak{m} = \mathfrak{m}_0 + R_+$. Suppose that I and $\mathfrak{a}$ are homogeneous ideals of R generated by homogeneous elements $f_1, \dots, f_r$ and $a_1, \dots, a_s$, respectively. Let $\deg f_t = i_t$ for all $1 \le t \le r$ with $i_1 \ge \cdots \ge i_r$ and $\deg a_t = l_t$ for $1 \le t \le s$. For a graded ideal $\mathfrak{b}$, the sum of the degrees of a minimal generating set of $\mathfrak{b}$ is denoted by $\sigma(\mathfrak{b})$.

Proposition 3.2. *With the same notation just introduced above, let* $(R, \mathfrak{m})$ *be a CM *local ring, let I be an ideal with* $\mathrm{ht}(I) = g \ge 2$ *and let* $s \ge g$. *Suppose that* $J = (\mathfrak{a} : I)$ *is an s-residual intersection of I and K is the disguised s-residual intersection of I with respect to* $\mathfrak{a}$. *Suppose that any of the following conditions hold for some* $0 \le k \le s - g + 1$:

(i) $r + k \le s$ *and I satisfies* SD_1, *or*

(ii) $r+k \ge s+1$, *I satisfies* SDC_2 *at level* $s-g-k$, *and* $\mathrm{depth}(Z_i(\boldsymbol{f})) \ge d-s+k+1$ *for* $0 \le i \le k$.

Then if $k \ge 1$,

$$\mathrm{reg}\big(\mathrm{Sym}_R^k(I/\mathfrak{a})\big) \le \mathrm{reg}(R) + \dim(R_0) + \sigma(\mathfrak{a}) - (s - g + 1 - k)\,\mathrm{indeg}(I/\mathfrak{a}) - s.$$

And if $k = 0$,

$$\mathrm{reg}(R/K) \le \mathrm{reg}(R) + \dim(R_0) + \sigma(\mathfrak{a}) - (s - g + 1)\,\mathrm{indeg}(I/\mathfrak{a}) - s.$$

Proof. The proof of this fact is essentially the same as that of [Hassanzadeh 2012, Theorem 3.6] wherein the case of $k = 0$ is verified. A crucial part of repeating that proof is that we have to change the structure of the families $\{{}_k\mathcal{Z}_\bullet^+\}$ from the beginning by substituting the tails with a free complex. Another way to prove this result in the case where I is SCM may be by applying the "$g = 2$" trick in Theorem 2.12. $\square$

The inequality in the above proposition becomes more interesting in particular cases; for example, if R is a polynomial ring over a field and $k = 1$, this inequality reads as

$$\operatorname{reg}_R(I/\mathfrak{a}) + (s-g)\operatorname{indeg}(I/\mathfrak{a}) \leq \sigma(\mathfrak{a}) - s.$$

Hence we have a numerical obstruction for $J = (\mathfrak{a} : I)$ to be an s-residual intersection.

At the other end of the spectrum we have $k = s - g + 1$. This case corresponds to the canonical module of R/J.

The next proposition improves [Hassanzadeh 2012, Proposition 3.5] by removing the G_s condition.

Proposition 3.3. *Suppose that* $(R, \mathfrak{m})$ *is a positively graded Gorenstein *local ring, over a Noetherian local ring* $(R_0, \mathfrak{m}_0)$, *with canonical module* $\omega_R = R(b)$ *for some integer b. Let I be a homogeneous SCM ideal with* $\operatorname{ht}(I) = g$, *let* $\mathfrak{a} \subset I$ *be homogeneous and let* $J = (\mathfrak{a} : I)$ *be an arithmetic s-residual intersection of I. Then* $\omega_{R/J} = \operatorname{Sym}_R^{s-g+1}(I/\mathfrak{a})(b + \sigma(\mathfrak{a}))$.

In particular, if $\dim(R_0) = 0$ *then*

$$\operatorname{reg}(R/J) = \operatorname{reg}(R) + \sigma(\mathfrak{a}) - \operatorname{indeg}\big(\operatorname{Sym}_R^{s-g+1}(I/\mathfrak{a})\big) - s.$$

Proof. We may and do assume that $g \geq 2$. First suppose that $g = 2$. We keep the notation defined at the beginning of Section 2 to define the complex ${}_k\mathcal{Z}_\bullet^+$. Considering R as a subalgebra of $S = R[T_1, \dots, T_r]$, we write the degrees of an element x of R as the 2-tuple $(\deg x, 0)$ with the second entry zero. Therefore, $\deg f_t = (i_t, 0)$ for all $1 \leq t \leq r$, $\deg a_t = (l_t, 0)$ for all $1 \leq t \leq s$, $\deg T_t = (i_t, 1)$ for all $1 \leq t \leq r$, and thus $\deg \boldsymbol{\gamma}_t = (l_t, 1)$ for all $1 \leq t \leq s$. With this notation the $\mathcal{Z}$-complex has the shape

$$\mathcal{Z}_\bullet : 0 \to Z_{r-1} \otimes_R S(0, -r+1) \to \cdots \to Z_1 \otimes_R S(0, -1) \to Z_0 \otimes_R S \to 0.$$

Consequently,

$$\mathcal{Z}_{r-1} = R\Big(-\sum_{t=1}^{r} i_t, 0\Big) \otimes_R S(0, -r+1),$$

and, taking into account the fact that $a_1, \dots, a_s$ is a minimal generating set of $\mathfrak{a}$ by Corollary 2.8,

$$D_{r+s-1} = S\Big(-\sum_{t=1}^{r} i_t - \sigma(\mathfrak{a}), -s-r+1\Big).$$

It then follows that ${}_{(s-1)}\mathcal{Z}_s^+ = H_{\mathfrak{g}}^r(D_{r+s-1})_{[*,s-1]}$ is isomorphic to

$$H_{\mathfrak{g}}^r(S)\Bigl(-\sum_{t=1}^{r} i_t - \sigma(\mathfrak{a}), 0\Bigr)_{[*,-r]} \simeq \Bigl(RT_1^{-1}\cdots T_r^{-1}\Bigl(-\sum_{t=1}^{r} i_t\Bigr)\Bigr)(-\sigma(\mathfrak{a})) \simeq R(-\sigma(\mathfrak{a})). \quad (3\text{-}1)$$

Notice that $\deg(T_j^{-1}) = (-i_j, -1)$.

We next turn to the proof of Theorem 2.12 and consider the spectral sequences therein for $k = s - g + 1 = s - 1$. We have ${}^1\mathrm{E}_{\mathrm{ver}}^{-s,-d} = \mathrm{Ext}_R^d(R/\mathfrak{m}, {}_{(s-1)}\mathcal{Z}_s^+)$ which is isomorphic to $\mathrm{Ext}_R^d(R/\mathfrak{m}, R(b))(-b - \sigma(\mathfrak{a}))$ by (3-1). The latter is in turn homogeneously isomorphic to $R/\mathfrak{m}(-b-\sigma(\mathfrak{a}))$, since R is Gorenstein. The inclusion

$$\mathrm{Ext}_R^{d-s}\bigl(R/\mathfrak{m}, \mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})\bigr) \simeq {}^{\infty}\mathrm{E}_{\mathrm{hor}}^{0,-(d-s)} \hookrightarrow {}^1\mathrm{E}_{\mathrm{ver}}^{-s,-d} = \mathrm{Ext}_R^d(R/\mathfrak{m}, {}_{(s-1)}\mathcal{Z}_s^+)$$

is indeed an isomorphism, since the latter is a vector space of dimension one. In the graded case, all of the homomorphisms in the proof of Theorem 2.12 are homogeneous; therefore, the above isomorphism implies $\mathrm{Ext}_R^{d-s}\bigl(R/\mathfrak{m}, \mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})\bigr) \simeq R/\mathfrak{m}(-b-\sigma(\mathfrak{a}))$. We already know that in the local case $\omega_{R/J} = \mathrm{Sym}_R^{s-2+1}(I/\mathfrak{a})$, by Theorem 2.14, hence in the graded case we have $\omega_{R/J} = \mathrm{Sym}_R^{s-2+1}(I/\mathfrak{a})(b+\sigma(\mathfrak{a}))$.

Now suppose that $g \geq 2$ and let $\boldsymbol{a} = a_1, \ldots, a_{g-2}$ be a regular sequence in $\mathfrak{a}$ which is a part of its minimal generating set. Then $\omega_{R/\boldsymbol{a}} = (R/\boldsymbol{a})(b + \sigma(\boldsymbol{a}))$ and, moreover, we go back to the case where $g = 2$. We then have

$$\omega_{R/J} = \mathrm{Sym}_{R/\boldsymbol{a}}^{(s-g+2)-2+1}\Bigl(\frac{I/\boldsymbol{a}}{\mathfrak{a}/\boldsymbol{a}}\Bigr)((b+\sigma(\boldsymbol{a})) + \sigma(\mathfrak{a}/\boldsymbol{a})) = \mathrm{Sym}_R^{s-g+1}(I/\mathfrak{a})(b+\sigma(\mathfrak{a})). \quad \square$$

Finally, one may ask whether under the G_s condition our methods can contribute more information about the structure of residual intersections than what is known so far. The next result shows that in the presence of the G_s condition the complex ${}_0\mathcal{Z}_\bullet^+$ approximates R/J for any (algebraic) s-residual intersection J. In particular, the above facts about regularity and the Hilbert function hold for R/J.

Proposition 3.4. *Let $(R, \mathfrak{m})$ be a Gorenstein local ring and suppose that I is SCM and evenly linked to an ideal which satisfies G_s. Then the disguised s-residual intersections of I, with respect to an ideal $\mathfrak{a}$, is the same as the algebraic residual intersection $J = \mathfrak{a} : I$.*

Proof. It is shown in [Huneke and Ulrich 1988, Theorem 5.3] that under the above assumptions on I there exists a deformation (R', I') of (R, I) such that I' is SCM and G_s. Let $I = (f_1, \ldots, f_r)$, $\mathfrak{a} = (a_1, \ldots, a_s)$, and suppose that

$a_i = \sum_{j=1}^{r} c_{ij} f_j$. Let s_{ij} be a lifting of c_{ij} to R' and let f_i' be a lifting of f_i to R'. Let $X = (x_{ij})$ be an $s \times r$ matrix of invariants and define $\boldsymbol{\alpha} := (\alpha_1, \dots, \alpha_s) = X \cdot (f_1', \dots, f_r')^t$. Let $\tilde{R} = R'[x_{ij}]_{(\mathfrak{m}+(x_{ij}-s_{ij}))}$ and $\tilde{J} := (\alpha_1, \dots, \alpha_s) :_{\tilde{R}} I'\tilde{R}$. By [Huneke and Ulrich 1988, Lemma 3.2], $\tilde{J}$ is a geometric s-residual intersection of $I'\tilde{R}$. We construct the complex ${}_0\mathcal{Z}_\bullet^+\big((\alpha_1, \dots, \alpha_s); (f_1', \dots, f_s')\big)$ in $\tilde{R}$ and set $\tilde{R}/\tilde{K} = H_0\big({}_0\mathcal{Z}_\bullet^+((\alpha_1, \dots, \alpha_s); (f_1', \dots, f_s'))\big)$. Since $\tilde{J}$ is a geometric residual, we have $\tilde{J} = \tilde{K}$.

Assume that π is the projection from $\tilde{R}$ to R which sends x_{ij} to c_{ij}. By [Hassanzadeh 2012, Proposition 2.13], $R/\pi(\tilde{K}) = H_0({}_0\mathcal{Z}_\bullet^+)$ and, by [Huneke and Ulrich 1988, Theorem 4.7], $\pi(\tilde{J}) = J$. Therefore, $R/\pi(\tilde{K}) = R/\pi(\tilde{J}) = R/J$. □

4. Arithmetic residual intersections

In this section we scrutinize the structure of the ${}_0\mathcal{Z}_\bullet^+$ complex in the case where $I/\mathfrak{a}$ is cyclic. We find that the homologies of ${}_0\mathcal{Z}_\bullet^+$ determine the uniform annihilator of nonzero Koszul homologies. The importance of this fact is that, although the Koszul complex is an old object, not much is known about the annihilator of higher homologies. Besides the rigidity and differential graded algebra structure of the homologies, there is an interesting paper of Corso, Huneke, Katz and Vasconcelos [Corso et al. 2006] wherein the authors make efforts to find out "whether the annihilators of nonzero Koszul homology modules of an unmixed ideal is contained in the integral closure of that ideal". However, it seems that the situation of mixed ideals is more involved; for example, for $\mathfrak{a} = (x^2 - xy, y^2 - xy, z^2 - zw, w^2 - zw)$, (taken from [loc. cit.]), we have $\mathrm{Ann}(H_1) = \bar{\mathfrak{a}}$, the integral closure, and $\mathrm{Ann}(H_2) = \sqrt{\mathfrak{a}}$. This example shows that the intersection of the annihilators has a better structure than the union. This is an interesting example for us since $(\mathfrak{a} : \bar{\mathfrak{a}})$ is a 4-residual intersection, with $\bar{\mathfrak{a}}/\mathfrak{a}$ cyclic and $\bar{\mathfrak{a}}$ satisfying SD.

Since a part of our results works well with arithmetic s-residual intersections, it is worth mentioning some general properties of these ideals.

Proposition 4.1. *Let R be a Noetherian ring, let $\mathfrak{a} = (a_1, \dots, a_s) \subset I$ be an ideal of R and let $J = \mathfrak{a} : I$ be an s-residual intersection. Then*:

(i) *J is an arithmetic s-residual intersection if and only if* $\mathrm{ht}\big(0 : \bigwedge^2(I/\mathfrak{a})\big) \geq s+1$.

(ii) *If* 2 *is unit in R, then* $I + J \subseteq \big(0 : \bigwedge^2(I/\mathfrak{a})\big)$.

(iii) *The following statements are equivalent*:

  (1) *For all $i \leq s$, an i-residual intersection of I exists.*

  (2) *For any prime ideal* $\mathfrak{p} \supset I$ *such that* $\mathrm{ht}(\mathfrak{p}) \leq s-1$, *we have* $\mu(I_\mathfrak{p}) \leq \mathrm{ht}(\mathfrak{p}) + 1$.

  (3) *There exist $a_1, \dots, a_s \in I$ such that, for $i \leq s-1$, $J_i = (a_1, \dots, a_i) : I$ is an arithmetic i-residual intersection and $J = (a_1, \dots, a_s) : I$ is an* (*algebraic*) *s-residual intersection.*

(iv) *J is an arithmetic s-residual intersection if and only if there exists an element $b \in I$ such that* $\operatorname{ht}((\mathfrak{a}, b) : I) \geq s+1$.

Proof. For (i), $J = \mathfrak{a} : I$ is an arithmetic s-residual if and only if $\operatorname{ht}(\operatorname{Fitt}_j(I/\mathfrak{a})) \geq s+j$ for $j = 0, 1$ [Eisenbud 1995, Proposition 20.6]. Moreover, $\operatorname{Fitt}_j(I/\mathfrak{a})$ and the annihilator of $\bigwedge^{j+1}(I/\mathfrak{a})$ have the same radical [Eisenbud 1995, Exercise 20.10] which yields the assertion. Part (ii) is straightforward. Part (iii) goes along the same line as the proof of [Ulrich 1994, Lemma 1.4]. Part (iv) follows from [Ulrich 1994, Lemma 1.3]. □

The concept of s-parsimonious ideals first appeared in [Chardin et al. 2001], where it is shown in Proposition 3.1 that weakly s-residually S_2 ideals which satisfy G_s are s-parsimonious. However, if I is an ideal which satisfies the conditions of Corollary 2.8 then it follows from [Chardin et al. 2001, Proposition 3.3(a)] that I is s-parsimonious.

Remark 4.2. Let $(R, \mathfrak{m})$ be a CM local ring of dimension d, and let I be an ideal such that $\operatorname{depth}(R/I) \geq d - s$ and such that I satisfies any of the conditions in Theorem 2.6 for $k = 1$. Then for any arithmetic s-residual intersection $J = (\mathfrak{a} : I)$, there exists an element $b \in I$ such that $J = (\mathfrak{a} : b)$.

This remark is further motivation to study the cases where $I/\mathfrak{a}$ is cyclic.

To exploit the structure of $H_\bullet({}_0\mathcal{Z}^+_\bullet(\mathfrak{a}; \mathfrak{a}, b))$, the following lemma is crucial.

Lemma 4.3. *Let* $\mathfrak{a} = (a_1, \dots, a_s)$, $I = (f) = (b, a_1, \dots, a_s)$, *and consider the complex $\mathcal{D}_\bullet$ defined in* (2-1). *Let $Z'_\bullet$, $B'_\bullet$ and $H'_\bullet$ be the cycles, boundaries and homologies of the Koszul complex $K_\bullet(a_1, \dots, a_s)$. Put $\tilde{B}_i = (B'_i :_{Z'_i} b)$ and $S = R[T_0, T_1, \dots, T_s]$ with standard grading. Then*:

(i) *We have*

$$H_i(\mathcal{D}_\bullet) \cong \frac{(Z'_i \otimes_R R[T_0])(-i)}{(\tilde{B}_i \otimes_R T_0 R[T_0])(-i)},$$

and, in particular,

$$(H_i(\mathcal{D}_\bullet)_{T_0})_{[0]} \cong b H_i(a_1, \dots, a_s) \quad \text{for all } i.$$

(ii) *If R is a graded ring and $b, a_1, \dots, a_s$ are homogeneous, then we have a homogeneous isomorphism*

$$(H_i(\mathcal{D}_\bullet)_{T_0})_{[0]} \cong b H_i(a_1, \dots, a_s)((i+1)\deg(b)) \quad \text{for all } i.$$

Proof. (i) To compute the homology modules of $\mathcal{D}_\bullet$, we consider the two spectral sequences arising from its double complex structure (2-1). We put this double complex in the second quadrant and locate $K_0(\boldsymbol{\gamma}) \otimes_S \mathcal{Z}_0(\boldsymbol{f})$ in the corner.

Set $S = R[T_0, T_1, \dots, T_s]$, where T_0 corresponds to b and T_i to a_i. Then $(\gamma_1 \cdots \gamma_s) = (T_1 \cdots T_s)$ is a regular sequence on $\mathcal{Z}_i = Z_i[T_0, T_1, \dots, T_s](-i)$, with

$Z_i = Z_i(b, a_1, \ldots, a_s)$. We then have ${}^1\mathrm{E}_{\mathrm{ver}}^{-i,j} = H_j(K_\bullet(\boldsymbol{\gamma}) \otimes \mathcal{Z}_i) = 0$ for all $j \geq 1$ and all i. We also have $(S/(T_1, T_2, \ldots, T_s)) \otimes_S \mathcal{Z}_i = R[T_0] \otimes_R Z_i$, in which we consider $R[T_0]$ as an S-module by the trivial multiplication $T_i R[T_0] = 0$, $1 \leq i \leq s$. It then follows that ${}^1\mathrm{E}_{\mathrm{ver}}^{0,\bullet}$ is the complex

$$0 \to Z_{s+1} \otimes_R R[T_0](-s-1) \xrightarrow{\partial_{s+1}} \cdots \to Z_1 \otimes_R R[T_0](-1) \xrightarrow{\partial_1} Z_0 \otimes_R R[T_0] \to 0, \quad (4\text{-}1)$$

wherein the differentials are induced by those in $\mathcal{Z}_\bullet(\boldsymbol{f})$. As well,

$$H_i\big({}^1\mathrm{E}_{\mathrm{ver}}^{0,\bullet}\big) = H_i\big(\mathcal{Z}_\bullet \otimes_S (S/(T_1, T_2, \ldots, T_s))\big)(-i). \quad (4\text{-}2)$$

Considering the convergence of the spectral sequences to the homology module of the total complex $\mathcal{D}_\bullet$, we have $H_i(\mathcal{D}_\bullet) = H_i\big({}^1\mathrm{E}_{\mathrm{ver}}^{0,\bullet}\big) = \mathrm{Ker}(\partial_i)/\mathrm{Im}(\partial_{i+1})$. We compute this homology in two steps.

Step 1: $(\mathrm{Ker}(\partial_i))$ Let $z_i = \big(\sum_j r_j e_{j_1} \wedge e_{j_2} \wedge \cdots \wedge e_{j_i}\big)\omega(T_0) \in Z_i[T_0](-i)$. We study two cases: first, if $j_1 \neq 0$ for every j which appears in the above presentation of z_i, then we have $z_i \in Z'_i[T_0]$ and, moreover,

$$\partial_i(z_i) = \sum_j r_j \left(\sum_{l=1}^{i} e_{j_1} \wedge \cdots \wedge e_{j_{l-1}} \wedge e_{j_{l+1}} \wedge \cdots \wedge e_{j_i} T_l \omega(T_0)\right) = 0,$$

where the last vanishing is due to the fact that $T_i R[T_0] = 0$ for all $1 \leq i \leq s$. It follows that $z_i \in \mathrm{Ker}(\partial_i)$.

Second, we show that if $j_1 = 0$ for some j that appears in z_i then $z_i \notin \mathrm{Ker}(\partial_i)$. Writing $z_i = z'_i + z''_i \wedge e_0$, where z'_i and z''_i do not involve e_0, one has

$$\partial_i(z_i) = \partial_i(z'_i) + \partial_{i-1}(z''_i) \wedge e_0 + (-1)^{i+1} T_0 z''_i$$

by the DG-algebra structure of $(\mathcal{Z}_\bullet, \partial'_\bullet)$. Since z'_i and z''_i do not involve e_0, it follows that $\partial_i(z'_i) = 0$ and $\partial_{i-1}(z''_i) = 0$ by the multiplication on $R[T_0]$. Thus $\partial_i(z_i) = (-1)^{i+1} T_0 z''_i \neq 0$. Therefore, $\ker(\partial_i) = Z_i(a_1, \ldots, a_s)[T_0](-i) = Z'_i[T_0](-i)$.

Step 2: $(\mathrm{Im}(\partial_{i+1}))$ To compute the image, we consider an arbitrary element

$$z_{i+1}\omega(T_0) = \left(\sum_j r_j e_{j_1} \wedge e_{j_2} \wedge \cdots \wedge e_{j_{i+1}}\right)\omega(T_0) \in Z_{i+1}[T_0](-i-1)$$

and write $z_{i+1} = z'_{i+1} + z''_{i+1} \wedge e_0$, where z'_{i+1} and z''_{i+1} do not involve e_0. Recall that $Z_{i+1} = Z_{i+1}(\boldsymbol{f})$. Consider the Koszul complex $(K_\bullet(\boldsymbol{f}), d_\bullet)$. We have

$$0 = d(z_{i+1}) = d(z'_{i+1}) + d(z''_{i+1}) \wedge e_0 + (-1)^i b z''_{i+1}.$$

Since z'_{i+1} and z''_{i+1} do not involve e_0, the above equation implies that $d(z''_{i+1}) = 0$ and $b z''_{i+1} = d((-1)^{i+1} z'_{i+1})$, hence $z''_{i+1} \in (B'_i :_{Z'_i} b)$.

Conversely, for any $z''_{i+1} \in (B'_i :_{Z'_i} b)$, we have $bz''_{i+1} = d((-1)^{i+1} z'_{i+1})$ for some $z'_{i+1} \in K'_{i+1}$, hence $z_{i+1} = (z'_{i+1} + z''_{i+1} \wedge e_0) \in Z_{i+1}$.

On the other hand, $\partial(z_{i+1}) = \partial(z'_{i+1}) + \partial(z''_{i+1} \wedge e_0) = 0 + \partial(z''_{i+1} \wedge e_0) = (-1)^i T_0 z''_{i+1}$.

It follows that $\operatorname{Im}(\partial_{i+1}) = (T_0 \tilde{B}_i[T_0])(-i)$. Finally, the above two steps show that

$$H_i(\mathcal{D}_\bullet) = \frac{(Z'_i \otimes_R R[T_0])(-i)}{(\tilde{B}_i \otimes_R T_0 R[T_0])(-i)}.$$

To see the last assertion, just notice that

$$((Z'_i[T_0]_{T_0})(-i))_{[0]} = Z'_i\big[T_0, T_0^{-1}\big]_{[-i]} = T_0^{-i} Z'_i,$$

and

$$((\tilde{B}_i T_0[T_0]_{T_0})(-i))_{[0]} = T_0^{-i} \tilde{B}_i.$$

Therefore,

$$(H_i(\mathcal{D}_\bullet)_{T_0})_{[0]} \simeq Z'_i/\tilde{B}_i = Z'_i/(B'_i :_{Z'_i} b) = H'_i/(0 :_{H'_i} b) \simeq bH_i(a_1, \dots, a_s). \quad (4\text{-}3)$$

(ii) In the graded case, every homomorphism in the above argument is homogeneous except the first and last isomorphisms in (4-3). To see the desired shift, notice that the bidegree of T_0 is $(\deg(b), 1)$. Hence, considering $(H_i(\mathcal{D}_\bullet)_{T_0})$ in bidegree $(0, 0)$, we have

$$(H_i(\mathcal{D}_\bullet)_{T_0})_{[(0,0)]} \simeq \frac{T_0^{-i}\big((Z'_i)_{[i \deg(b)]}\big)}{T_0^{-i}\big((\tilde{B}_i)_{[i \deg(b)]}\big)} \simeq \left(\frac{Z'_i}{\tilde{B}_i}\right)_{[i \deg(b)]}.$$

Moreover, there exists the homogeneous exact sequence

$$0 \to \tilde{B}_i \to Z'_i \to bH_i(a_1, \dots, a_s)(\deg(b)) \to 0,$$

which yields the assertion. $\square$

Now, we are ready to explain our main theorem in this section. The dependence of the homology modules of ${}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})$ on the generating sets becomes clear by this theorem; furthermore, it motivates the study of the uniform annihilator of the Koszul homology modules.

Theorem 4.4. *Let R be a (Noetherian) ring, and let $\mathfrak{a} = (\boldsymbol{a}) = (a_1, \dots, a_s)$ and $I = (\boldsymbol{f}) = (b, a_1, \dots, a_s)$. Then $H_i({}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})) \simeq bH_i(a_1, \dots, a_s)$ for all $i \geq 1$ and $H_0({}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})) \simeq R/(\mathfrak{a} : b)$; furthermore, in the graded case we have $H_i({}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})) \simeq bH_i(a_1, \dots, a_s)((i+1)\deg(b))$ for $i \geq 1$.*

Proof. We study the spectral sequences arising from the third-quadrant double complex $C^\bullet_{\mathfrak{g}} \otimes \mathcal{D}_\bullet$ with $C^0_{\mathfrak{g}} \otimes D_0$ in the center. Since $T_0, T_1, \dots, T_s$ is a regular sequence on the D_i, we have $({}^1\mathrm{E}_{\mathrm{ver}})_{[0]} = H^{s+1}_{\mathfrak{g}}(\mathcal{D}_\bullet)_{[0]}$. The latter is $\mathcal{Z}^+_\bullet$ by definition (2-5), which is the truncated complex ${}_0\mathcal{Z}^+_\bullet(\boldsymbol{a}, \boldsymbol{f})_{\geq 1}$.

On the other hand, in the horizontal spectral sequence we deal with $H^j_{\mathfrak{g}}(H_i(\mathcal{D}_\bullet))$. $H_i(\mathcal{D}_\bullet)$ is annihilated by $(T_1, \dots, T_s)$ by (4-2), which implies that $H^j_{\mathfrak{g}}(H_i(\mathcal{D}_\bullet)) = H^j_{(T_0)}(H_i(\mathcal{D}_\bullet))$ for all i and j. Also notice that $H^0_{(T_0)}(H_i(\mathcal{D}_\bullet))_{[0]} = 0$ for $i \geq 1$, since $\operatorname{indeg}(D_i) = i$. Hence

$$\left({}^2\mathrm{E}^{-i,-j}_{\mathrm{hor}}\right)_{[0]} = \begin{cases} H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]}, & i = 0 \text{ and } j = 0, \\ H^1_{(T_0)}(H_i(\mathcal{D}_\bullet))_{[0]} = (H_i(\mathcal{D}_\bullet)_{T_0})_{[0]}, & i \geq 1 \text{ and } j = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4-4}$$

Considering both spectrals at the same time, we have:

$$\begin{array}{lcccc}
({}^2\mathrm{E}_{\mathrm{hor}})_{[0]} : & & 0 & 0 & H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} \\
 & \cdots & H^1_{(T_0)}(H_2(\mathcal{D}_\bullet))_{[0]} & H^1_{(T_0)}(H_1(\mathcal{D}_\bullet))_{[0]} & H^1_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} \\
 & \cdots & 0 & 0 & 0 \\
\mathcal{Z}^+_\bullet = ({}^1\mathrm{E}_{\mathrm{ver}})_{[0]} : & & \vdots & \vdots & \vdots \\
\end{array}$$

$$H^{s+1}_{\mathfrak{g}}(\mathcal{D}_{s+1})_{[0]} \longrightarrow H^{s+1}_{\mathfrak{g}}(\mathcal{D}_s)_{[0]} \longrightarrow H^{s+1}_{\mathfrak{g}}(\mathcal{D}_{s-1})_{[0]} \longrightarrow \cdots$$

By the convergence of the spectral sequences, we have $H_{i+1}({}_0\mathcal{Z}^+_\bullet) = H_i(\mathcal{Z}^+_\bullet) = H^1_{(T_0)}(H_{i+1}(\mathcal{D}_\bullet))_{[0]}$ for $i \geq 1$. Quite generally, $H^1_{(T_0)}(H_{i+1}(\mathcal{D}_\bullet))_{[0]}$ is isomorphic to $(H_{i+1}(\mathcal{D}_\bullet)_{T_0})_{[0]}$, which in turn is isomorphic to $bH_{i+1}(a_1, \dots, a_s)$ by Lemma 4.3. In the case where $i = 0$, the convergence of the spectral sequences provides the exact sequence

$$0 \to H^1_{(T_0)}(H_1(\mathcal{D}_\bullet))_0 \to H_0(\mathcal{Z}^+_\bullet) \xrightarrow{\psi} H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_0 \to 0. \tag{4-5}$$

According to (2-5), $H_0(\mathcal{Z}^+_\bullet) = \operatorname{Coker}(\phi_0)$, and ψ is used to define τ_0 in (2-7). Therefore, $\operatorname{Ker}(\psi) = H_1({}_0\mathcal{Z}^+_\bullet)$ and its image determines $H_0({}_0\mathcal{Z}^+_\bullet)$, which we will compute below.

Since the horizontal spectral stabilizes in the second step, we have ${}^\infty\mathrm{E}^{0,0}_{\mathrm{hor}} = {}^2\mathrm{E}^{0,0}_{\mathrm{hor}} = H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))$. Moreover, $\operatorname{end}(H^{s+1}_{\mathfrak{g}}(\mathcal{D}_{s+1})) = 0$, thus $\left({}^\infty\mathrm{E}^{0,0}_{\mathrm{hor}}\right)_{[i]} = 0$ for $i \geq 1$, which in particular implies that $H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[1]} = 0$. Specifically, we then have $T_0 H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} = 0$. Recall that $H_0(\mathcal{D}_\bullet) = \operatorname{Sym}(I)/\boldsymbol{\gamma} \operatorname{Sym}(I)$ and that we consider $\mathcal{S}_I := \operatorname{Sym}_R(I)$ as an $S = R[T_0, \dots, T_s]$-module via the ring homomorphism $S \to \mathcal{S}_I$ sending T_0 to b and T_i to a_i as an element of $(\mathcal{S}_I)_{[1]} = I$. Hence $T_0 H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} = 0$ as an element in $(\mathrm{Sym}(I)/\boldsymbol{\gamma}\,\mathrm{Sym}(I))_{[1]} = I/\mathfrak{a}$ is equivalent to saying that $H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} \subseteq (\mathfrak{a} : b)$.

On the other hand, $H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} = (\mathfrak{a} : b) \cup \big(\bigcup_{i=1}^{\infty} \big(\mathfrak{a}\,\mathrm{Sym}^i_R(I) :_R \mathrm{Sym}^{i+1}_R(I)\big)\big)$, hence $H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_0 \supseteq (\mathfrak{a} : b)$, which yields $H^0_{(T_0)}(H_0(\mathcal{D}_\bullet))_{[0]} = (\mathfrak{a} : b)$.

Further, by Lemma 4.3, $H^1_{(T_0)}(H_1(\mathcal{D}_\bullet))_{[0]} = (H_1(\mathcal{D}_\bullet)_{T_0})_{[0]} = bH_1(f_1, f_2, \dots, f_s)$. Substituting these facts in the short exact sequence (4-5), the assertion follows. □

As a summary we have the following corollary.

Corollary 4.5. *Let R be a (Noetherian) ring, and let $\mathfrak{a} = (\boldsymbol{a}) = (a_1, \dots, a_s)$ and $I = (\boldsymbol{f}) = (b, a_1, \dots, a_s)$. If $bH_i(a_1, \dots, a_s) = 0$ for all $i \geq 1$, then there exists an exact complex*

$$0 \to \mathcal{Z}^+_s \to \mathcal{Z}^+_{s-1} \to \cdots \to \mathcal{Z}^+_0 \to (\mathfrak{a} :_R b) \to 0,$$

wherein $\mathcal{Z}^+_i = \bigoplus_{j=i+1}^{r} Z_j(\boldsymbol{f})^{\oplus n_j}$ for some positive integers n_j and r. In particular, the index of the lowest cycle which appears in the components increases along the complex.

Theorem 4.4, in conjunction with other conditions which imply the acyclicity of the complex ${}_0\mathcal{Z}^+_\bullet$, yields nontrivial facts about the uniform annihilators of nonzero Koszul homologies.

Corollary 4.6. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d, and suppose that $I = (b, a_1, \dots, a_s)$ satisfies SD. If $\dim(bH_i(a_1, \dots, a_s)) \leq d - s$ for all $i \geq 1$ then*

- *$bH_i(a_1, \dots, a_s) = 0$ for all $i \geq 1$, and*
- *$\mathrm{depth}(R/((a_1, \dots, a_s) : I)) \geq d - s$.*

Proof. We consider the complex ${}_0\mathcal{Z}^+_\bullet(\mathfrak{a}; (b, \mathfrak{a}))$. By Theorem 4.4, the hypothesis on the dimensions implies that this complex is acyclic locally at codimension $s - 1$. Then, appealing to the acyclicity lemma, the sliding depth hypothesis shows that the complex is acyclic on the punctured spectrum and, finally, acyclic (the same technique which we used to prove Proposition 2.5). Once more applying Theorem 4.4, we conclude the first assertion. The depth inequality follows from the proof of Theorem 2.6. □

As a final remark in this section, we notice that if $(\mathfrak{a} : I)$ is an s-residual intersection then $\dim(IH_i(\mathfrak{a})) \leq d - s$ for all i. Thus, the latter property may be considered as a generalization of algebraic residual intersection.

5. Uniform annihilator of Koszul homologies

Motivated by the results in the previous section, we concentrate on the uniform annihilator of Koszul homologies in this section. In the main theorem of this section, we see that in any residual intersection $J = \mathfrak{a} : I$ the sliding depth condition passes from I to $\mathfrak{a}$. This fact was known to experts only in the presence of the G_∞ condition [Ulrich 1994, 1.8(3)+1.12], which is not that surprising since I is then generated by a d-sequence. However, it was unknown even for perfect ideals of height 2 which are not G_∞; see [Eisenbud and Ulrich 2016] for such an example.

The next interesting result in this section is Corollary 5.6, in which we show that, for any residual intersection $J = \mathfrak{a} : I$ with I satisfying sliding depth, I is contained in the uniform annihilator of the nonzero Koszul homologies of $\mathfrak{a}$. Hence this is a kind of universal property for such ideals.

The next proposition is fundamental to proving the aforementioned results, though it is also interesting by itself.

Proposition 5.1. *Let* $(R, \mathfrak{m})$ *be a CM local ring of dimension* d, *and let* $\mathfrak{a} = (f_1, f_2, \dots, f_s)$ *and* $I = (f_1, f_2, \dots, f_r)$ *be ideals with* $s \le r$. *Let* $J = (\mathfrak{a} : I)$ *and* $g = \mathrm{ht}(I) = \mathrm{ht}(\mathfrak{a}) \ge 1$. *Let* n *be an integer and suppose that* $\mathrm{ht}(J) \ge n \ge 1$ *and that* $\mathrm{depth}(H_l(f_1, f_2, \dots, f_r)) \ge d - r + j$ *for* $l \ge r - n + 2$. *Then, for any* k *with* $r \ge k \ge s$ *and any* $j \ge k - n + 2$:

(i) $\mathrm{depth}(H_j(f_1, f_2, \dots, f_k) \ge d - k + j$, *and*

(ii) $I \subseteq \mathrm{Ann}(H_j(f_1, f_2, \dots, f_k))$.

Both assertions hold for sequences of the form $(f_1, \dots, f_s, f_{i_1}, \dots, f_{i_{k-s}})$, *where* $\{f_{i_1}, \dots, f_{i_{k-s}}\} \subseteq \{f_{s+1}, \dots, f_r\}$.

Proof. The proof is by a *recursive* induction on k. By assumption, the result is true for $k = r$. Suppose that it is true for $k \le r$. We now apply a recursive induction on $j \ge (k-1) - n + 2$ to show the following claims:

(1) $f_k H_j(f_1, f_2, \dots, f_{k-1}) = 0$;

(2) $\mathrm{depth}(H_j(f_1, f_2, \dots, f_{k-1})) \ge d - (k-1) + j$;

(3) $\mathrm{depth}(Z_j(f_1, f_2, \dots, f_{k-1})) \ge d - (k-1) + j$;

(4) $\mathrm{depth}(B_j(f_1, f_2, \dots, f_{k-1})) \ge d - (k-1) + j$.

Since these claims are clear for $j > k - g$, we consider an integer $j \ge (k-1) - n + 3$ and we suppose, by induction hypothesis, that the result holds for j. A simple depth-chasing between cycles, boundaries and homologies shows that (2)–(4) for $j-1$ follow from (1) for $j-1$ and the induction hypotheses of (2)–(4) for j. We just notice that condition (1) provides the exact sequence

$$0 \to H_j(f_1, \dots, f_{k-1}) \to H_j(f_1, \dots, f_k) \to H_{j-1}(f_1, \dots, f_{k-1}) \to 0.$$

We then proceed to prove (1) for $j-1$ while assuming the four claims hold for j. To do this, we show that $f_k H_{j-1}(f_1, f_2, \dots, f_{k-1})_\mathfrak{p} = 0$ by an induction on $\mathrm{ht}(\mathfrak{p})$. For $\mathrm{ht}(\mathfrak{p}) < n$ it is clear, since $I_\mathfrak{p} = a_\mathfrak{p}$. Now let $\mathfrak{q}$ be a prime ideal with $\mathrm{ht}(\mathfrak{q}) \ge n$

and suppose that (1)–(4) hold locally for all $\mathfrak{p}$ such that $\mathfrak{p} \subset \mathfrak{q}$. We may, and will, replace $(R, \mathfrak{m})$ by $(R_{\mathfrak{q}}, \mathfrak{q}R_{\mathfrak{q}})$. Considering the exact sequence

$$0 \to Z_j(f_1, \dots, f_{k-1}) \to \bigwedge^j (R^{k-1}) \to B_{j-1}(f_1, \dots, f_{k-1}) \to 0,$$

out of the Koszul complex, we obtain

$$\operatorname{depth}(B_{j-1}) \geq \min\Bigl\{\operatorname{depth}\Bigl(\bigwedge^j (R^{k-1})\Bigr), \operatorname{depth}(Z_j) - 1\Bigr\} \geq d - (k-1) + j - 1.$$

By assumption, $j \geq k - n + 2$, which implies that $\operatorname{depth}(B_{j-1}(f_1, \dots, f_{k-1})) \geq d - n + 2 \geq 2$. This time,

$$0 \to B_{j-1}(f_1, \dots, f_{k-1}) \to Z_{j-1}(f_1, \dots, f_{k-1}) \to H_{j-1}(f_1, \dots, f_{k-1}) \to 0$$

implies that $\operatorname{depth}(H_{j-1}) \geq 1$, since $\operatorname{depth}(Z_{j-1}) \geq 1$, as $\dim(R_{\mathfrak{q}}) \geq n \geq 1$.

We now consider the exact sequence

$$0 \to H_j(f_1, \dots, f_{k-1}) \to H_j(f_1, \dots, f_{k-1}, f_k) \to \Gamma_{j-1} \to 0,$$

where $\Gamma_{j-1} = (0 :_{H_{j-1}(f_1, \dots, f_{k-1})} f_k)$.

The induction hypotheses imply that

$$\begin{aligned}\operatorname{depth}(\Gamma_{j-1}) &\geq \min\{\operatorname{depth}(H_j(f_1, \dots, f_{k-1}, f_k)), \operatorname{depth}(H_j(f_1, \dots, f_{k-1})) - 1\}\\ &\geq \min\{d - k + j, d - (k-1) + j - 1\} \geq d - k + j\\ &\geq d - k + (k - n + 2) = d - n + 2 \geq 2.\end{aligned}$$

At last, we have the inclusion $\Gamma_{j-1} \subseteq H_{j-1}$, where the former has depth at least 2 and the latter has depth at least 1, while the equality holds on the punctured spectrum, by the induction hypothesis. Therefore, this inclusion must be an equality, which proves claim (1) for $j - 1$.

Applying the above argument for $k = r$, we have $\operatorname{depth}(H_j(f_1, \dots, f_{r-1})) \geq d - (r-1) + j$ for all $j \geq (r-1) - n + 2$, and $f_r H_j(f_1, f_2, \dots, f_{r-1}) = 0$. However, as long as we fix $f_1, \dots, f_s$ we may change the role of f_r with any one of $\{f_{s+1}, \dots, f_r\}$. In other words, we have $\operatorname{depth}(H_j(f_1, \dots, f_s, \dots, \hat{f_i}, \dots, f_r)) \geq d - (r-1) + j$ for all $j \geq (r-1) - n + 2$, and $f_i H_j(f_1, \dots, f_s, \dots, \hat{f_i}, \dots, f_r) = 0$. notice that, for any ideal $\mathfrak{a} \subseteq I' \subseteq I$, we have $\operatorname{ht}(\mathfrak{a} : I') \geq \operatorname{ht}(\mathfrak{a} : I) \geq n$. Therefore, if (i) holds for $(f_1, \dots, f_s, f_{i_1}, \dots, f_{i_{k-s}})$, where $\{f_{i_1}, \dots, f_{i_{k-s}}\} \subseteq \{f_{s+1}, \dots, f_r\}$, then the above argument shows that (i) holds for $(f_1, \dots, f_s, f_{i_1}, \dots, f_{i_{k-s-1}})$ and, moreover, that $f_{i_{k-s}} H_j(f_1, \dots, f_s, f_{i_1}, \dots, f_{i_{k-s-1}}) = 0$ for $j \geq k - 1 - n + 2$. Since $f_{i_{k-s}}$ varies over the set $\{f_{s+1}, \dots, f_r\}$, we must have $IH_j(f_1, \dots, f_s, f_{i_1}, \dots, f_{i_{k-s-1}}) = 0$ for $j \geq k - 1 - n + 2$, which completes the proof. □

We next want to show, in Lemma 5.3, that under the sliding depth condition there exists a similar depth inequality for cycles of the Koszul complex. For this, we need the following lemma which is reminiscent of the inductive Koszul long exact sequence.

Lemma 5.2. *Let R be a commutative ring, let $\{f_0, f_1, \dots, f_r\} \subseteq R$, $Z_\bullet$, $B_\bullet$ be the Koszul cycles and boundaries with respect to the sequence $\{f_1, f_2, \dots, f_r\}$, let $Z'_\bullet$ be the Koszul cycles with respect to the sequence $(f_0, f_1, \dots, f_r)$ and let $\Gamma_\bullet = (B_\bullet :_{Z_\bullet} f_0)$. Then, for any $j \geq 0$, there exists an exact sequence*

$$0 \to Z_j \to Z'_j \to \Gamma_{j-1} \to 0.$$

Proof. The Koszul complex $K_\bullet(f_1, f_2, \dots, f_r)$ is canonically a subcomplex of $K_\bullet(f_0, f_1, \dots, f_r)$. We will denote the differential of both complexes by d. Let $\{e_0, e_1, \dots, e_r\}$ be the canonical basis of R^{r+1}. Every $\theta \in Z'_j$ can be written uniquely as a sum $\theta = e_0 \wedge w + v$, where e_0 does not appear in terms w and v. Since $\theta \in Z'_j$,

$$0 = d(\theta) = d(e_0 \wedge w + v) = f_0 w - e_0 \wedge d(w) + d(v).$$

It follows that

$$\begin{cases} e_0 \wedge d(w) = 0 \\ f_0 w = -d(v) \end{cases} \quad \text{hence} \quad \begin{cases} d(w) = 0 \\ f_0 w = -d(v) \end{cases} \quad \text{therefore} \quad w \in \Gamma_{j-1}.$$

We define $\varphi : Z'_j \to \Gamma_{j-1}$ by $\varphi(\theta) = \varphi(e_0 \wedge w + v) = w$. The homomorphism φ is well-defined, as the expression is unique. It is onto, since for any $w \in \Gamma_{j-1}$ there is a $v \in K_j(f_1, f_2, \dots, f_r)$ such that $f_0 w = -d(v)$. Hence $\theta = e_0 \wedge w + v \in K_j(f_0, f_1, \dots, f_r)$ is the preimage of w.

Moreover, $Z_j = \ker(\varphi)$. On one hand, if $\theta \in Z'_j$ belongs to Z_j, then there is no e_0 in its decomposition in terms of the canonical basis of R^{r+1}. Consequently, $\theta = e_0 \wedge w + v = v$ and $w = 0$. On the other hand, if $0 = d(\theta) = d(e_0 \wedge w + v) = w$, then $\theta = v$. Therefore, $\theta \in Z'_j \cap K_j(f_1, f_2, \dots, f_r) = Z_j$. □

The next lemma is needed in the proof of Theorem 5.4. Notice that the sliding depth condition does not pass automatically from homologies to cycles because there is a level restriction; see Proposition 2.4.

Lemma 5.3. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d. Let $\mathfrak{a} = (f_1, f_2, \dots, f_s)$ and $I = (f_1, f_2, \dots, f_r)$ be ideals with $s \leq r$, and let $J = (\mathfrak{a} : I)$ and $g = \mathrm{ht}(I) = \mathrm{ht}(\mathfrak{a}) \geq 1$. Suppose that $ht(J) \geq n \geq 1$ and that I satisfies* SD. *Then, for $k \geq s$ and $j \geq k - n + 2$,*

$$\mathrm{depth}(Z_j(f_1, f_2, \dots, f_k)) \geq d - k + j + 1.$$

Proof. By Proposition 5.1, $IH_j(f_1, f_2, \dots, f_k) = 0$ whenever $s \leq k \leq r$ and $j \geq k - n + 2$. Therefore, by Lemma 5.2, we have the following exact sequence for $k - 1 \geq s$ and $j - 1 \geq (k-1) - n + 2$:

$$0 \to Z_j(f_1, f_2, \dots, f_{k-1}) \to Z_j(f_1, f_2, \dots, f_k) \to Z_{j-1}(f_1, f_2, \dots, f_{k-1}) \to 0.$$

We will prove the result by a recursive induction on k. If $k = r$, then it follows from Proposition 2.4. Suppose that the result holds for $k \geq s + 1$. Since $Z_j(f_1, f_2, \dots, f_{k-1})$ satisfies SDC_1 for $j \geq (k-1) - g + 1$, we can use a new induction on j to conclude that

$$\operatorname{depth}(Z_{j-1}(f_1, f_2, \dots, f_{k-1})) \geq d - (k-1) + (j-1) + 1 = d - k + j + 1. \quad \square$$

We are now ready to prove the main theorem of this section. As we mentioned at the beginning of the section, this fact was known only in the presence of the G_∞ condition, under which the problem reduced to studying the properties of ideals generated by d-sequences. Here the only assumption is the sliding depth and that $\operatorname{depth}(R/I) \geq d - s$.

Theorem 5.4. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d and let $J = (\mathfrak{a} : I)$ be an s-residual intersection with $s \geq \operatorname{ht}(I) \geq 1$. If I satisfies* SD *and* $\operatorname{depth}(R/I) \geq d - s$, *then $\mathfrak{a}$ satisfies* SD *as well.*

Proof. We apply Lemma 5.3 for $n = s$. It then follows that

$$\operatorname{depth}(Z_j(f_1, f_2, \dots, f_s)) \geq d - s + j + 1$$

for $j \geq 2$. The hypothesis $\operatorname{depth}(R/I) \geq d - s$ enables us to apply Corollary 2.8, which in turn implies that $\operatorname{depth}(R/\mathfrak{a}) \geq d - s$. Therefore, $\operatorname{depth}(\mathfrak{a}) \geq d - s + 1$ and hence $\operatorname{depth}(Z_1(\mathfrak{a})) \geq d - s + 2$; that is, $\mathfrak{a}$ satisfies SDC_1, which is equivalent to SD by Proposition 2.4. $\square$

In the next proposition we prove a partial converse of Theorem 5.4.

Proposition 5.5. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d, and let $\mathfrak{a} = (f_1, f_2, \dots, f_s)$ and $I = (f_1, f_2, \dots, f_r)$ with $s \leq r$. Let $J = (\mathfrak{a} : I)$ be an s-residual intersection. If $\mathfrak{a}$ satisfies* SD, *then, for any $s \leq k \leq r$ and $j \geq k - s + 1$:*

(i) $\operatorname{depth}(H_j(f_1, f_2, \dots, f_k)) \geq d - k + j$, *and*

(ii) $\operatorname{depth}(Z_j(f_1, f_2, \dots, f_k)) \geq d - k + j + 1$.

Proof. We will use induction on k. If $k = s$, then (ii) is clear by Proposition 2.4. Suppose that the result holds for $k \geq s + 1$. Then, by the induction hypothesis, we have

$$\operatorname{depth}(H_j(f_1, f_2, \dots, f_k)) \geq d - k + j \geq d - s + 1.$$

By a similar argument as in the proof of Proposition 5.1, $IH_j(f_1, f_2, \dots, f_k) = 0$ for all $j \geq k - s + 1$. Then we have the short exact sequences

$$0 \to H_j(f_1, f_2, \dots, f_k) \to H_j(f_1, f_2, \dots, f_{k+1}) \to H_{j-1}(f_1, f_2, \dots, f_k) \to 0$$

and

$$0 \to Z_j(f_1, f_2, \dots, f_k) \to Z_j(f_1, f_2, \dots, f_{k+1}) \to Z_{j-1}(f_1, f_2, \dots, f_k) \to 0,$$

provided by Lemma 5.2 for any $j \geq k+1-s+1$.

A depth-chasing then completes the proof. □

The next corollary of Theorem 5.4 shows a tight relation between the uniform annihilator of nonzero Koszul homology modules and residual intersections.

Corollary 5.6. *Let $(R, \mathfrak{m})$ be a CM local ring of dimension d, and let I satisfy* SD *and* $\operatorname{depth}(R/I) \geq d-s$ *with* $s \geq \operatorname{ht}(I) \geq 1$. *Let $J = (\mathfrak{a} : I)$ be an s-residual intersection and use $H_j(\mathfrak{a})$ to denote the j-th Koszul homology module with respect to a minimal generating set of $\mathfrak{a}$. Then*

$$I \subseteq \bigcap_{j \geq 1} \operatorname{Ann}(H_j(\mathfrak{a})). \qquad (\star)$$

Furthermore, the equality happens in the following cases, if $\operatorname{depth}(R/I) \geq d-s+1$ *(hence $s \geq g+1$):*

- *If $\mu(I_\mathfrak{p}) \leq s-i$ for a fixed positive integer i and for all $\mathfrak{p} \in \operatorname{Ass}(R/I)$, then $I = \operatorname{Ann}(H_j(\mathfrak{a}))$ for all $1 \leq j \leq i$.*
- *If R is Gorenstein and $\operatorname{Ass}(R/I) = \min(R/I)$, then $I = \bigcap_{j \geq 1} \operatorname{Ann}(H_j(\mathfrak{a}))$.*

Proof. By Corollary 2.8, $\mathfrak{a}$ is minimally generated by s elements. According to Theorem 5.4, for any j, we have $\operatorname{depth}(H_j(\mathfrak{a})) \geq d-s+j$. In particular, $\operatorname{ht}(\mathfrak{p}) \leq s-1$ for any $\mathfrak{p} \in \operatorname{Ass}(H_j(\mathfrak{a}))$ with $j \geq 1$; see [Bruns and Herzog 1998, 1.2.13].

To show that $IH_j(\mathfrak{a}) = 0$, we show that it has no associated prime. Notice that $IH_j(\mathfrak{a}) \subseteq H_j(\mathfrak{a})$, hence, for any $\mathfrak{p} \in \operatorname{Ass}(IH_j(\mathfrak{a}))$, we have $\operatorname{ht}(\mathfrak{p}) \leq s-1$ for which $I_\mathfrak{p} = \mathfrak{a}_\mathfrak{p}$. Therefore, $(IH_j(\mathfrak{a}))_\mathfrak{p} = 0$ as desired.

To see the equality, we localize at associated primes of I which have height at most $s-1$, so that $I_\mathfrak{p} = \mathfrak{a}_\mathfrak{p}$ for them. Now let $(a_1, \ldots, a_s)$ be a minimal generating set of $\mathfrak{a}$. If $\mu(I_\mathfrak{p}) \leq s-i$ then $H_j(a_1, \ldots, a_s)_\mathfrak{p}$ has a direct summand of $(R/\mathfrak{a})_\mathfrak{p}$ for all $1 \leq j \leq i$, and thus its annihilator is $\mathfrak{a}_\mathfrak{p} = I_\mathfrak{p}$. In the case where R is Gorenstein and $\operatorname{Ass}(R/I) = \min(R/I)$, we see that, locally at the associated primes of R/I, $I_\mathfrak{p}$ is unmixed and, moreover, $(a_1, \ldots, a_s)_\mathfrak{p}$ is not a regular sequence since $s > \operatorname{ht}(\mathfrak{p})$. Hence, for $t = \operatorname{ht}(I_\mathfrak{p})$, $\operatorname{Ann}(H_{s-t}(\mathfrak{a}_\mathfrak{p}))$ is the unmixed part of $\mathfrak{a}_\mathfrak{p} = I_\mathfrak{p}$ which is $I_\mathfrak{p}$. □

The next amazing example shows a benefit of relaxing the G_s condition in our theorems. However, we took this example from [Eisenbud and Ulrich 2016] wherein the authors used it to show the necessity of the G_s condition in their theorems.

Example 5.7. Let $(R, \mathfrak{m}) = \mathbb{Q}[x_1, \ldots, x_5]_{(x_1, \ldots, x_5)}$ and let $I = I_4(N)$, where

$$N = \begin{pmatrix} x_2 & x_3 & x_4 & x_5 \\ x_1 & x_2 & x_3 & x_4 \\ 0 & x_1 & x_2 & 0 \\ 0 & 0 & x_1 & x_2 \\ 0 & 0 & 0 & x_1 \end{pmatrix}.$$

Let Δ_i be the minors obtained by omitting the i-th row of N. We then consider $\mathfrak{a}$ to be the ideal generated by entries of

$$(\Delta_1 \ -\Delta_2 \ \ \Delta_3 \ -\Delta_4 \ \ \Delta_5)\begin{pmatrix} x_1 & 0 & 0 & 0 & 0 \\ 0 & x_5 & 0 & 0 & 0 \\ x_3 & x_4 & x_5 & x_3 & 0 \\ 0 & x_3 & x_4 & x_5 & 0 \\ 0 & x_2 & x_3 & x_4 & x_5 \end{pmatrix}.$$

$J=(\mathfrak{a}:I)$ is a 5-residual intersection of I, and I is a perfect ideal of height 2 hence it is SCM. Easily, one can see that any prime that contains I must contain $\mathfrak{p}=(x_1,x_2)$; that is, I is $\mathfrak{p}$-primary, since it is unmixed. The minimal presentation of I is given by N. Therefore, $\mathrm{Fitt}_4(I)=\mathfrak{m}$, and $\mathrm{Fitt}_3(I)\not\subset\mathfrak{p}$ because $x_4^2-x_3x_5\notin\mathfrak{p}$, and $\mathrm{Fitt}_2(I)\subseteq\mathfrak{p}$ because any combination of 3 “diagonal” elements contains either 0 or x_1 or x_2. Thus $\mu(I_{\mathfrak{p}})=3$; in particular, I does not satisfy G_3. However, it completely suits our Corollary 5.6 for $s=5$ and $i=2$, hence $\mathrm{Ann}(H_1(\mathfrak{a}))=\mathrm{Ann}(H_2(\mathfrak{a}))=I$. Also, since $\mathrm{ht}(J)=5$, any prime ideal of height 2 containing $\mathfrak{a}$ must contain I, which means $I=\mathfrak{a}^{unm}$. Since the ring is Gorenstein, we then have $I=\mathfrak{a}^{unm}=\mathrm{Ann}(H_3(\mathfrak{a}))$ (our *Macaulay2* system was not able to calculate $\mathrm{Ann}(H_1(\mathfrak{a}))$ directly!).

We notice that Corollary 5.6 presents a universal property for any ideal $\mathfrak{a}$ such that $J=(\mathfrak{a}:I)$ is an s-residual intersection of I. For instance, in the above example, for any other $\mathfrak{a}$ such that $J=(\mathfrak{a}:I)$ is a 5-residual intersection of I, we have $\mathrm{Ann}(H_1(\mathfrak{a}))=\mathrm{Ann}(H_2(\mathfrak{a}))=\mathrm{Ann}(H_3(\mathfrak{a}))=\mathfrak{a}^{unm}=I$.

An unpublished result of G.Levin [Vasconcelos 2005, Theorem 5.26] yields in [Corso et al. 2006, Corollary 2.7] that *if R is a Noetherian ring and $\mathfrak{a}$ is an ideal of finite projective dimension n, then* $(\mathrm{Ann}(H_1(\mathfrak{a})))^{n+1}\subseteq\mathfrak{a}$. Combining this result with Corollary 5.6, one faces the strange fact that $\mathfrak{a}$ and I have the same radical if one adopts the conditions of Corollary 5.6. However, the following simple example (taken from [Huneke 1983]) shows that [Corso et al. 2006, Corollary 2.7] is not correct, and thus Levin's result is not true.

Example 5.8. Let

$$X=\begin{pmatrix} x_1 & x_3 & x_5 & x_7 \\ x_2 & x_4 & x_6 & x_8 \end{pmatrix}$$

be a generic 2×4 matrix. Let $J=I_2(X)$, $\mathfrak{a}=(x_1x_4-x_2x_3,\, x_1x_6-x_2x_5,\, x_1x_8-x_2x_7)$ and $I=(x_1,x_2)$. Then $J=\mathfrak{a}:I$ is a geometric 3-residual intersection of I and is prime, and $\mathfrak{a}=I\cap J$, hence $\mathfrak{a}$ is radical. Furthermore, $\mathrm{Ann}(H_1(\mathfrak{a}))=\mathfrak{a}^{unm}=I$, which is not contained in $\mathfrak{a}$.

A simpler example would be the following degeneration of the above example. Let $\mathfrak{a}=\big(y_1^2-y_2^2,\, y_1y_3,\, y_2y_4\big)\in k[y_1,\dots,y_4]$. Then $\mathrm{Ann}(H_1(\mathfrak{a}))=(y_1,y_2)\neq\mathrm{rad}(\mathfrak{a})$.

At the end, we make a conjecture and ask a question. The conjecture will imply in turn the Cohen–Macaulayness of algebraic residual intersections for ideals with the sliding depth condition.

Conjecture 5.9. *If* $(R, \mathfrak{m})$ *is a CM local ring of dimension* d *and* I *satisfies* SD *with* $\operatorname{depth}(R/I) \geq d - s$, *then the disguised* s*-residual intersection of* I *coincides with the algebraic* s*-residual intersection.*

Here is a peculiar question that arises from Corollary 5.6.

Question 5.10. Does the equality hold in $(\star)$, if $s > \operatorname{ht}(I)$?

Finally, we present, without proof (a proof is based on Proposition 5.1), the following proposition which is an effort toward Question 5.10.

Proposition 5.11. *Let* $(R, \mathfrak{m})$ *be a CM local ring of dimension* d, *and let* $\mathfrak{a} = (f_1, f_2, \ldots, f_s)$ *and* $I = (f_1, f_2, \ldots, f_r)$ *be ideals with* $s \leq r$ *and* $\operatorname{ht}(I) = \operatorname{ht}(\mathfrak{a}) \geq 1$. *Suppose that* $\operatorname{ht}(\mathfrak{a} : I) \geq n \geq 1$.

- *If* $s + 1 \leq r$ *and* $\operatorname{depth}(R/I) \geq d - n + 1$, *then* $\operatorname{Ann}(H_j(f_1, f_2, \ldots, f_r)) = I$ *for* $0 \leq j \leq r - s$.
- *If* I *satisfies* SD, *then, for* $\mu \geq s$ *and* $\mu - g \geq j \geq \mu - n + 2$,

$$\operatorname{Ann}(H_j(f_1, f_2, \ldots, f_\mu)) = \bigcap_{i=0}^{\mu-s} \operatorname{Ann}(H_{j-i}(f_1, f_2, \ldots, f_s)).$$

Acknowledgment

The authors would like to thank Bernd Ulrich for encouragement and illuminating questions during his stays in Brazil in 2013 and 2014. A part of this work was done while Hassanzadeh was visiting the University of Utah, supported by the PDE-fellowship 233035/2014-1 from CNPq-Brazil. He expresses his gratitude to CNPq and thanks the University of Utah for the hospitality. This work is a part of Naéliton's Ph.D. thesis at IMPA, Rio de Janeiro; he would like to thank IMPA for all of its support during the preparation of this work. Thanks to E. Tavanfar, A. Boocher and S. Iyengar for reading the draft and correcting some important typos. We also thank the referee for his/her careful reading of the manuscript and useful comments.

References

[Artin and Nagata 1972] M. Artin and M. Nagata, "Residual intersections in Cohen–Macaulay rings", *J. Math. Kyoto Univ.* **12**:2 (1972), 307–323. MR 0301006 Zbl 0263.14019

[Bruns and Herzog 1998] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, 2nd ed., Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1998. MR 1251956 Zbl 0909.13005

[Chardin et al. 2001] M. Chardin, D. Eisenbud, and B. Ulrich, "Hilbert functions, residual intersections, and residually S_2 ideals", *Compositio Math.* **125**:2 (2001), 193–219. MR 1815393 Zbl 0983.13005

[Chardin et al. 2015] M. Chardin, D. Eisenbud, and B. Ulrich, "Hilbert series of residual intersections", *Compos. Math.* **151**:9 (2015), 1663–1687. MR 3406441 Zbl 1329.13023

[Corso et al. 2006] A. Corso, C. L. Huneke, D. Katz, and W. V. Vasconcelos, "Integral closure of ideals and annihilators of homology", pp. 33–48 in *Commutative algebra*, edited by A. Corso et al., Lecture Notes in Pure and Applied Mathematics **244**, Chapman & Hall/CRC, Boca Raton, FL, 2006. MR 2184788 Zbl 1119.13006

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, NY, 1995. MR 1322960 Zbl 0819.13001

[Eisenbud and Ulrich 2016] D. Eisenbud and B. Ulrich, "Duality and socle generators for residual intersections", preprint, 2016. arXiv 1309.2050v3

[Eisenbud et al. 2004] D. Eisenbud, C. L. Huneke, and B. Ulrich, "Order ideals and a generalized Krull height theorem", *Math. Ann.* **330**:3 (2004), 417–439. MR 2099187 Zbl 1087.13009

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **2**, Springer, Berlin, 1998. MR 1644323 Zbl 0885.14002

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, NY, 1977. MR 0463157 Zbl 0367.14001

[Hassanzadeh 2012] S. H. Hassanzadeh, "Cohen–Macaulay residual intersections and their Castelnuovo–Mumford regularity", *Trans. Amer. Math. Soc.* **364**:12 (2012), 6371–6394. MR 2958940 Zbl 1277.13008

[Herzog et al. 1983] J. Herzog, A. Simis, and W. V. Vasconcelos, "Koszul homology and blowing-up rings", pp. 79–169 in *Commutative algebra* (Trento, 1981), edited by S. Greco and G. Valla, Lecture Notes in Pure and Applied Mathematics **84**, Dekker, New York, NY, 1983. MR 686942 Zbl 0499.13002

[Huneke 1983] C. L. Huneke, "Strongly Cohen–Macaulay schemes and residual intersections", *Trans. Amer. Math. Soc.* **277**:2 (1983), 739–763. MR 694386 Zbl 0514.13011

[Huneke and Ulrich 1988] C. L. Huneke and B. Ulrich, "Residual intersections", *J. Reine Angew. Math.* **390** (1988), 1–20. MR 953673 Zbl 0732.13004

[Stanley 1980] R. P. Stanley, "Weyl groups, the hard Lefschetz theorem, and the Sperner property", *SIAM J. Algebraic Discrete Methods* **1**:2 (1980), 168–184. MR 578321 Zbl 0502.05004

[Ulrich 1992] B. Ulrich, "Remarks on residual intersections", pp. 133–138 in *Free resolutions in commutative algebra and algebraic geometry* (Sundance, UT, 1990), edited by D. Eisenbud and C. L. Huneke, Research Notes in Mathematics **2**, Jones and Bartlett, Boston, MA, 1992. MR 1165323 Zbl 0790.13005

[Ulrich 1994] B. Ulrich, "Artin–Nagata properties and reductions of ideals", pp. 373–400 in *Commutative algebra: syzygies, multiplicities, and birational algebra* (South Hadley, MA, 1992), edited by W. J. Heinzer et al., Contemporary Mathematics **159**, American Mathematical Society, Providence, RI, 1994. MR 1266194 Zbl 0821.13008

[Vasconcelos 1994] W. V. Vasconcelos, *Arithmetic of blowup algebras*, London Mathematical Society Lecture Note Series **195**, Cambridge University Press, 1994. MR 1275840 Zbl 0813.13008

[Vasconcelos 2005] W. V. Vasconcelos, *Integral closure: Rees algebras, multiplicities, algorithms*, Springer, Berlin, 2005. MR 2153889 Zbl 1082.13006

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR 1269324 Zbl 0797.18001

Communicated by Craig Huneke
Received 2015-01-21 Revised 2016-01-13 Accepted 2016-03-22

hamid@im.ufrj.br *Instituto de Matemática, Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos 149, Centro de Tecnologia - Bloco C, Cidade Universitária - Ilha do Fundão, 68530 21941-909 Rio de Janeiro, Brazil*

jnaeliton@yahoo.com.br *Departamento de Matemática, CCEN, Campus I - sn - Cidade Universitária, Universidade Federal de Paraíba, 58051-090 João Pessoa, Brazil*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.771

The Prym map of degree-7 cyclic coverings

Herbert Lange and Angela Ortega

We study the Prym map for degree-7 étale cyclic coverings over a curve of genus 2. We extend this map to a proper map on a partial compactification of the moduli space and prove that the Prym map is generically finite onto its image of degree 10.

1. Introduction

Consider an étale finite covering $f : Y \to X$ of degree p of a smooth complex projective curve X of genus $g \geq 2$. Let $\mathrm{Nm}_f : JY \to JX$ denote the norm map of the corresponding Jacobians. One can associate to the covering f its Prym variety

$$P(f) := (\mathrm{Ker}\,\mathrm{Nm}_f)^0,$$

the connected component containing 0 of the kernel of the norm map, which is an abelian variety of dimension

$$\dim P(f) = g(Y) - g(X) = (p-1)(g-1).$$

The variety $P(f)$ carries a natural polarization, namely, the restriction of the principal polarization Θ_Y of JY to $P(f)$. Let D denote the type of this polarization. If, moreover, $f : Y \to X$ is a cyclic covering of degree p, then the group action induces an action on the Prym variety. Let $\mathcal{B}_D$ denote the moduli space of abelian varieties of dimension $(p-1)(g-1)$ with a polarization of type D and an automorphism of order p compatible with the polarization. If $\mathcal{R}_{g,p}$ denotes the moduli space of étale cyclic coverings of degree p of curves of genus g, we get a map

$$\mathrm{Pr}_{g,p} : \mathcal{R}_{g,p} \to \mathcal{B}_D$$

associating to every covering in $\mathcal{R}_{g,p}$ its Prym variety, called the *Prym map*.

Particularly interesting are the cases where $\dim \mathcal{R}_{g,p} = \dim \mathcal{B}_D$. For instance, for $p = 2$ this occurs only if $g = 6$. In this case the Prym map $\mathrm{Pr}_{6,2} : \mathcal{R}_6 \to \mathcal{A}_5$ is generically finite of degree 27 (see [Donagi and Smith 1981]) and the fibers carry

Ortega was supported by Deutsche Forschungsgemeinschaft, SFB 647.
MSC2010: primary 14H40; secondary 14H30.
Keywords: Prym variety, Prym map.

the structure of the 27 lines on a smooth cubic surface. For $(g, p) = (4, 3)$, it is also known that $\mathrm{Pr}_{4,3}$ is generically finite of degree 16 onto its 9-dimensional image $\mathcal{B}_D$ (see [Faber 1988]).

In this paper we investigate the case $(g, p) = (2, 7)$, where $\dim \mathcal{R}_{g,p} = \dim \mathcal{B}_D$. The main result of the paper is the following theorem. Let G be the cyclic group of order 7.

Theorem 1.1. *For any cyclic étale G-cover $f : \widetilde{C} \to C$ of a curve C of genus* 2, *the Prym variety* $\mathrm{Pr}(f)$ *is an abelian variety of dimension* 6 *with a polarization of type* $D = (1, 1, 1, 1, 1, 7)$ *and a G-action. The Prym map*

$$\mathrm{Pr}_{2,7} : \mathcal{R}_{2,7} \to \mathcal{B}_D$$

is generically finite of degree 10.

The paper is organized as follows. First we compute in Section 2 the dimension of the moduli space $\mathcal{B}_D$ when $(g, p) = (2, 7)$. In Sections 3–5, we extend the Prym map to a partial compactification of admissible coverings $\widetilde{\mathcal{R}}_{2,7}$ such that $\mathrm{Pr}_{2,7} : \widetilde{\mathcal{R}}_{2,7} \to \mathcal{B}_D$ is a proper map. We prove the generic finiteness of the Prym map in Section 6 by specializing to a curve in the boundary. In order to compute the degree of the Prym map, we describe in Section 7 a complete fiber over a special abelian sixfold with polarization type $(1, 1, 1, 1, 1, 7)$, and in Section 8 we give a basis for the Prym differentials for the different types of admissible coverings appearing in the special fiber. Finally, in Section 9 we determine the degree of the Prym map by computing the local degrees along the special fiber.

2. Dimension of the moduli space $\mathcal{B}_D$

As in the introduction, let $\mathcal{R}_{2,7}$ denote the moduli space of nontrivial cyclic étale coverings $f : \widetilde{C} \to C$ of degree 7 of curves of genus 2. The Hurwitz formula gives $g(\widetilde{C}) = 8$. Hence the Prym variety $P = P(f)$ is of dimension 6 and the canonical polarization of the Jacobian $J\widetilde{C}$ induces a polarization of type $(1, 1, 1, 1, 1, 7)$ on P (see [Lange and Ortega 2011, p. 397]). Let σ denote an automorphism of $J\widetilde{C}$ generating the group of automorphisms of $\widetilde{C}/C$. It induces an automorphism of P, also of order 7, which is compatible with the polarization. The Prym map $\mathrm{Pr}_{2,7} : \mathcal{R}_{2,7} \to \mathcal{B}_D$ is the morphism defined by $f \mapsto P(f)$. Here $\mathcal{B}_D$ is the moduli space of abelian varieties of dimension 6 with a polarization of type $(1, 1, 1, 1, 1, 7)$ and an automorphism of order 7 compatible with the polarization. The main result of this section is the following proposition.

Proposition 2.1. $\dim \mathcal{B}_D = \dim \mathcal{R}_{2,7} = 3.$

Proof. Clearly $\dim \mathcal{R}_{2,7} = \dim \mathcal{M}_2 = 3$. So we have to show that also $\dim \mathcal{B}_D = 3$. For this we use Shimura's theory of abelian varieties with endomorphism structure (see [Shimura 1963] or [Birkenhake and Lange 2004, Chapter 9]).

Let $K = \mathbb{Q}(\rho_7)$ denote the cyclotomic field generated by a primitive 7-th root of unity ρ_7. Clearly $\mathcal{B}_D$ coincides with one of Shimura's moduli spaces of polarized abelian varieties with endomorphism structure in K. The field K is a totally complex quadratic extension of a totally real number field of degree $e_0 = 3$. Define

$$m := \frac{\dim P}{e_0} = 2.$$

The polarization of P depends on the lattice of P and a matrix $T \in M_m(\mathbb{Q}(\rho_7))$. The signature of T (see [Birkenhake and Lange 2004, p. 264]) is an e_0-tuple of nonnegative integers $((r_1, s_1), \dots, (r_{e_0}, s_{e_0}))$ satisfying

$$r_\nu + s_\nu = m = 2$$

for all ν, where e_0 is the number of real embeddings of the totally real subfield of $\mathbb{Q}(\rho_7)$. Recall that for each embedding $\mathbb{Q}(\rho_7) \hookrightarrow \mathbb{C}$, the matrix T is skew-hermitian, and the (r_ν, s_ν) are the signatures of the corresponding skew-hermitian matrices. Then, according to [Shimura 1963, p. 162] or [Birkenhake and Lange 2004, p. 266, lines 6–8], we have

$$\dim \mathcal{B}_D = \sum_{\nu=1}^{e_0} r_\nu s_\nu \leq 3, \tag{2-1}$$

with equality if and only if $r_\nu = s_\nu = 1$ for all ν.

On the other hand, in Section 6 we will see that the map $\mathrm{Pr}_{2,7}$ is generically injective. This implies that

$$\dim \mathcal{B}_D \geq \dim \mathcal{R}_{2,7} = 3,$$

which completes the proof of the proposition. □

Remark 2.2. According to [Ortega 2003], we know that P is isogenous to the product of a Jacobian of dimension 3 with itself. Then $\mathrm{End}_{\mathbb{Q}}(P)$ is not a simple algebra. Hence, if one knows that $\mathrm{Pr}_{2,7}$ is dominant onto the component $\mathcal{B}_D$, then [Birkenhake and Lange 2004, Proposition 9.9.1] implies that $r_\nu = s_\nu = 1$ for $\nu = 1, 2, 3$, which also gives $\dim \mathcal{B}_D = 3$.

Remark 2.3. It is claimed in [Faber 1988] that the Prym map $\mathrm{Pr}_{2,6} : \mathcal{R}_{2,6} \to \mathcal{B}_D$ satisfies $\dim \mathcal{B}_D = \dim \mathcal{R}_{2,6} = 3$. In a subsequent paper [Lange and Ortega $\geq$ 2016], we show that this does not occur. Moreover, we prove that there are no further examples with this property in the case of étale cyclic coverings of degree $2p$ for a prime p, but there are 3 more cases for cyclic ramified coverings of these degrees.

3. The condition (∗)

In this section we study the Prym map for coverings of degree 7 between stable curves. Let $G = \mathbb{Z}/7\mathbb{Z}$ be the cyclic group of order 7 with generator σ and let

$f: \widetilde{C} \to C$ be a G-cover of a connected stable curve C of arithmetic genus g. We fix in the sequel a primitive 7-th root of unity ρ. We assume the following condition for the covering f:

(∗) The fixed points of σ are exactly the nodes of $\widetilde{C}$ and at each node one local parameter is multiplied by ρ^{δ} and the other by $\rho^{-\delta}$ for some δ, $1 \le \delta \le 3$.

As in [Beauville 1977], we have $f^*\omega_C \simeq \omega_{\widetilde{C}}$, which implies

$$p_a(\widetilde{C}) = 7g - 6.$$

Let $\widetilde{N}$ and N be the normalizations of $\widetilde{C}$ and C, respectively, and let $\tilde{f}: \widetilde{N} \to N$ be the induced map. At each node s of $\widetilde{C}$ we make the usual identification

$$\mathcal{K}_s^*/\mathcal{O}_s^* \simeq \mathbb{C}^* \times \mathbb{Z} \times \mathbb{Z}.$$

Then the action of σ on $\mathcal{K}_s^*/\mathcal{O}_s^*$ is

$$\sigma^*((z, m, n)_s) = (\rho^{\delta(m-n)} z, m, n)_s$$

for some δ, $1 \le \delta \le 3$. Here we label the branches at the node s such that a local parameter at the first branch (corresponding to m) is multiplied by ρ^{δ} with $1 \le \delta \le 3$. Then we have

$$f_*((z, m, n)_s) = (z^7, m, n)_{f(s)}$$

since $f_*((z, m, n)_s) = \sum_{k=0}^{6} (\sigma^*)^k (z, m, n)_s$, viewed as a divisor at $f(s)$, and

$$\begin{aligned}
\sum_{k=0}^{6} (\sigma^*)^k (z, m, n)_s &= \Big(\prod_{k=0}^{6} (\rho^{\delta(m-n)})^k z, 7m, 7n \Big)_s \\
&= \big(\rho^{\sum_{k=0}^{6} \delta(m-n)k} z^7, 7m, 7n \big)_s \\
&= \big(\rho^{\delta(m-n)\frac{6\cdot 7}{2}} z^7, 7m, 7n \big)_s \\
&= (z^7, 7m, 7n)_s \\
&= f^*(z^7, m, n)_{f(s)}.
\end{aligned}$$

We define the multidegree of a line bundle L on $\widetilde{C}$ by

$$\deg L = (d_1, \dots, d_v),$$

where v is the number of components of $\widetilde{C}$ and d_i is the degree of L on the i-th component of $\widetilde{C}$.

Lemma 3.1. *Let $L \in \operatorname{Pic} \widetilde{C}$ with $\operatorname{Nm} L \simeq \mathcal{O}_{\widetilde{C}}$. Then*

$$L \simeq M \otimes \sigma^* M^{-1}$$

for some $M \in \operatorname{Pic} \widetilde{C}$. *Moreover*, M *can be chosen of multidegree* $(k, 0, \dots, 0)$ *with* $0 \le k \le 6$.

Proof. As in [Mumford 1971, Lemma 1], using Tsen's theorem, there is a divisor D such that $L \simeq \mathcal{O}_{\widetilde{C}}(D)$ and $f_* D = 0$. We write

$$D = \underbrace{\sum_{x \in \widetilde{C}_{\mathrm{reg}}} x}_{D_{\mathrm{reg}}} + \underbrace{\sum_{s \in \widetilde{C}_{\mathrm{sing}}} (z_s, m_s, n_s)_s}_{D_{\mathrm{sing}}} . \tag{3-1}$$

If $x \in \widetilde{C}_{\mathrm{reg}}$ is in the support of D_{reg}, then $\sigma^k(x)$ is in the support of D_{reg} for some $1 \le k \le 6$. Since

$$x - \sigma^k(x) = (x + \sigma(x) + \cdots + \sigma^{k-1}(x)) - \sigma(x + \sigma(x) + \cdots + \sigma^{k-1}(x)),$$

there is a divisor E_{reg} such that $D_{\mathrm{reg}} = E_{\mathrm{reg}} - \sigma^* E_{\mathrm{reg}}$. From (3-1) one sees that the divisor D_{sing} is the sum of divisors of the form $(\rho^{a_s}, 0, 0)_s$ for some $1 \le a_s \le 6$. Choosing an integer i_s such that $-i_s \delta_s \equiv a_s \bmod 7$, we have

$$(1, i_s, 0) - \sigma^*(1, i_s, 0) = (\rho^{-\delta_s i_s}, 0, 0) = (\rho^{a_s}, 0, 0)_s.$$

Then there is a divisor E_{sing} such that $D_{\mathrm{sing}} = E_{\mathrm{sing}} - \sigma^* E_{\mathrm{sing}}$. Thus $D = E - \sigma^* E$ with $E = E_{\mathrm{reg}} + E_{\mathrm{sing}}$. Set $M = \mathcal{O}_{\widetilde{C}}(E)$. By replacing M with $M \otimes f^* N$, where N is a line bundle on C, we may assume that the multidegree of M is $(\epsilon_1, \dots, \epsilon_\nu)$ with $0 \le \epsilon_i \le 6$. Using the fact that $\widetilde{C}$ is connected, the multidegree can be accumulated on one of the components by further tensoring M with line bundles associated to divisors of the form $(1, 1, -1)_s$, since $(1, 1, -1)_s - \sigma^*(1, 1, -1)_s = (1, 0, 0)_s = 0$. □

Let P denote the Prym variety of $f : \widetilde{C} \to C$, i.e., the connected component containing 0 of the kernel of norm map $\mathrm{Nm} : J\widetilde{C} \to JC$. By definition, it is a connected commutative algebraic group. Lemma 3.1 implies that P is the variety of line bundles in $\ker \mathrm{Nm}$ of the form $M \otimes \sigma^* M^{-1}$ with M of multidegree $(0, \dots, 0)$. Let $\nu : N \to C$ and $\tilde{\nu} : \widetilde{N} \to \widetilde{C}$ denote the normalizations of C and $\widetilde{C}$, respectively. Hence there is a map $\tilde{f}$ making the following diagram commutative:

$$\begin{tikzcd} \widetilde{N} \arrow[r, "\tilde{\nu}"] \arrow[d, "\tilde{f}"'] & \widetilde{C} \arrow[d, "f"] \\ N \arrow[r, "\nu"] & C \end{tikzcd}$$

Proposition 3.2. *Suppose* $p_a(C) = g$. *Then* P *is an abelian variety of dimension* $6g - 6$.

Proof. (As in [Beauville 1977] and [Faber 1988].) Consider the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \widetilde{T} & \longrightarrow & J\widetilde{C} & \xrightarrow{\tilde{\nu}^*} & J\widetilde{N} & \longrightarrow & 0 \\ & & \Big\downarrow{\scriptstyle \mathrm{Nm}} & & \Big\downarrow{\scriptstyle \mathrm{Nm}} & & \Big\downarrow{\scriptstyle \mathrm{Nm}} & & \\ 0 & \longrightarrow & T & \longrightarrow & JC & \xrightarrow{\nu^*} & JN & \longrightarrow & 0 \end{array} \tag{3-2}$$

of commutative algebraic groups, where the vertical arrows are the norm maps and T and $\widetilde{T}$ are the groups of classes of divisors of multidegree $(0,\dots,0)$ with singular support. Since f^* restricted to T is an isomorphism and $\mathrm{Nm}\circ f^* = 7$, the norm on $\widetilde{T}$ is surjective and

$$\ker \mathrm{Nm}_{|\widetilde{T}} \simeq \widetilde{T}[7] = \{\text{points of order 7 in } \widetilde{T}\}.$$

On the other hand, Lemma 3.1 implies that $\ker \mathrm{Nm}$ consists either of 7 components or is connected. Let R be the kernel of $\mathrm{Nm} : J\widetilde{N} \to JN$. Then one obtains an exact sequence

$$0 \to \widetilde{T}[7] \to \widetilde{R} \to R \to 0 \tag{3-3}$$

with $\widetilde{R} = P$ if $\ker \mathrm{Nm}$ is connected, and $\widetilde{R} = P \times \mathbb{Z}/7\mathbb{Z}$ if $\ker \mathrm{Nm}$ is not connected. We will see that in both cases P is an abelian variety.

Suppose first that C and hence $\widetilde{C}$ are nonsingular. Then $\widetilde{C} = \widetilde{N}$ and hence $\widetilde{T} = 0$. Since $\ker(J\widetilde{N} \to JN)$ has 7 components, P is an abelian variety.

Suppose that C and thus also $\widetilde{C}$ have $s > 0$ singular points. Then $\dim \widetilde{T} = \dim T = s$. Then R is an abelian variety, since $\tilde{f}$ is ramified. We get a surjective homomorphism $P \to R$ with kernel consisting of 7^s elements when $\mathrm{Ker}\,\mathrm{Nm}$ is connected and 7^{s-1} when it is not. Hence also P is an abelian variety. Moreover,

$$\dim P = \dim R = \dim J\widetilde{C} - \dim JC.$$

Now, if C has s nodes and N has t connected components, then also $\widetilde{C}$ has s nodes and $\widetilde{N}$ has t connected components. This implies

$$\dim J\widetilde{C} - \dim JC = p_a(\widetilde{C}) - p_a(C) = 6g - 6. \qquad \square$$

Let $\widetilde{\Theta}$ denote the canonical polarization of the generalized Jacobian $J\widetilde{C}$ (see [Beauville 1977]). It restricts to a polarization Σ on the abelian subvariety P. We denote the isogeny $P \to \widehat{P}$ associated to Σ by the same letter.

Proposition 3.3. *The polarization Σ on P is of type $D := (1,\dots,1,7,\dots,7)$, where* 7 *occurs*

$$\begin{cases} g \textit{ times} & \textit{if } \ker \mathrm{Nm} \textit{ is connected}, \\ g-1 \textit{ times} & \textit{if } \ker \mathrm{Nm} \textit{ is not connected}. \end{cases}$$

Proof. (Similar to [Faber 1988, Proposition 2.4]; we use also [Beauville 1977, Corollary 2.3, p. 156].) If C is smooth, this is well known. In this case $\ker f$ consists of 7 components and 7 occurs $g-1$ times in the polarization D (see [Birkenhake and Lange 2004, §12.3]). So suppose C is not smooth and f is a covering of type $(*)$. In this case the maps f^* and $\tilde{f}^*$ are injective.

Consider the isogeny

$$h : P \times JN \to J\widetilde{N}, \quad (L, M) \mapsto \tilde{\nu}^*(L) \otimes \tilde{f}^* M.$$

We first claim that $\ker h \subset P[7] \times JN[7]$. To see this, let $(L, M) \in \ker h$, i.e., $L \in P$ and $M \in JN$ and $\tilde{\nu}^*(L) \otimes \tilde{f}^* M \simeq \mathcal{O}_{\widetilde{N}}$. Choose $M' \in JC$ with

$$M \simeq \nu^* M'. \tag{3-4}$$

Then

$$\tilde{\nu}^*(L \otimes f^* M') \simeq \tilde{\nu}^* L \otimes \tilde{f}^* \nu^* M' \simeq \tilde{\nu}^* L \otimes \tilde{f}^* M \simeq \mathcal{O}_{\widetilde{N}}, \tag{3-5}$$

so $L \otimes f^* M' \in \ker \tilde{\nu}^* \simeq \widetilde{T}$. Since $f^* : T \to \widetilde{T}$ is an isomorphism, there is a unique $N \in T$ such that

$$L \simeq f^*(N \otimes M'^{-1}). \tag{3-6}$$

This implies

$$(N \otimes M'^{-1})^{\otimes 7} \simeq \mathrm{Nm} \circ f^*(N \otimes M'^{-1}) \simeq \mathrm{Nm}\, L \simeq \mathcal{O}_C$$

and hence $L \in P[7]$ and using (3-5) we get

$$M^{\otimes 7} \simeq \mathrm{Nm}\, \tilde{f}^* M \simeq \mathrm{Nm}\, \tilde{f}^* \nu^* M' \simeq \mathrm{Nm}\, \tilde{\nu}^* L^{-1} \simeq \nu^* \mathrm{Nm}\, L^{-1} \simeq \nu^* \mathcal{O}_N \simeq \mathcal{O}_C.$$

This completes the proof of the claim.

Now, since the map $\nu^* : JC[7] \to JN[7]$ is surjective, we can choose M' in (3-4) even as an element of $JC[7]$. Moreover, from (3-6) we also get $N \in T[7]$, since f^* is injective. Now consider the following extension of the map h to the whole of the kernel of the norm map of f:

$$\tilde{h} : \ker \mathrm{Nm} \times JN \to J\widetilde{N}, \quad (L, M) \mapsto \tilde{\nu}^*(L) \otimes \tilde{f}^* M.$$

We claim that

$$\ker \tilde{h} = \{(f^*(N \otimes M'^{-1}), \nu^* M') \mid M' \in JC[7],\ N \in T[7]\}. \tag{3-7}$$

So $\ker h$ consists of those elements of the right-hand side of (3-7) for which $f^*(N \otimes M'^{-1})$ is contained in the connected component containing 0 of $\ker \mathrm{Nm}_f$.

Proof of (3-7). The inclusion of $\ker \tilde{h}$ in the left-hand side of the equation follows from (3-4) and (3-6), which are valid for the extension $\tilde{h}$.

For the converse inclusion, suppose that $M' \in JC[7]$ and $N \in T[7]$. First note that $f^*(N \otimes M'^{-1}) \subset \ker \mathrm{Nm}$, since $\mathrm{Nm}\, f^*(N \otimes M'^{-1}) \simeq (N \otimes M'^{-1})^{\otimes 7} \simeq \mathcal{O}_C$. Moreover,

$$\tilde{\nu}^* f^*(N \otimes M'^{-1}) \otimes \tilde{f}^* \nu^* M' \simeq \tilde{\nu}^* f^*(N \otimes M'^{-1}) \otimes \tilde{\nu}^* f^* M' \simeq \tilde{\nu}^* f^* N \simeq \mathcal{O}_{\widetilde{N}}.$$

This completes the proof of (3-7). □

Since for any $\alpha \in T[7]$ we have $(N \otimes \alpha) \otimes (\alpha^{-1} \otimes M'^{-1}) \simeq N \otimes M'^{-1}$ and $\nu^*(M' \otimes \alpha) \simeq \nu^* M'$, we conclude that

$$\dim_{\mathbb{F}_7} \ker \tilde{h} = \dim_{\mathbb{F}_7} JC[7] = 2g - t,$$

with $t = \dim T = \dim \widetilde{T}$. This implies

$$\dim_{\mathbb{F}_7} \ker h = \begin{cases} 2g - t & \text{if ker is connected,} \\ 2g - t - 1 & \text{if ker is not connected,} \end{cases}$$

since in the not-connected case $\ker \mathrm{Nm}$ consists of 7 components.

Now $\ker h$ is a maximal isotropic subgroup of the kernel of the polarization of $P \times JN$, since this polarization is the pullback under h of the principal polarization of $J\widetilde{N}$. This implies $\dim_{\mathbb{F}_7}(\ker \Sigma \times JN[7]) = 2(\dim_{\mathbb{F}_7} \ker h)$. Since $\dim_{\mathbb{F}_7}(JN[7]) = 2g - 2t$, it follows that

$$\dim_{\mathbb{F}_7}(\ker \Sigma) = \begin{cases} 2g & \text{if } \ker \mathrm{Nm} \text{ is connected,} \\ 2g - 2 & \text{if } \ker \mathrm{Nm} \text{ is not connected.} \end{cases}$$

Since $\ker \Sigma \subset P[7]$, this gives the assertion. □

Corollary 3.4. $\ker \mathrm{Nm}$ *consists of* 7 *components.*

Proof. Consider $\widetilde{C}_t \to C_t$, a family of coverings, where the central fiber $\widetilde{C}_0 \to C_0$ satisfies condition $(*)$ and all the other fibers are coverings of smooth curves. The fibers in the associated family of abelian varieties $\ker \mathrm{Nm}_t$ have 7 components for $t \neq 0$ and, according to Proposition 3.3, they have a polarization of type $(1, \dots, 1, 7, \dots, 7)$, where 7 appears $g - 1$ times. The type of a polarization in a family of abelian varieties is constant since it is given by integers; therefore, $\ker \mathrm{Nm}_0$ has the same polarization type as the nearby fibers and, again by Proposition 3.3, $\ker \mathrm{Nm}_0$ is nonconnected, so it consists of 7 components. □

4. The condition $(**)$

As in the last section, let $f : \widetilde{C} \to C$ be a G-covering of stable curves. Recall that a node $z \in \widetilde{C}$ is either

- of index 1, i.e., $|\mathrm{Stab}\, z| = 1$, in which case $f^{-1}(f(z))$ consists of 7 nodes which are cyclically permuted under σ, or

- of index 7, i.e., $|\mathrm{Stab}\, z| = 7$, in which case z is the only preimage of the node $f(z)$ and f is totally ramified at both branches of z. Since σ is of order 7, the two branches of z are not exchanged.

We also say a node of C is *of index i* if a preimage (and hence every preimage) under f is a node of index i. We assume the following conditions for the G-covering $f : \widetilde{C} \to C$ of connected stable curves:

$$(**)\quad \begin{cases} p_a(C) = g \text{ and } p_a(\widetilde{C}) = 7g-6; \\ \sigma \text{ is not the identity on any irreducible component of } \widetilde{C}; \\ \text{if at a fixed node of } \sigma \text{ one local parameter is multiplied by } \rho^i, \text{ the other} \\ \quad \text{is multiplied by } \rho^{-i}, \text{ where } \rho \text{ denotes a fixed 7-th root of unity;} \\ P := \mathrm{Pr}(f) \text{ is an abelian variety.} \end{cases}$$

Under these assumptions the nodes of $\widetilde{C}$ are exactly the preimages of the nodes of C. We define for $i = 1$ and 7:

- $n_i :=$ the number of nodes of C of index i, i.e., nodes whose preimage consists of $7/i$ nodes of $\widetilde{C}$,
- $c_i :=$ the number irreducible components of C whose preimage consists of $7/i$ irreducible components of $\widetilde{C}$,
- $r :=$ the number of fixed nonsingular points under σ.

Lemma 4.1. *The covering satisfies* $(**)$ *if and only if* $r = 0$ *and* $c_1 = n_1$.

In particular, any covering satisfying $(**)$ is an admissible G-cover (for the definition, see Section 5), and coverings satisfying condition $(*)$ also verify $(**)$.

Proof. (As in [Beauville 1977] and [Faber 1988].) Let $\widetilde{N}$ and N be the normalizations of $\widetilde{C}$ and C, respectively. The covering $\tilde{f} : \widetilde{N} \to N$ is ramified exactly at the points lying over the fixed points of $\sigma : \widetilde{C} \to \widetilde{C}$. Hence the Hurwitz formula says

$$p_a(\widetilde{N}) - 1 = 7(p_a(N) - 1) + 3r + 6n_7.$$

So

$$p_a(\widetilde{C}) - 1 = p_a(\widetilde{N}) - 1 + 7n_1 + n_7 = 7(p_a(N) - 1) + 3r + 7n_1 + 7n_7.$$

Moreover,

$$p_a(C) - 1 = p_a(N) - 1 + n_1 + n_7,$$

which altogether gives

$$p_a(\widetilde{C}) - 1 = 7(p_a(C) - 1) + 3r.$$

Hence the first condition in $(**)$ is equivalent to $r = 0$.

Now we discuss the condition that P is an abelian variety. For this consider again the diagram (3-2). From the surjectivity of the norm maps it follows that P is an abelian variety if and only if $\dim \widetilde{T} = \dim T$. Now

$$\dim J\widetilde{N} = p_a(\widetilde{N}) - n_7 - 7n_1 + c_7 + 7c_1 - 1$$

and thus

$$\dim \widetilde{T} = (n_7 - c_7) + 7(n_1 - c_1) + 1 \quad \text{and} \quad \dim T = (n_7 - c_7) + (n_1 - c_1) + 1.$$

Hence $\dim \widetilde{T} = \dim T$ if and only if $c_1 = n_1$. □

Let $f : \widetilde{C} \to C$ be a G-covering satisfying condition $(**)$ with generating automorphism σ. We denote by B the union of the components of $\widetilde{C}$ fixed under σ and write

$$\widetilde{C} = A_1 \cup \cdots \cup A_7 \cup B$$

with $\sigma(A_i) = A_{i+1}$, where $A_8 = A_1$. Observe that the covering $B \to B/\sigma$ satisfies condition $(*)$.

Proposition 4.2. (i) *If $B = \varnothing$, then $\widetilde{C} = A_1 \cup \cdots \cup A_7$, where A_1 can be chosen connected and tree-like, and $\# A_i \cap A_{i+1} = 1$ for $i = 1, \ldots, 7$.*

(ii) *If $B \neq \varnothing$, then $A_i \cap A_{i+1} = \varnothing$ for $i = 1, \ldots, 7$. Each connected component of A_1 is tree-like and meets B at only one point. Also B is connected.*

For the proof we need the following elementary lemma (the analogue of [Beauville 1977, Lemma 5.3] and [Faber 1988, Lemma 2.6]), which will be applied to the dual graph of $\widetilde{C}$.

Lemma 4.3. *Let Γ be a connected graph with a fixed-point free automorphism σ of order 7. Then there exists a connected subgraph S of γ such that $\sigma^i(S) \cap \sigma^{i+1}(S)$ is empty for $i = 0, \ldots, 6$ and $\bigcup_{i=0}^{6} \sigma^i(S)$ contains every vertex of Γ.*

Proof of Proposition 4.2. (As in [Beauville 1977] and [Faber 1988].) Let Γ denote the dual graph of $\widetilde{C}$. If $B = \varnothing$, let A_1 correspond to the subgraph S of Lemma 4.3. Let v be the number of vertices of S, e the number of edges of S and s the number of nodes of A_1 which belong to only one component. The equality $c_1 = n_1$ implies

$$v = e + s - \# A_1 \cap A_2.$$

Since $1 - v + e \geq 0$ and $\# A_1 \cap A_2 \geq 1$ give $s = 0$, we have $\# A_1 \cap A_2 = 1$ and $1 - v + e = 0$. So A_1 is tree-like. This proves (i).

Assume $B \neq \varnothing$ and define

- $t := \# A_1 \cap A_2$,
- $m := \# A_i \cap B$ for $i = 1, \ldots, 7$,
- $i_{A_1} := \#$ irreducible components of A_1,

- $c_{A_1} :=$ # of connected components of A_1,
- $n_{A_1} :=$ # nodes of A_1.

Recall that assumption $(**)$ implies that B does not contain any node which moves under σ. Then

$$c_1 = i_{A_1} \quad \text{and} \quad n_1 = n_{A_1} + r + m.$$

For any curve we have $n_{A_1} - i_{A_1} + c_{A_1} \geq 0$ (see [Beauville 1977, Proof of Lemma 5.3]). Thus, if $c_1 = n_1$,

$$0 = n_{A_1} + t + m - i_{A_1} \geq t + m - c_{A_1}.$$

Since $\widetilde{C}$ is connected, any connected component of $\bigcup_{i=1}^{7} A_i$ meets B. But then any connected component of A_1 meets B, which implies $m \geq c_{A_1}$. Hence

$$0 \geq t + m - c_{A_1} \geq t \geq 0.$$

Hence $t = 0$, $m = c_{A_1}$ and $n_{A_1} - i_{A_1} + c_{A_1} = 0$. So $A_i \cap A_{i+1} = \varnothing$ and B is connected. $\square$

In the next section we will extend the Prym map to the G-covers satisfying $(**)$, thereby obtaining a proper map. Theorem 4.4 describes the associated Pryms by reducing to the easier situation of coverings verifying $(*)$.

Theorem 4.4. *Suppose that $f : \widetilde{C} \to C$ satisfies condition* $(**)$. *Then there exist the following isomorphisms of polarized abelian varieties*:

- *In case* (i) *of Proposition* 4.2, $(P, \Sigma) \simeq \ker((JA_1)^7 \to JA_1)$ *with the polarization induced by the principal polarization on* $(JA_1)^7$.
- *In case* (ii) *of Proposition* 4.2, $(P, \Sigma) \simeq \ker((JA_1)^7 \to JA_1) \times Q$, *where Q is the generalized Prym variety associated to the covering $B \to B/\sigma$.*

Proof. (As in [Beauville 1977, Theorem 5.4].) In case (i), $\widetilde{C}$ is obtained from the disjoint union of 7 copies of A_1 by fixing 2 smooth points p and q of A_1 and identifying q in the i-th copy with p in the $(i+1)$-st copy of A_1 cyclically. The curve $C = \widetilde{C}/G$ is obtained from A_1 by identifying p and q, and $f : \widetilde{C} \to C$ is an étale covering. Note that JA_1 is an abelian variety, since A_1 is tree-like.

Consider the diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{C}^* \arrow[r] \arrow[d, "\mathrm{Nm}"] & J\widetilde{C} \arrow[r] \arrow[d, "\mathrm{Nm}"] & (JA_1)^7 \arrow[r] \arrow[d, "m"] & 0 \\
0 \arrow[r] & \mathbb{C}^* \arrow[r] & JC \arrow[r] & JA_1 \arrow[r] & 0
\end{tikzcd}$$

in which m is the addition map. One checks immediately that $\mathrm{Nm} : \mathbb{C}^* \to \mathbb{C}^*$ is an isomorphism. This implies the assertion. Notice that the polarization of P is of type

$$(1, \ldots, 1, \underbrace{7, \ldots, 7}_{g-1}).$$

In case (ii), we have

$$J\widetilde{C} \simeq (JA_1)^7 \times JB \quad \text{and} \quad P = (\ker \mathrm{Nm})^0 \simeq \ker((JA_1)^7 \to JA_1) \times Q,$$

which immediately implies the assertion. □

5. The extension of the Prym map to a proper map

Let $\mathcal{R}_{g,7}$ denote the moduli space of nontrivial étale G-covers $f : \widetilde{C} \to C$ of smooth curves C of genus g, and let $\mathcal{B}_D$ denote the moduli space of polarized abelian varieties of dimension $6g-6$ with polarization of type D, with D as in Proposition 3.3 and compatible with the G-action. As in the introduction we denote by

$$\mathrm{Pr}_{g,7} : \mathcal{R}_{g,7} \to \mathcal{B}_D$$

the corresponding Prym map associating to the covering f the Prym variety $\mathrm{Pr}(f)$. In order to extend this map to a proper map, we consider the compactification $\overline{\mathcal{R}}_{g,7}$ of $\mathcal{R}_{g,7}$ consisting of admissible G-coverings of stable curves of genus g introduced in [Abramovich et al. 2003].

Let $\mathcal{X} \to S$ be a family of stable curves of arithmetic genus g. A *family of admissible G-covers* of $\mathcal{X}$ over S is a finite morphism $\mathcal{Z} \to \mathcal{X}$ such that

(1) the composition $\mathcal{Z} \to \mathcal{X} \to S$ is a family of stable curves;

(2) every node of a fiber of $\mathcal{Z} \to S$ maps to a node of the corresponding fiber of $\mathcal{X} \to S$;

(3) $\mathcal{Z} \to \mathcal{X}$ is a principal G-bundle away from the nodes;

(4) if z is a node of index 7 in a fiber of $\mathcal{Z} \to S$ and ξ and η are local coordinates of the two branches near z, any element of the stabilizer $\mathrm{Stab}_G(z)$ acts as

$$(\xi, \eta) \mapsto (\rho\xi, \rho^{-1}\eta),$$

where ρ is a primitive 7-th root of unity.

In the case of $S = \operatorname{Spec} \mathbb{C}$ we just speak of an admissible G-cover. In this case the ramification index at any node z over x equals the order of the stabilizer of z and depends only on x. It is called the *index* of the G-cover $\mathcal{Z} \to \mathcal{X}$ at x. Since 7 is a prime, the index of a node is either 1 or 7. Note that, for any admissible G-cover $Z \to X$, the curve Z is stable if and only if X is stable.

As shown in [Abramovich et al. 2003] or [Arbarello et al. 2011, Chapter 16], the moduli space $\overline{\mathcal{R}}_{g,7}$ of admissible G-covers of stable curves of genus g is a natural compactification of $\mathcal{R}_{g,7}$. Clearly the coverings satisfying condition (∗∗) are admissible and form an open subspace $\widetilde{\mathcal{R}}_{g,7}$ of $\overline{\mathcal{R}}_{g,7}$.

Theorem 5.1. *The map* $\mathrm{Pr}_{g,7} : \mathcal{R}_{g,7} \to \mathcal{B}_D$ *extends to a proper map* $\widetilde{\mathrm{Pr}}_{g,7} : \widetilde{\mathcal{R}}_{g,7} \to \mathcal{B}_D$.

Proof. The proof is the same as the proof of [Faber 1988, Theorem 2.8] just replacing 3-fold covers by 7-fold covers. So we will omit it. □

6. Generic finiteness of $\mathrm{Pr}_{2,7}$

From now on we consider only the case $g = 2$, i.e., of G-covers of curves of genus 2. So $\dim \mathcal{R}_{2,7} = \dim \mathcal{M}_2 = 3$ and $\mathcal{B}_D$ is the moduli space of polarized abelian varieties of type $(1, 1, 1, 1, 1, 7)$ with G-action which is also of dimension 3. Let $[f : \widetilde{C} \to C] \in \mathcal{R}_{2,7}$ be a general point and let the covering f be given by the 7-division point $\eta \in JC$. The next lemma is a particular case of the results in [Lange and Ortega 2011, p. 397–398] that we include here for the sake of completeness.

Lemma 6.1. (i) *The cotangent space of* $\mathcal{B}_D$ *at the point* $\mathrm{Pr}_{2,7}([f : \widetilde{C} \to C]) \in \mathcal{B}_D$ *is identified with the vector space* $\bigoplus_{i=1}^{3} \big(H^0(\omega_C \otimes \eta^i) \otimes H^0(\omega_C \otimes \eta^{7-i})\big)$.

(ii) *The codifferential of the map* $\mathrm{Pr}_{2,7} : \mathcal{R}_{2,7} \to \mathcal{B}_D$ *at the point* (f, η) *is given by the sum of the multiplication maps*

$$\bigoplus_{i=1}^{3} \big(H^0(\omega_C \otimes \eta^i) \otimes H^0(\omega_C \otimes \eta^{7-i})\big) \longrightarrow H^0(\omega_C^2).$$

Proof. (i) Consider the composed map $\mathcal{R}_{g,7} \xrightarrow{\mathrm{Pr}_{2,7}} \mathcal{B}_D \xrightarrow{\pi} \mathcal{A}_D$, where $\mathcal{A}_D$ denotes the moduli space of abelian varieties with polarization of type D. The cotangent space of the image of $[f : \widetilde{C} \to C]$ in $\mathcal{A}_D$ is by definition the cotangent at the Prym variety P of f. It is well known that the cotangent space $T^*_{P,0}$ at 0 is

$$T^*_{P,0} = H^0(\widetilde{C}, \omega_{\widetilde{C}})^- = \bigoplus_{i=1}^{6} H^0(C, \omega_C \otimes \eta^i). \tag{6-1}$$

According to [Welters 1983] the cotangent space of $\mathcal{A}_D$ at the point P can be identified with the second symmetric product of $H^0(\widetilde{C}, \omega_{\widetilde{C}})^-$. This gives

$$T^*_{\mathcal{A}_D,P} = \bigoplus_{i=1}^{6} S^2 H^0(\omega_C \otimes \eta^i) \oplus \bigoplus_{i=1}^{3} \big(H^0(\omega_C \otimes \eta^i) \otimes H^0(\omega_C \otimes \eta^{7-i})\big). \tag{6-2}$$

Since the map $\pi : \mathcal{B}_D \to \mathcal{A}_D$ is finite onto its image and the group G acts on the cotangent space of $\mathcal{B}_D$ at the point, we conclude that this space can be identified with a 3-dimensional G-subspace of the G-space $T^*_{\mathcal{A}_D,P}$, which is defined over the rationals. But there is only one such subspace, namely,

$$\bigoplus_{i=1}^{3}\big(H^0(\omega_C\otimes\eta^i)\otimes H^0(\omega_C\otimes\eta^{7-i})\big).$$

This gives (i).

(ii) It is well known that the cotangent space of $\mathcal{R}_{2,7}$ at a point (C,η) without automorphism is given by $H^0(\omega_C^2)$ and the codifferential of $\mathrm{Pr}_{2,7}:\mathcal{R}_{2,7}\to\mathcal{A}_D$ at (C,η) by the natural map $S^2(H^0(\widetilde{C},\omega_{\widetilde{C}})^-)\longrightarrow H^0(\omega_C^2)$. The assertion follows immediately from Lemma 6.1(i) and equations (6-1) and (6-2). □

Theorem 6.2. *The map* $\widetilde{\mathrm{Pr}}_{2,7}:\widetilde{\mathcal{R}}_{2,7}\to\mathcal{B}_D$ *is surjective and hence of finite degree.*

Proof. Since the extension $\widetilde{\mathrm{Pr}}_{2,7}$ is proper according to Theorem 5.1, it suffices to show that the map $\mathrm{Pr}_{2,7}$ is generically finite. Now $\mathrm{Pr}_{2,7}$ is generically finite as soon as its differential at the generic point $[f:\widetilde{C}\to C]\in\mathcal{R}_{2,7}$ is injective. Let f be given by the 7-division point η. According to Lemma 6.1, the codifferential of $\mathrm{Pr}_{2,7}$ at $[f:\widetilde{C}\to C]$ is given by (the sum of) the multiplication of sections

$$\mu_{C,\eta}:\bigoplus_{j=1}^{3}H^0(C,\omega_C\otimes\eta^j)\otimes H^0(C,\omega_C\otimes\eta^{7-j})\longrightarrow H^0(C,\omega_C^2).$$

Note that the surjectivity of $\mu_{C,\eta}$ is an open condition in $\overline{\mathcal{R}}_{2,7}$. Since $\overline{\mathcal{R}}_{2,7}$ is irreducible and $\widetilde{\mathcal{R}}_{2,7}$ is open and dense in $\overline{\mathcal{R}}_{2,7}$, it suffices to show that the map $\mu_{X,\eta}$ is surjective at a point (X,η) in the compactification $\overline{\mathcal{R}}_{2,7}$, even if $\mathrm{Pr}_{2,7}$ is not defined at (X,η). So if $\mu_{X,\eta}$ is surjective at this point, it will be surjective at a general point of $\mathcal{R}_{2,7}$. Moreover, it suffices to show that $\mu_{X,\eta}$ is injective, since both sides of the map are of dimension 3.

Consider the curve

$$X=Y\cup Z,$$

the union of two rational curves intersecting in 3 points q_1,q_2,q_3 which we can assume to be $[1,0]$, $[0,1]$, $[1,1]$, respectively. The line bundle $\eta_X=(\eta_Y,\eta_Z)$ on X of degree 0 is uniquely determined by the gluing of the fiber over the nodes

$$\mathcal{O}_{Y_{|q_i}}\xrightarrow{\cdot c_i}\mathcal{O}_{Z_{|q_i}},$$

given by the multiplication by a nonzero constant c_i. We may assume $c_3=1$ and since $\eta_X^7\simeq\mathcal{O}_X$ we have $c_1^7=c_2^7=1$. Notice that $\omega_{X_{|Y}}=\mathcal{O}_Y(1)$ and $\omega_{X_{|Z}}=\mathcal{O}_Z(1)$ and the restrictions η_Y,η_Z are trivial line bundles.

From the exact sequence

$$0\to\mathcal{O}_Z(-2)\to\omega_X\otimes\eta_X^i\xrightarrow{\beta_i}\mathcal{O}_Y(1)\to 0$$

we have $h^0(X, \omega_X \otimes \eta_X^i) = 1$ for $i = 1, \dots, 6$. Moreover, since $H^0(Z, \mathcal{O}_Z(-2)) = 0$, the map β_i induces an inclusion $H^0(X, \omega_X \otimes \eta_X^i) \hookrightarrow H^0(Y, \mathcal{O}_Y(1))$ for $i = 1, \dots, 6$.

Therefore, to study the injectivity of $\mu_{X,\eta}$, it is enough to check whether the projection of $\bigoplus_{i=1}^3 H^0(X, \omega_X \otimes \eta_X^i) \otimes H^0(X, \omega_X \otimes \eta_X^{7-i})$ to $H^0(Y, \mathcal{O}_Y(1)) \otimes H^0(Y, \mathcal{O}_Y(1))$ is contained in the kernel of the multiplication map

$$H^0(Y, \mathcal{O}_Y(1)) \otimes H^0(Y, \mathcal{O}_Y(1)) \longrightarrow H^0(Y, \mathcal{O}_Y(2)).$$

We claim that the line bundle $\omega_X = (\omega_{|Y}, \omega_{|Z})$ is uniquely determined and one can choose the gluing c_i at the nodes q_i to be the multiplication by the same constant. To see this, first notice that, since (X, ω_X) is a limit linear series of canonical line bundles, the nodes of X are necessary Weierstrass points of X. Let $s_3 \in H^0(X, \omega_X(-2q_3))$ be a section giving a trivialization of ω_Y and ω_Z away from q_3. For $i = 1, 2$, we have

$$\mathcal{O}_Y(1)_{|q_i} \xrightarrow{s_3^{-1}} \mathcal{O}_{X\,|q_i} \xrightarrow{s_3} \mathcal{O}_Z(1)_{|q_i},$$

which implies that $c_1 = c_2$. Similarly, by using a section in $H^0(X, \omega_X(-2q_2))$ one shows that $c_1 = c_3$.

A section of $\omega_{X_{|Y}} \otimes \eta_Y^i \simeq \mathcal{O}_Y(1)$ for $i = 1, 2, 3$ is of the form $f_i(x, y) = a_i x + b_i y$, with a_i, b_i constants. Suppose that the sections f_i are in the image of the inclusion

$$H^0(X, \omega_X \otimes \eta_X^i) \hookrightarrow H^0(Y, \mathcal{O}_Y(1)).$$

By evaluating the section at the points q_i and using the gluing conditions, one gets $a_i = c_1^i - 1$ and $b_i = c_2^i - 1$. One obtains a similar condition for the image of the sections of $H^0(X, \omega_X \otimes \eta_X^{7-i})$ in $H^0(Y, \mathcal{O}_Y(1))$. Set $j = 7 - i$. By multiplying the corresponding sections of $\omega_X \otimes \eta_X^i$ and $\omega_X \otimes \eta_X^j$ we have that an element in the image of $\mu_{X,\eta}$ is of the form

$$\bigl(2 - c_1^i - c_1^j\bigr)x^2 + \bigl(2 - c_2^i - c_2^j\bigr)y^2 - \bigl(2 - c_1^j - c_2^i + c_1^j c_2^i - c_1^i - c_2^j + c_1^i c_2^j\bigr)xy.$$

Hence, after taking the sum of such sections for $i = 1, 2, 3$ we conclude that there is a nontrivial element in the kernel of $\mu_{X,\eta}$ if and only if there is a nontrivial solution for the linear system $A\boldsymbol{x} = 0$ with

$$A = \begin{pmatrix} 2 - c_1 - c_1^6 & 2 - c_1 c_2^6 - c_1^6 c_2 & 2 - c_2 - c_2^6 \\ 2 - c_1^2 - c_1^5 & 2 - c_1^2 c_2^5 - c_1^5 c_2^2 & 2 - c_2^2 - c_2^5 \\ 2 - c_1^3 - c_1^4 & 2 - c_1^3 c_2^4 - c_1^4 c_2^3 & 2 - c_2^3 - c_2^4 \end{pmatrix}.$$

Clearly, if $c_i = 1$ for some i or $c_1 = c_2$, the determinant of A vanishes. We compute

$$\tfrac{1}{7}\det A = c_1^6(c_2^3 - c_2^5) + c_1^5(c_2^6 - c_2^3) + c_1^4(c_2^2 - c_2) + c_1^3(c_2^5 - c_2^6) + c_1^2(c_2 - c_2^4) + c_1(c_2^4 - c_2^2).$$

Suppose that $c_i \neq 1$ and $c_2 = c_1^k$ for some $2 \leq k \leq 6$. Then a straightforward computation shows that $\det A \neq 0$ if and only if $k = 3$ or $k = 5$.

In conclusion, we can find a limit linear series (X, η_X) with $\eta_X^7 \simeq \mathcal{O}_X$, for suitable values of the c_i, such that the composition map in the commutative diagram

$$\begin{tikzcd}
\bigoplus_{i=1}^3 H^0(X, \omega_X \otimes \eta^i) \otimes H^0(X, \omega_X \otimes \eta^{7-i}) \arrow[r] \arrow[d, hook] & H^0(X, \omega_X^2) \arrow[d, "\simeq"] \\
H^0(Y, \mathcal{O}_Y(1)) \otimes H^0(Y, \mathcal{O}_Y(1)) \arrow[r] & H^0(Y, \mathcal{O}_Y(2))
\end{tikzcd}$$

is an isomorphism. □

7. A complete fiber of $\widetilde{\mathbf{Pr}}_{2,7}$

For a special point of $\mathcal{B}_D$, consider a smooth curve E of genus 1. Then the kernel of the addition map

$$X = X(E) := \ker(m : E^7 \to E) \quad \text{with} \quad m(x_1, \dots, x_7) = x_1 + \cdots + x_7$$

is an abelian variety of dimension 6, isomorphic to E^6. The kernel of the induced polarization of the canonical principal polarization of E^7 is $\{(x, \dots, x) : x \in E[7]\}$, which consists of 7^2 elements. So the polarization on X induced by the canonical polarization of E^7 is of type $D = (1, 1, 1, 1, 1, 7)$. Since the symmetric group $\mathcal{S}_7$ acts on E^7 in the obvious way, X admits an automorphism of order 7. Hence X with the induced polarization is an element of $\mathcal{B}_D$. To be more precise, the group $\mathcal{S}_7$ admits exactly 120 subgroups of order 7. Hence to every elliptic curve there exist exactly 120 abelian varieties X as above with G-action. All of them are isomorphic to each other, since the corresponding subgroups are conjugate to each other according the Sylow theorems. We want to determine the complete preimage $\widetilde{\mathrm{Pr}}_{2,7}^{-1}(X)$ of X. We need some lemmas. For simplicity we denote by $\mathrm{Pr}(f)$ the Prym variety of a covering f in $\widetilde{\mathcal{R}}_{2,7}$.

Lemma 7.1. *Let $f : \widetilde{C} \to C$ be a covering satisfying* $(**)$ *with $g = 2$ such that* $\mathrm{Pr}(f) \simeq X$. *Then C contains a node of index* 1.

Proof. Suppose that either C is smooth or all the nodes of C are of index 7. Then the exact sequence (3-3) gives an isogeny $j : \mathrm{Pr}(f) \to \mathrm{Pr}(\tilde{f})$ onto the Prym variety of the normalization $\tilde{f}$ of f. Actually, in the smooth case, j is an isomorphism and, if there is a node of index 1, the kernel $\widetilde{T}[7]$ is positive dimensional. The isomorphism $\mathrm{Pr}(f) = X$ implies that the kernel of j is of the form $\{(x, \dots, x)\}$ with $x \in X[7]$. Hence the action of the symmetric group S_7 on X descends to a nontrivial action on $\mathrm{Pr}(\tilde{f})$.

We can extend this action to $J\widetilde{N}$ by combining it with the identity on JN. Namely, $J\widetilde{N} \simeq (JN \times \mathrm{Pr}(\tilde{f}))/H$, where H is constructed as follows. Let $\langle \eta \rangle \subset JN[7]$

be the subgroup defining the covering $\tilde{f}$ and let $H_1 \subset JN[7]$ be its orthogonal complement with respect to the Weil pairing. Then $H = \{(\alpha, -f^*\alpha) : \alpha \in H_1\}$. Since $f^*H_1 = \{(x, \dots, x) : x \in E[7]\} \subset \mathrm{Pr}(f)$, we get an S_7-action on $J\widetilde{N}$ which is clearly nontrivial.

If C is smooth, then $\widetilde{N} \simeq \widetilde{C}$. On the other hand, if all the nodes of C are of index 7, then $\widetilde{N}$ consists of at least two components. In any case, for each component $\widetilde{N}_i$ of $\widetilde{N}$ we have

$$\#\operatorname{Aut}(J\widetilde{N}_i) \geq \tfrac{1}{2}\# S_7 = 2520.$$

Moreover, according to a classical theorem of Weil, $\operatorname{Aut}\widetilde{N}_i$ embeds into $\operatorname{Aut}J\widetilde{N}_i$ with quotient of order ≤ 2. So

$$\#\operatorname{Aut}\widetilde{N}_i \geq \tfrac{1}{2}\#\operatorname{Aut}(J\widetilde{N}_i).$$

On the other hand, $\widetilde{N}_i$ is a smooth curve of genus ≤ 8. So Hurwitz's theorem implies

$$\#\operatorname{Aut}(\widetilde{N}_i) \leq 84 \cdot (8-1) = 588.$$

Together, this gives a contradiction. $\square$

Lemma 7.2. *Let $f : \widetilde{C} \to C$ be a covering satisfying $(**)$ such that C has a component containing nodes of index* 1 *and* 7*. Then any node of index* 1 *is the intersection with another component of C.*

Proof. Suppose x and y are nodes of C in a component C_i of index 1 and 7, respectively. Then the preimage $f^{-1}(C_i)$ is a component, since over y the map f is totally ramified. Since $f^{-1}(x)$ consists of 7 nodes, the equality $n_1 = c_1$ implies that x is the intersection of 2 components. $\square$

Theorem 7.3. *Let $X = \ker(m : E^7 \to E)$ be a polarized abelian variety as above. The fiber $\widetilde{\mathrm{Pr}}_{2,7}^{-1}(X)$ consists of the following* 4 *types of elements of $\widetilde{\mathcal{R}}_{2,7}$* (see Figure 1).

(i) *$C = E/p \sim q$ and $\widetilde{C} = \bigsqcup_{i=1}^{7} E_i/p_i \sim q_{i+1}$ with $E_i \simeq E$ for all i and $q_8 = q_1$ and we can enumerate in such a way that the preimages of p and q are $p_i, q_i \in E_i$.*

(ii) *$C = E_1 \cup_p E_2$ consists of* 2 *elliptic curves intersecting in one point p. Then up to exchanging E_1 and E_2 we have: $\widetilde{C}$ consists of an elliptic curve F_1, which is a 7-fold cover of E_1 and 7 copies of $E_2 \simeq E$ not intersecting each other and intersecting F_1 each in one point.*

(iii) *$C = E_1 \cup_p E_2$ with E_2 elliptic and E_1 rational with a node at q. Then $E_2 \simeq E$ and $f^{-1}(E_2)$ consists of* 7 *disjoint curves all isomorphic to E and $f^{-1}(E_1)$ is a rational curve with one node lying* 7:1 *over E_1 and intersecting each component of $f^{-1}(E_2)$ in a point over p.*

(iv) *$C = E_1 \cup_p E_2$ as in* (iii) *and $\widetilde{C}$ is an étale G-cover over C.*

We call the coverings of the theorem *of type* (i), (ii), (iii) and (iv), respectively. Theorem 7.3 will be used in Proposition 7.5 to describe the complete fibers of the Prym map over $X(E)$.

Proof. There are 7 types of stable curves of genus 2. We determine the coverings $f : \widetilde{C} \to C$ in $\widetilde{\Pr}_{2,7}^{-1}(X)$ in each case separately.

(1) There is no étale G-cover $f : \widetilde{C} \to C$ of a smooth curve C of genus 2 such that $P = \Pr(f) \simeq X$. This is a direct consequence of Lemma 7.1.

(2) If $C = E/p \sim q$ then the singular point of C is of index 1 and $f : \widetilde{C} \to C$ is a G-covering satisfying $(**)$ such that $P = \Pr(f) \simeq X$. Then

$$\widetilde{C} = \bigsqcup_{i=1}^{7} E_i/(p_i \sim q_{i+1})$$

with $E_i \simeq E$ and we enumerate in such a way that the preimages of p and q are p_i and q_i with $q_8 = q_1$. In this case $\Pr(f) \simeq X$.

Proof: According to Lemma 7.1 the node is necessarily of index 1 and thus the map $\tilde{f} : \widetilde{C} \to C$ is étale. The exact sequence (3-3) together with Corollary 3.4 gives an isomorphism $P \simeq R$ with R the Prym variety of the map $\tilde{f}$. Clearly we can enumerate the components of $\widetilde{N}$ in such a way that $\widetilde{C}$ is as above and R is the kernel of the map $m : \times_{i=1}^{7} E \to E$, i.e., $R \simeq X$. We are in case (i) of the theorem.

(3) There is no rational curve C with 2 nodes admitting a G-cover $f : \widetilde{C} \to C$ satisfying $(**)$ such that $P = \Pr(f) \simeq X$.

Proof: Suppose there is such a covering. By Lemmas 7.1 and 7.2 both nodes are of index 1 and hence the map $\widetilde{C} \to C$ is étale. Then all components of $\widetilde{C}$ are rational. This implies that $P \simeq \mathbb{C}^{*6}$ is not an abelian variety, a contradiction.

(4) Let $C = E_1 \cup_p E_2$ consist of 2 elliptic curves intersecting in one point and let $f : \widetilde{C} \to C$ be a covering satisfying $(**)$ such that $P = \Pr(f) \simeq X$. Then, up to exchanging E_1 and E_2, we have that $\widetilde{C}$ consists of an elliptic curve F_1, which is a 7-fold cover of E_1 and 7 copies of $E_2 \simeq E$ not intersecting each other and intersecting F_1 each in one point. So $X = \ker(m : E_2^7 \to E_2)$.

Proof: By Lemma 7.1 the node is of index 1 and the map $f : \widetilde{C} \to C$ is étale. Since there is no connected graph with 14 vertices and 7 edges, we are necessarily in case (ii) of the theorem.

(5) Suppose $C = E_1 \cup_p E_2$ with components E_2 elliptic and E_1 rational with a node q and $f : \widetilde{C} \to C$ a G-covering satisfying $(**)$ such that $P = \Pr(f) \simeq X$. Then $E_2 \simeq E$ and either $f : \widetilde{C} \to C$ is étale and connected or $\widetilde{C}$ consists of 7 components all isomorphic to E and a rational component F_2 over E_2 totally ramified exactly over q and intersecting each E_i exactly in one point lying over p. So $\Pr(f) \simeq X$.

Proof: According to Lemma 7.1 at least one node of C is of index 1. Suppose first that both nodes are of index 1. Then clearly f is étale and we are in case (iv) of the theorem. If only one node is of index 1, then according to Lemma 7.2 q is of index 7 and p of index 1. This gives case (iii) of the theorem.

(6) There is no curve C consisting of 2 rational components intersecting in one point p admitting a G-cover $f : \widetilde{C} \to C$ satisfying $(**)$ such that $P = \mathrm{Pr}(f) \simeq X$.

Proof: According to Lemmas 7.1 and 7.2 the node p is of index 1 and the nodes q_1 and q_2 of the rational components of C are of index 1 or 7. By the Hurwitz formula all components of $\widetilde{C}$ are rational. This implies $\mathrm{Pr}(f) \simeq \mathbb{C}^{*6}$, contradicting $(**)$.

(7) If C is the union of 2 rational curves intersecting in 3 points, there is no cover $f : \widetilde{C} \to C$ satisfying $(**)$ such that $P = \mathrm{Pr}(f) \simeq X$.

Proof: By Lemma 7.1 at least one of the 3 nodes of C is of index 1. So $\widetilde{C}$ consists of at least 8 components. But then the other nodes also are of index 1, because if one node is of index 7, the curve $\widetilde{C}$ consists of 2 components only. Hence all 3 nodes are of index 1. But then all components of $\widetilde{C}$ are rational. So $P = \mathrm{Pr}(f)$ cannot be an abelian variety, contradicting $(**)$.

Together, steps (1)–(7) prove the theorem. □

Corollary 7.4. *Let E be a general elliptic curve. Then the fiber $\widetilde{\mathrm{Pr}}_{2,7}^{-1}(X(E))$ consists of two irreducible components S_1 and S_2. The component S_1 is a covering over E whose points correspond to coverings of type* (i), *except for one point*, *which corresponds to a covering of type* (iv). *The component S_2 is a finite covering of the moduli space of elliptic curves*, *and every point corresponds to a covering of type* (ii), *except for two points*, *one of which corresponds to a covering of type* (iii) *and the other to a covering of type* (iv). *The components S_1 and S_2 intersect at a point corresponding to a covering of type* (iv) (see Figure 1).

Proof. It is known that for 2 elliptic curves $E_1 \neq E_2$ we can have $X(E_1) \simeq X(E_2)$ as abelian varieties, but not necessarily as polarized abelian varieties. Hence $X(E)$ determines E (which can be seen also from Theorem 7.3).

We claim that the coverings of type (iv) are contained in S_1 and S_2, whereas the coverings of type (iii) are contained in S_2 only: it is known that a curve $\widetilde{C}$ degenerates to a curve $\widetilde{C}'$ of some other type if and only if the dual graph of $\widetilde{C}'$ can be contracted to the dual graph of $\widetilde{C}$. On the other hand, the locus of curves covering some curve of genus ≥ 2 of some fixed degree is closed in the moduli space of curves. Now considering the dual graphs of the curves $\widetilde{C}$ of the coverings of the different types gives the assertion.

In the case of coverings of type (i) we have $C = E/p_1 \sim p_2$. We can use the translations of E to fix p_1, and then p_2 is free, which gives the assertion, since there are only finitely many étale coverings of C.

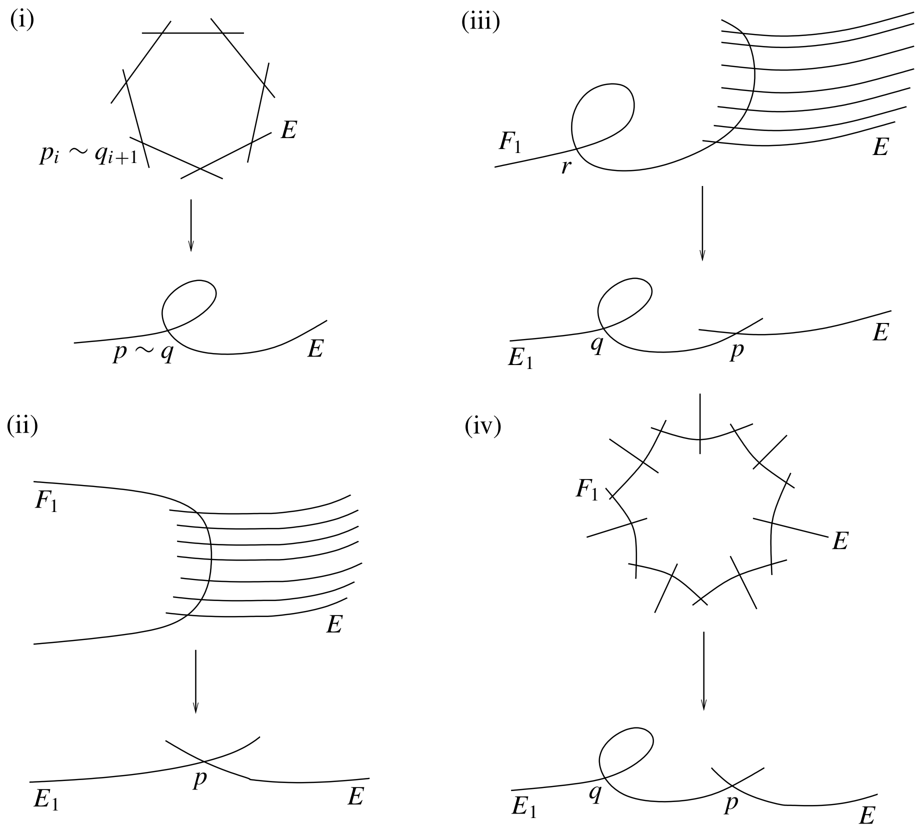

Figure 1. Admissible coverings on the fiber of $X(E)$.

In case (ii) we have $C = E_1 \cup_p E_2$, where E_1 is an arbitrary elliptic curve and $E_2 \simeq E$. Since p may be fixed with an isomorphism of E and E_2, this gives the isomorphism of $\widetilde{\mathrm{Pr}}_{2,7}^{-1}(E)$ with a finite covering of the moduli space of elliptic curves, again since there are only finitely many coverings $\widetilde{C}$ of type (ii) of C.

Finally, in cases (iii) and (iv), the 3 points of the normalization of E_1 given by p and the 2 preimages of the node, which we can assume to be 1, 0, and ∞, respectively, determine the curve C uniquely. For type (iii) the induced map on the normalization of F_1 is a 7:1 map $h : \mathbb{P}^1 \to \mathbb{P}^1$ totally ramified at 2 points, which we assume to be ∞ and 0. So h can be expressed as a polynomial in one variable of degree 7, with vanishing order 7 at 0 and such that $h(1) = 1$, that is, $h(x) = x^7$. Then the map h, and hence the covering, is uniquely determined. For a covering of type (iv) over C we consider 7 copies of $\mathbb{P}^1$, where the point 1 on every rational component is identified to the point ∞ of another rational component, and we attach elliptic curves isomorphic to E_2 at each point 0. The number of étale coverings is the number of subgroups of order 7 in $JE_1[7] \simeq \mathbb{Z}/7\mathbb{Z}$ (the 7-torsion points in the nodal curve E_1 are determined by a 7-th root of unity). So there is only one such covering up to isomorphism. $\square$

Varying the elliptic curve E, we obtain a one-dimensional locus $\mathcal{E} \subset \mathcal{B}_D$ consisting of the polarized abelian varieties $X(E)$ with G-action as above. Let $\mathcal{S}$ denote the preimage of $\mathcal{E}$ under the extended Prym map $\widetilde{\mathrm{Pr}}_{2,7} : \widetilde{\mathcal{R}}_{2,7} \to \mathcal{B}_D$. The next proposition is a direct consequence of Corollary 7.4.

Proposition 7.5. *The scheme $\mathcal{S}$ has dimension* 2 *and is the union of* 2 *closed subschemes*

$$\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2,$$

where $\mathcal{S}_1$ parametrizes coverings of type (i) *and* (iv), *and $\mathcal{S}_2$ parametrizes coverings of type* (ii), (iii) *and* (iv). *In particular*, *they intersect exactly in the points parametrizing coverings of type* (iv).

8. The codifferential on the boundary divisors

In this section we will give bases of the Prym differentials and an explicit description of the codifferential of the Prym map.

Let $f : \widetilde{C} \to C$ be a covering corresponding to a point of $\mathcal{S}$. We want to compute the rank of the codifferential of the Prym map $\mathrm{Pr}_{2,7} : \widetilde{\mathcal{R}}_{2,7} \to \mathcal{B}_D$ at the point $[f : \widetilde{C} \to C] \in \widetilde{\mathcal{R}}_{2,7}$. According to [Donagi and Smith 1981] this codifferential is the map

$$\mathcal{P}^* : S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)^G \to H^0(C, \Omega_C \otimes \omega_C),$$

where Ω_C is the sheaf of Kähler differentials on C and $S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)^G$ is the cotangent space to $\mathcal{B}_D$ at the Prym variety of the covering $f : \widetilde{C} \to C$. Letting $j : \Omega_C \to \omega_C$ denote the canonical map, we first compute the rank of the composed map

$$S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)^G \xrightarrow{\mathcal{P}^*} H^0(C, \Omega_C \otimes \omega_C) \xrightarrow{j} H^0(C, \omega^2).$$

Suppose first that f is of type (i), (ii) or (iv). In these cases the covering f is étale and hence given by a 7-division point η of JC. According to the analogue of [Lange and Ortega 2011, Equation (3.4)] we have

$$H^0(\widetilde{C}, \omega_C)^- = \bigoplus_{i=1}^{6} H^0(C, \omega_C \otimes \eta^i) \tag{8-1}$$

and hence

$$S^2\big(H^0(\widetilde{C}, \omega_C)^-\big)^G = \bigoplus_{i=1}^{3} \big(H^0(C, \omega_C \otimes \eta^i) \otimes H^0(C, \omega_C \otimes \eta^{7-i})\big). \tag{8-2}$$

Using this, the above composed map is just the sum of the cup product map

$$\phi : \bigoplus_{i=1}^{3} \big(H^0(C, \omega_C \otimes \eta^i) \otimes H^0(C, \omega_C \otimes \eta^{7-i})\big) \longrightarrow H^0(C, \omega_C^2), \tag{8-3}$$

whose rank we want to compute first.

We shall give a suitable basis for the space of Prym differentials. First we consider a covering of type (i) constructed as follows. Let E be a smooth curve of genus 1 and let $q \neq q'$ be two fixed points of E. Then

$$C := E/q \sim q'$$

is a stable curve of genus 2 with normalization $n : E \to C$ and node $p := n(q) = n(q')$. Let $f : \widetilde{C} \to C$ be a cyclic étale covering with Galois group $G = \langle \sigma \rangle \simeq \mathbb{Z}/7\mathbb{Z}$. The normalization $\tilde{n} : \widetilde{N} \to \widetilde{C}$ consists of 7 components $N_i \simeq E$ with $\sigma(N_i) = N_{i+1}$ for $i = 1, \dots, 7$ and $N_8 = N_1$. Let q_i and q_i' be the elements of N_i corresponding to q and q'. Then $\tilde{n}(q_i) = \tilde{n}(q_{i+1}') =: p_i$ for $i = 1, \dots, 7$ with $q_8' = q_1'$. Clearly $\sigma(p_i) = p_{i+1}$ for all i.

Recall that $\omega_{\widetilde{C}}$ is the subsheaf of $\tilde{n}_*\big(\mathcal{O}_{\widetilde{N}} \sum (q_i + q_i')\big)$ consisting of (local) sections φ which considered as sections of $\mathcal{O}_{\widetilde{N}} \sum (q_i + q_i')$ satisfy the condition

$$\mathrm{Res}_{q_i}(\varphi) + \mathrm{Res}_{q_{i+1}'}(\varphi) = 0$$

for $i = 1, \dots, 7$. Here we use the fact that $\omega_{\widetilde{N}} = \mathcal{O}_{\widetilde{N}}$. Consider the following elements of $H^0\big(\widetilde{C}, \tilde{n}_*\big(\mathcal{O}_{\widetilde{N}} \sum (q_i + q_i')\big)\big)$, regarded as sections on $\widetilde{N}$:

$$\omega_1 := \begin{cases} \text{nonzero section of } \mathcal{O}_{N_1}(q_1 + q_1') \text{ vanishing at } q_1 \text{ and } q_1', \\ 0 \text{ elsewhere}, \end{cases}$$

$$\omega_i := (\sigma^{-i})^*(\omega_1) \quad \text{for } i = 2, \dots, 7.$$

Note that ω_i is nonzero on N_i vanishing at q_i and q_i' and zero elsewhere.

Now we construct similar differentials for coverings of type (ii). Let

$$C = E_1 \cup_p E_2$$

consist of 2 elliptic curves E_1 and E_2 intersecting transversally in one point p, and let $f : \widetilde{C} \to C$ be a covering of type (ii). So $\widetilde{C}$ consists of an elliptic curve F_1, which is an étale cyclic cover of E_1 of degree 7 and 7 disjoint curves $E_2^1, \dots, E_2^7$ all isomorphic to E_2. The curve E_i intersects F_1 transversally in a point p_i, such that the group G permutes the curves E_i and the points p_i cyclically, i.e., $\sigma(E_2^i) = E_2^{i+1}$ and $\sigma(p_i) = p_{i+1}$ with $E_2^8 = E_2^1$ and $p_8 = p_1$.

Let $\tilde{n} : \widetilde{N} \to \widetilde{C}$ denote the normalization map. Then $\widetilde{N}$ is the disjoint union of the 8 elliptic curves $F_1, E_2^1, \dots, E_2^7$. We denote the point p_i by the same letter when considered as a point of F_1 and E_2^i. Consider the line bundle

$$L = \mathcal{O}_{F_1}(p_1 + \cdots + p_7) \sqcup \mathcal{O}_{E_2^1}(p_1) \sqcup \cdots \sqcup \mathcal{O}_{E_2^7}(p_7).$$

Then $\omega_{\widetilde{C}}$ is the subsheaf of $\tilde{n}_*(L)$ consisting of (local) sections φ which considered as sections of L satisfy the condition

$$(\mathrm{Res}_{p_i})_{|F_1}(\varphi) + (\mathrm{Res}_{p_i})_{|E_2^i}(\varphi) = 0 \tag{8-4}$$

for $i = 1, \dots, 7$. Consider the following sections of $\tilde{n}_*(L)$, regarded as sections on $\widetilde{N}$:

$$\omega_1 := \begin{cases} \text{nonzero sections of } \mathcal{O}_{F_1}(p_1) \text{ and } \mathcal{O}_{E_2^1}(p_1) \text{ satisfying (8-4) at } p_1, \\ 0 \text{ elsewhere,} \end{cases}$$

$$\omega_i := (\sigma^{-i})^*(\omega_1) \quad \text{for } i = 2, \dots, 7.$$

Thus ω_i is nonzero on $F_1(p_i) \sqcup E_2^i(p_i)$ vanishing at p_i and zero elsewhere.

We construct the analogous differentials for the covering f of type (iv), which is uniquely determined according to Proposition 7.5. So let

$$C = E_1 \cup_p E_2$$

with E_2 elliptic and E_1 a rational curve with one node q and let $f : \widetilde{C} \to C$ be the covering of type (iv). Then $\widetilde{C}$ consists of 14 components $F_1, \dots, F_7$ isomorphic to $\mathbb{P}^1$ with $f_{|F_i} : F_i \to E_1$ the normalization and $E_2^1, \dots, E_2^7$ all isomorphic to E_2 with $f_{|E_2^i} : E_2^i \to E_2$ the isomorphism. Then E_2^i intersects F_i in the point p_i lying over p and no other component of $\widetilde{C}$. If q_i and q_i' are the points of F_i lying over q, the F_i and F_{i+1} intersect transversally in the points q_i and q_{i+1}' for $i = 1, \dots, 7$, where $q_8' = q_1$. The group G permutes the components and points cyclically, i.e., $\sigma(F_1) = F_{i+1}$ and similarly for E_2^i, p_i, q_i and q_i'.

The normalization $\tilde{n} : \widetilde{N} \to \widetilde{C}$ of the curve $\widetilde{C}$ is the disjoint union of the components F_i and E_2^i. We denote also the point p_i by the same letter when considered as a point of F_i and E_2^i. Consider the following line bundle on $\widetilde{N}$:

$$L = \bigsqcup_{i=1}^{7} \mathcal{O}_{F_i}(q_i + q_i' + p_i) \sqcup \bigsqcup_{i=1}^{7} \mathcal{O}_{E_2^i}(p_i).$$

Then $\omega_{\widetilde{C}}$ is the subsheaf of $\tilde{n}_*(L)$ consisting of (local) sections φ which viewed as sections of L satisfy the conditions

$$(\mathrm{Res}_{p_i})_{|F_i}(\varphi) + (\mathrm{Res}_{p_i})_{|E_2^i}(\varphi) = 0, \quad (\mathrm{Res}_{q_i})_{|F_i}(\varphi) + (\mathrm{Res}_{q_{i+1}'})_{|F_{i+1}}(\varphi) = 0 \tag{8-5}$$

for $i = 1, \dots, 7$. Consider the following sections of $\tilde{n}_*(L)$, regarded as sections on $\widetilde{N}$:

$$\omega_1 := \begin{cases} \text{nonzero sections of } \omega_{F_1}(q_1 + q_1' + p_1) \text{ and } \mathcal{O}_{E_2^1}(p_1) \text{ satisfying (8-5)} \\ \quad \text{at } p_1 \text{ and vanishing at } q_1 \text{ and } q_1', \\ 0 \text{ elsewhere,} \end{cases}$$

$$\omega_i := (\sigma^{-i})^*(\omega_1) \quad \text{for } i = 2, \dots, 7.$$

Note that up to a multiplicative constant there is exactly one such section ω_1, since $h^0(\omega_{F_1}(q_1 + q_1' + p_1)) = 2$ and $h^0(\mathcal{O}_{E_2^1}(p_1)) = 1$.

Finally, we consider coverings of type (iii). Let $C = E_1 \cup_p E_2$, as for the covering of type (iv) above, and let $f : \widetilde{C} \to C$ be a covering of type (iii). So $\widetilde{C}$ consists of a rational curve F_1 with a node r lying over the node q of C and 7 components $E_2^1, \dots, E_2^7$ all isomorphic to E_2. Then E_2^i intersects F_1 in the point p_i lying over p and intersects no other component of $\widetilde{C}$. The group G acts on F_1 with only a fixed point r and permutes the E_i^2 and p_i cyclically as above. We use the following partial normalization $\tilde{n} : \widetilde{N} \to \widetilde{C}$ of $\widetilde{C}$:

$$\widetilde{N} := F_1 \sqcup \bigsqcup_{i=1}^{7} E_2^i.$$

Consider the following line bundle on $\widetilde{N}$:

$$L = \mathcal{O}_{F_1}(p_1 + \cdots + p_7) \sqcup \bigsqcup_{i=1}^{7} \mathcal{O}_{E_2^i}(p_i).$$

Since the canonical bundles of F_1 and E_2^i are trivial, it is clear that $\omega_{\widetilde{C}}$ is the subsheaf of $\tilde{n}_*(L)$ consisting of (local) sections φ which regarded as sections of L satisfy the relations

$$(\mathrm{Res}_{p_i})_{|F_1}(\varphi) + (\mathrm{Res}_{p_i})_{|E_2^i}(\varphi) = 0 \tag{8-6}$$

for $i = 1, \dots, 7$. As before, define a section

$$\omega_1 := \begin{cases} \text{nonzero sections of } \mathcal{O}_{F_1}(p_1 + \cdots + p_7) \text{ and } \mathcal{O}_{E_2^i}(p_1) \\ \quad \text{vanishing at } p_1, \dots, p_7, \\ 0 \text{ elsewhere,} \end{cases}$$

of $\tilde{n}_*(L)$, considered as a section of $\widetilde{N}$, and define the sections ω_i, for $i = 2, \dots, 7$, as in the previous cases. Note that up to a multiplicative constant there is exactly one such section ω_1.

From now on, $f : \widetilde{C} \to C$ will be a covering of type (i)–(iv) as above. We fix a primitive 7-th root of unity, for example, $\rho := e^{2\pi i/7}$, and define for $i = 0, \dots, 6$

the section

$$\Omega_i := \sum_{j=1}^{7} \rho^{ij} \omega_j.$$

Clearly Ω_i is a global section of L that defines a section of $\omega_{\widetilde{C}}$, which we denote with the same symbol.

Lemma 8.1. *$\sigma^*(\Omega_i) = \rho^i \Omega_i$ for $i = 0, \dots, 6$. In particular, $\Omega_0 \in H^0(\widetilde{C}, \omega_{\widetilde{C}})^+$ and $\{\Omega_1, \dots, \Omega_6\}$ is a basis of $H^0(\widetilde{C}, \omega_{\widetilde{C}})^-$.*

Proof. The first assertion follows from a simple calculation using the definition of ω_i. So clearly $\Omega_0 \in H^0(\widetilde{C}, \omega_{\widetilde{C}})^+$ and $\Omega_i \in H^0(\widetilde{C}, \omega_{\widetilde{C}})^-$ for $i = 1, \dots, 6$. Since $\Omega_1, \dots, \Omega_6$ are in different eigenspaces of σ, they are linearly independent and since $H^0(\widetilde{C}, \omega_{\widetilde{C}})^-$ is of dimension 6, they form a basis. □

Remark 8.2. In cases (i), (ii) and (iv), $H^0(C, \omega_C \otimes \eta^{7-i})$ is the eigenspace of σ^i and Ω_i is a generator for $i = 1, \dots, 6$.

Proposition 8.3. *The map*

$$\phi : S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)^G \longrightarrow H^0(C, \omega_C^2)$$

is of rank 1.

Proof. We have to show that the kernel of ϕ is 2-dimensional. A basis of $S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)^G$ is given by $\{\Omega_1 \otimes \Omega_6, \Omega_2 \otimes \Omega_5, \Omega_3 \otimes \Omega_4\}$. Let a, b, c be complex numbers with

$$\phi(a\Omega_1 \otimes \Omega_6 + b\Omega_2 \otimes \Omega_5 + c\Omega_3 \otimes \Omega_4) = 0.$$

Define, for $i = 1, \dots, 7$,

$$\psi_i := \sum_{i=1}^{7} \omega_j \otimes \omega_{j+i-1}.$$

An easy but tedious computation gives

$$\Omega_1 \otimes \Omega_6 = \psi_1 + \rho\psi_7 + \rho^2\psi_6 + \rho^3\psi_5 + \rho^4\psi_4 + \rho^5\psi_3 + \rho^6\psi_2,$$
$$\Omega_2 \otimes \Omega_5 = \psi_1 + \rho\psi_4 + \rho^2\psi_7 + \rho^3\psi_3 + \rho^4\psi_6 + \rho^5\psi_2 + \rho^6\psi_5,$$
$$\Omega_3 \otimes \Omega_4 = \psi_1 + \rho\psi_3 + \rho^2\psi_5 + \rho^3\psi_7 + \rho^4\psi_2 + \rho^5\psi_4 + \rho^6\psi_6.$$

So we get

$$\begin{aligned} 0 &= \phi\big((a+b+c)\psi_1 + (a\rho^6+b\rho^5+c\rho^4)\psi_2 + (a\rho^5+b\rho^3+c\rho)\psi_3 \\ &\qquad + (a\rho^4+b\rho+c\rho^5)\psi_4 + (a\rho^3+b\rho^6+c\rho^2)\psi_5 \\ &\qquad + (a\rho^2+b\rho^4+c\rho^6)\psi_6 + (a\rho+b\rho^2+c\rho^3)\psi_7\big) \\ &= (a+b+c)(\omega_1^2+\cdots+\omega_7^2) \\ &\qquad + \big(a(\rho+\rho^6)+b(\rho^2+\rho^5)+c(\rho^3+\rho^4)\big)\sum_{j=1}^{7}\omega_j\omega_{j+1} \\ &\qquad + \big(a(\rho^2+\rho^5)+b(\rho^3+\rho^4)+c(\rho+\rho^6)\big)\sum_{j=1}^{7}\omega_j\omega_{j+2} \\ &\qquad + \big(a(\rho^3+\rho^4)+b(\rho+\rho^6)+c(\rho^2+\rho^5)\big)\sum_{j=1}^{7}\omega_j\omega_{j+3}. \end{aligned}$$

This section is zero if and only if its restriction to any component is zero. Now the restriction to N_i for all i gives

$$0 = a+b+c+(6a+6b+6c)\sum_{j=1}^{6}\rho^j = -5(a+b+c).$$

So $\phi(a\Omega_1\otimes\Omega_6 + b\Omega_2\otimes\Omega_5 + c\Omega_3\otimes\Omega_4) = 0$ if and only if $a+b+c=0$. Hence the kernel of ϕ is of dimension 2, which proves the proposition. □

Proposition 8.3 shows that the codifferential map along the divisors $\mathcal{S}_1$ and $\mathcal{S}_2$ in Proposition 7.5 is not surjective. In fact, as we will see later, the kernel of ϕ coincides with the conormal bundle of the image of these divisors in $\mathcal{B}_D$. In order to compute the degree we will perform a blow-up along these divisors.

Let $\mathcal{E} \subset \mathcal{B}_D$ denote the one-dimensional locus consisting of the abelian varieties which are of the form $X = \ker(m : E^7 \to E)$ with $m(x_1,\ldots,x_7) = x_1+\cdots+x_7$ for a given elliptic curve E. As we saw in Section 7, the induced polarization is of type D. Note that $\mathcal{E}$ is a closed subset of $\mathcal{B}_D$. The aim is to compute the degree of $\mathrm{Pr}_{7,2}$ above a point $X \in \mathcal{E}$. We denote by $\mathcal{S} \subset \widetilde{\mathcal{R}}_{2,7}$ the inverse image of $\mathcal{E}$ under $\widetilde{\mathrm{Pr}}_{2,7}$. According to Proposition 7.5, $\mathcal{S}$ is a divisor consisting of 2 irreducible components in the boundary $\widetilde{\mathcal{R}}_{2,7} \setminus \mathcal{R}_{2,7}$. We have $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, where a general point of $\mathcal{S}_1$ corresponds to the G-covers with base an irreducible nodal curve of genus 1, and a point of $\mathcal{S}_2$ corresponds to a product of elliptic curves intersecting in a point. Moreover, for any fixed elliptic curve E, $\mathcal{S}_1$ and $\mathcal{S}_2$ intersect in the unique point given by the covering of type (iv).

As in [Donagi and Smith 1981], let $\widetilde{\mathcal{B}}_D$ be the blow-up of $\mathcal{B}_D$ along $\mathcal{E}$ and $\widetilde{\mathcal{R}} \simeq \widetilde{\mathcal{R}}_{2,7}$ the blow-up of $\widetilde{\mathcal{R}}_{2,7}$ along the divisor $\mathcal{S} = \widetilde{\mathrm{Pr}}_{2,7}^{-1}(\mathcal{E})$. We then obtain the commutative diagrams

$$\begin{tikzcd} \widetilde{\mathcal{R}} \arrow[r, "\widetilde{\mathcal{P}}"] \arrow[d, "\simeq"'] & \widetilde{\mathcal{B}}_D \arrow[d] \\ \widetilde{\mathcal{R}}_{2,7} \arrow[r, "\widetilde{\mathrm{Pr}}_{2,7}"] & \mathcal{B}_D \end{tikzcd} \qquad \begin{tikzcd} \widetilde{\mathcal{S}} \arrow[r, "\widetilde{\mathcal{P}}"] \arrow[d, "\simeq"'] & \widetilde{\mathcal{E}} \arrow[d, "\mathbb{P}^1\text{-bundle}"] \\ \mathcal{S} \arrow[r, "\widetilde{\mathrm{Pr}}_{2,7}"] & \mathcal{E} \end{tikzcd}$$

in which $\widetilde{\mathcal{S}}$ and $\widetilde{\mathcal{E}}$ are the exceptional loci.

Lemma I.3.2 of [Donagi and Smith 1981] guarantees that the local degree of $\widetilde{\mathrm{Pr}}_{2,7}$ along a component of $\mathcal{S}$ equals the degree of the induced map on the exceptional divisors $\widetilde{\mathcal{P}}_{|\mathcal{S}_i} : \mathcal{S}_i \to \widetilde{\mathcal{E}}$ if the codifferential map $\mathcal{P}^*$ is surjective on the respective conormal bundles. Recall that the fibers of the conormal bundles at the point X are given by

$$\mathcal{N}^*_{X,\mathcal{E}/\mathcal{B}_D} = \ker(T^*_X\mathcal{B}_D \to T^*_X\mathcal{E}),$$
$$\mathcal{N}^*_{(C,\eta),\mathcal{S}_i/\widetilde{\mathcal{R}}_{2,7}} = \ker(T^*_{(C,\eta)}\widetilde{\mathcal{R}}_{2,7} \to T^*_{(C,\eta)}\mathcal{S}_i)$$

for $i = 1, 2$. As in [Donagi and Smith 1981], by taking level structures on the moduli spaces we can assume we are working on fine moduli spaces, which allows us to identify the tangent space to $\mathcal{S}_1$ at the G-admissible cover $[\widetilde{C} \to C]$, where $C = E/(p \sim q)$, with the tangent space to $\overline{\mathcal{M}}_2$ at C. Thus the conormal bundle $\mathcal{N}^*_{(C,\eta),\mathcal{S}_1/\widetilde{\mathcal{R}}_{2,7}}$ can be identified with the conormal bundle $\mathcal{N}^*_{C,\Delta_0/\overline{\mathcal{M}}_2} \subset H^0(\Omega_C \otimes \omega_C)$, where Δ_0 is the divisor of irreducible nodal curves in $\overline{\mathcal{M}}_2$. Similarly, we can identify the tangent space to $\mathcal{S}_2$ at $[\widetilde{C} \to C]$, where $C = E_1 \cup_p E_2$, with the tangent space to $\overline{\mathcal{M}}_2$ at C. Thus $\mathcal{N}^*_{(C,\eta),\mathcal{S}_2/\widetilde{\mathcal{R}}_{2,7}}$ can be identified with $\mathcal{N}^*_{C,\Delta_1/\overline{\mathcal{M}}_2} \subset H^0(\Omega_C \otimes \omega_C)$, where Δ_1 is the divisor of reducible nodal curves in $\overline{\mathcal{M}}_2$.

Using the fact that $X = \mathrm{Pr}(\widetilde{C}, C)$ we can identify $(T_X\mathcal{A}_D)^* \simeq \bigoplus_{i=1}^{6} \Omega_i\mathbb{C}$ and $S^2\big(H^0(\widetilde{C}, \omega_C)^-\big)^G$ as in (8-2). Then for a covering (C, η) the conormal bundles fit in the commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{N}^*_{X,\mathcal{E}/\mathcal{B}_D} \arrow[r] \arrow[d, "n^*"] & S^2\big(H^0(\widetilde{C}, \omega_C)^-\big)^G \arrow[r] \arrow[d, "\mathcal{P}^*"] & T^*_X\mathcal{E} \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & \mathcal{N}^*_{(C,\eta),\mathcal{S}_i/\widetilde{\mathcal{R}}_{2,7}} \arrow[r] & H^0(C, \Omega_C \otimes \omega_C) \arrow[r] \arrow[d, "H^0(j)"] & T^*_{(C,\eta)}\mathcal{S}_i \arrow[r] & 0 \\ & & H^0(C, \omega_C^2) & & \end{tikzcd} \tag{8-7}$$

in which n^* is the conormal map, $i = 1, 2$ and j is induced by the canonical map $\Omega_C \to \omega_C$ (see [Donagi and Smith 1981, IV, 2.3.3]).

Lemma 8.4. *The kernel of $H^0(j)$ is one-dimensional.*

Proof. As a map of sheaves, the canonical map $j : \Omega_C \otimes \omega_C \to \omega_C^2$ has one-dimensional kernel, namely, the one-dimensional torsion sheaf with support the node of C (see [Donagi and Smith 1981, IV, 2.3.3]). On the other hand, the map $H^0(j)$ is the composition of the pullback to the normalization with the push forward to C. This implies that the kernel of $H^0(j)$ consists exactly of the sections of the skyscraper sheaf supported at the node, and hence is one-dimensional. □

Proposition 8.5. *For coverings of type* (i)–(iv) *the restricted codifferential map* $n^* : \mathcal{N}^*_{X,\mathcal{E}/\mathcal{B}_D} \to \mathcal{N}^*_{(C,\eta),\mathcal{S}_i/\widetilde{\mathcal{R}}_{2,7}}$ *is surjective, for $i = 1, 2$.*

Proof. First notice that from the "local-global" exact sequence (see [Bardelli 1989]) we have

$$\operatorname{Ker} H^0(j) = \mathcal{N}^*_{C,\Delta_0/\overline{\mathcal{M}}_2} \subset H^0(\Omega_C \otimes \omega_C).$$

Therefore, $\operatorname{Ker} H^0(j) \subset \operatorname{Im} \mathcal{P}^*$. Since $\dim \operatorname{Ker} H^0(j) = 1$ and, by Proposition 8.3, $\dim \operatorname{Ker}(H^0(j) \circ \mathcal{P}^*) = 2$, we have $\dim \operatorname{Ker}(\mathcal{P}^*) = 1$. By diagram (8-7) this implies that the kernel of n^* is of dimension ≤ 1. Since $\mathcal{N}^*_{X,\mathcal{E}/\mathcal{B}_D}$ is a vector space of dimension 2 and $\mathcal{N}^*_{(C,\eta),\mathcal{S}_i/\widetilde{\mathcal{R}}_{2,7}}$ a vector space of dimension 1, it follows that n^* is surjective. □

9. Local degree of $\mathrm{Pr}_{2,7}$ over the boundary divisors

First we compute the local degree of the Prym map $\widetilde{\mathrm{Pr}}_{2,7}$ along the divisor $\mathcal{S}_1$. Since the conormal map of $\mathrm{Pr}_{2,7}$ along $\mathcal{S}_1$ is surjective according to Proposition 8.5, [Donagi and Smith 1981, I, Lemma 3.2] implies that the local degree along $\mathcal{S}_1$ is given by the degree of the induced map $\widetilde{\mathcal{P}} : \widetilde{\mathcal{S}}_1 \to \widetilde{\mathcal{E}}$ on the exceptional divisor $\widetilde{\mathcal{S}}_1$. Now the polarized abelian variety $X(E)$ is uniquely determined by the elliptic curve E, according to its definition. Hence the curve $\mathcal{E}$ can be identified with the moduli space of elliptic curves, i.e., with the affine line. The exceptional divisor $\widetilde{\mathcal{E}}$ is then a $\mathbb{P}^1$-bundle over $\mathcal{E}$. On the other hand, $\mathcal{S}_1$ is a divisor in $\widetilde{\mathcal{R}}_{2,7}$, so $\widetilde{\mathcal{S}}_1$ is isomorphic to $\mathcal{S}_1$. Clearly $\widetilde{\mathcal{P}}$ maps the fibers $\widetilde{\mathrm{Pr}}^{-1}(X(E)) \cap \widetilde{\mathcal{S}}_1$ onto the fibers $\mathbb{P}^1$ over the elliptic curves E.

Now $\widetilde{\mathrm{Pr}}^{-1}(X(E)) \cap \widetilde{\mathcal{S}}_1$ consists of coverings of type (i) and one covering of type (iv), which we denote by $\mathcal{C}_E^{(\mathrm{iv})}$. The coverings of type (i) have as base a nodal curve of the form $C = E/p \sim q$ and we can assume that $p = 0$, thus $\widetilde{\mathrm{Pr}}^{-1}(X(E)) \cap \widetilde{\mathcal{S}}_1$ is parametrized by E itself (the point $q = 0$ corresponds to the covering of type (iv)). Hence the induced conormal map on the exceptional divisors $\widetilde{\mathcal{P}} : \widetilde{\mathcal{S}}_1 \to \widetilde{\mathcal{E}}$ restricted to the fiber over $X(E)$ is a map $\phi : E \to \mathbb{P}^1$. Combining everything, we conclude that the local degree of the Prym map along $\mathcal{S}_1$ coincides with the degree of the induced map $\phi : E \to \mathbb{P}^1$.

Proposition 9.1. *The local degree of the Prym map* $\widetilde{\mathrm{Pr}}_{2,7}$ *along* $\mathcal{S}_1$ *is* 2.

Proof. According to what we have written above, it is sufficient to show that the map $\phi : E \to \mathbb{P}^1$ induced by $\widetilde{\mathcal{P}}$ is a double covering. We use again the identification (8-2) (and its analogue for coverings of type (iii)). As in [Donagi and Smith 1981], let x, y be local coordinates at 0 and q, and let dx, dy be the corresponding differentials. If $(a, b, c) \in S^2\big(H^0(\widetilde{C}, \omega_{\widetilde{C}})^-\big)$ are coordinates in the basis of Lemma 8.1, then $\mathbb{P}(\operatorname{Ker}(H^0(j) \circ \mathcal{P}^*)) \simeq \mathbb{P}^1$ has coordinates $[a, b]$ and its dual is identified with $\mathbb{P}(\operatorname{Im} \widetilde{\mathcal{P}}_{|\text{fiber}})$. In order to describe the kernel of $\mathcal{P}^*$, we look at the multiplication on the stalk over the node $p = (q \sim 0)$. Around p the line bundles η^i are trivial; therefore, the element $(a, b, c) \in \bigoplus_{i=1}^3 (\omega_{C,p} \otimes \omega_{C,p})$ (in coordinates $a, b, c \in \mathcal{O}_p$ for a fixed basis of $\omega_{C,p} \otimes \omega_{C,p}$) is sent to $a+b+c \in (\Omega_C \otimes \omega_C)_p$ under $\widetilde{\mathcal{P}}$. Thus the germ $a+b+c \in \mathcal{O}_p$ is zero if it is in the kernel of $\mathcal{P}^*$. In particular, the coefficient of $dx\,dy$ must vanish. Set $\alpha = a_0\,dx$, $\beta = b_0\,dx$, $\gamma = c_0\,dx$, and let $\alpha = a_q\,dy$, $\beta = b_q\,dy$, $\gamma = c_q\,dy$ be the local description of the differentials. Then the coefficient of $dx\,dy$ must satisfy

$$a_0 a_q + b_0 b_q + c_0 c_q = 0. \tag{9-1}$$

Now, by looking at the dual picture, we consider $\mathbb{P}^1 = \mathbb{P}(\operatorname{Ker} \mathcal{P}^*)^* \subset \mathbb{P}^{2*}$. Let E be embedded in $\mathbb{P}^{2*}$ by the linear system $|3\cdot 0|$. The coordinate functions $[a, b, c] \in \mathbb{P}^{2*}$ satisfy condition (9-1) for all $q \in E$. Then the points on the fiber over $[a_q, b_q, c_q]$ are points in E over the line passing through the origin $0 \in E \subset \mathbb{P}^{2*}$ and q. Hence the map $E \to \mathbb{P}^1$ corresponds to the restriction to E of the projection $\mathbb{P}^{2*} \to \mathbb{P}^1$ from the origin, which is the double covering $E \to \mathbb{P}^1$ determined by the divisor $0+q$ of E and thus of degree two. □

We now turn our attention to the Prym map on $\mathcal{S}_2$. By the surjectivity of the conormal map of $\mathrm{Pr}_{2,7}$ on $\mathcal{S}_2$ (Proposition 8.5), the local degree along $\mathcal{S}_2$ is computed by the degree of the map $\widetilde{\mathcal{P}} : \widetilde{S}_2 \to \widetilde{\mathcal{E}}$ on the divisor $\widetilde{\mathcal{S}}_2$, which is a $\mathbb{P}^1$-bundle over $\mathcal{E}$. Given an elliptic curve E, the fiber of $\widetilde{\mathrm{Pr}}^{-1}(X(E))$ intersected with the divisor $\mathcal{S}_2$ consists of coverings of type (ii), one covering of type (iii), denoted by $\mathcal{C}_E^{(\mathrm{iii})}$, and one covering of type (iv), $\mathcal{C}_E^{(\mathrm{iv})}$, which lies in the intersection with the divisor $\mathcal{S}_1$.

Recall that the type (ii) coverings have base curve $C = E_1 \cup E$ intersecting at one point, which we can assume to be 0, and with E_1 an arbitrary elliptic curve. The covering over C is the union of a degree-7 étale cyclic covering F_1 over E_1 and 7 elliptic curves E_i attached to F_1 mapping each one of them isomorphically to E. So the type (ii) coverings on the fiber over E are parametrized by pairs $(E_1, \langle\eta\rangle)$, where E_1 is an elliptic curve and $\langle\eta\rangle \subset E_1$ is a subgroup of order 7.

It is known that the parametrization space of the pairs $(E_1, \langle\eta\rangle)$ is the modular curve $Y_0(7) := \Gamma_0(7) \setminus \mathbb{H}$. The natural projection $(E_1, \eta) \mapsto E_1$ defines a map $\pi_0 : Y_0(7) \to \mathbb{C}$. Moreover, the curve $Y_0(7)$ admits a compactification $X_0(7) := \overline{Y_0(7)}$ such that the map π_0 extends to a map $\pi : X_0(7) \to \mathbb{P}^1$ (see [Silverman 1986]). The genus of $X_0(7)$ can be computed by the Hurwitz formula using the fact that π is of degree 8, and it is ramified over the points corresponding to elliptic curves with j-invariant 0 and 12^3 (with ramification degree 4 on each fiber) and over ∞, where the inverse image consists of two cusps, one étale and the other of ramification index 7. The two cusps over ∞ represent the coverings $\mathcal{C}_E^{(\mathrm{iii})}$ and $\mathcal{C}_E^{(\mathrm{iv})}$ above $X(E)$ (see Remark 9.4). This gives that $X_0(7)$ is of genus zero. Thus, we can identify $\widetilde{\mathcal{S}}_2 \cap \widetilde{\mathrm{Pr}}^{-1}(X(E))$ with $X_0(7) \simeq \mathbb{P}^1$.

Then, since $\widetilde{\mathcal{S}}_2 \simeq \mathcal{S}_2$, the restriction of the conormal map $\widetilde{\mathcal{P}}$ to a fiber over the point $[E] \in \mathcal{E}$ is a map $\psi : \mathbb{P}^1 \to \mathbb{P}^1$.

Proposition 9.2. *The map π coincides with the map $\psi : \mathbb{P}^1 \to \mathbb{P}^1$ of the fibers of $\widetilde{\mathcal{S}}_2 \to \widetilde{\mathcal{E}}$ over a point $[E] \in \mathcal{E}$.*

Proof. Let o and o' be the zero elements of E and E_1, respectively, with local coordinates x and y, respectively. Set $\alpha = a_o dx$, $\beta = b_o dx$, $\gamma = c_o dx$, and let $\alpha = a_{o'} dy$, $\beta = b_{o'} dy$, $\gamma = c_{o'} dy$ be the local description of elements of $(\omega_{C,p} \otimes \omega_{C,p})$ around the node $p = (o \sim o')$. As in the proof of Proposition 9.1, for an element (a, b, c) in the kernel of $\mathcal{P}^*$, the coefficient of $dx\,dy$ must vanish, i.e., it satisfies

$$a_o a_{o'} + b_o b_{o'} + c_o c_{o'} = 0 \tag{9-2}$$

in a neighborhood of the node. Considering the dual map, one sees that the fiber of $\widetilde{\mathcal{P}}$ over a point $[a_o, b_o, c_o] \in \mathbb{P}(\mathrm{Ker}\,\mathcal{P}^*)^* \subset \mathbb{P}^2$ with $c_o = -a_o - b_o$ corresponds to the pairs $(E_1, \langle\eta\rangle)$ such that the local functions $a, b, c \in \mathcal{O}_p$ take the values $[a_{o'}, b_{o'}, c_{o'}]$ around the $o' \in E_1$ with $c_{o'} = -a_{o'} - b_{o'}$ and such that they verify (9-2). This determines completely the triple $[a_{o'}, b_{o'}, c_{o'}]$, which depends only on the values at the node $o' \in E_1$ of the base curve. Note that a, b, c are elements of the local ring $\mathcal{O}_p$, which determines the curve E_1 uniquely. In fact, its quotient ring is the direct product of the function fields of E_1 and E, which in turn determines the curves. Therefore, the map ψ can be identified with the projection $(E_1, \langle\eta\rangle) \mapsto E_1$. □

As an immediate consequence we have:

Corollary 9.3. *The local degree of the Prym map $\widetilde{\mathrm{Pr}}_{2,7}$ along the divisor $\mathcal{S}_2$ is* 8.

Using Proposition 9.1 and Corollary 9.3 we conclude that the degree of the Prym map $\widetilde{\mathrm{Pr}}_{2,7}$ is 10, which finishes the proof of Theorem 1.1

Remark 9.4. The moduli interpretation of $X_0(7) \setminus Y_0(7)$ is given by the *Néron polygons*: one of the cusps represents a 1-*gon*, that is, a nodal cubic curve, corresponding to the covering $\mathcal{C}_E^{(\mathrm{iii})}$, and the other represents a 7-*gon*, that is, 7 copies of $\mathbb{P}^1$ with the point 0 of one attached to the point ∞ of the other in a closed chain, which corresponds to the covering $\mathcal{C}_E^{(\mathrm{iv})}$ (see [Silverman 1994, IV, §8]).

Acknowledgements

We thank E. Esteves for his useful suggestions for the proof of Theorem 6.2 and the referee for pointing out an error and suggesting some improvements in the presentation. Ortega is thankful to G. Farkas for stimulating discussions.

References

[Abramovich et al. 2003] D. Abramovich, A. Corti, and A. Vistoli, "Twisted bundles and admissible covers", *Comm. Algebra* **31**:8 (2003), 3547–3618. MR Zbl

[Arbarello et al. 2011] E. Arbarello, M. Cornalba, and P. A. Griffiths, *Geometry of algebraic curves*, vol. II, Grundlehren der Mathematischen Wissenschaften **268**, Springer, Heidelberg, 2011. MR Zbl

[Bardelli 1989] F. Bardelli, "Lectures on stable curves", pp. 648–704 in *Lectures on Riemann surfaces* (Trieste, 1987), edited by M. Cornalba et al., World Scientific, Teaneck, NJ, 1989. MR Zbl

[Beauville 1977] A. Beauville, "Prym varieties and the Schottky problem", *Invent. Math.* **41**:2 (1977), 149–196. MR Zbl

[Birkenhake and Lange 2004] C. Birkenhake and H. Lange, *Complex abelian varieties*, 2nd ed., Grundlehren der Mathematischen Wissenschaften **302**, Springer, Berlin, 2004. MR Zbl

[Donagi and Smith 1981] R. Donagi and R. C. Smith, "The structure of the Prym map", *Acta Math.* **146**:1 (1981), 25–102. MR Zbl

[Faber 1988] C. Faber, "Prym varieties of triple cyclic covers", *Math. Z.* **199**:1 (1988), 61–79. MR Zbl

[Lange and Ortega 2011] H. Lange and A. Ortega, "Prym varieties of cyclic coverings", *Geom. Dedicata* **150** (2011), 391–403. MR Zbl

[Lange and Ortega $\geq$ 2016] H. Lange and A. Ortega, "On the Prym map of cyclic coverings", in preparation.

[Mumford 1971] D. Mumford, "Theta characteristics of an algebraic curve", *Ann. Sci. École Norm. Sup.* (4) **4**:2 (1971), 181–192. MR Zbl

[Ortega 2003] A. Ortega, "Variétés de Prym associées aux revêtements n-cycliques d'une courbe hyperelliptique", *Math. Z.* **245**:1 (2003), 97–103. MR Zbl

[Shimura 1963] G. Shimura, "On analytic families of polarized abelian varieties and automorphic functions", *Ann. of Math.* (2) **78** (1963), 149–192. MR Zbl

[Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, NY, 1986. MR Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, New York, NY, 1994. MR Zbl

[Welters 1983] G. E. Welters, "Polarized abelian varieties and the heat equations", *Compositio Math.* **49**:2 (1983), 173–194. MR Zbl

Communicated by Ravi Vakil
Received 2015-02-06 Revised 2016-03-11 Accepted 2016-04-26

lange@math.fau.de — *Department Mathematik, Universität Erlangen, Cauerstr. 11, D-91058 Erlangen, Germany*

ortega@math.hu-berlin.de — *Institut für Mathematik, Humboldt Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.803

Local bounds for L^p norms of Maass forms in the level aspect

Simon Marshall

We apply techniques from harmonic analysis to study the L^p norms of Maass forms of varying level on a quaternion division algebra. Our first result gives a candidate for the local bound for the sup norm in terms of the level, which is new when the level is not squarefree. The second result is a bound for L^p norms in the level aspect that is analogous to Sogge's theorem on L^p norms of Laplace eigenfunctions.

1. Introduction

Let ϕ be a cuspidal newform of level $\Gamma_0(N)$ on $GL_2/\mathbb{Q}$ or a quaternion division algebra over $\mathbb{Q}$, which we shall assume is L^2-normalised with respect to the measure that gives $\Gamma_0(N)\backslash\mathbb{H}^2$ mass 1. There has recently been interest in bounding the sup norm $\|\phi\|_\infty$ in terms of N and the infinite component of ϕ, see [Blomer and Holowinsky 2010; Harcos and Templier 2013; 2012; Saha 2015b; Templier 2010; 2014; 2015]. The "trivial" bound in the level aspect (with the infinite component remaining bounded) is generally considered to be $\|\phi\|_\infty \ll_\epsilon N^{1/2+\epsilon}$, provided N is squarefree; see [Abbes and Ullmo 1995] or any of the previously cited papers. Our first result is a candidate for the generalisation of this to arbitrary N.

Theorem 1. *Let $D/\mathbb{Q}$ be a quaternion division algebra that is split at infinity. Let ϕ be an L^2-normalised newform of level $K_0(N)$ on $\mathrm{PGL}_1(D)$, where N is odd and coprime to the primes that ramify in D. Assume that ϕ is spherical at infinity with spectral parameter t, which is related to the Laplace eigenvalue by the equation $(\Delta + 1/4 + t^2)\phi = 0$. Let $N_0 \geq 1$ be the smallest number with $N | N_0^2$. We have*

$$\|\phi\|_\infty \ll (1+|t|)^{1/2} N_0^{1/2} \prod_{p|N} (1+1/p)^{1/2}. \tag{1}$$

Supported by NSF grants DMS-1509331 and DMS-1501230.
MSC2010: primary 35P20; secondary 11F41, 11F25.
Keywords: L^p norms, sup norms, Maass forms, automorphic forms.

Notation is standard, and specified below. When t is bounded, Theorem 1 gives a bound of $N^{1/2+\epsilon}$ for N squarefree, but roughly $N^{1/4+\epsilon}$ for powerful N. While the theorem is restricted to compact quotients, Section 3.1 gives a weaker result in the case of $\mathrm{PGL}_2/\mathbb{Q}$. This has been strengthened by Saha [2015a], who proves a bound on $\mathrm{PGL}_2/\mathbb{Q}$ that combines (1) in the level aspect with the $t^{5/12+\epsilon}$ bound of Iwaniec and Sarnak [1995] in the eigenvalue aspect. He does this by combining the methods of Iwaniec and Sarnak with bounds for Whittaker functions and an amplification argument.

Our second result is the analogue in the level aspect of a classical theorem of Sogge [1988], which we now recall. Let M be a compact Riemannian surface with Laplacian Δ, and let ψ be a function on M satisfying $(\Delta+\lambda^2)\psi=0$ and $\|\psi\|_2=1$. Define $\delta:[2,\infty]\to\mathbb{R}$ by

$$\delta(p)=\begin{cases}\frac{1}{2}-\frac{2}{p}, & 0\le\frac{1}{p}\le\frac{1}{6},\\ \frac{1}{4}-\frac{1}{2p}, & \frac{1}{6}\le\frac{1}{p}\le\frac{1}{2}.\end{cases}\tag{2}$$

Sogge's theorem states that

$$\|\psi\|_p\ll\lambda^{\delta(p)}\quad\text{for}\quad 2\le p\le\infty.\tag{3}$$

In particular, this is stronger than the bound obtained by interpolating between bounds for the L^2 and L^∞ norms. Our next theorem demonstrates that something similar is possible in the level aspect.

Theorem 2. *Let $D/\mathbb{Q}$ be a quaternion division algebra that is split at infinity. Let ϕ be an L^2-normalised newform of level $K_0(q^2)$ on $\mathrm{PGL}_1(D)$, where q is an odd prime that does not ramify in D. Assume that ϕ is principal series at q, that ϕ is spherical at infinity with spectral parameter t, and that $|t|\le A$ for some $A>0$. We have*

$$\|\phi\|_p\ll_A q^{\delta(p)}.$$

It should be possible to give some extension of Theorem 2 to general ϕ, although in some cases the method may not give any improvement over the bound given by interpolating between L^2 and L^∞ norms. In particular, this seems to occur when ϕ is special at q. We have chosen to work in the simplest case where the method gives a nontrivial result.

Theorems 1 and 2 are an attempt to prove the correct local bounds for L^p norms of eigenfunctions in the level aspect, in the same way that (3) is the local bound in the eigenvalue aspect. The term "local bound" means the best bound that may be proved by only considering the behaviour of ϕ in one small open set at a time, without taking the global structure of the space into account.

The bound (3) is sharp on the round sphere. In the same way, one may obtain limited evidence that Theorem 1, and Theorem 2 for $p\ge 6$, are the sharp local bounds by comparing them with what may be proved on the "compact form" of the arithmetic quotient being considered. In the case of Theorem 2, this means taking an L^2-normalised function ψ on $\mathrm{PGL}(2, \mathbb{Z}_q)$ of the same type as ϕ' defined below — in other words, invariant under the group $K(q, q)$ defined in Section 2.1, and generating an irreducible representation of the same type as ϕ' under right translation — and proving bounds for $\|\psi\|_p$. We may prove that $\|\psi\|_p \ll q^{\delta(p)}$ in the same way as Theorem 2, after which Equation (4) and Lemma 11 imply that this is sharp for $p \geq 6$. An analogous statement may be proved for Theorem 1 when N is a growing power of a fixed prime. However, we do not yet know if Theorem 2 is sharp in this sense for $2 \leq p \leq 6$. We expect the bound of Theorem 1 to have a natural expression as the square root of the Plancherel density around the representation of ϕ.

Because the proofs do not make use of the global structure of the arithmetic quotient, it should be possible to improve the exponents by using arithmetic amplification.

2. Notation

2.1. ***Adelic groups.*** Let $\mathbb{A}$ and $\mathbb{A}_f$ be the adeles and finite adeles of $\mathbb{Q}$. Let $D/\mathbb{Q}$ be a quaternion division algebra that is split at infinity. Let S be the set containing 2 and all primes that ramify in D, and let $S_\infty = S \cup \{\infty\}$. Let $G = \mathrm{PGL}_1(D)$. If v is a place of $\mathbb{Q}$, let $G_v = G(\mathbb{Q}_v)$. Let $X = G(\mathbb{Q}) \backslash G(\mathbb{A})$. Let $\mathcal{O} \subset D$ be a maximal order. Let $K = \bigotimes_p K_p \subset G(\mathbb{A}_f)$ be a compact subgroup with the properties that K_p is open in G_p for $p \in S$, and K_p is isomorphic to the image of $\mathcal{O}_p^\times$ in G_p when $p \notin S$. This allows us to choose isomorphisms $K_p \simeq \mathrm{PGL}(2, \mathbb{Z}_p)$ when $p \notin S$. When $M, N \geq 1$ are prime to S, we shall use these isomorphisms to define the upper triangular congruence subgroup $K_0(N)$, principal congruence subgroup $K(N)$, and

$$K(M, N) = \left\{ k \in K \, : \, k \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} (M), \; k \equiv \begin{pmatrix} * & 0 \\ * & * \end{pmatrix} (N) \right\}$$

in the natural way. We choose a maximal compact subgroup $K_\infty \subset G_\infty$.

We fix a Haar measure on $G(\mathbb{A})$ by taking the product of the measures on G_p assigning mass 1 to K_p, and any Haar measure on G_∞. We use this measure to define convolution of functions on $G(\mathbb{A})$, which we denote by $*$, and if $f \in C_0^\infty(G(\mathbb{A}))$ we use it to define the operator $R(f)$ by which f acts on $L^2(X)$. If H is a group and f is a function on H, we define the function $f^\vee$ by $f^\vee(h) = \overline{f}(h^{-1})$. If $f \in C_0^\infty(G(\mathbb{A}))$, the operators $R(f)$ and $R(f^\vee)$ are adjoints.

2.2. ***Newforms.*** Let $N \geq 1$ be prime to S. We shall say that $\phi \in L^2(X)$ is a newform of level $K_0(N)$ if ϕ lies in an automorphic representation $\pi = \bigotimes_v \pi_v$ of G, ϕ is invariant under $K_0(N)$, and we have a factorisation $\phi = \bigotimes_v \phi_v$ where ϕ_v is a local newvector of level N for $v \notin S_\infty$. We shall say that ϕ is spherical with spectral parameter $t \in \mathbb{C}$ if π_∞ satisfies these conditions, and ϕ is invariant under K_∞. Note that our normalisation of t is such that the tempered principal series corresponds to $t \in \mathbb{R}$.

2.3. ***The Harish-Chandra transform.*** Given $k \in C_0^\infty(G_\infty)$, we define its Harish-Chandra transform by

$$\hat{k}(t) = \int_{G_\infty} k(g)\varphi_t(g)\,dg$$

for $t \in \mathbb{C}$, where φ_t is the standard spherical function with spectral parameter t. We will use the following standard result on the existence of a K_∞-biinvariant function with concentrated spectral support.

Lemma 3. *There is a compact set $B \subset G_\infty$ such that for any $t \in \mathbb{R} \cup [0, i/2]$, there is a K_∞-biinvariant function $k \in C_0^\infty(G_\infty)$ with the following properties*:

(a) *The function k is supported in B, and $\|k\|_\infty \ll 1 + |t|$.*

(b) *The Harish-Chandra transform $\hat{k}$ is nonnegative on $\mathbb{R} \cup [0, i/2]$, and satisfies $\hat{k}(t) \geq 1$.*

Proof. When $t \in \mathbb{R}$ and $|t| \geq 1$, this is, e.g., Lemma 2.1 of [Templier 2015]. When $|t| \leq 1$, one may fix a K_∞-biinvariant real bump function k_0 supported near the identity and define $k = k_0 * k_0$. □

Note that condition (b) implies that $k = k^\vee$.

2.4. ***Inner products of matrix coefficients.*** The following lemma is known as Schur orthogonality, see Theorem 2.4 and Proposition 2.11 of [Bump 2013] for the proof.

Lemma 4. *Let H be a finite group, and let dh be the Haar measure of mass 1 on H. Let (ρ, V) be an irreducible representation of H, and let $\langle \cdot, \cdot \rangle$ be a positive definite H-invariant Hermitian form on V. If $v_i \in V$ for $1 \leq i \leq 4$, we have*

$$\int_H \langle \rho(h)v_1, v_2\rangle \overline{\langle \rho(h)v_3, v_4\rangle}\,dh = \frac{\langle v_1, v_3\rangle \overline{\langle v_2, v_4\rangle}}{\dim V}. \tag{4}$$

3. Proof of Theorem 1

Choose $h \in G(\mathbb{A})$ by setting $h_v = 1$ for $v \in S_\infty$, and

$$h_v = \begin{pmatrix} N & \\ & N_0 \end{pmatrix}$$

when $v \notin S_\infty$. We define $\phi' = R(h)\phi$, so that ϕ' is invariant under $K(N_0, N/N_0)$. Let $V = \bigotimes V_v \subset \pi$ be the space generated by ϕ' under the action of K.

Lemma 5. *V is an irreducible representation of K, and*

$$\dim V \leq N_0 \prod_{p|N} (1 + 1/p).$$

Proof. It suffices to prove the analogous statement for the tensor factors V_p, and we may assume that $p \notin S$. If we could write $V_p = V^1 + V^2$, where V^i are nontrivial K_p-invariant subspaces, then the projection of ϕ'_p to each subspace would be invariant under $K_p(N_0, N/N_0)$. However, this contradicts the uniqueness of the newvector.

As V_p is irreducible and factors through $K_p/K_p(N_0)$, the lemma now follows from the results of Silberger [1970, §3.4], in particular the remarks on pages 96–97. Note that we use our assumption that $2 \in S$ at this point. □

We define $k_f \in C_0^\infty(G(\mathbb{A}_f))$ to be $\overline{\langle R(g)\phi', \phi' \rangle}$ for $g \in K$ and 0 otherwise. Choose a function $k_\infty \in C_0^\infty(G_\infty)$ as in Lemma 3, and define $k = k_\infty k_f$. It may be seen that $k = k^\vee$, which implies that $R(k)$ is self-adjoint. Lemma 5 and Equation (4) imply that $k_f = \dim V k_f * k_f$, and combining this with Lemma 3(b) gives that $R(k)$ is nonnegative. Lemmas 3 and 5 and Equation (4) imply that $R(k)\phi' = \lambda\phi'$, where $\lambda > 0$ and

$$\lambda^{-1} \leq \dim V \leq N_0 \prod_{p|N} (1 + 1/p). \tag{5}$$

Extend ϕ' to an orthonormal basis $\{\phi_i\}$ of eigenfunctions for $R(k)$ with eigenvalues $\lambda_i \geq 0$. The pretrace formula associated to k is

$$\sum_i \lambda_i |\phi_i(x)|^2 = \sum_{\gamma \in G(\mathbb{Q})} k(x^{-1}\gamma x)$$

and dropping all terms from the left hand side but ϕ' gives

$$\lambda |\phi'(x)|^2 \leq \sum_{\gamma \in G(\mathbb{Q})} k(x^{-1}\gamma x). \tag{6}$$

The compactness of X and uniformly bounded support of k implies that the number of nonzero terms on the right hand side is bounded independently of x, and combining (5) and Lemma 3(a) completes the proof.

3.1. ***A result in the noncompact case.*** If we set $G = \mathrm{PGL}_2 /\mathbb{Q}$, it may be seen that we have the following analogue of Theorem 1.

Proposition 6. *Let $\Omega \subset G(\mathbb{Q})\backslash G(\mathbb{A})$ be compact. Let ϕ be an L^2-normalised newform of level $K_0(N)$ on G, where N is odd. Assume that ϕ is spherical at infinity with spectral parameter t. Let $N_0 \geq 1$ be the smallest number with $N | N_0^2$. If ϕ' is related to ϕ as above, we have $\|\phi'|_\Omega\|_\infty \ll (1+|t|)^{1/2} N_0^{1/2} \prod_{p|N}(1+1/p)^{1/2}$.*

See [Saha 2015a] for a strengthening of this result.

4. Proof of Theorem 2

We maintain the notation ϕ', V, and k_f from Section 3. Our assumption $2 \in S$ implies that $q \geq 3$.

Lemma 7. *We have* $\dim V = q$ *or* $q+1$.

Proof. We have assumed that π_q is isomorphic to an irreducible principal series representation $\mathcal{I}(\chi, \chi^{-1})$, for some character χ of $\mathbb{Q}_q^\times$ with conductor q. By considering the compact model of $\mathcal{I}(\chi, \chi^{-1})$, and applying the fact that V factors through $K/K(q) \simeq \mathrm{PGL}(2, \mathbb{F}_q)$, we see that V must be a subrepresentation of a principal series representation of $\mathrm{PGL}(2, \mathbb{F}_q)$. (Here $\mathbb{F}_q$ denotes the field with q elements.) It follows that $\dim V$ must be either 1, q, or $q+1$. The possibility $\dim V = 1$ is ruled out because any one-dimensional representation of K that is trivial on $K(q,q)$ must be trivial, and this contradicts our assumption that ϕ is new at $K_0(q^2)$. □

Let $k_\infty^0 \in C_0^\infty(G_\infty)$ be a real-valued K_∞-biinvariant function, so that $k_\infty^0 = (k_\infty^0)^\vee$. If we choose k_∞^0 to be a nonnegative bump function with sufficiently small support, we may assume that its Harish-Chandra transform satisfies $\hat{k}_\infty^0(t) \geq 1$ for $t \in [0, A] \cup [0, i/2]$. We define $k_\infty = k_\infty^0 * k_\infty^0$. Let $k_0 = k_\infty^0 k_f$, and $k = k_\infty k_f$. Let $T_0 = R(k_0)$ and $T = R(k)$. We see that T_0 is self-adjoint, and Equation (4) implies that $T = \dim V T_0^2$. Let $W \subset K_q$ be the subgroup $\{1, w\}$, where

$$w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Let $k_{1,f} \in C_0^\infty(G(\mathbb{A}_f))$ be k_f times the characteristic function of $WK(q,q)$, and let $k_{2,f} = k_f - k_{1,f}$. Let $k_i = k_\infty k_{i,f}$, and $T_i = R(k_i)$. The proof of Theorem 2 works by combining the decomposition $T = T_1 + T_2$ with interpolation between the following bounds.

Lemma 8. *We have*

$$\|T_1 f\|_\infty \ll \|f\|_1, \quad \|T_2 f\|_\infty \ll q^{-1/2} \|f\|_1,$$

for any $f \in C^\infty(X)$.

Proof. The integral kernels of T_i are given by

$$\sum_{\gamma \in G(\mathbb{Q})} k_i(x^{-1} \gamma y).$$

The result now follows from the compactness of $G(\mathbb{Q}) \backslash G(\mathbb{A})$, the bound $\|k_1\|_\infty \ll 1$, and the bound $\|k_2\|_\infty \ll q^{-1/2}$ which follows from Lemma 11 below. □

Lemma 9. *We have*

$$\|T_1 f\|_2 \ll q^{-2}\|f\|_2, \quad \|T_2 f\|_2 \ll q^{-1}\|f\|_2,$$

for any $f \in C^\infty(X)$.

Proof. The choice of k_∞ and the identity $k_f = \dim V k_f * k_f$ imply that the $L^2 \to L^2$ norm of T is $\ll (\dim V)^{-1} \le q^{-1}$. Lemma 11 implies that $k_{1,f} = k_{1,f}^\vee$ and

$$k_{1,f} = [K : WK(q,q)] k_{1,f} * k_{1,f} = (q(q+1)/2) k_{1,f} * k_{1,f},$$

and this implies that the $L^2 \to L^2$ norm of T_1 is $\ll q^{-2}$. The bound for T_2 follows from the triangle inequality. □

Interpolating between these bounds gives the following.

Lemma 10. *We have* $\|Tf\|_p \ll q^{2\delta(p)-1}\|f\|_{p'}$ *for any* $2 \le p \le \infty$, *where* p' *is the dual exponent to* p.

Proof. Applying the Riesz–Thorin interpolation theorem [Folland 1999, Theorem 6.27] to the bounds

$$\|T_1 f\|_\infty \ll \|f\|_1, \quad \|T_1 f\|_2 \ll q^{-2}\|f\|_2$$

gives $\|T_1 f\|_p \ll q^{-4/p}\|f\|_{p'}$ for $2 \le p \le \infty$, and applying it to

$$\|T_2 f\|_\infty \ll q^{-1/2}\|f\|_1, \quad \|T_2 f\|_2 \ll q^{-1}\|f\|_2$$

gives $\|T_2 f\|_p \ll q^{-1/2-1/p}\|f\|_{p'}$ for $2 \le p \le \infty$. Applying Minkowski's inequality then gives

$$\|Tf\|_p \le \|T_1 f\|_p + \|T_2 f\|_p \ll (q^{-4/p} + q^{-1/2-1/p})\|f\|_{p'},$$

and the observation $2\delta(p) - 1 = \max(-4/p, -1/2 - 1/p)$ completes the proof. □

We now combine Lemma 10 with the usual adjoint-square argument: we have

$$\begin{aligned} \langle \dim V T_0^2 f, f\rangle = \langle Tf, f\rangle & \\ \ll q^{2\delta(p)-1}\|f\|_{p'}^2 & \\ \langle T_0 f, T_0 f\rangle \ll q^{2\delta(p)-2}\|f\|_{p'}^2 & \\ \|T_0 f\|_2 \ll q^{\delta(p)-1}\|f\|_{p'}. & \end{aligned}$$

Taking adjoints gives $\|T_0 f\|_p \ll q^{\delta(p)-1}\|f\|_2$. Applying this with $f = \phi'$ and estimating the eigenvalue of T_0 on ϕ' as in Section 3 completes the proof.

Lemma 11. *Let π_q be isomorphic to an irreducible principal series representation $\mathcal{I}(\chi, \chi^{-1})$, for some character χ of $\mathbb{Q}_q^\times$ with conductor q. When $g \in K_q$, the matrix coefficient $\langle \pi_q(g)\phi'_q, \phi'_q\rangle$ satisfies*

$$\langle \pi_q(g)\phi'_q, \phi'_q\rangle = 1, \qquad g \in K_q(q, q), \tag{7}$$

$$\langle \pi_q(g)\phi'_q, \phi'_q\rangle = \chi(-1), \quad g \in wK_q(q, q), \tag{8}$$

$$\langle \pi_q(g)\phi'_q, \phi'_q\rangle \ll q^{-1/2}, \quad g \notin WK_q(q, q), \tag{9}$$

where the implied constant is absolute.

Proof. We may reduce the problem to one for the group $\mathrm{PGL}(2, \mathbb{F}_q)$ as in Lemma 7. We let T and B be the usual diagonal and upper triangular subgroups of $\mathrm{PGL}(2, \mathbb{F}_q)$. We now think of χ as a nontrivial character of $\mathbb{F}_q^\times$, and let (ρ, H) denote the corresponding induced representation of $\mathrm{PGL}(2, \mathbb{F}_q)$. We realise H as the space of functions $f : \mathrm{PGL}(2, \mathbb{F}_q) \to \mathbb{C}$ satisfying

$$f\left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} g\right) = \chi(a/d) f(g)$$

with the invariant Hermitian form

$$\langle f_1, f_2\rangle = \sum_{g \in B\backslash \mathrm{PGL}(2,\mathbb{F}_q)} f_1(g) \overline{f}_2(g). \tag{10}$$

It may be seen that there is a unique function $f_0 \in H$ that is invariant under T, up to scaling, and we may choose it to be

$$f_0 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{cases} \chi(\det/cd)/\sqrt{q-1}, & cd \neq 0, \\ 0, & cd = 0. \end{cases} \tag{11}$$

It follows that $\|f_0\| = 1$, and so $\langle \pi_q(g)\phi'_q, \phi'_q\rangle = \langle \rho(g) f_0, f_0\rangle$. Equation (7) is immediate, and (8) follows from $\rho(w) f_0 = \chi(-1) f_0$. To prove (9), we assume that

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \notin WT.$$

We may write $\langle \rho(g) f_0, f_0\rangle$ as a sum over $\mathbb{F}_q$ as follows.

Lemma 12. *We have*

$$\langle \rho(g) f_0, f_0\rangle = (q-1)^{-1} \chi(\det(g)) \sum_n \chi^{-1}((c + an)(d + bn)) \chi(n). \tag{12}$$

Proof. We choose a set of coset representatives for $B\backslash \mathrm{PGL}(2, \mathbb{F}_q)$ consisting of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}, \quad n \in \mathbb{F}_q.$$

Applying (11) for these representatives gives

$$f_0\left(\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right)=0 \quad\text{and}\quad f_0\left(\begin{pmatrix}1 & 0\\ n & 1\end{pmatrix}\right)=\chi^{-1}(n)/\sqrt{q-1}.$$

The first coset representative therefore makes no contribution to $\langle \rho(g)f_0, f_0\rangle$. For the others, we calculate

$$\begin{aligned}[\rho(g)f_0]\left(\begin{pmatrix}1 & 0\\ n & 1\end{pmatrix}\right) &= f_0\left(\begin{pmatrix}1 & 0\\ n & 1\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right)\\ &= f_0\left(\begin{pmatrix}a & b\\ c+an & d+bn\end{pmatrix}\right)\\ &= \chi(\det(g))\chi^{-1}((c+an)(d+bn))/\sqrt{q-1}.\end{aligned}$$

Substituting these into (10) completes the proof. □

We bound the sum (12) by rewriting it as

$$\sum_n \chi^{-1}\big((c+an)(d+bn)n^{q-2}\big)$$

and applying [Schmidt 2004, Chapter 2, Theorem 2.4] (see also [Iwaniec and Kowalski 2004, Theorem 11.23]). We must first check that $(c+an)(d+bn)n^{q-2}$ is not a proper power. The assumption $g\notin WT$ implies that one or both of $a+cn$ and $b+dn$ have a root distinct from 0. If they both have the same root distinct from 0, this contradicts the invertability of g. Therefore $(c+an)(d+bn)n^{q-2}$ must have at least one root of multiplicity 1, so it cannot be a power. As $(c+an)(d+bn)n^{q-2}$ has at most 3 distinct roots, [Schmidt 2004] or [Iwaniec and Kowalski 2004] therefore give

$$\left|\sum_n \chi\big((c+an)(d+bn)n^{q-2}\big)\right| \le 2\sqrt{q},$$

which completes the proof of (9). □

Acknowledgements

We would like to thank Valentin Blomer, Abhishek Saha, and Nicolas Templier for helpful discussions, and the referee for a careful reading of the paper.

References

[Abbes and Ullmo 1995] A. Abbes and E. Ullmo, "Comparaison des métriques d'Arakelov et de Poincaré sur $X_0(N)$", *Duke Math. J.* **80**:2 (1995), 295–307. MR 1369394 Zbl 0895.14007

[Blomer and Holowinsky 2010] V. Blomer and R. Holowinsky, "Bounding sup-norms of cusp forms of large level", *Invent. Math.* **179**:3 (2010), 645–681. MR 2587342 Zbl 1243.11059

[Bump 2013] D. Bump, *Lie groups*, 2nd ed., Graduate Texts in Mathematics **225**, Springer, New York, 2013. MR 3136522 Zbl 1279.22001

[Folland 1999] G. B. Folland, *Real analysis*, 2nd ed., Wiley, New York, 1999. MR 1681462 Zbl 0924.28001

[Harcos and Templier 2012] G. Harcos and N. Templier, "On the sup-norm of Maass cusp forms of large level: II", *Int. Math. Res. Not.* **2012**:20 (2012), 4764–4774. MR 2989618 Zbl 1332.11048

[Harcos and Templier 2013] G. Harcos and N. Templier, "On the sup-norm of Maass cusp forms of large level, III", *Math. Ann.* **356**:1 (2013), 209–216. MR 3038127 Zbl 1332.11049

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, American Mathematical Society, Providence, RI, 2004. MR 2061214 Zbl 1059.11001

[Iwaniec and Sarnak 1995] H. Iwaniec and P. Sarnak, "L^∞ norms of eigenfunctions of arithmetic surfaces", *Ann. of Math.* (2) **141**:2 (1995), 301–320. MR 1324136 Zbl 0833.11019

[Saha 2015a] A. Saha, "Hybrid sup-norm bounds for Maass newforms of powerful level", preprint, 2015. arXiv 1509.07489

[Saha 2015b] A. Saha, "On sup-norms of cusp forms of powerful level", preprint, 2015. arXiv 1404.3179

[Schmidt 2004] W. Schmidt, *Equations over finite fields: an elementary approach*, 2nd ed., Kendrick Press, Heber City, UT, 2004. MR 2121285 Zbl 1128.11032

[Silberger 1970] A. J. Silberger, PGL_2 *over the p-adics: its representations, spherical functions, and Fourier analysis*, Lecture Notes in Mathematics **166**, Springer, Berlin, 1970. MR 0285673 Zbl 0204.44102

[Sogge 1988] C. D. Sogge, "Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds", *J. Funct. Anal.* **77**:1 (1988), 123–138. MR 930395 Zbl 0641.46011

[Templier 2010] N. Templier, "On the sup-norm of Maass cusp forms of large level", *Selecta Math. (N.S.)* **16**:3 (2010), 501–531. MR 2734340 Zbl 1260.11032

[Templier 2014] N. Templier, "Large values of modular forms", *Camb. J. Math.* **2**:1 (2014), 91–116. MR 3272013 Zbl 1307.11062

[Templier 2015] N. Templier, "Hybrid sup-norm bounds for Hecke–Maass cusp forms", *J. Eur. Math. Soc. (JEMS)* **17**:8 (2015), 2069–2082. MR 3372076 Zbl 06481085

Communicated by Philippe Michel
Received 2015-03-09 Revised 2015-12-31 Accepted 2016-04-23

marshall@math.wisc.edu *Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53706, United States*

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.813

Hasse principle for Kummer varieties

Yonatan Harpaz and Alexei N. Skorobogatov

The existence of rational points on the Kummer variety associated to a 2-covering of an abelian variety A over a number field can sometimes be established through the variation of the 2-Selmer group of quadratic twists of A. In the case when the Galois action on the 2-torsion of A has a large image, we prove, under mild additional hypotheses and assuming the finiteness of relevant Shafarevich–Tate groups, that the Hasse principle holds for the associated Kummer varieties. This provides further evidence for the conjecture that the Brauer–Manin obstruction controls rational points on K3 surfaces.

1. Introduction

The principal aim of this paper is to give some evidence in favour of the conjecture that the Brauer–Manin obstruction is the only obstruction to the Hasse principle for rational points on K3 surfaces over number fields; see [Skorobogatov 2009, p. 1859] and [Skorobogatov and Zarhin 2008, p. 484]. Conditionally on the finiteness of relevant Shafarevich–Tate groups we establish the Hasse principle for certain families of Kummer surfaces. These surfaces are quotients of 2-coverings of an abelian surface A by the antipodal involution, where

(a) A is the product of elliptic curves $A = E_1 \times E_2$, or

(b) A is the Jacobian of a curve C of genus 2 with a rational Weierstrass point.

Both cases are treated by the same method which allows us to prove a more general result for the Kummer varieties attached to 2-coverings of an abelian variety A over a number field k, provided certain conditions are satisfied. By a 2-covering we understand a torsor Y for A such that the class $[Y] \in \mathrm{H}^1(k, A)$ has order at most 2. Thus Y is the twist of A by a 1-cocycle with coefficients in $A[2]$ acting on A by translations. The antipodal involution $\iota_A = [-1] : A \to A$ induces an involution $\iota_Y : Y \to Y$ and we define the Kummer variety $X = \mathrm{Kum}(Y)$ as the minimal desingularisation of Y/ι_Y; see Section 6 for details.

In this introduction we explain the results pertaining to cases (a) and (b) above and postpone the statement of a more general theorem until the next section. In case (a)

MSC2010: primary 14G05; secondary 11J95.
Keywords: Kummer varieties, Hasse principle.

we have the following result, whose proof can be found at the end of Section 2. We denote by $\Delta(f)$ the discriminant of a (not necessarily monic) polynomial $f(x)$; see (3) for the classical formula for $\Delta(f)$ in the case $\deg(f)=4$.

Theorem A. *Let $g_1(x)$ and $g_2(x)$ be irreducible polynomials of degree 4 over a number field k, each with the Galois group S_4. Let w_1 and w_2 be distinct primes of k not dividing 6 such that for all $i, j \in \{1,2\}$ the coefficients of $g_i(x)$ are integral at w_j and $\mathrm{val}_{w_j}(\Delta(g_i))=\delta_{ij}$. Let E_i be the Jacobian of the curve $y^2=g_i(x)$, where $i=1,2$. For $i=1,2$, assume the finiteness of the 2-primary torsion subgroup of the Shafarevich–Tate group for each quadratic twist of E_i whose 2-Selmer group has rank 1. If the Kummer surface with the affine equation*

$$z^2 = g_1(x)g_2(y) \tag{1}$$

is everywhere locally soluble, then it has a Zariski dense set of k-points.

We expect that the conditions of Theorem A are in a certain sense ‘generic’. To illustrate this, let $\mathbb{Z}[t]_{\deg=4} \subset \mathbb{Z}[t]$ be the set of polynomials of degree 4 ordered by the maximal height of their coefficients. By a theorem of van der Waerden, 100% of polynomials in $\mathbb{Z}[t]_{\deg=4}$ have the Galois group S_4 (see [Cohen 1979, Theorem 1] for a statement over an arbitrary number field). By [Hall 2011, Proposition A.2], for 100% of monic polynomials $g \in \mathbb{Z}[t]_{\deg=4}$ there exists a place w such that $\mathrm{val}_w(\Delta(g))=1$. Similar arguments are likely to show that 100% of pairs $g_1, g_2 \in \mathbb{Z}[t]_{\deg=4}$ satisfy the condition $\mathrm{val}_{w_j}(\Delta(g_i))=\delta_{ij}$ for some w_1, w_2. The finiteness of the Shafarevich–Tate group is a well-known conjecture, established by M. Bhargava, C. Skinner and W. Zhang [Bhargava et al. 2014, Theorem 2] for a majority of elliptic curves over $\mathbb{Q}$ ordered by naïve height. Note finally that using [Bright et al. 2016, Theorem 1.4] one can show that the Kummer surface (1) is everywhere locally soluble for a positive proportion of pairs $g_1, g_2 \in \mathbb{Z}[t]_{\deg=4}$.

To give an explicit description of our results in case (b) we need to recall the realisation of Kummer surfaces attached to the Jacobian of a genus-2 curve as smooth complete intersections of three quadrics in $\mathbb{P}^5_k$. We mostly follow [Skorobogatov 2010, Section 3]; for the classical theory over an algebraically closed field, see [Dolgachev 2012, Chapter 10].

Let $f(x)$ be a separable monic polynomial of degree 5 over a field k of characteristic different from 2. Let C be the hyperelliptic curve with the affine equation $y^2 = f(x)$ and let A be the Jacobian of C. Let L be the étale k-algebra $k[x]/(f(x))$ and let $\theta \in L$ be the image of x. The 2-torsion $\mathrm{Gal}(\bar{k}/k)$-module $A[2]$ is isomorphic to $\mathrm{R}_{L/k}(\mu_2)/\mu_2$, where $\mathrm{R}_{L/k}$ is the Weil restriction of scalars. Since $[L:k]$ is odd, $A[2]$ is a direct summand of $\mathrm{R}_{L/k}(\mu_2)$. It follows that the map $\mathrm{H}^1(k, \mathrm{R}_{L/k}(\mu_2)) = L^*/L^{*2} \to \mathrm{H}^1(k, A[2])$ is surjective and induces an isomorphism $\mathrm{H}^1(k, A[2]) = L^*/k^*L^{*2}$.

Let $\lambda \in L^*$. Let $W_\lambda \subset \mathrm{R}_{L/k}(\mathbb{G}_{m,L})$ be the closed subscheme given by $z^2 = \lambda$. It is clear that W_λ is a k-torsor for $\mathrm{R}_{L/k}(\mu_2)$ whose class in $\mathrm{H}^1(k, \mathrm{R}_{L/k}(\mu_2)) = L^*/L^{*2}$ is given by λ. Let $Z_\lambda = W_\lambda/\{\pm 1\}$ be the subscheme of $\mathrm{R}_{L/k}(\mathbb{G}_{m,L})/\{\pm 1\}$ given by the same equation. We obtain that Z_λ is the k-torsor for $A[2]$ whose class in $\mathrm{H}^1(k, A[2]) = L^*/k^*L^{*2}$ is defined by λ.

Now let $Y_\lambda = (A \times Z_\lambda)/A[2]$ be the 2-covering of A obtained by twisting A by Z_λ. Then $\mathrm{Kum}(Y_\lambda)$ is the following smooth complete intersection of three quadrics in $\mathbb{P}(\mathrm{R}_{L/k}(\mathbb{A}^1_L) \times \mathbb{A}^1_k) \simeq \mathbb{P}^5_k$:

$$\mathrm{Tr}_{L/k}\left(\lambda \frac{u^2}{f'(\theta)}\right) = \mathrm{Tr}_{L/k}\left(\lambda \frac{\theta u^2}{f'(\theta)}\right) = \mathrm{Tr}_{L/k}\left(\lambda \frac{\theta^2 u^2}{f'(\theta)}\right) - \mathrm{N}_{L/k}(\lambda) u_0^2 = 0, \quad (2)$$

where u is an L-variable, u_0 is a k-variable, and $f'(x)$ is the derivative of $f(x)$ (cf. equations (7) and (8) in [Skorobogatov 2010]). If $\lambda \in k^*L^{*2}$, then an easy change of variable reduces (2) to the same system of equations with $\lambda = 1$. As $Y_1 \cong A$ has a rational point, this case can be excluded for the purpose of establishing the Hasse principle.

Theorem B. *Let $f(x)$ be a monic irreducible polynomial of degree 5 over a number field k, and let $L = k[x]/(f(x))$. Let w be an odd prime of k such that the coefficients of $f(x)$ are integral at w and $\mathrm{val}_w(\Delta(f)) = 1$. Let A be the Jacobian of the hyperelliptic curve $y^2 = f(x)$. Assume the finiteness of the 2-primary torsion subgroup of the Shafarevich–Tate group for each quadratic twist of A whose 2-Selmer group has rank 1. Let $\lambda \in L^*$ be such that for some $r \in k^*$ the valuation of λr at each completion of L over w is even, but $\lambda \notin k^*L^{*2}$. If the Kummer surface given by (2) is everywhere locally soluble, then it has a Zariski dense set of k-points.*

Let $[\lambda] \in \mathrm{H}^1(k, A[2])$ be the class defined by λ. The conditions imposed on λ in Theorem B are equivalent to the condition that $[\lambda] \neq 0$ and $[\lambda]$ is unramified at w. Equivalently, the k-torsor Z_λ defined above has a k_w^{un}-point, where k_w^{un} is the maximal unramified extension of k_w, but no k-point.

Any Kummer surface (2) can be mapped to $\mathbb{P}^3_k$ by a birational morphism that contracts 16 disjoint rational curves onto singular points. The image of $\mathrm{Kum}(Y_\lambda)$ is a singular quartic surface $S \subset \mathbb{P}^3_k$ which is the classical Kummer surface with 16 nodes. (See [Dolgachev 2012, Section 10.3.3] and [Gonzalez-Dorrego 1994] for a modern account of the geometry of S over an algebraically closed field.) The group $A[2]$ acts on S by projective automorphisms and the singular locus S_{sing} is a k-torsor for $A[2]$. Then S is identified with the twist of A/ι_A by S_{sing}. The condition $\lambda \notin k^*L^{*2}$, which we use to prove the Zariski density of $S(k)$, is precisely the condition that the torsor S_{sing} is nontrivial, that is, no singular point of S is a k-point.

Theorem B is proved at the end of Section 2. The main idea of the proof of Theorems A and B is due to Swinnerton-Dyer. Let $\alpha \in \mathrm{H}^1(k, A[2])$ be the class of a 1-cocycle used to obtain Y from A. The group $\mu_2 = \{\pm 1\}$ acts on A by multiplication. As this action commutes with the action of $A[2]$ by translations, we have an induced action of μ_2 on Y. For an extension F/k of degree at most 2, let T_F be the torsor for μ_2 defined by F. The quadratic twists A^F and Y^F are defined as the quotients of $A \times_k T_F$ and $Y \times_k T_F$, respectively, by the diagonal action of μ_2. We identify $A^F[2] = A[2]$ and consider Y^F as a torsor for A^F defined by the same 1-cocycle with the class $\alpha \in \mathrm{H}^1(k, A^F[2]) = \mathrm{H}^1(k, A[2])$. The projection to the first factor defines a morphism $Y^F = (Y \times_k T_F)/\mu_2 \longrightarrow Y/\mu_2$. Thus in order to find a rational point on the Kummer variety $X = \mathrm{Kum}(Y)$ it is enough to find a rational point on Y^F for some F. At the first step of the proof, using a fibration argument, one produces a quadratic extension F such that Y^F is everywhere locally soluble. Equivalently, $\alpha \in \mathrm{H}^1(k, A^F[2])$ is in the 2-Selmer group of A^F. At the second step one modifies F so that the 2-Selmer group of A^F is spanned by α and the image of $A^F[2](k)$ under the Kummer map. (In the cases considered in this paper, $A^F[2](k) = A[2](k) = 0$.) This implies that $Ш(A^F)[2]$ is $\mathbb{Z}/2$ or 0. In previous applications of the method [Swinnerton-Dyer 2001; Skorobogatov and Swinnerton-Dyer 2005], as well as in Theorem A above, A is a product of two elliptic curves, in which case the Cassels–Tate pairing on $Ш(A^F)$ is alternating. The assumption that $Ш(A^F)$ is finite then implies that the order of $Ш(A^F)[2]$ is a square and hence $Ш(A^F)[2] = 0$. In particular, Y^F has a k-point, so that $Y^F \simeq A^F$. In this paper we consider more general principally polarised abelian varieties. The theory developed by Poonen and Stoll [1999] ensures that in the cases considered here the Cassels–Tate pairing on $Ш(A^F)$ defined using the principal polarisation is still alternating, so the proof can be concluded as before.

Swinnerton-Dyer's method was used in combination with Schinzel's Hypothesis (H) in [Colliot-Thélène et al. 1998; Swinnerton-Dyer 2000; Wittenberg 2007]. It is in [Swinnerton-Dyer 2001] that the method was applied without Hypothesis (H) for the first time, using Dirichlet's theorem on primes in an arithmetic progression, the only known case of (H). That work tackled diagonal cubic surfaces, which are dominated by a product of two elliptic curves with complex multiplication. The immediate precursor of our Theorem A is [Skorobogatov and Swinnerton-Dyer 2005], which treats Kummer surfaces attached to products of elliptic curves, again without assuming Hypothesis (H). Central to Swinnerton-Dyer's method is a linear algebra construction that represents the Selmer group as the kernel of a symmetric bilinear form. The difficulty of operating this machinery makes implementation of the method a rather delicate task. In the present paper this linear algebra machinery is not used. Instead we use the ideas from [Mazur and Rubin 2007] and especially from [Mazur and Rubin 2010].

Let us note that given an elliptic curve E over a number field k it is not always possible to find a quadratic extension F/k such that the 2-Selmer group of E^F is spanned by a fixed class $\alpha \in \mathrm{H}^1(k, E[2])$ and the image of $E^F[2](k)$. Firstly, the parity of the rank of the 2-Selmer group of E^F can be the same for all F: this happens precisely when k is totally imaginary and E acquires everywhere good reduction over an abelian extension of k; see [Dokchitser and Dokchitser 2011, Remark 4.9]. Secondly, over any number field k there are elliptic curves E such that, for any quadratic extension F/k, the difference between the 2-Selmer rank of E^F and the dimension of the $\mathbb{F}_2$-vector space $E[2](k)$ is at least the number of complex places of k; see [Klagsbrun 2012a; 2012b]. Such examples can occur when $E[2](k) \cong \mathbb{Z}/2$ and E has a cyclic isogeny of degree 4 defined over $k(E[2])$ but not over k.

In this paper we do not discuss the conjecture (for which see [Skorobogatov 2009, p. 1859; Skorobogatov and Zarhin 2008, p. 484]) that rational points on a K3 surface are dense in its Brauer–Manin set.[1] Nevertheless we make the following simple observation in the direction of Mazur's conjectures [1992; 1995]:

Proposition 1.1. *Let $E_1, \dots, E_n$ be elliptic curves over $\mathbb{Q}$ such that $E_i[2](\mathbb{Q}) = 0$ for $i = 1, \dots, n$. Let $X = \mathrm{Kum}\big(\prod_{i=1}^n Y_i\big)$, where Y_i is a 2-covering of E_i defined by a class in $\mathrm{H}^1(\mathbb{Q}, E_i[2])$ that restricts to a nonzero class in $\mathrm{H}^1(\mathbb{R}, E_i[2])$, for $i = 1, \dots, n$. Then the real topological closure of $X(\mathbb{Q})$ in $X(\mathbb{R})$ is a union of connected components of $X(\mathbb{R})$.*

This can be compared with the result of M. Kuwata and L. Wang [1993]. See the end of Section 7 for the proof of Proposition 1.1.

Our main technical result is Theorem 2.3. It is stated in Section 2 where we also show that Theorem 2.3 implies Theorems A and B. In Section 3 we systematically develop the Galois-theoretic aspect of the approach of Mazur and Rubin. We recall the necessary facts about the Kummer map for quadratic twists of abelian varieties over local fields in Section 4. In Section 5 we discuss the Selmer group and the Cassels–Tate pairing over a number field. A reduction to everywhere soluble 2-coverings is carried out in Section 6 using a known case of the fibration method. We finish the proof of Theorem 2.3 in Section 7.

2. Main results

Let k be a field of characteristic different from 2 with a separable closure $\bar{k}$ and the Galois group $\Gamma_k = \mathrm{Gal}(\bar{k}/k)$.

Let A be an abelian variety over k. Let $K = k(A[2]) \subset \bar{k}$ be the field of definition of $A[2]$, that is, the smallest field such that $A[2](K) = A[2](\bar{k})$. Let $G = \mathrm{Gal}(K/k)$. Consider the following conditions:

[1] A recent result of D. Holmes and R. Pannekoek [2015] shows that if this conjecture is extended to all Kummer varieties, then the ranks of quadratic twists of any given abelian variety over a given number field are not bounded.

(a) $A[2]$ is a simple G-module and $\mathrm{End}_G(A[2]) = \mathbb{F}_2$;

(b) $\mathrm{H}^1(G, A[2]) = 0$;

(c) there exists a $g \in G$ such that $A[2]/(g-1) = \mathbb{F}_2$;

(d) there exists an $h \in G$ such that $A[2]/(h-1) = 0$.

Lemma 2.1. *Let A be the Jacobian of a smooth projective curve with the affine equation $y^2 = f(x)$, where $f(x) \in k[x]$ is an irreducible separable polynomial of odd degree $m \geq 3$. If the Galois group of $f(x)$ is the symmetric group S_m on m letters, then A satisfies conditions* (a), (b), (c), (d).

Proof. It is well-known that the Γ_k-module $A[2]$ is the zero-sum submodule of the vector space $(\mathbb{F}_2)^m$ freely generated by the roots of $f(x) = 0$ with the natural permutation action of Γ_k. Since m is odd, the permutation Γ_k-module $(\mathbb{F}_2)^m$ is the direct sum of $A[2]$ and the $\mathbb{F}_2$-vector space spanned by the vector $(1, \dots, 1)$.

If an S_m-submodule of $(\mathbb{F}_2)^m$ contains a vector with at least one coordinate 0 and at least one coordinate 1, then it contains the zero-sum submodule. Hence $A[2]$ is a simple S_m-module. A direct calculation with matrices shows that the $m \times m$ matrices commuting with all permutation matrices are the linear combinations of the identity and the all-1 matrix. We deduce that $\mathrm{End}_{S_m}(A[2]) = \mathbb{F}_2$, thus (a) holds.

The permutation S_m-module $(\mathbb{F}_2)^m$ is isomorphic to $\mathbb{F}_2[S_m/S_{m-1}]$. By Shapiro's lemma we have

$$\mathrm{H}^1(S_m, \mathbb{F}_2[S_m/S_{m-1}]) = \mathrm{H}^1(S_{m-1}, \mathbb{F}_2) = \mathrm{Hom}(S_{m-1}, \mathbb{F}_2) = \mathbb{F}_2.$$

Since $\mathrm{H}^1(S_m, \mathbb{F}_2) = \mathbb{F}_2$, we obtain $\mathrm{H}^1(S_m, A[2]) = 0$, so (b) holds.

If g is a cycle of length $m-1$, then $A[2]/(g-1) = \mathbb{F}_2$, so (c) is satisfied. If h is a cycle of length m, then $A[2]/(h-1) = 0$, so (d) is satisfied. $\square$

Remark 2.2. There are other natural cases when the Galois module $A[2]$ satisfies conditions (a) through (d). Let $\dim(A) = n > 1$. We only deal with the case when the Cassels–Tate pairing on $Ш(A)$ defined by a polarisation $\lambda \in \mathrm{NS}(\bar{A})^{\Gamma_k}$ is alternating. According to the results of Poonen, Stoll and Rains recalled in Section 5, this holds when λ lifts to a symmetric element of $\mathrm{Pic}(A)$. (This happens, for example, when A is as in Lemma 2.1.) In this case the pairing $A[2] \times A[2] \to \mathbb{Z}/2$ induced by λ and the Weil pairing admits a Galois-invariant quadratic refinement $q : A[2] \to \mathbb{Z}/2$; see [Poonen and Rains 2011, Proposition 3.2(c)]. The 'generic' Galois action compatible with this assumption is when G is the corresponding orthogonal group $O(q) \subset \mathrm{GL}(A[2])$. It can be shown that conditions (a), (c) and (d) are always satisfied for $G = O(q)$. Condition (b) is satisfied for all $n \neq 2, 3$ when q is split (i.e., isomorphic to a direct sum of copies of the rank-2 hyperbolic space) and for all $n \neq 3, 4$ if q is nonsplit (see [Sah 1977, Proposition 2.1]). We do not elaborate on these statements here, as we will not use them.

Let $A_1, \dots, A_r$ be abelian varieties over k. For each $i = 1, \dots, r$, let $K_i = k(A_i[2])$ and $G_i = \mathrm{Gal}(K_i/k)$. We assume the following condition:

(e) The fields $K_1, \dots, K_r$ are linearly disjoint over k.

By definition this means that $[K_1 \cdots K_r : k] = \prod_{i=1}^r [K_i : k]$. Thus the Galois group of $K_1 \cdots K_r$ over k is $\prod_{i=1}^r G_i$.

When k is a *number field* we shall also assume the following condition:

(f) There exist distinct odd primes $w_1, \dots, w_r$ of k such that for each $i = 1, \dots, r$ the abelian variety A_i has bad reduction at w_i and the number of geometric connected components of the Néron model of A_i at w_i is odd, whereas each A_j for $j \neq i$ has good reduction at w_i.

Let k_i^{ab} be the maximal abelian subextension of $k \subset K_i$. Equivalently, $\mathrm{Gal}(k_i^{\mathrm{ab}}/k)$ is the maximal abelian quotient G_i^{ab} of G_i. Let us finally assume:

(g) For each $i = 1, \dots, r$, the field k_i^{ab} is totally ramified at w_i. Equivalently, k_i^{ab} has a unique prime ideal above w_i, and G_i^{ab} coincides with the inertia subgroup of this ideal.

Let F be a field extension of k of degree at most 2. As in the introduction, we denote by A^F the quadratic twist of A by F, that is, the abelian variety over k obtained by twisting A by the quadratic character of F/k with respect to the action of μ_2 on A by multiplication. For example, if A is an elliptic curve with the Weierstrass equation $y^2 = f(x)$, then A^F is given by $y^2 = cf(x)$, where $c \in k^*$ is such that $F = k(\sqrt{c})$.

We are now ready to state the main theorem. Recall that a class in $\mathrm{H}^1(k, A[2])$ is said to be *unramified* at an odd non-Archimedean place v of k if it goes to zero under the restriction map $\mathrm{H}^1(k, A[2]) \to \mathrm{H}^1(k_v^{\mathrm{nr}}, A[2])$, where k_v^{nr} is the maximal unramified extension of the completion k_v of k at v.

Theorem 2.3. *Let k be a number field. Let $A = \prod_{i=1}^r A_i$, where each A_i is a principally polarised abelian variety satisfying conditions* (a), (b), (c) *and* (d). *Assume in addition that conditions* (e), (f) *and* (g) *are satisfied. Assume that the* 2*-primary subgroup of the Shafarevich–Tate group* $\text{Ш}(A_i^F)\{2\}$ *is finite for all $i = 1, \dots, r$ and all extensions F of k with $[F : k] \leq 2$ for which the* 2*-Selmer group of A_i^F has rank* 1. *Consider the classes in* $\mathrm{H}^1(k, A[2])$ *that are unramified at $w_1, \dots, w_r$ and whose projection to* $\mathrm{H}^1(k, A_i[2])$ *is nonzero for each $i = 1, \dots, r$. If the Kummer variety of A defined by such a class is everywhere locally soluble, then it has a Zariski dense set of k-points.*

Remarks. (1) If $r = 1$, then condition (d) is not needed and condition (e) is vacuous.

(2) The Brauer–Manin obstruction does not appear in the conclusion of the theorem. In fact, the purely algebraic conditions (a), (b) and (e) imply that a certain part of

the Brauer group is trivial; see Proposition 6.1. The problem of calculation of the full Brauer group of a Kummer variety will be addressed in a separate paper.

(3) If the 2-primary torsion subgroup $Ш(A_i^F)\{2\}$ is finite, then condition (b) implies that the nondegenerate Cassels–Tate pairing on $Ш(A_i^F)\{2\}$ is alternating. See Proposition 5.2, which is based on the work of [Poonen and Stoll 1999] and [Poonen and Rains 2011]. In the proof of Theorem 2.3 we use a well-known consequence of this result that the number of elements of $Ш(A_i^F)[2]$ is a square.

We employ the following standard notation:

- k_{w_i} is the completion of k at w_i,
- $\mathcal{O}_{w_i}$ is the ring of integers of k_{w_i},
- $\mathfrak{m}_{w_i}$ is the maximal ideal of $\mathcal{O}_{w_i}$, and
- $\mathbb{F}_{w_i} = \mathcal{O}_{w_i}/\mathfrak{m}_{w_i}$ is the residue field.

Corollary 2.4. *Let k be a number field. For $i = 1, \dots, r$, let $f_i(x) \in k[x]$ be a monic irreducible polynomial of odd degree $n_i \geq 3$ whose Galois group is the symmetric group S_{n_i}, and let A_i be the Jacobian of the hyperelliptic curve $y^2 = f_i(x)$. Assume the existence of distinct odd primes $w_1, \dots, w_r$ of k such that $f_i(x) \in \mathcal{O}_{w_j}[x]$ and $\mathrm{val}_{w_i}(\Delta(f_j)) = \delta_{ij}$ for any $i, j \in \{1, \dots, r\}$. Assume that $Ш(A_i^F)\{2\}$ is finite for all $i = 1, \dots, r$ and all extensions F of k with $[F : k] \leq 2$ for which the 2-Selmer group of A_i^F has rank 1. Consider the classes in $\mathrm{H}^1(k, A[2])$ that are unramified at $w_1, \dots, w_r$ and whose projection to $\mathrm{H}^1(k, A_i[2])$ is nonzero for each $i = 1, \dots, r$. If the Kummer variety of A defined by such a class is everywhere locally soluble, then it has a Zariski dense set of k-points.*

Proof. Each A_i is a canonically principally polarised abelian variety which satisfies conditions (a) through (d) by Lemma 2.1.

Let C_i be the proper, smooth and geometrically integral curve over k given by the affine equation $y^2 = f_i(x)$, so that $A_i = \mathrm{Jac}(C_i)$. As in [Liu 1996, Section 4.3], a proper and flat *Weierstrass model* $\mathcal{C}_i$ over $\mathrm{Spec}(\mathcal{O}_{w_i})$ is defined as the normalisation in $C_i \times_k k_{w_i}$ of the projective line $\mathbb{P}^1_{\mathcal{O}_{w_i}}$ with the affine coordinate x. Since $2 \in \mathcal{O}^*_{w_i}$, the integral closure of $\mathcal{O}_{w_i}[x]$ in $k_{w_i}(C_i)$ is $\mathcal{O}_{w_i}[x, y]/(y^2 - f_i(x))$. The condition $\mathrm{val}_{w_i}(\Delta(f_i)) = 1$ implies that $\mathcal{C}_i$ is regular and the special fibre $\mathcal{C}_i \times_{\mathcal{O}_{w_i}} \mathbb{F}_{w_i}$ is geometrically integral with a unique singular point, which is an ordinary double point; see Corollary 6 and Remark 18 on p. 4602 of [Liu 1996]. In particular, the reduction of $f_i(x)$ modulo $\mathfrak{m}_{w_i}$ has one rational double root and $n_i - 2$ simple roots. (This can also be checked directly using Sylvester's formula for the discriminant.) Now [Bosch et al. 1990, Theorem 9.6.1] implies that the special fibre of the Néron model of $A_i \times_k k_{w_i}$ over $\mathrm{Spec}(\mathcal{O}_{w_i})$ is connected. If $j \neq i$, then $\mathrm{val}_{w_i}(\Delta(f_j)) = 0$, and this implies that A_j has good reduction at w_i. We conclude that (f) holds.

For each $i = 1, \dots, r$, the field $K_i = k(A_i[2])$ is the splitting field of $f_i(x)$. Since $\mathrm{Gal}(K_i/k) \cong S_{n_i}$, the alternating group is the unique nontrivial normal subgroup of $\mathrm{Gal}(K_i/k)$. Its invariant subfield is $k(\sqrt{\Delta(f_i)})$. Thus if k' is a Galois extension of k such that $k \subsetneq k' \subsetneq K_i$, then $k' = k(\sqrt{\Delta(f_i)}) = k_i^{\mathrm{ab}}$. The extension $k(\sqrt{\Delta(f_i)})$ of k is ramified at w_i, so (g) holds.

Let K_i' be the compositum of the fields K_j for $j \neq i$. Since each K_i is a Galois extension of k, the field $K_i \cap K_i'$ is also a Galois extension of k. To verify (e) we need to check that $K_i \cap K_i' = k$ for each $i = 1, \dots, r$. Otherwise, $K_i \cap K_i'$ contains $k(\sqrt{\Delta(f_i)})$ which is ramified at w_i. However, this contradicts the criterion of Néron–Ogg–Shafarevich according to which K_i' is unramified at the odd place w_i, where each of the abelian varieties A_j for $j \neq i$ has good reduction. Thus (e) holds. □

Proof of Theorem A assuming Theorem 2.3. For $i = 1, 2$, let C_i be the curve of genus 1 given by $y^2 = g_i(x)$. Write $g_i(x) = ax^4 + bx^3 + cx^2 + dx + e$. The classical SL(2)-invariants of the corresponding quartic binary form $G_i(u, v) = v^4 g_i(u/v)$ are

$$\begin{aligned} I &= 12ae - 3bd + c^2, \\ J &= 72ace + 9bcd - 27ad^2 - 27eb^2 - 2c^3, \\ \Delta &= (4I^3 - J^2)/27. \end{aligned} \tag{3}$$

Then the Jacobian of C_i is the elliptic curve E_i with the equation $u^2 = p_i(t)$, where $p_i(t) = t^3 - 27Ix - 27J$ is the resolvent cubic polynomial of $g_i(x)$; see [Skorobogatov 2001, Proposition 3.3.6(a)]. The 0-dimensional scheme $g_i(x) = 0$ is a k-torsor Z_i for $E_i[2]$. Then C_i can be viewed as the twist of E_i by Z_i; that is, $C_i = (E_i \times Z_i)/E_i[2]$, where $E_i[2]$ acts simultaneously on both factors. The antipodal involution acts on C_i by changing the sign of y, so the Kummer surface $\mathrm{Kum}(C_1 \times C_2)$ is the minimal desingularisation of the quotient of $C_1 \times C_2$ by the involution that acts on each component as $(x, y) \mapsto (x, -y)$. Thus $z^2 = g_1(x)g_2(y)$ defines an affine surface birationally equivalent to $\mathrm{Kum}(C_1 \times C_2)$.

Since the polynomials $g_1(x)$ and $g_2(x)$ have no roots in k, each of the torsors Z_1 and Z_2 is nontrivial. The field of definition $K_i = k(E_i[2])$ of $E_i[2]$ is the splitting field of $p_i(t)$. Hence the condition $\mathrm{Gal}(g_1) \simeq S_4$ implies $\mathrm{Gal}(K_i/k) = \mathrm{Gal}(p_i) \simeq S_3$, for $i = 1, 2$. The discriminant of the quartic $g_i(x)$ is equal to the discriminant of its resolvent cubic $p_i(t)$ up to a power of 3, and $g_i(x) \in \mathcal{O}_{w_j}[x]$ implies $p_i(t) \in \mathcal{O}_{w_j}[t]$, so the primes w_1 and w_2 satisfy the assumption in Corollary 2.4. To be in a position to appeal to that corollary we now show that Z_i is unramified at both w_1 and w_2.

Indeed, let $\mathcal{Z}_{ij} \subset \mathbb{P}^1_{\mathcal{O}_{w_j}}$ be the closed subscheme given by $G_i(u, v) = 0$, where $G_i(u, v) = v^4 g_i(u/v)$. For $j \neq i$, the discriminant of $G_i(u, v)$ is a unit in $\mathcal{O}_{w_j}$. Thus $\mathcal{Z}_{ij}$ is a finite and étale $\mathcal{O}_{w_j}$-scheme of degree 4 with the generic fibre $Z_i \times_k k_{w_j}$, hence Z_i is unramified at w_j. For $i = j$, the discriminant of $G_i(u, v)$ is a generator of the maximal ideal of $\mathcal{O}_{w_i}$. This implies that the fibre $\mathcal{Z}_{ii} \times_{\mathcal{O}_{w_i}} \mathbb{F}_{w_i}$ at the closed point of $\operatorname{Spec}(\mathcal{O}_{w_i})$ is the disjoint union of a double $\mathbb{F}_{w_i}$-point and a reduced 2-point $\mathbb{F}_{w_i}$-scheme. The latter gives rise to two sections of the morphism

$$\mathcal{Z}_{ii} \times_{\mathcal{O}_{w_i}} \mathcal{O}_{w_i}^{\mathrm{nr}} \longrightarrow \operatorname{Spec}(\mathcal{O}_{w_i}^{\mathrm{nr}}).$$

Hence Z_i is unramified at w_i. An application of Corollary 2.4 finishes the proof. □

Proof of Theorem B assuming Theorem 2.3. The condition $\mathrm{val}_w(\Delta(f)) = 1$ implies that $k(\sqrt{\Delta(f)})$ has degree 2 over k. Hence the Galois group of $f(x)$ is not a subgroup of the alternating group A_5. Any proper subgroup of S_5 which acts transitively on $\{1, 2, 3, 4, 5\}$ and is not contained in A_5 is conjugate to $\mathrm{Aff}_5 = \mathbb{F}_5 \rtimes \mathbb{F}_5^*$, the group of affine transformations of the affine line over the finite field $\mathbb{F}_5$; see [Burnside 1911, Chapter XI, §166, p. 215]. Let us show that this case cannot occur. Indeed, in the proof of Corollary 2.4 we have seen that the reduction of $f(x)$ modulo $\mathfrak{m}_w$ has one rational double root and three simple roots, whereas the integral model defined by $y^2 = f(x)$ is regular. It follows that over the maximal unramified extension of k_w the polynomial $f(x)$ is the product of three linear and one irreducible quadratic polynomials. Hence the image of the inertia subgroup in Aff_5 is generated by a cycle of length 2. This is a contradiction because the elements of order 2 in Aff_5 are always products of two cycles, as they are given by affine transformations of the form $x \mapsto -x + a$.

We conclude that the Galois group of $f(x)$ is S_5. The theorem now follows from Corollary 2.4, provided we check that the relevant class in $\mathrm{H}^1(k, A[2])$ is nonzero and unramified at w.

For this it is enough to prove that the corresponding k-torsor for $A[2]$ has no k-points but has a k_w^{un}-point. This torsor is the subset $Z_\lambda \subset \mathrm{R}_{L/k}(\mathbb{G}_{m,L})/\{\pm 1\}$ given by $z^2 = \lambda$. The natural surjective map

$$\mathrm{R}_{L/k}(\mathbb{G}_{m,L}) \longrightarrow \mathrm{R}_{L/k}(\mathbb{G}_{m,L})/\{\pm 1\}$$

is a torsor for μ_2. Thus $Z_\lambda(k)$ is the disjoint union of the images of k-points of the torsors $tz^2 = \lambda$ for $\mathrm{R}_{L/k}(\mu_2)$, where $t \in k^*$. Hence $Z_\lambda(k) \neq \varnothing$ if and only if $\lambda \in k^* L^{*2}$, but this is excluded by one of the assumptions of Theorem B. Next, the group $\mathrm{H}^1(k_w^{\mathrm{un}}, \mu_2)$ consists of the classes of 1-cocycles defined by 1 and π, where π is a generator of $\mathfrak{m}_w$. Hence $Z_\lambda(k_w^{\mathrm{un}})$ is the disjoint union of the images of k_w^{un}-points of the torsors $z^2 = \lambda$ and $z^2 = \pi\lambda$ for $\mathrm{R}_{L/k}(\mu_2)$. By assumption there exists an $\varepsilon \in \{0, 1\}$ such that the valuation of $\pi^\varepsilon \lambda$ at each completion of L over w is even. Then the torsor for $\mathrm{R}_{L/k}(\mu_2)$ given by $z^2 = \pi^\varepsilon \lambda$ has a k_w^{un}-point, because any unit is a square, as the residue field of k_w^{un} is separably closed and of characteristic different from 2. It follows that $Z_\lambda(k_w^{\mathrm{un}}) \neq \varnothing$. □

3. Galois theory of finite torsors

This section develops some ideas from [Mazur and Rubin 2010]; see Lemma 3.5 in particular.

We shall work with groups that are semidirect products of a group G with a semisimple G-module M. Recall that a G-module M is *simple* if it has no G-submodules except 0 and M. A G-module M is *semisimple* if M is a direct sum of simple G-modules $M = \bigoplus_i M_i$. The simple G-modules M_i are called the *simple factors* of M. Their isomorphism types do not depend on the presentation of M as a direct sum. Indeed, one can characterise the simple factors of M as the simple G-modules that admit a nonzero map to M or from M.

Remark 3.1. If M is a semisimple G-module, then each G-submodule of M is a direct summand of M; see, e.g., [Wisbauer 1991, 20.2]. Furthermore, each G-submodule $N \subseteq M$ is semisimple and each simple factor of N is a simple factor of M. Similarly, each quotient G-module M/N is semisimple and each simple factor of M/N is a simple factor of M.

Lemma 3.2. *Let G be a group with more than one element and let M be a semisimple G-module such that the action of G on each simple factor of M is faithful. Let $H \subseteq M \rtimes G$ be a normal subgroup. Then*

(i) *either $H \subseteq M$ or $M \subseteq H$;*

(ii) *if $(M \rtimes G)/H$ is abelian, then $M \subseteq H$.*

Proof. (i) Suppose that M is not contained in H. The subgroup $K = H \cap M$ is normal in $M \rtimes G$, thus K is a proper G-submodule of M. The quotient G-module $N = M/K \neq 0$ is semisimple by Remark 3.1. Moreover, each simple factor of N is a simple factor of M, hence N is a faithful G-module. We identify K with the kernel of the natural surjective group homomorphism $\rho : M \rtimes G \to N \rtimes G$. Then $\rho(H)$ and N are normal subgroups of $N \rtimes G$ such that $\rho(H) \cap N = \{1\}$, hence they centralise each other. Thus the image of H in G acts trivially on N. But N is a faithful G-module, so the image of H in G is trivial, hence $H \subseteq M$.

(ii) By the result of (i) we just need to show that the case $H \subsetneq M$ is not possible. Indeed, since H is normal in $M \rtimes G$, in this case H is a proper G-submodule of M, so that $(M \rtimes G)/H = N \rtimes G$, where $N = M/H \neq 0$. The same argument as in the proof of (i) shows that N is a faithful G-module. By assumption $(M \rtimes G)/H$ is abelian, so G acts trivially on N. This contradicts the fact that G contains an element other than the unit of the group law. □

Let us now set up notation and terminology for this section.

Let k be a field and $\bar{k}$ a separable closure of k, and let $\Gamma_k = \mathrm{Gal}(\bar{k}/k)$. Let M be a finite Γ_k-module such that the order of M is not divisible by $\mathrm{char}(k)$. We denote by

$\varphi : \Gamma_k \to \mathrm{Aut}(M)$ the action of Γ_k on M. We identify M with the group of $\bar{k}$-points of a finite étale commutative group k-scheme $\mathcal{G}_M$. A cocycle $c : \Gamma_k \to M = \mathcal{G}_M(\bar{k})$ gives rise to a twisted action of Γ_k on $\mathcal{G}_M(\bar{k})$, defined as the original action of Γ_k on M followed by the translation by c. The quotient of $\mathrm{Spec}(\bar{k}[\mathcal{G}_M])$ by the twisted action is a k-torsor of $\mathcal{G}_M$. It comes equipped with a $\bar{k}$-point corresponding to the neutral element of $\mathcal{G}_M$. Conversely, suppose we are given a k-torsor Z for $\mathcal{G}_M$. For any $z_0 \in Z(\bar{k})$, the map $c : \Gamma_k \to M = \mathcal{G}_M(\bar{k})$ determined by the condition $c(\gamma)z_0 = {}^{\gamma}z_0$ is a cocycle $\Gamma_k \to M$. These constructions induce a bijection between $\mathrm{H}^1(k, M)$ and the set of isomorphisms classes of k-torsors for $\mathcal{G}_M$. See [Skorobogatov 2001, Section 2.1], and also [Bosch et al. 1990, Chapter 6]. For $\alpha \in \mathrm{H}^1(k, M)$, we denote by Z_α the torsor for $\mathcal{G}_M$ obtained by twisting $\mathcal{G}_M$ by a 1-cocycle representing α; such a torsor is well-defined up to an isomorphism of $\mathcal{G}_M$-torsors.

Definition 3.3. Let K be the smallest extension of k such that Γ_K acts trivially on M. For $\alpha \in \mathrm{H}^1(k, M)$, let K_α be the smallest extension of k such that Γ_{K_α} acts trivially on $Z_\alpha(\bar{k})$. Write $G = \mathrm{Gal}(K/k)$ and $G_\alpha = \mathrm{Gal}(K_\alpha/k)$.

Note that $K \subset K_\alpha$, which follows from the surjectivity of the difference map $Z_\alpha \times Z_\alpha \to \mathcal{G}_M$. Write $W_\alpha = \mathrm{Gal}(K_\alpha/K)$. Then there is an exact sequence

$$1 \longrightarrow W_\alpha \longrightarrow G_\alpha \xrightarrow{\varphi} G \longrightarrow 1. \tag{4}$$

The group G of Definition 3.3 is identified with $\varphi(\Gamma_k)$, which makes M a faithful G-module. Let $\alpha \in \mathrm{H}^1(k, M)$ be a class represented by a 1-cocycle $c : \Gamma_k \to M$. If Z_α is the twist of $\mathcal{G}_M$ by c, then the semidirect product $M \rtimes G$ acts on $Z_\alpha(\bar{k}) \cong \mathcal{G}_M(\bar{k}) \cong M$ by affine transformations, and Γ_k acts on $Z_\alpha(\bar{k})$ by the homomorphism $(c, \varphi) : \Gamma_k \to M \rtimes G$. By the definition of K_α this homomorphism factors through an injective homomorphism $G_\alpha \to M \rtimes G$. Since M is a trivial Γ_K-module, the restriction of α to W_α defines an injective homomorphism of G-modules $\tilde{\alpha} : W_\alpha \to M$, and we have the commutative diagram

$$\begin{tikzcd}
1 \arrow[r] & W_\alpha \arrow[r] \arrow[d, hook, "\tilde{\alpha}"'] & G_\alpha \arrow[r] \arrow[d, hook, "(c,\varphi)"'] & G \arrow[r] \arrow[d, "="] & 1 \\
1 \arrow[r] & M \arrow[r] & M \rtimes G \arrow[r] & G \arrow[r] & 1
\end{tikzcd}$$

Let $R = \mathrm{End}_G(M) = \mathrm{End}_{\Gamma_k}(M)$ be the endomorphism ring of the Γ_k-module M.

Definition 3.4. Let N be an R-module. We say that $\alpha \in N$ is *nondegenerate* if the annihilator of α in R is zero, i.e., if, for $r \in R$, $r\alpha = 0$ implies $r = 0$. Equivalently, α is nondegenerate if $R\alpha \subset N$ is a free R-module.

Remark 3.5. For any Γ_k-module M, the group $\mathrm{H}^1(k, M)$ is naturally an R-module. If M is a simple G-module, then R is a division ring by Schur's lemma, hence a finite field by Wedderburn's theorem. Then an element $\alpha \in \mathrm{H}^1(k, M)$ is nondegenerate if and only if $\alpha \neq 0$. When $M = \bigoplus_{i=1}^r M_i$, where the G-modules M_i are simple and pairwise nonisomorphic, $R = \bigoplus_{i=1}^r \mathrm{End}_G(M_i)$ is a direct sum of fields. We have $\mathrm{H}^1(k, M) = \bigoplus_{i=1}^r \mathrm{H}^1(k, M_i)$. If we write $\alpha = \sum \alpha_i$ with $\alpha_i \in \mathrm{H}^1(k, M_i)$, then α is nondegenerate if and only if each $\alpha_i \neq 0$. When $M = N^{\oplus r}$ for a simple G-module N, the ring R is the algebra of matrices of size r with entries in the field $\mathrm{End}_G(N)$. In this case $\alpha = \sum \alpha_i$ is nondegenerate if and only if $\alpha_1, \dots, \alpha_r$ are linearly independent in the $\mathrm{End}_G(N)$-vector space $\mathrm{H}^1(k, N)$.

In the following proposition we consider M as a G_α-module via the surjective homomorphism $\varphi : G_\alpha \to G$.

Proposition 3.6. *With the above notation assume that M is a semisimple Γ_k-module such that $\mathrm{H}^1(G, M) = 0$. Let $\alpha \in \mathrm{H}^1(k, M)$ be a class represented by a 1-cocycle c. The following conditions are equivalent*:

(i) *the map (c, φ) is an isomorphism of groups $G_\alpha \xrightarrow{\sim} M \rtimes G$*;

(ii) *the map $\tilde{\alpha} : W_\alpha \xrightarrow{\sim} M$ is an isomorphism of G-modules*;

(iii) *$\mathrm{H}^1(G_\alpha, M)$ is a free R-module generated by α*;

(iv) *α is nondegenerate in $\mathrm{H}^1(k, M)$.*

Proof. (i) $\Rightarrow$ (ii) Since (c, φ) is an isomorphism, for each $m \in M$ there exists a $\gamma \in G_\alpha$ such that $(c, \varphi)(\gamma) = (m, 1)$. Then γ goes to $1 \in G$ and hence $\gamma \in W_\alpha$ and $c(\gamma) = \tilde{\alpha}(\gamma) = m$. It follows that the map $\tilde{\alpha} : W_\alpha \longrightarrow M$ is surjective. Since it is also injective by construction, we conclude that it is an isomorphism of G-modules.

(ii) $\Rightarrow$ (iii) Assume that $\tilde{\alpha} : W_\alpha \xrightarrow{\sim} M$ is an isomorphism of G-modules. Then $\mathrm{Hom}_G(W_\alpha, M)$ is a free R-module with generator $\tilde{\alpha}$. The inflation-restriction exact sequence

$$0 \to \mathrm{H}^1(G, M) \to \mathrm{H}^1(G_\alpha, M) \to \mathrm{H}^1(W_\alpha, M)^G = \mathrm{Hom}_G(W_\alpha, M) \tag{5}$$

is an exact sequence of R-modules. We note that $Z_\alpha(K_\alpha) \neq \varnothing$, so $\alpha \in \mathrm{H}^1(k, M)$ belongs to the kernel $\mathrm{H}^1(G_\alpha, M)$ of the restriction map to $\mathrm{H}^1(K_\alpha, M)$. By assumption $\mathrm{H}^1(G, M) = 0$, hence the map $\mathrm{H}^1(G_\alpha, M) \to \mathrm{Hom}_G(W_\alpha, M)$ is injective. This map of R-modules sends α to the generator $\tilde{\alpha}$ of the R-module $\mathrm{Hom}_G(W_\alpha, M) = R\tilde{\alpha}$, so it is surjective, hence an isomorphism. We obtain that $\mathrm{H}^1(G_\alpha, M)$ is a free R-module generated by α.

(iii) $\Rightarrow$ (iv) Assume that $\mathrm{H}^1(G_\alpha, M)$ is a free R-module with generator α. By the inflation-restriction exact sequence for $\Gamma_{K_\alpha} \subseteq \Gamma_k$, the map $\mathrm{H}^1(G_\alpha, M) \to \mathrm{H}^1(k, M)$ is injective, and so (iv) holds.

(iv) $\Rightarrow$ (i) Suppose that α is nondegenerate in $\mathrm{H}^1(k, M)$ and assume for contradiction that the map (c, φ) is not an isomorphism. Since (c, φ) is injective by construction, we conclude that it is not surjective. The intersection of the image of (c, φ) with M is then a proper G-submodule $\tilde{\alpha}(W_\alpha) \subsetneq M$. Since M is semisimple, $\tilde{\alpha}(W_\alpha)$ is a direct summand of M; see Remark 3.1. It follows that there exists a nonzero element $r \in R$ such that $r\tilde{\alpha}(W_\alpha) = 0$, so that $r\tilde{\alpha} = 0$ in $\mathrm{Hom}_G(W_\alpha, M)$. From (5) we see that $r\alpha = 0$ in $\mathrm{H}^1(G_\alpha, M)$. But this is a contradiction because the map $\mathrm{H}^1(G_\alpha, M) \to \mathrm{H}^1(k, M)$ is injective and α is nondegenerate in $\mathrm{H}^1(k, M)$. □

We record an amusing corollary of this proposition.

Corollary 3.7. *Under the assumptions of Proposition 3.6, let $\alpha, \beta \in \mathrm{H}^1(k, M)$ be nondegenerate. Then the associated torsors Z_α, Z_β for $\mathcal{G}_M$ are integral k-schemes. Furthermore, the following conditions are equivalent:*

(i) *there exists an $r \in R^*$ such that $r\alpha = \beta$;*

(ii) *$R\alpha = R\beta \subset \mathrm{H}^1(k, M)$;*

(iii) *Z_α and Z_β are isomorphic as abstract k-schemes.*

Proof. Let c be a cocycle representing α. By Proposition 3.6 the group G_α acts on $Z_\alpha(\bar{k}) \simeq M$ via the isomorphism $(c, \varphi) : G_\alpha \xrightarrow{\sim} M \rtimes G$. Hence G_α acts transitively on $Z_\alpha(\bar{k})$, because already the subgroup $M \subset M \rtimes G$ acts (simply) transitively on M. Hence Z_α is integral. The same argument proves that Z_β is integral.

Let us now establish the equivalence of (i), (ii) and (iii). The implication (i) $\Rightarrow$ (ii) is clear. Conversely, if $R\alpha = R\beta$, then there exist $r, s \in R$ such that $r\alpha = \beta$ and $\alpha = s\beta$. Then $sr\alpha = \alpha$ and $rs\beta = \beta$. Since α and β are nondegenerate, we obtain that r and s are invertible in R, so (ii) implies (i).

We now show that (i) is equivalent to (iii). Assume (i) and take a cocycle $c : \Gamma_k \to M$ which represents α. Then rc represents β. We identify $Z_\alpha(\bar{k})$ with $\mathcal{G}_M(\bar{k})$ such that Γ_k acts via its original action on $\mathcal{G}_M(\bar{k})$ followed by the translation by c. Then $Z_\beta(\bar{k})$ can be identified with $\mathcal{G}_M(\bar{k})$ such that Γ_k acts via its original action on $\mathcal{G}_M(\bar{k})$ followed by the translation by rc. Under these identifications the map $r : \mathcal{G}_M(\bar{k}) \to \mathcal{G}_M(\bar{k})$ becomes a Γ_k-equivariant map $Z_\alpha(\bar{k}) \to Z_\beta(\bar{k})$. Thus Z_α and Z_β are isomorphic as 0-dimensional k-schemes.

Finally, assume that Z_α and Z_β are isomorphic as k-schemes. Since α and β are nondegenerate, we see from Proposition 3.6 that the maps $\tilde{\alpha} : W_\alpha \xrightarrow{\sim} M$ and $\tilde{\beta} : W_\beta \xrightarrow{\sim} M$ are isomorphisms of G-modules. The splitting fields K_α and K_β coincide as subfields of $\bar{k}$, so there exists an isomorphism of Γ_k-modules represented by the dotted arrow in the diagram

$$\begin{tikzcd}
 & W_\alpha \arrow[r, "\tilde{\alpha}"] \arrow[d, "\cong"] & M \arrow[d, dashed, "r"] \\
\Gamma_K \arrow[ur, two heads] \arrow[dr, two heads] & W_\beta \arrow[r, "\tilde{\beta}"] & M \\
 & &
\end{tikzcd}$$

It is obtained as the action of an invertible element $r \in R^*$. It follows that $r\alpha$ and β have the same image in $\mathrm{H}^1(K, M)$. By assumption $\mathrm{H}^1(G, M) = 0$, hence the restriction-inflation exact sequence implies that the map $\mathrm{H}^1(k, M) \to \mathrm{H}^1(K, M)$ is injective. Thus $r\alpha = \beta$, as desired. $\square$

A continuous action of the procyclic group $\hat{\mathbb{Z}}$ on a discrete module N is determined by the homomorphism $g : N \to N$ which is the action of the generator $1 \in \hat{\mathbb{Z}}$. There is a canonical isomorphism

$$\mathrm{H}^1(\hat{\mathbb{Z}}, N) \cong N/(g-1)$$

induced by sending the class of a cocycle ξ to the class of $\xi(1)$ in $N/(g-1)$.

An element $\gamma \in G_\alpha$ determines a map $f_\gamma : \hat{\mathbb{Z}} \to G_\alpha$ which sends 1 to γ, and hence an induced map

$$f_\gamma^* : \mathrm{H}^1(G_\alpha, M) \longrightarrow \mathrm{H}^1(\hat{\mathbb{Z}}, M) = M/(g-1).$$

Here we denote by g the image of γ in G (which acts on M) under the natural surjective map $G_\alpha \to G$. In particular, if $c : G_\alpha \to M$ is a cocycle representing $\alpha \in \mathrm{H}^1(G_\alpha, M)$, then $f_\gamma^*(\alpha)$ is equal to the class of $c(\gamma)$ in $M/(g-1)$.

Corollary 3.8. *In the assumptions of Proposition 3.6 let $\alpha \in \mathrm{H}^1(k, M)$ be nondegenerate. Take any $g \in G$ and any $x \in M/(g-1)$. Then g has a lifting $\gamma \in G_\alpha$ such that $f_\gamma(\alpha) = x$.*

Proof. Let $c : G_\alpha \to M$ be a cocycle representing α and let $m \in M$ be an element whose class in $M/(g-1)$ is x. By Proposition 3.6 the map $(c, \varphi) : G_\alpha \xrightarrow{\sim} M \rtimes G$ is an isomorphism. Hence there exists an element $\gamma \in G_\alpha$ such that $(c, \varphi)(\gamma) = (m, g)$. Then γ is a lifting of g and $c(\gamma) = m$ so that $f_\gamma(\alpha) = x$, as desired. $\square$

Corollary 3.9. *Let M be a semisimple Γ_k-module such that G contains more than one element, the action of G on each simple factor of M is faithful and $\mathrm{H}^1(G, M) = 0$. Let $\alpha \in \mathrm{H}^1(k, M)$ be nondegenerate. Then*

(i) *each subfield of K_α that is Galois over k is either contained in K or contains K;*

(ii) *each subfield of K_α that is abelian over k is contained in K.*

Proof. By Proposition 3.6 we have $G_\alpha \simeq M \rtimes G$. The desired result now follows directly from Lemma 3.2. $\square$

Until the end of this section we assume that k is a field of characteristic different from 2. Let $A_1, \dots, A_r$ be abelian varieties satisfying conditions (a) and (b) of Section 2 and let $A = \prod_{i=1}^r A_i$. Let K_i be the splitting field of $A_i[2]$. The compositum $K = K_1 \cdots K_r$ is the field of definition of $A[2]$. Assume that condition (e) of Section 2 holds, i.e., the fields $K_1, \dots, K_r$ are linearly disjoint over k.

Remark 3.10. Condition (a) implies that each K_i is a nontrivial extension of k, so that G_i has more than one element.

We now present two applications of the results above. In the first one we consider the semisimple Γ_k-module $M = A[2] = \bigoplus_{i=1}^r A_i[2]$.

Proposition 3.11. *Suppose that abelian varieties $A_1, \dots, A_r$ satisfy conditions* (a) *and* (b), *and that condition* (e) *holds. Let Z_i be a nontrivial k-torsor for $A_i[2]$, for each $i = 1, \dots, r$, and let $Z = \prod_{i=1}^r Z_i$. Let L be the étale k-algebra $k[Z]$, so that $Z \cong \mathrm{Spec}(L)$. Then L is a field which contains no quadratic extension of k.*

Proof. Let $M = A[2]$ and let $\alpha \in \mathrm{H}^1(k, M)$ be the class of Z. Write $\alpha = \sum_{i=1}^r \alpha_i$, where each $\alpha_i \in \mathrm{H}^1(k, A_i[2])$ is nonzero. By condition (a) each $A_i[2]$ is simple and hence M is semisimple with simple factors $A_1[2], \dots, A_r[2]$. By condition (e) the fields $K_1, \dots, K_r$ are linearly disjoint over k, so that the Galois group $G = \mathrm{Gal}(K/k)$ is the product $G = \prod_{i=1}^r G_i$, and the $A_i[2]$ are pairwise nonisomorphic Γ_k-modules. From Remark 3.5 we see that α is nondegenerate.

For each $i = 1, \dots, r$, we have $A_i[2]^{G_i} = 0$ and $\mathrm{H}^1(G_i, A_i[2]) = 0$ by conditions (a) and (b). The inflation-restriction exact sequence for $G_i \subset G$ then gives $\mathrm{H}^1(G, A_i[2]) = 0$, and so $\mathrm{H}^1(G, M) = 0$. Let $c : \Gamma_k \to M$ be a cocycle representing α. By Proposition 3.6 the map $(c, \varphi) : G_\alpha \xrightarrow{\sim} M \rtimes G$ is an isomorphism. Let $s : G \to G_\alpha$ be the section corresponding to the canonical section $G \to M \rtimes G$ under the isomorphism (c, φ).

By Corollary 3.7 the scheme Z_α is integral, and hence $L = k[Z_\alpha]$ is a field whose Galois closure is K_α by definition. Moreover, $L \cong (K_\alpha)^{s(G)}$. If L contains a quadratic extension of k, then $s(G)$ is contained in a normal subgroup $H \subset G_\alpha$ of index 2. Since s is a section, the induced homomorphism $H \to G$ is surjective, so its kernel is a G-submodule of M which is a subgroup of M of index 2. But this is a contradiction since M is semisimple and the simple factors $A_i[2]$ of M have size 4. □

In the second application we consider the semisimple module $M = A_1[2]^{\oplus r}$.

Proposition 3.12. *Suppose that abelian varieties $A_1, \dots, A_r$ satisfy conditions* (a) *and* (b), *and that condition* (e) *holds. Let $M = A_1[2]^{\oplus r}$ be a direct sum of copies of $A_1[2]$ and let $\alpha \in \mathrm{H}^1(k, M)$ be nondegenerate. Then the fields $(K_1)_\alpha, K_2, \dots, K_r$ are linearly disjoint.*

Proof. Write $E = (K_1)_\alpha \cap K_2 \cdots K_r$. In view of condition (e) it is enough to show that $E = k$. Indeed, E is a Galois subfield of $(K_1)_\alpha$, so by Remark 3.10 and Corollary 3.9 we have $E \subset K_1$ or $K_1 \subset E$. In the first case, $E = k$ because E is contained in $K_1 \cap K_2 \cdots K_r = k$, where the equality holds by condition (e). By the same condition the second case cannot actually occur, because then $K_1 \subset E \subset K_2 \cdots K_r$ which contradicts the linear disjointness of $K_1, \dots, K_r$. □

4. Kummer map over a local field

Let A be an abelian variety over a local field k of characteristic zero. The Kummer exact sequence gives rise to a map $\delta : A(k) \to \mathrm{H}^1(k, A[2])$, called the Kummer map. For $x \in A(k)$, choose $\bar{x} \in A(\bar{k})$ such that $2\bar{x} = x$. Then $\delta(x)$ is represented by the cocycle that sends $\gamma \in \Gamma_k$ to ${}^{\gamma}\bar{x} - \bar{x} \in A[2]$.

The Weil pairing is a nondegenerate pairing of Γ_k-modules $A[2] \times A^t[2] \to \mathbb{Z}/2$. The induced pairing on cohomology followed by the local invariant of local class field theory gives a nondegenerate pairing of finite abelian groups [Milne 1986, Corollary I.2.3]

$$\mathrm{H}^1(k, A[2]) \times \mathrm{H}^1(k, A^t[2]) \longrightarrow \mathrm{Br}(k)[2] \xrightarrow{\mathrm{inv}} \tfrac{1}{2}\mathbb{Z}/\mathbb{Z}.$$

The local Tate duality implies that $\delta(A(k))$ and $\delta(A^t(k))$ are the orthogonal complements to each other under this pairing (see, e.g., the first commutative diagram in the proof of [Milne 1986, I.3.2]).

When A is principally polarised, we combine the last pairing with the principal polarisation $A \xrightarrow{\sim} A^t$ and obtain a nondegenerate symmetric pairing

$$\mathrm{inv}(\alpha \cup \beta) : \mathrm{H}^1(k, A[2]) \times \mathrm{H}^1(k, A[2]) \longrightarrow \mathrm{Br}(k)[2] \xrightarrow{\mathrm{inv}} \tfrac{1}{2}\mathbb{Z}/\mathbb{Z}.$$

It is well-known that $\delta(A(k))$ is a maximal isotropic subspace of $\mathrm{H}^1(k, A[2])$; see [Poonen and Rains 2012, Proposition 4.11]. Note that the pairing $\mathrm{inv}(\alpha \cup \beta)$ is also defined for $k = \mathbb{R}$ and the above statements carry over to this case; cf. [Milne 1986, Theorem I.2.13(a), Remark I.3.7].

Let us recall a well-known description of $\delta(A(k))$ when A has good reduction. Let κ be the residue field of k, and assume $\mathrm{char}(\kappa) = \ell \neq 2$. Then $\delta(A(k))$ is the unramified subgroup

$$\mathrm{H}^1_{\mathrm{nr}}(k, A[2]) = \mathrm{Ker}[\mathrm{H}^1(\Gamma_k, A[2]) \longrightarrow \mathrm{H}^1(I, A[2])],$$

where $I \subset \Gamma_k$ is the inertia subgroup. By Néron–Ogg–Shafarevich the inertia acts trivially on $A[2]$, so that $\mathrm{H}^1_{\mathrm{nr}}(k, A[2]) = \mathrm{H}^1(\kappa, A[2])$. The absolute Galois group $\mathrm{Gal}(\bar{\kappa}/\kappa) = \Gamma_k/I$ is isomorphic to $\hat{\mathbb{Z}}$ with the Frobenius element as a topological generator. Thus we have a canonical isomorphism

$$\delta(A(k)) = A[2]/(\mathrm{Frob} - 1). \tag{6}$$

Since $\hat{\mathbb{Z}}$ has cohomological dimension 1, the spectral sequence

$$\mathrm{H}^p(\hat{\mathbb{Z}}, \mathrm{H}^q(I, A[2])) \Rightarrow \mathrm{H}^{p+q}(k, A[2])$$

gives rise to the exact sequence

$$0 \to A[2]/(\mathrm{Frob} - 1) \to \mathrm{H}^1(k, A[2]) \to \mathrm{Hom}(I, A[2])^{\mathrm{Frob}} \to 0.$$

The maximal abelian pro-2-quotient of I is isomorphic to $\mathbb{Z}_2$, and Frob acts on it by multiplication by ℓ. Thus $\mathrm{Hom}(I, A[2]) = A[2]$ with the natural action of Frob, so that

$$\mathrm{Hom}(I, A[2])^{\mathrm{Frob}} = A[2]^{\mathrm{Frob}} = \mathrm{Ker}(\mathrm{Frob} - 1 : A[2] \to A[2]).$$

It follows that the dimension of the $\mathbb{F}_2$-vector space $A[2]/(\mathrm{Frob} - 1)$ equals the dimension of $A[2]^{\mathrm{Frob}}$, and therefore

$$\dim \mathrm{H}^1(k, A[2]) = 2 \dim A[2]/(\mathrm{Frob} - 1). \tag{7}$$

Let us now return to the general case, where A does not necessarily have good reduction. If F/k is a quadratic extension, we write $\delta^F : A^F(k) \to \mathrm{H}^1(k, A[2])$ for the Kummer map of A^F. In the rest of this section we summarise some known results relating δ, δ^F and the norm map $\mathrm{N} : A(F) \to A(k)$.

Lemma 4.1. *We have* $\delta(\mathrm{N}(A(F))) = \delta(A(k)) \cap \delta^F(A^F(k)) \subset \mathrm{H}^1(k, A[2])$.

Proof. Cf. [Kramer 1981, Proposition 7] or [Mazur and Rubin 2007, Proposition 5.2]. Let $\chi : \Gamma_k \to \{\pm 1\}$ be the quadratic character associated to F. We choose $\sigma \in \Gamma_k$ such that $\chi(\sigma) = -1$.

Suppose that $x \in A(k)$ and $y \in A^F(k)$ are such that $\delta(x) = \delta^F(y)$. Using the embedding $A^F(k) \subset A(F)$ we can consider y as a point in $A(F)$ such that ${}^{\sigma}y = -y$. If $\bar{y} \in A(\bar{k})$ is such that $2\bar{y} = y$, then $\delta^F(y)$ is represented by the cocycle that sends $\gamma \in \Gamma_k$ to

$$\chi(\gamma)\, {}^{\gamma}\bar{y} - \bar{y} = {}^{\gamma}\bar{y} - \chi(\gamma)\bar{y} \in A[2].$$

Since $\delta(x) = \delta^F(y)$, we can choose $\bar{x} \in A(\bar{k})$ such that $2\bar{x} = x$ and such that

$$\chi(\gamma)\, {}^{\gamma}\bar{y} - \bar{y} = {}^{\gamma}\bar{x} - \bar{x}.$$

We deduce that ${}^{\gamma}(\bar{x} - \bar{y}) = \bar{x} - \chi(\gamma)\bar{y}$ for every $\gamma \in \Gamma_k$. It follows that $\bar{x} - \bar{y} \in A(F)$ and ${}^{\sigma}(\bar{x} - \bar{y}) = \bar{x} + \bar{y}$. Therefore, $x = 2\bar{x} = \mathrm{N}(\bar{x} - \bar{y})$ is a norm from $A(F)$.

Conversely, suppose that $x = \mathrm{N}(z) = z + {}^{\sigma}z$ for some $z \in A(F)$. Let $y = {}^{\sigma}z - z$. Then $y \in A^F(k)$ and we claim that $\delta(x) = \delta^F(y)$. Choose $\bar{x} \in A(\bar{k})$ such that $2\bar{x} = x$ and set $\bar{y} = \bar{x} - z$. Then $2\bar{y} = x - 2z = y$ and we have $\bar{x} - \bar{y} = z$ and $\bar{x} + \bar{y} = {}^{\sigma}z$. It follows that for each $\gamma \in \Gamma_k$ we have ${}^{\gamma}(\bar{x} - \bar{y}) = \bar{x} - \chi(\gamma)\bar{y}$, and hence

$${}^{\gamma}\bar{x} - \bar{x} = {}^{\gamma}\bar{y} - \chi(\gamma)\bar{y}.$$

This implies $\delta(x) = \delta^F(y)$, as desired. □

Lemma 4.2. *Let A be a principally polarised abelian variety over k with bad reduction such that the number of geometric connected components of the Néron model of A is odd. If F is an unramified quadratic extension of k, then* $\delta(A(k)) = \delta^F(A^F(k))$.

Proof. Since A is principally polarised, it is isomorphic to its dual abelian variety. It follows from [Mazur 1972, Propositions 4.2 and 4.3] that the norm map $\mathrm{N}: A(F) \to A(k)$ is surjective. By Lemma 4.1 we see that $\delta(A(k)) \subset \delta^F(A^F(k))$. Since F is unramified, the quadratic twist A^F also satisfies the assumptions of the lemma, and the same argument applied to A^F gives the opposite inclusion. □

Lemma 4.3. *Assume that the residue characteristic of k is not 2. If A is an abelian variety over k with good reduction and F is a ramified quadratic extension of k, then $\delta(A(k)) \cap \delta^F(A^F(k)) = 0$.*

Proof. In this case we have $\mathrm{N}(A(F)) = 2A(k)$. If $\dim(A) = 1$ this is proved in [Mazur and Rubin 2007, Lemma 5.5(ii)], and the same proof works in the general case. It remains to apply Lemma 4.1. □

5. Selmer group and Cassels–Tate pairing

Let A be an abelian variety over a field k of characteristic zero. Let $\mathrm{NS}(\bar{A})$ be the Néron–Severi group of $\bar{A}$. The dual abelian variety A^t represents the functor Pic^0_A. In particular, we have an exact sequence of Γ_k-modules

$$0 \longrightarrow A^t(\bar{k}) \longrightarrow \mathrm{Pic}(\bar{A}) \longrightarrow \mathrm{NS}(\bar{A}) \longrightarrow 0. \tag{8}$$

The antipodal involution $\iota_A = [-1] : A \to A$ induces an action of $\mathbb{Z}/2$ on $\mathrm{Pic}(\bar{A})$ which turns (8) into an exact sequence of $\mathbb{Z}/2$-modules. The induced action on $\mathrm{NS}(\bar{A})$ is trivial; see [Skorobogatov and Zarhin 2012, p. 119]. The involution ι_A induces the involution ι_{A^t} on A^t. Since $A^t(\bar{k})$ is divisible, we obtain $\mathrm{H}^1(\mathbb{Z}/2, A^t(\bar{k})) = 0$. Thus the long exact sequence of cohomology gives an exact sequence

$$0 \longrightarrow A^t[2] \longrightarrow \mathrm{Pic}(\bar{A})^{[-1]^*} \longrightarrow \mathrm{NS}(\bar{A}) \longrightarrow 0; \tag{9}$$

cf. [Poonen and Rains 2011, Section 3.2]. It is well-known that $\mathrm{NS}(\bar{A})$ is canonically isomorphic to the group $\mathrm{Hom}(\bar{A}, \bar{A}^t)^{\mathrm{sym}}$ of self-dual homomorphisms of abelian varieties $\bar{A} \to \bar{A}^t$; see, e.g., [Polishchuk 2003, Theorem 13.7]. Hence $\mathrm{NS}(\bar{A})^{\Gamma_k}$ is canonically isomorphic to the group $\mathrm{Hom}(A, A^t)^{\mathrm{sym}}$ of self-dual k-homomorphisms of abelian varieties $A \to A^t$; cf. [Poonen and Rains 2011, Remark 3.3]. A *polarisation* on A is an element $\lambda \in \mathrm{NS}(\bar{A})^{\Gamma_k}$ that comes from an ample line bundle on $\bar{A}$. The polarisation is called *principal* if the associated morphism $\varphi_\lambda : A \to A^t$ is an isomorphism. Following [Poonen and Rains 2011] we shall write c_λ for the image of λ under the differential $\mathrm{NS}(\bar{A})^{\Gamma_k} \to \mathrm{H}^1(k, A^t[2])$ attached to (9). In particular, c_λ vanishes if and only if λ lifts to an element of $(\mathrm{Pic}(\bar{A})^{[-1]^*})^{\Gamma_k} = \mathrm{Pic}(A)^{[-1]^*}$. For example, if A is the Jacobian of a smooth projective curve C and λ is the canonical principal polarisation of A, then c_λ is the image of the class of the theta characteristics torsor of C under the isomorphism $\varphi_{\lambda *} : \mathrm{H}^1(k, A[2]) \xrightarrow{\sim} \mathrm{H}^1(k, A^t[2])$; see

[Poonen and Rains 2011, Theorem 3.9]. In this case, $c_\lambda = 0$ when C has a rational Weierstrass point.

Lemma 5.1. *Let A be an abelian variety over a field k of characteristic* 0 *with polarisation λ. Let $K = k(A[2])$. Then c_λ belongs to the kernel of the restriction map* $\mathrm{H}^1(k, A^t[2]) \to \mathrm{H}^1(K, A^t[2])$.

Proof. This is a particular case of [Poonen and Rains 2011, Lemma 3.6(a)]. □

Now let k be a number field. For a place v of k, let

$$\mathrm{loc}_v : \mathrm{H}^1(k, A[2]) \longrightarrow \mathrm{H}^1(k_v, A[2])$$

be the natural restriction map. The 2-*Selmer group* $\mathrm{Sel}_2(A) \subset \mathrm{H}^1(k, A[2])$ is defined as the set of elements x such that $\mathrm{loc}_v(x) \in \delta(A(k_v))$ for all places v of k. If v is a place of good reduction, then (6) allows us to write the restriction map at v as

$$\mathrm{loc}_v : \mathrm{Sel}_2(A) \longrightarrow A[2]/(\mathrm{Frob}_v - 1).$$

For every quadratic extension F/k, we have $A^F[2] = A[2]$ and hence we may consider the 2-Selmer groups $\mathrm{Sel}_2(A^F)$ of all quadratic twists A^F as subgroups of $\mathrm{H}^1(k, A[2])$. We have the well-known exact sequence

$$0 \longrightarrow A(k)/2 \longrightarrow \mathrm{Sel}_2(A) \longrightarrow Ш(A)[2] \longrightarrow 0. \tag{10}$$

The Cassels–Tate pairing is a bilinear pairing

$$\langle\, , \rangle : Ш(A) \times Ш(A^t) \longrightarrow \mathbb{Q}/\mathbb{Z}.$$

If $Ш(A)$ is finite, then $Ш(A^t)$ is finite too and the Cassels–Tate pairing is nondegenerate; see [Milne 1986, Theorem I.6.26]. A polarisation λ on A induces a homomorphism $\varphi_{\lambda *} : Ш(A) \to Ш(A^t)$.

Proposition 5.2. *Let A be an abelian variety over a number field k with a principal polarisation λ. Then condition* (b) *of Section 2 implies that the Cassels–Tate pairing $\langle x, \varphi_{\lambda *} y\rangle$ on $Ш(A)\{2\}$ is alternating. In particular, if the* 2*-primary subgroup $Ш(A)\{2\}$ is finite, then the cardinality of $Ш(A)[2]$ is a square.*

Proof. By [Poonen and Stoll 1999, Corollary 2] we know that $c_\lambda \in \mathrm{Sel}_2(A^t)$. If c'_λ is the image of c_λ in $Ш(A^t)[2]$, then [Poonen and Stoll 1999, Theorem 5] says that $\langle x, \varphi_{\lambda *} x + c'_\lambda\rangle = 0$ for any $x \in Ш(A)$. Thus it is enough to prove that $c_\lambda = 0$. Lemma 5.1 implies that c_λ belongs to the image of the inflation map $\mathrm{H}^1(G, A^t[2]) \to \mathrm{H}^1(k, A^t[2])$, where $G = \mathrm{Gal}(k(A[2])/k)$ is the image of $\Gamma_k \to \mathrm{GL}(A[2])$. Since λ is a principal polarisation, φ_λ induces an isomorphism of Γ_k-modules $A[2] \xrightarrow{\sim} A^t[2]$. Now condition (b) implies $\mathrm{H}^1(G, A^t[2]) = \mathrm{H}^1(G, A[2]) = 0$, hence $c_\lambda = 0$. □

6. Kummer varieties

Let A be an abelian variety over a field k of characteristic different from 2. Let Z be a k-torsor for the group k-scheme $A[2]$. Recall that the 2-covering $f : Y \to A$ associated to Z is a k-torsor for A defined as the quotient of $A \times_k Z$ by the diagonal action of $A[2]$. In other words, Y is the twisted form of A by Z with respect to the action of $A[2]$ by translations. The morphism f is induced by the first projection, and we have $Z = f^{-1}(0)$. Let L be the étale k-algebra $k[Z]$, so that $Z \cong \mathrm{Spec}(L)$.

Let $\tilde{Y}$ be the blowing-up of Z in Y. The antipodal involution $\iota_A : A \to A$ induces the map $(\iota_A, \mathrm{Id}) : A \times_k Z \to A \times_k Z$ which commutes with the action of $A[2]$ and hence induces an involution $\iota_Y : Y \to Y$. As ι_Y fixes $Z = f^{-1}(0) \subseteq Y$ it extends to an involution $\iota_{\tilde{Y}} : \tilde{Y} \to \tilde{Y}$ whose fixed-point set is precisely the exceptional divisor. It is easy to see that the quotient $X = \mathrm{Kum}(Y) = \tilde{Y}/\iota_{\tilde{Y}}$ is smooth. We call X the *Kummer variety* attached to A and Z. We note that the branch locus of $\tilde{Y} \to X$ is $Z \times_k \mathbb{P}^{d-1}_k$, where $d = \dim(A)$.

Let F be an extension of k of degree at most 2. Recall that A^F denotes the quadratic twist of A by F, that is, the abelian variety over k obtained by twisting A by the quadratic character of F with respect to the action of μ_2 via the antipodal involution ι_A. Similarly, Y^F denotes the quadratic twist of Y with respect to the involution ι_Y; see Section 1. Since ι_A commutes with translations by the elements of $A[2]$, the quadratic twist Y^F of Y is a k-torsor for A^F. We have a natural embedding $i_F : Z \to Y^F$. Then $\tilde{Y}^F$, defined as the blowing-up of $i_F(Z)$ in Y^F, is the quadratic twist of $\tilde{Y}$ by the quadratic character of F with respect to the action of μ_2 on $\tilde{Y}$ via $\iota_{\tilde{Y}}$. We can also consider $\tilde{Y}^F$ as a quadratic twist of the 2-covering $\tilde{Y} \to X$, and consequently consider every $\tilde{Y}^F$ as a (ramified) 2-covering of X. It is clear that Y^F, and hence X, has a K-point for any extension K/k such that α is in the kernel of the natural map $\mathrm{H}^1(k, A[2]) \to \mathrm{H}^1(K, A^F)$.

We now recall a construction from [Skorobogatov and Swinnerton-Dyer 2005, §5]. Let $\mathcal{Y}$ be the quotient of $\tilde{Y} \times \mathbb{G}_{m,k}$ by the action of μ_2 in which the generator $-1 \in \mu_2$ acts as the multiplication by -1 on $\mathbb{G}_m$ and by $\iota_{\tilde{Y}}$ on $\tilde{Y}$. The fibre of $\mathcal{Y}$ over $a \in \mathbb{G}_{m,k}(k)$ can be naturally identified with the quadratic twist $\tilde{Y}^F$ where $F = k(\sqrt{a})$. As in [Skorobogatov and Swinnerton-Dyer 2005, §5] one may consider a smooth compactification $\mathcal{Y} \subset \mathcal{X}$ that fits into the commutative diagram

$$\begin{tikzcd} \mathcal{Y} \arrow[r] \arrow[d] & \mathcal{X} \arrow[d, "p"] \\ \mathbb{G}_{m,k} \arrow[r] & \mathbb{P}^1_k \end{tikzcd}$$

Proposition 6.1. *Let $A = \prod_{i=1}^r A_i$ be a product of abelian varieties over k satisfying conditions* (a) *and* (b) *of Section 2 such that condition* (e) *holds. Assume in addition* *that the class* $\alpha \in \mathrm{H}^1(k, A[2])$ *of* Z *is nondegenerate* (see Definition 3.4). *Then the vertical Brauer group of* $\mathcal{X}$ *over* $\mathbb{P}^1_k$ *is the image of* $\mathrm{Br}(k)$ *in* $\mathrm{Br}(\mathcal{X})$.

Proof. Let t be a coordinate on $\mathbb{P}^1$ invertible on $\mathbb{G}_{m,k} \subset \mathbb{P}^1_k$. According to [Skorobogatov and Swinnerton-Dyer 2005, Theorem 3] the vertical Brauer group of $\mathcal{X}$ is generated by the image of $\mathrm{Br}(k)$ and the pullbacks of the classes $(t, c) \in \mathrm{Br}(k(\mathbb{P}^1_k))$, where $c \in k^*$ becomes a square in $L = k[Z]$. By Proposition 3.11 the element c is already a square in k, hence the result. □

Proposition 6.2. *Let* k *be a number field. Let* $A = \prod_{i=1}^r A_i$ *be a product of abelian varieties over* k *satisfying conditions* (a) *and* (b) *of Section* 2, *and such that conditions* (e) *and* (f) *hold. Let* Z *be a* k*-torsor for* $A[2]$ *whose class* $\alpha \in \mathrm{H}^1(k, A[2])$ *is unramified at the places* $w_1, \dots, w_r$ *and nondegenerate. Let* Y *be the attached* 2*-covering of* A *and let* $X = \mathrm{Kum}(Y)$. *If* X *is everywhere locally soluble*, *then there exists an extension* F *of* k *of degree at most* 2 *such that* Y^F *is everywhere locally soluble and* F *is split at* $w_1, \dots, w_r$.

Proof. This is proved in [Skorobogatov and Swinnerton-Dyer 2005, Lemma 6], but we give a detailed proof for the convenience of the reader. Let w be one of the places $w_1, \dots, w_r$. By assumption $\alpha \in \mathrm{H}^1(k, A[2])$ goes to zero under the composed map

$$\mathrm{H}^1(k, A[2]) \longrightarrow \mathrm{H}^1(k_w, A[2]) \longrightarrow \mathrm{H}^1(k_w^{\mathrm{nr}}, A[2]).$$

Hence the class $[Y] \in \mathrm{H}^1(k, A)[2]$ goes to zero under the composed map

$$\mathrm{H}^1(k, A) \longrightarrow \mathrm{H}^1(k_w, A) \longrightarrow \mathrm{H}^1(k_w^{\mathrm{nr}}, A). \tag{11}$$

The second arrow in (11) is the restriction map $\mathrm{H}^1(\Gamma_{k_w}, A) \to \mathrm{H}^1(I_w, A)$, where $\Gamma_{k_w} = \mathrm{Gal}(\overline{k_w}/k_w)$ and $I_w \subset \Gamma_{k_w}$ is the inertia subgroup. By the inflation-restriction sequence we see that the class $[Y \times_k k_w] \in \mathrm{H}^1(\Gamma_{k_w}, A)$ belongs to the subgroup $\mathrm{H}^1(\Gamma_{k_w}/I_w, A(k_w^{\mathrm{nr}}))$. Let $\mathcal{A} \to \mathrm{Spec}(\mathcal{O}_w)$ be the Néron model of $A \times_k k_w$. By [Milne 1986, Proposition I.3.8] we have an isomorphism

$$\mathrm{H}^1(\Gamma_{k_w}/I_w, A(k_w^{\mathrm{nr}})) = \mathrm{H}^1(\Gamma_{k_w}/I_w, \pi_0(\mathcal{A} \times_{\mathcal{O}_w} \mathbb{F}_w)),$$

where $\pi_0(\mathcal{A} \times_{\mathcal{O}_w} \mathbb{F}_w)$ is the group of connected components of the special fibre of $\mathcal{A} \to \mathrm{Spec}(\mathcal{O}_w)$. Since $2[Y] = 0$, condition (f) implies that $[Y \times_k k_w] = 0$, hence Y has a k_w-point $P_w \in Y(k_w)$. We view P_w as a point $(P_w, 1) \in \mathcal{Y}$ above $1 \in \mathbb{G}_{m,k}(k) \subset \mathbb{P}^1_k$.

For each place v of k and for each point $Q_v \in X(k_v)$, there exists an extension F_v/k_v of degree at most 2 such that Q_v lifts to $\tilde{Y}^{F_v}(k_v)$. Since X is everywhere locally soluble, we can use this observation to extend the collection of local points $(P_w, 1)$, $w \in \{w_1, \dots, w_r\}$, to an adelic point $(P_v) \in \mathcal{Y}(\mathbb{A}_k) \subseteq \mathcal{X}(\mathbb{A}_k)$. The fibration $\mathcal{X} \to \mathbb{P}^1_k$ has only two bad fibres at 0 and ∞ (both of which are geometrically split).

By Proposition 6.1 the vertical Brauer group of $\mathcal{X}$ over $\mathbb{P}^1_k$ is generated by the image of $\mathrm{Br}(k)$; therefore the desired result can now be obtained by applying the fibration method. More precisely, one proceeds as in the proof of [Colliot-Thélène and Skorobogatov 2000, Theorem A]. (As a more recent reference one can apply [Harpaz and Wittenberg 2016, Theorem 9.17] with $B = 0$ and $U = \mathbb{G}_{m,k}$, which is justified in the light of [Harpaz and Wittenberg 2016, Theorem 9.11].) We obtain that there exists an adelic point $(P'_v) \in \mathcal{X}(\mathbb{A}_k)$ arbitrarily close to (P_v) such that the image of (P'_v) in $\mathbb{P}^1_k(\mathbb{A}_k)$ is a k-point. Let us call it a. By the construction of (P_v) we can assume that $a \in \mathbb{G}_{m,k}(k)$ and that a is arbitrarily close to 1 in the w-adic topology for $w \in \{w_1, \dots, w_r\}$. The quadratic extension $F = k(\sqrt{a})$ now satisfies the desired properties. □

7. Proof of Theorem 2.3

Suppose that our Kummer variety is $X = \mathrm{Kum}(Y)$, where Y is the k-torsor for A defined by a class $\alpha \in \mathrm{H}^1(k, A[2])$. To prove the existence of a k-point on X, it is enough to find a quadratic (or trivial) extension F of k such that α goes to 0 in $\mathrm{H}^1(k, A^F)$. We write $\alpha = \sum_{i=1}^r \alpha_i$, where $\alpha_i \in \mathrm{H}^1(k, A_i[2])$ is nonzero for each $i = 1, \dots, r$. Let $K_i = K(A_i[2])$. For each $i = 1, \dots, r$, we fix $g_i, h_i \in \mathrm{Gal}(K_i/k)$ satisfying conditions (c) and (d), respectively, for A_i.

By Proposition 6.2 there is a quadratic extension F of k split at $w_1, \dots, w_r$ such that $\alpha \in \mathrm{Sel}_2(A^F)$. Replacing A with A^F, we can assume without loss of generality that $\alpha \in \mathrm{Sel}_2(A)$. By doing so we preserve conditions (a), (b), (c), (d), (e) and (g), which are not affected by quadratic twisting. The extension F/k is split at $w_1, \dots, w_r$, so replacing A by A^F also preserves condition (f) for each A_i.

Let S_0 be the set of places of k that contains all the Archimedean places and the places above 2.

Lemma 7.1. *Let S be the set of places of k which is the union of S_0 and all the places of bad reduction of A excluding $w_1, \dots, w_r$. For each $i = 1, \dots, r$, let $\alpha_i \in \mathrm{Sel}_2(A_i)$ be nonzero. Let $\beta \in \mathrm{Sel}_2(A_1)$ be a nonzero class such that $\beta \neq \alpha_1$. Then there exists a $q \in k^*$ such that $\mathfrak{q} = (q)$ is a prime ideal of k with the following properties*:

(1) *all the places in S (including the Archimedean places) are split in $F = k(\sqrt{q})$; in particular, $\mathfrak{q} \notin S$;*

(2) *A has good reduction at $\mathfrak{q}$;*

(3) $\mathrm{Frob}_{\mathfrak{q}}$ *acts on $A_1[2]$ as g_1;*

(4) $\mathrm{Frob}_{\mathfrak{q}}$ *acts on $A_i[2]$ as h_i, for each $i \neq 1$;*

(5) $\mathrm{loc}_{\mathfrak{q}}(\alpha_1) = 0$, *but* $\mathrm{loc}_{\mathfrak{q}}(\beta) \neq 0$.

Proof. We adapt the arguments from the proof of Proposition 5.1 in [Mazur and Rubin 2010].

Let $M = A_1[2]^{\oplus 2}$ be the direct sum of two copies of $A_1[2]$. Let

$$\alpha = \alpha_1 + \beta \ \in \ \mathrm{H}^1(k, A_1[2]) \oplus \mathrm{H}^1(k, A_1[2]) = \mathrm{H}^1(k, M).$$

The splitting field of M is K_1 and the Galois action on M factors through $G_1 = \mathrm{Gal}(K_1/k)$. Let $(K_1)_\alpha$ and $(G_1)_\alpha$ be as in Definition 3.3. By Corollary 3.8 we can find a lift $\gamma \in (G_1)_\alpha$ of $g \in G_1$ such that the associated map

$$f_\gamma : \mathrm{H}^1(k, M) \longrightarrow M/(g-1) = (A_1[2]/(g-1)) \oplus (A_1[2]/(g-1)) = \mathbb{Z}/2 \oplus \mathbb{Z}/2$$

sends α to the class $(0, 1) \in \mathbb{Z}/2 \oplus \mathbb{Z}/2$. That is, $f_\gamma(\alpha_1, 0) = 0$, whereas $f_\gamma(0, \beta) \neq 0$.

The fields $(K_1)_\alpha, K_2, \dots, K_r$ are Galois extensions of k that are linearly disjoint by condition (e) and Proposition 3.12. Let $\mathcal{K}$ be the compositum of $(K_1)_\alpha, K_2, \dots, K_r$. This is a Galois extension of k with the Galois group $\mathrm{Gal}(\mathcal{K}/k) = (G_1)_\alpha \times \prod_{i=2}^r G_i$.

Let the modulus $\mathfrak{m}$ be the formal product of the real places of k, 8 and all the odd primes in S. Let $k_\mathfrak{m}$ be the ray class field associated to the modulus $\mathfrak{m}$. This is an abelian extension of k which is unramified away from $\mathfrak{m}$. We claim that $k_\mathfrak{m}$ and $\mathcal{K}$ are linearly disjoint over k. Indeed, $k' = k_\mathfrak{m} \cap \mathcal{K}$ is a subfield of $\mathcal{K}$ that is abelian over k and unramified at $w_1, \dots, w_r$. We note that $\mathrm{Gal}(\mathcal{K}/k)^{\mathrm{ab}} = (G_1)_\alpha^{\mathrm{ab}} \times \prod_{i=2}^r G_i^{\mathrm{ab}}$. By Corollary 3.9 (applicable in the light of Remark 3.10) we have $(G_1)_\alpha^{\mathrm{ab}} = (G_1)^{\mathrm{ab}}$. Therefore, $\mathrm{Gal}(\mathcal{K}/k)^{\mathrm{ab}} = \prod_{i=1}^r G_i^{\mathrm{ab}}$, so that k' is contained in the compositum $L = k_1^{\mathrm{ab}} \cdots k_r^{\mathrm{ab}}$ of linearly disjoint abelian extensions $k_1^{\mathrm{ab}}, \dots, k_r^{\mathrm{ab}}$, where, as in Section 2, k_i^{ab} denotes the maximal abelian subextension of K_i/k.

Write $M = k_1^{\mathrm{ab}} \cdots k_{r-1}^{\mathrm{ab}}$. The extension k_r^{ab}/k is totally ramified at w_r by condition (g), whereas k'/k and M/k are unramified at w_r (the latter by the criterion of Néron–Ogg–Shafarevich). Hence L/M is totally ramified at each prime v of M over w_r. Since $M \subset k'M \subset L$, where $k'M/M$ is unramified over v, we must have $k' \subset M$. Continuing by induction, we prove that $k' = k$, as required.

It follows that $k_\mathfrak{m}\mathcal{K}$ is a Galois extension of k with the Galois group

$$\mathrm{Gal}(k_\mathfrak{m}\mathcal{K}/k) = \mathrm{Gal}(k_\mathfrak{m}/k) \times (G_1)_\alpha \times \prod_{i=2}^r G_i.$$

By Chebotarev's density theorem we can find a place $\mathfrak{q}$ of k such that the corresponding Frobenius element in $\mathrm{Gal}(k_\mathfrak{m}\mathcal{K}/k)$ is the conjugacy class of $(1, \gamma, h_2, \dots, h_r)$. Then $\mathfrak{q}$ is a principal prime ideal with a totally positive generator $q \equiv 1 \mod 8$, hence q is a square in each completion of k at a prime over 2. We also have $q \equiv 1 \mod \mathfrak{p}$ for any odd $\mathfrak{p} \in S$. Thus all the places of S including the Archimedean places are split in $F = k(\sqrt{q})$. All other conditions are satisfied by construction. $\square$

Proposition 7.2. *For any* $\beta \in \mathrm{Sel}_2(A_1)$, $\beta \neq 0$, $\beta \neq \alpha_1$, *there exists a quadratic extension* F/k, *unramified at the places of* S_0 *and all the places of bad reduction of* A, *such that*

$$\mathrm{Sel}_2(A_1^F) \subset \mathrm{Sel}_2(A_1),$$
$$\alpha_1 \in \mathrm{Sel}_2(A_1^F),$$
$$\beta \notin \mathrm{Sel}_2(A_1^F),$$
$$\mathrm{Sel}_2(A_i^F) = \mathrm{Sel}_2(A_i) \quad \textit{for } i \neq 1.$$

Proof. Let $F = k(\sqrt{q})$ be as in Lemma 7.1. Let $i \in \{1, \ldots, r\}$. Since F is split at each $v \in S$, we have $A_i^F \times_k k_v \cong A_i \times_k k_v$, so that the Selmer conditions at S are identical for A_i and A_i^F. These conditions are also identical for all primes where both A_i and A_i^F have good reduction, and this includes the primes w_j if $j \neq i$. At w_i the extension F/k is unramified, and by condition (f) we can apply Lemma 4.2, so we obtain $\delta(A_i(k_{w_i})) = \delta^F(A_i^F(k_{w_i}))$.

It remains to check the behaviour at $\mathfrak{q}$, which is a prime of good reduction for A_i. If $i \neq 1$ then $\mathrm{Frob}_{\mathfrak{q}} = h$, and from condition (d) and formula (7) we deduce $\mathrm{H}^1(k_{\mathfrak{q}}, A_i[2]) = 0$, so the Selmer conditions for both A_i and A_i^F at $\mathfrak{q}$ are vacuous. This proves that $\mathrm{Sel}_2(A_i^F) = \mathrm{Sel}_2(A_i)$ whenever $i \neq 1$.

In the rest of the proof we work with A_1. The Selmer conditions for A_1 and A_1^F are the same at each place $v \neq \mathfrak{q}$. Thus $\mathrm{loc}_{\mathfrak{q}}(\alpha_1) = 0$ implies $\alpha_1 \in \mathrm{Sel}_2(A_1^F)$. Moreover, $\delta(A_1(k_{\mathfrak{q}})) \cap \delta^F(A_1^F(k_{\mathfrak{q}})) = 0$ by Lemma 4.3. By property (5) in Lemma 7.1 we have $\mathrm{loc}_{\mathfrak{q}}(\beta) \neq 0$, so we conclude that $\beta \notin \mathrm{Sel}_2(A_1^F)$.

To prove that $\mathrm{Sel}_2(A_1^F) \subset \mathrm{Sel}_2(A_1)$, it is enough to show that for A_1 the Selmer condition at $\mathfrak{q}$ is implied by the Selmer conditions at the other places of k. Indeed, let $\xi \in \mathrm{H}^1(k, A_1[2])$ be an element satisfying the Selmer condition at each place $v \neq \mathfrak{q}$, but not necessarily at $\mathfrak{q}$. By global reciprocity the sum of $\mathrm{inv}_v(\beta \cup \xi) \in \frac{1}{2}\mathbb{Z}/\mathbb{Z}$ over all places of k, including the Archimedean places, is 0. Since the images of ξ and β in $\mathrm{H}^1(k_v, A_1[2])$ belong to $\delta(A_1(k_v))$ for all $v \neq \mathfrak{q}$, we obtain $\mathrm{inv}_v(\beta \cup \xi) = 0$. By the global reciprocity we deduce $\mathrm{inv}_{\mathfrak{q}}(\beta \cup \xi) = 0$. The nonzero element $\mathrm{loc}_{\mathfrak{q}}(\beta)$ generates $\delta(A_1(k_{\mathfrak{q}}))$, because

$$\delta(A_1(k_{\mathfrak{q}})) = A_1[2]/(\mathrm{Frob}_q - 1) = A_1[2]/(g-1) = \mathbb{Z}/2,$$

where we used (6) and the fact that $\mathrm{Frob}_{\mathfrak{q}}$ acts on $A_1[2]$ as the element g of condition (c). Since A_1 is principally polarised, $\delta(A_1(k_{\mathfrak{q}}))$ is a maximal isotropic subspace of $\mathrm{H}^1(k_{\mathfrak{q}}, A_1[2])$ (see the beginning of Section 4). Therefore, $\mathrm{inv}_{\mathfrak{q}}(\beta \cup \xi) = 0$ implies that the image of ξ in $\mathrm{H}^1(k_{\mathfrak{q}}, A_1[2])$ lies in $\delta(A_1(k_{\mathfrak{q}}))$. □

End of proof of Theorem 2.3. The extension F/k is unramified at all the places where A has bad reduction, so replacing A by A^F preserves condition (f) for each A_i. Conditions (a), (b), (c), (d), (e) and (g) are not affected by quadratic twisting. By repeated applications of Proposition 7.2 we can find a quadratic extension F/k such that α_i is the only nonzero element in $\mathrm{Sel}_2(A_i^F)$, for all $i = 1, \dots, r$. The exact sequence (10) for A_i^F shows that $Ш(A_i^F)[2]$ is of size at most 2. If the 2-primary subgroup of $Ш(A_i^F)$ is finite, then, by Proposition 5.2, the number of elements in $Ш(A_i^F)[2]$ is a square. Thus $Ш(A_i^F)[2] = 0$, so that the image of α_i in $\mathrm{H}^1(k, A_i^F)$ is 0. Then the image of α in $\mathrm{H}^1(k, A^F)$ is 0, so that $Y^F \cong A^F$ and hence $Y^F(k) \neq \varnothing$. This implies that $\tilde{Y}^F(k) \neq \varnothing$ and hence $X = \tilde{Y}/\iota_{\tilde{Y}}$ has a k-point.

It remains to prove that k-points are Zariski dense in X. Since $Y^F(k) \neq \varnothing$, we have $Y^F \simeq A^F$, so we may identify X with $\mathrm{Kum}(A^F)$. Hence it will suffice to show that $A^F(k)$ is Zariski dense in A^F. For each i, the exact sequence (10) for A_i^F shows that $A_i^F(k)/2 \neq 0$. Since $A_i^F[2](k) = 0$ by condition (a), we see that $A_i^F(k)$ is infinite. The neutral connected component of the Zariski closure of $A_i^F(k)$ in A_i^F is an abelian subvariety $B \subset A_i^F$ of positive dimension. By condition (a) we must have $B = A_i^F$. Thus the set $A_i^F(k)$ is Zariski dense in A_i^F for each $i = 1, \dots, r$, so that $A^F(k)$ is Zariski dense in A^F. □

Proof of Proposition 1.1. For each $i = 1, \dots, n$, we have the exact sequence

$$0 \longrightarrow E_i(\mathbb{Q})/2 \longrightarrow \mathrm{H}^1(\mathbb{Q}, E_i[2]) \longrightarrow \mathrm{H}^1(\mathbb{Q}, E_i)[2] \longrightarrow 0.$$

By assumption there is a class $\alpha_i \in \mathrm{H}^1(\mathbb{Q}, E_i[2])$ that goes to the class of the torsor Y_i in $\mathrm{H}^1(\mathbb{Q}, E_i)[2]$. The restriction of α_i to $\mathrm{H}^1(\mathbb{R}, E_i[2])$ is nonzero, hence $\alpha_i \neq 0$. Recall from Section 6 that the fixed-point set of the antipodal involution ι_Y on $Y = \prod_{i=1}^n Y_i$ is $Z = \prod_{i=1}^n Z_i$, where Z_i is a torsor for $E_i[2]$ defined by α_i.

Consider the double covering of smooth projective varieties $\pi : \tilde{Y} \to X = \tilde{Y}/\iota_{\tilde{Y}}$ whose branch locus is $Z \times_k \mathbb{P}_k^{n-1} \subset X$. Let $V \subset X$ be the complement to the branch locus, and let $U = \pi^{-1}(V)$. Then $\pi : U \to V$ is a torsor with the structure group μ_2.

We need to show that a real point $M \in X(\mathbb{R})$ path-connected with a rational point $P \in X(\mathbb{Q})$ can be approximated by a point in $X(\mathbb{Q})$. In our assumptions we have $Z(\mathbb{R}) = \varnothing$, hence $P \in V(\mathbb{Q})$, $M \in V(\mathbb{R})$ and the path connecting P and M is contained in $V(\mathbb{R})$. There exists a unique extension F of $\mathbb{Q}$ of degree $[F : \mathbb{Q}] \leq 2$ such that P lifts to a $\mathbb{Q}$-point $\tilde{P}$ on the quadratic twist U^F. Moreover, M lifts to an $\mathbb{R}$-point $\tilde{M}$ in U^F which is path connected with $\tilde{P}$.

We note that U^F is naturally a subset of the quadratic twist $Y^F = \prod_{i=1}^n Y_i^F$. Recall from the introduction that each Y_i^F is a torsor for E_i^F defined by the image of α_i under the map

$$\mathrm{H}^1(\mathbb{Q}, E_i[2]) = \mathrm{H}^1(\mathbb{Q}, E_i^F[2]) \longrightarrow \mathrm{H}^1(\mathbb{Q}, E_i^F).$$

Now $U^F(\mathbb{Q}) \neq \varnothing$ implies $Y^F(\mathbb{Q}) \neq \varnothing$, thus Y^F is a trivial torsor, i.e., $Y^F \cong \prod_{i=1}^n E_i^F$. Therefore, α_i goes to zero in $\mathrm{H}^1(\mathbb{Q}, E_i^F)$, so α_i is a nonzero element of $E_i^F(\mathbb{Q})/2$. Thus E_i^F has a $\mathbb{Q}$-point not divisible by 2. By assumption $E_i^F[2](\mathbb{Q}) = E_i[2](\mathbb{Q}) = 0$, so this point has infinite order in E_i^F. It follows that $\mathbb{Q}$-points of E_i^F are dense in the neutral connected component of $E_i^F(\mathbb{R})$ for each $i = 1, \dots, n$. But we have $Y^F \cong \prod_{i=1}^n E_i^F$, so $\mathbb{Q}$-points are dense in the connected component of $Y^F(\mathbb{R})$ containing $\tilde{P}$ and $\tilde{M}$. Hence we can find a $\mathbb{Q}$-point on $X = \mathrm{Kum}(Y) = \mathrm{Kum}(Y^F)$ which is as close as we wish to M. □

Acknowledgements

While working on this paper Harpaz was supported by the Fondation Sciences Mathématiques de Paris. The paper was finalised when Skorobogatov was at the Institute for Advanced Study in Princeton where he was supported by the Charles Simonyi Endowment. We are grateful to Jean-Louis Colliot-Thélène, Tim Dokchitser, Evis Ieronymou, and René Pannekoek for helpful discussions and to the referee for a careful reading and detailed comments.

References

[Bhargava et al. 2014] M. Bhargava, C. Skinner, and W. Zhang, "A majority of elliptic curves over $\mathbb{Q}$ satisfy the Birch and Swinnerton-Dyer conjecture", preprint, 2014. arXiv 1407.1826

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Results in Mathematics and Related Areas (3) **21**, Springer, Berlin, 1990. MR 1045822 Zbl 0705.14001

[Bright et al. 2016] M. Bright, T. D. Browning, and D. Loughran, "Failures of weak approximation in families", *Compos. Math.* (online publication April 2016), 1–41.

[Burnside 1911] W. Burnside, *Theory of groups of finite order*, 2nd ed., Cambridge University Press, 1911. MR 0069818 Zbl 42.0151.02

[Cohen 1979] S. D. Cohen, "The distribution of the Galois groups of integral polynomials", *Illinois J. Math.* **23**:1 (1979), 135–152. MR 516576 Zbl 0402.12005

[Colliot-Thélène and Skorobogatov 2000] J.-L. Colliot-Thélène and A. N. Skorobogatov, "Descent on fibrations over $\mathbf{P}_k^1$ revisited", *Math. Proc. Cambridge Philos. Soc.* **128**:3 (2000), 383–393. MR 1744112 Zbl 1024.14003

[Colliot-Thélène et al. 1998] J.-L. Colliot-Thélène, A. N. Skorobogatov, and P. Swinnerton-Dyer, "Hasse principle for pencils of curves of genus one whose Jacobians have rational 2-division points", *Invent. Math.* **134**:3 (1998), 579–650. MR 1660925 Zbl 0924.14011

[Dokchitser and Dokchitser 2011] T. Dokchitser and V. Dokchitser, "Root numbers and parity of ranks of elliptic curves", *J. Reine Angew. Math.* **658** (2011), 39–64. MR 2831512 Zbl 1314.11041

[Dolgachev 2012] I. V. Dolgachev, *Classical algebraic geometry: A modern view*, Cambridge University Press, 2012. MR 2964027 Zbl 1252.14001

[Gonzalez-Dorrego 1994] M. R. Gonzalez-Dorrego, "(16, 6) configurations and geometry of Kummer surfaces in $\mathbf{P}^3$", *Mem. Amer. Math. Soc.* **107**:512 (1994), vi+101. MR 1182682 Zbl 0809.14032

[Hall 2011] C. Hall, "An open-image theorem for a general class of abelian varieties", *Bull. Lond. Math. Soc.* **43**:4 (2011), 703–711. With an appendix by E. Kowalski. MR 2820155 Zbl 1225.11083

[Harpaz and Wittenberg 2016] Y. Harpaz and O. Wittenberg, "On the fibration method for zero-cycles and rational points", *Ann. of Math.* (2) **183**:1 (2016), 229–295. MR 3432584 Zbl 06541586

[Holmes and Pannekoek 2015] D. Holmes and R. Pannekoek, "The Brauer–Manin obstruction on Kummer varieties and ranks of twists of abelian varieties", *Bull. Lond. Math. Soc.* **47**:4 (2015), 565–574. MR 3375924 Zbl 1326.11032

[Klagsbrun 2012a] Z. Klagsbrun, "Elliptic curves with a lower bound on 2-Selmer ranks of quadratic twists", *Math. Res. Lett.* **19**:5 (2012), 1137–1143. MR 3039836 Zbl 1285.11085

[Klagsbrun 2012b] Z. Klagsbrun, "Selmer ranks of quadratic twists of elliptic curves with partial rational two-torsion", preprint, 2012. arXiv 1201.5408

[Kramer 1981] K. Kramer, "Arithmetic of elliptic curves upon quadratic extension", *Trans. Amer. Math. Soc.* **264**:1 (1981), 121–135. MR 597871 Zbl 0471.14020

[Kuwata and Wang 1993] M. Kuwata and L. Wang, "Topology of rational points on isotrivial elliptic surfaces", *Internat. Math. Res. Notices* 4 (1993), 113–123. MR 1214702 Zbl 0804.14008

[Liu 1996] Q. Liu, "Modèles entiers des courbes hyperelliptiques sur un corps de valuation discrète", *Trans. Amer. Math. Soc.* **348**:11 (1996), 4577–4610. MR 1363944 Zbl 0926.11043

[Mazur 1972] B. Mazur, "Rational points of abelian varieties with values in towers of number fields", *Invent. Math.* **18** (1972), 183–266. MR 0444670 Zbl 0245.14015

[Mazur 1992] B. Mazur, "The topology of rational points", *Experiment. Math.* **1**:1 (1992), 35–45. MR 1181085 Zbl 0784.14012

[Mazur 1995] B. Mazur, "Speculations about the topology of rational points: an update", pp. 165–182 in *Columbia University Number Theory Seminar* (New York, 1992), Astérisque **228**, Société Mathématique de France, Paris, 1995. MR 1330932 Zbl 0851.14009

[Mazur and Rubin 2007] B. Mazur and K. Rubin, "Finding large Selmer rank via an arithmetic theory of local constants", *Ann. of Math.* (2) **166**:2 (2007), 579–612. MR 2373150 Zbl 1219.11084

[Mazur and Rubin 2010] B. Mazur and K. Rubin, "Ranks of twists of elliptic curves and Hilbert's tenth problem", *Invent. Math.* **181**:3 (2010), 541–575. MR 2660452 Zbl 1227.11075

[Milne 1986] J. S. Milne, *Arithmetic duality theorems*, Perspectives in Mathematics **1**, Academic Press, Boston, 1986. MR 881804 Zbl 0613.14019

[Polishchuk 2003] A. Polishchuk, *Abelian varieties, theta functions and the Fourier transform*, Cambridge Tracts in Mathematics **153**, Cambridge University Press, 2003. MR 1987784 Zbl 1018.14016

[Poonen and Rains 2011] B. Poonen and E. Rains, "Self cup products and the theta characteristic torsor", *Math. Res. Lett.* **18**:6 (2011), 1305–1318. MR 2915483 Zbl 1297.18005

[Poonen and Rains 2012] B. Poonen and E. Rains, "Random maximal isotropic subspaces and Selmer groups", *J. Amer. Math. Soc.* **25**:1 (2012), 245–269. MR 2833483 Zbl 1294.11097

[Poonen and Stoll 1999] B. Poonen and M. Stoll, "The Cassels–Tate pairing on polarized abelian varieties", *Ann. of Math.* (2) **150**:3 (1999), 1109–1149. MR 1740984 Zbl 1024.11040

[Sah 1977] C. H. Sah, "Cohomology of split group extensions, II", *J. Algebra* **45**:1 (1977), 17–68. MR 0463264 Zbl 0352.18021

[Skorobogatov 2001] A. Skorobogatov, *Torsors and rational points*, Cambridge Tracts in Mathematics **144**, Cambridge University Press, 2001. MR 1845760 Zbl 0972.14015

[Skorobogatov 2009] A. Skorobogatov, "Diagonal quartic surfaces", (2009). Abstract from Belabas et al., "Explicit methods in number theory", *Oberwolfach Rep.* **6**:3 (2009), 1843–1920. MR 2664991 Zbl 1192.00048

[Skorobogatov 2010] A. Skorobogatov, "del Pezzo surfaces of degree 4 and their relation to Kummer surfaces", *Enseign. Math.* (2) **56**:1-2 (2010), 73–85. MR 2674855 Zbl 1198.14035

[Skorobogatov and Swinnerton-Dyer 2005] A. Skorobogatov and P. Swinnerton-Dyer, "2-descent on elliptic curves and rational points on certain Kummer surfaces", *Adv. Math.* **198**:2 (2005), 448–483. MR 2183385 Zbl 1085.14021

[Skorobogatov and Zarhin 2008] A. N. Skorobogatov and Y. G. Zarhin, "A finiteness theorem for the Brauer group of abelian varieties and $K3$ surfaces", *J. Algebraic Geom.* **17**:3 (2008), 481–502. MR 2395136 Zbl 1157.14008

[Skorobogatov and Zarhin 2012] A. N. Skorobogatov and Y. G. Zarhin, "The Brauer group of Kummer surfaces and torsion of elliptic curves", *J. Reine Angew. Math.* **666** (2012), 115–140. MR 2920883 Zbl 1256.14036

[Swinnerton-Dyer 2000] P. Swinnerton-Dyer, "Arithmetic of diagonal quartic surfaces, II", *Proc. London Math. Soc.* (3) **80**:3 (2000), 513–544. MR 1744774 Zbl 1066.11029

[Swinnerton-Dyer 2001] P. Swinnerton-Dyer, "The solubility of diagonal cubic surfaces", *Ann. Sci. École Norm. Sup.* (4) **34**:6 (2001), 891–912. MR 1872424 Zbl 1003.11028

[Wisbauer 1991] R. Wisbauer, *Foundations of module and ring theory*, Algebra, Logic and Applications: A handbook for study and research **3**, Gordon and Breach Science Publishers, Philadelphia, PA, 1991. MR 1144522 Zbl 0746.16001

[Wittenberg 2007] O. Wittenberg, *Intersections de deux quadriques et pinceaux de courbes de genre 1/Intersections of two quadrics and pencils of curves of genus 1*, Lecture Notes in Mathematics **1901**, Springer, Berlin, 2007. MR 2307807 Zbl 1122.14001

Communicated by Jean-Louis Colliot-Thélène
Received 2015-05-08 Revised 2016-02-08 Accepted 2016-03-12

harpaz@dma.ens.fr *Département de Mathématiques et Applications, École Normale Supérieure, 45 rue d'Ulm, Paris 75005, France*

a.skorobogatov@imperial.ac.uk *Department of Mathematics, Imperial College, South Kensington Campus, London, SW7 2BZ, United Kingdom*

Institute for Information Transmission Problems, Russian Academy of Sciences, 19 Bolshoi Karetnyi, Moscow, 127994, Russia

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.843

Analytic continuation on Shimura varieties with μ-ordinary locus

Stéphane Bijakowski

We study the geometry of unitary Shimura varieties without assuming the existence of an ordinary locus. We prove, by a simple argument, the existence of canonical subgroups on a strict neighborhood of the μ-ordinary locus (with an explicit bound). We then define the overconvergent modular forms (of classical weight) as well as the relevant Hecke operators. Finally, we show how an analytic continuation argument can be adapted to this case to prove a classicality theorem, namely that an overconvergent modular form which is an eigenform for the Hecke operators is classical under certain assumptions.

Introduction

A modular form is defined as a global section of a certain sheaf on the modular curve. To study congruences between modular forms, one is led to introduce new objects, namely p-adic and overconvergent modular forms. These are sections of a sheaf on the ordinary locus of the modular curve and on a strict neighborhood of the ordinary locus, respectively. A lot of work has been done using these objects; one can for example construct families of overconvergent modular forms.

It is possible to generalize the definition of p-adic and overconvergent modular forms to other varieties, for example the Hilbert modular variety or the Siegel variety. The natural definition of an overconvergent modular form is a section of a certain sheaf on a strict neighborhood of the ordinary locus. In greater generality, one can consider Shimura varieties with a nonempty ordinary locus.

MSC2010: primary 11G18; secondary 11F55, 14G35.
Keywords: Shimura variety, overconvergent modular forms, μ-ordinary locus, canonical subgroups.

But for Shimura varieties without ordinary locus, this definition fails. Recall that, for Shimura varieties of PEL type, the criterion for the existence of an ordinary locus at p is that p splits completely in its reflex field. For Shimura varieties of PEL type (C) (associated to symplectic groups), the reflex field is $\mathbb{Q}$, so there is always an ordinary locus. But, if one considers Shimura varieties of type (A) (associated to unitary groups), then the reflex field may not be $\mathbb{Q}$ and the ordinary locus may be empty.

Let us look at an example. Let F be a CM-field with totally real subfield F_0, and consider the special fiber of the Shimura variety associated to a unitary group with signature (a_σ, b_σ) at each real place σ of F_0. Suppose for simplicity that p is a prime number inert in F_0 that splits as $\pi^+\pi^-$ in F. The choice of π^+ gives an order for the elements of the pair (a_σ, b_σ) for each real place σ of F_0. The existence of the ordinary locus at p is then equivalent to the fact that there exist integers a and b such that $a_\sigma = a$ and $b_\sigma = b$ for each real place σ. The structure of the p-torsion of the abelian variety is then well known on the ordinary locus. In our setting, the p-divisible group $A[p^\infty]$ splits as $A[(\pi^+)^\infty] \oplus A[(\pi^-)^\infty]$, and $A[(\pi^-)^\infty]$ is the dual of $A[(\pi^+)^\infty]$. On the ordinary locus, the p-divisible group $A[(\pi^+)^\infty]$ is an extension of a multiplicative part of height da and an étale part of height db (d is the degree of F_0 over $\mathbb{Q}$). In particular, there exists a multiplicative subgroup $H_a \subset A[\pi^+]$ of height da. Actually, this property characterizes the ordinary locus: if there exists a multiplicative subgroup H_a of height ha, then the abelian variety is ordinary at p. On the rigid space associated to the Shimura variety (i.e., the generic fiber of its formal completion along its special fiber), one can define the ordinary locus. There is still a multiplicative subgroup of $A[\pi^+]$ of height da on this locus, and work of Fargues [2011] shows that this subgroup extends to a canonical subgroup on a strict neighborhood of the ordinary locus.

If the (a_σ) are not equal to a certain integer a, then the ordinary locus is empty. There is always a special locus in the special fiber of the Shimura variety, called the μ-ordinary locus, but the situation is more involved. Suppose we are in the same setting as before, and let us order the elements $a_1 \leq a_2 \leq \cdots \leq a_d$. Then the μ-ordinary locus is characterized by the fact that the p-divisible group $A[(\pi^+)^\infty]$ has a filtration $0 \subset X_1 \subset X_2 \subset \cdots \subset X_{d+1} = A[(\pi^+)^\infty]$ such that X_{i+1}/X_i is a p-divisible group of height $d(a_{i+1} - a_i)$ with its structure explicitly described (by convention we set $a_0 = 0$ and $a_{d+1} = a_d + b_d$). The fact that A is μ-ordinary is then also equivalent to the existence of subgroups $0 \subset H_{a_1} \subset \cdots \subset H_{a_d} \subset A[\pi^+]$, with H_{a_i} of height da_i and the structure of $H_{a_{i+1}}/H_{a_i}$ explicitly described.

If one looks at the rigid space of the Shimura variety, then one can define the μ-ordinary locus. There are several canonical subgroups in $A[\pi^+]$ on this locus. We lack a good theory for these canonical subgroups, which should be analogous to the one given by Fargues [2011]. However, by simple arguments, one can prove the following fact:

Theorem. *On a strict neighborhood of the μ-ordinary locus, there exist canonical subgroups $H_{a_1} \subset \cdots \subset H_{a_d}$ in $A[\pi^+]$. These subgroups are characterized by the fact that their degree* (in the sense of [Fargues 2010]) *is maximal among the subgroups of the same height of $A[\pi^+]$.*

The proof is actually very simple: Let us consider X_{Iw}, the variety with Iwahori level at p, and let $f : X_{\mathrm{Iw}} \to X$ be the projection. The μ-ordinary locus is the image under f of the locus where the subgroups H_{a_i} are of maximal degree. Let $0 < \varepsilon < \frac{1}{2}$, and consider the locus in X_{Iw} where the degree of H_{a_i} is bigger than the maximal degree minus ε. The image under f of this locus is a strict neighborhood of the μ-ordinary locus. The existence of canonical subgroups follows from the definition. Their uniqueness is a simple computation using the properties of the degree function (see Proposition 1.24).

If we consider the space X_{Iw}, then one can call the locus where the degree of each H_{a_i} is maximal the μ-ordinary-multiplicative locus. It then makes sense to define an overconvergent modular form as the section of a certain sheaf on a strict neighborhood of the μ-ordinary-multiplicative locus.

The Hecke algebra at p acts both on the rigid space and on the space of modular forms. In the case of the existence of the ordinary locus, there is one relevant Hecke operator, parametrizing complements of the canonical subgroup. In the general case, there will be as many relevant Hecke operators as the number of canonical subgroups. We will denote by U_{p,a_i} these Hecke operators. One can show that these operators increase the degrees of all the subgroups of $A[\pi^+]$, then act on the space of overconvergent modular forms.

We can now state the main result of the paper, namely that an overconvergent modular form, which is an eigenform for the Hecke operators U_{p,a_i}, can be analytically continued to the whole variety under a certain assumption, and thus is classical. Let κ be a weight; explicitly, it is a collection of integers $(\kappa_{i,1} \geq \cdots \geq \kappa_{i,a_i}, \lambda_{i,a} \geq \cdots \geq \lambda_{i,b_i})_{1 \leq i \leq d}$. Let $S = \{a_1, \ldots, a_d\} \cap [1, a_d + b_d - 1]$. The cardinality of S is exactly the number of canonical subgroups. Let us write $\Sigma_i = \{j : a_j = a_i\}$ for all i.

Theorem 3.17. *Let f be an overconvergent modular form of weight κ. Suppose that f is an eigenform for the Hecke operators U_{p,a_i}, with eigenvalue α_i for $a_i \in S$, and that we have the relations*

$$n_i + v(\alpha_i) < \inf_{j \in \Sigma_i} (\kappa_{j,a_j} + \lambda_{j,b_j}).$$

Then f is classical.

Here n_i is a constant depending on the variety. It actually comes from the normalization factor of the Hecke operator U_{p,a_i}. This theorem is a classicality result, analogous to the one proved by Coleman [1996] in the case of the modular curve. Actually, this result has also been proven by Buzzard [2003] and Kassaei [2006] using an analytic continuation method, from which we take inspiration here. Note that there has been extensive work on the classicality problem in the case of the existence of the ordinary locus. We can cite the work of Sasaki [2010], Johansson [2013], Tian and Xiao [2013] and Pilloni and Stroh [2011] in the case of Hilbert varieties, and the work of the author with Pilloni and Stroh [Bijakowski et al. 2016] for more general PEL Shimura varieties (see also [Bijakowski 2014] for Shimura varieties with ramification).

In this introduction, we have assumed that p is inert in F_0 and splits in F. The group associated to the Shimura variety is then a linear group at p. If p is inert in F, then the group is a unitary group at p. Everything we have said adapts to that context: the description of the μ-ordinary locus, the existence of the canonical subgroups, and the analytic continuation theorem. Note that the geometry is more involved in that case; for example, to define Hecke operators, one has to deal with subgroups of $A[p^2]$.

Of course, the assumption that p is inert in F_0 is for simplicity, so what we have said can be formulated for any prime p unramified in F. In the writing of the paper, we have tried to formulate propositions valid in both the linear and unitary cases as much as possible, but of course we often had to treat the proofs separately.

Let us now talk briefly about the text. In Section 1, we introduce the varieties we are dealing with, define the μ-ordinary locus and study the canonical subgroups. In Section 2 we define the classical and overconvergent modular, as well as the Hecke operators. In Section 3 we prove the analytic continuation result. For simplicity, we suppose that the prime p is inert in Sections 2 and 3, and Section 4 briefly shows how to handle the general case.

1. Shimura varieties of type (A)

1A. *The moduli space.*

1A1. *Shimura datum.* We will introduce the objects needed to define the Shimura variety of unitary type we will work with. We refer to [Kottwitz 1992, Section 5] for more details.

Let F_0 be a totally real field of degree d, and F a CM-extension of F_0. Let $(U_{\mathbb{Q}}, \langle\,,\rangle)$ be a nondegenerate hermitian F-module and G its automorphism group. For any $\mathbb{Q}$-algebra R, we have

$$G(R)=\big\{(g,c)\in\mathrm{GL}_F(U_{\mathbb{Q}}\otimes_{\mathbb{Q}}R)\times R^*:\langle gx,gy\rangle=c\langle x,y\rangle\ \text{ for all }\ x,y\in U_{\mathbb{Q}}\otimes_{\mathbb{Q}}R\big\}.$$

Let $\tau_1,\dots,\tau_d$ be the embeddings of F_0 into $\mathbb{R}$, and let σ_i and $\overline{\sigma}_i$ be the two embeddings of F into $\mathbb{C}$ extending τ_i. The choice of σ_i gives an isomorphism $F\otimes_{F_0,\tau_i}\mathbb{R}\simeq\mathbb{C}$. Let $U_i=U_{\mathbb{Q}}\otimes_{F_0,\tau_i}\mathbb{R}$. We write (a_i,b_i) for the signature of the antihermitian structure on U_i. Then $G_{\mathbb{R}}$ is isomorphic to

$$\mathrm{G}\Big(\prod_{i=1}^{d} \mathrm{U}(a_i, b_i)\Big),$$

where $a_i + b_i$ is independent of i, and is equal to $(1/2d)\dim_{\mathbb{Q}} U_{\mathbb{Q}}$. We'll call this quantity $a+b$.

We also give ourselves a morphism of $\mathbb{R}$-algebras $h : \mathbb{C} \to \mathrm{End}_F U_{\mathbb{R}}$ such that $\langle h(z)v, w\rangle = \langle v, h(\bar{z})w\rangle$ and $(v, w) \mapsto \langle v, h(i)w\rangle$ is positive definite. This morphism gives a complex structure on $U_{\mathbb{R}}$; let $U_{\mathbb{C}}^{1,0}$ be the subspace of $U_{\mathbb{C}}$ on which $h(z)$ acts by multiplication by z.

We then have $U_{\mathbb{C}}^{1,0} \simeq \prod_{i=1}^{d} \mathbb{C}^{a_i} \oplus \overline{\mathbb{C}}^{b_i}$ as $F \otimes_{\mathbb{Q}} \mathbb{R}$-modules and $F \otimes_{\mathbb{Q}} \mathbb{R} \simeq \bigoplus_{i=1}^{d} \mathbb{C}$ (where the action of $\mathbb{C}$ on $\mathbb{C}^{a_i} \oplus \overline{\mathbb{C}}^{b_i}$ is the standard action on the first factor and the conjugated action on the second).

Let U be an O_F-stable lattice of $U_{\mathbb{Q}}$, and assume that the pairing $\langle\,,\rangle$ induces a pairing $U \times U \to \mathbb{Z}$ which is perfect at p.

The ring O_F is a free $\mathbb{Z}$-module. Let $\alpha_1, \ldots, \alpha_t$ be a basis of this module and

$$\det_{U^{1,0}} = f(X_1, \ldots, X_t) = \det\big(X_1\alpha_1 + \cdots + X_t\alpha_t;\, U_{\mathbb{C}}^{1,0} \otimes_{\mathbb{C}} \mathbb{C}[X_1, \ldots, X_t]\big).$$

We can show that f is polynomial with algebraic coefficients. The number field E generated by its coefficients is called the reflex field.

Remark 1.1. We chose for simplicity to work with a central algebra. One can easily adapt the arguments here, replacing F by a simple algebra B with center F.

1A2. *The Shimura variety.* Let us now define the PEL Shimura variety of type (A) associated to G. Let K be an extension of $\mathbb{Q}_p$ containing the images of all the embeddings $F \hookrightarrow \overline{\mathbb{Q}}_p$. Assume that p is unramified in F. We fix an embedding $E \hookrightarrow K$, so that the coefficients of the polynomial $\det_{U^{1,0}}$ can be seen as elements of O_K. We also fix an integer $N \geq 3$ prime to p.

Definition 1.2. Let X be the moduli space over O_K whose S-points are the isomorphism classes of $(A, \lambda, \iota, \eta)$, where

- $A \to S$ is an abelian scheme;
- $\lambda : A \to A^t$ is a prime-to-p polarization;
- $\iota : O_F \to \mathrm{End}\, A$ is compatible with complex conjugation and the Rosati involution, and the polynomials $\det_{U^{1,0}}$ and $\det_{\mathrm{Lie}(A)}$ are equal;
- $\eta : A[N] \to U/NU$ is an O_F-linear symplectic similitude, which lifts locally for the étale topology to an O_F-linear symplectic similitude

$$H_1(A, \mathbb{A}_f^p) \to U \otimes_{\mathbb{Z}} \mathbb{A}_f^p.$$

The moduli space X is representable by a quasiprojective scheme over O_K. We will make a slight assumption on the variety we consider.

Hypothesis 1.3. We suppose that we are not in the case $d = 1$ and $(a, b) = (1, 1)$.

This condition is technical, and will ensure that we can neglect the cusps in the definition of the modular forms. The case we exclude corresponds essentially to the modular curve, which is well known.

1A3. *Iwahori level.* Let π be a prime of F_0 above p; we will call f the residual degree and write $q := p^f$. Since p is unramified in F, we have two possibilities for the behavior of p in F:

- π splits as $\pi^+\pi^-$ in F. We say that π is in case (L).
- π is inert in F. We say that π is in case (U).

The terminology (L) and (U) comes from the fact that the group G at π is a linear or a unitary group, respectively. To define the Iwahori structure, we will break into the two cases.

Definition 1.4. Let $X_{\mathrm{Iw},\pi}$ be the moduli space of isomorphism classes of tuples $(A, \lambda, \iota, \eta, H_\bullet)$, where

- $(A, \lambda, \iota, \eta)$ is a point of X;
- $0 \subset H_1 \subset \cdots \subset H_{a+b} = A[\pi^+]$, where each H_i is a subgroup of $A[\pi^+]$ which is locally isomorphic for the étale topology to $(O_F/\pi^+)^i$ in the case (L);
- $0 \subset H_1 \subset \cdots \subset H_{a+b} = A[\pi]$, where each H_i is a subgroup of $A[\pi]$ which is locally isomorphic for the étale topology to $(O_F/\pi)^i$ and such that $H_{a+b-i} = H_i^\perp$ in the case (U).

The moduli spaces $X_{\mathrm{Iw},\pi}$ are representable by quasiprojective schemes over O_K. We also define the full Iwahori space by $X_{\mathrm{Iw}} = X_{\mathrm{Iw},\pi_1} \times_X X_{\mathrm{Iw},\pi_2} \times_X \cdots \times_X X_{\mathrm{Iw},\pi_g}$, where $\pi_1, \ldots, \pi_g$ are the primes of F_0 above p and the maps $X_{\mathrm{Iw},\pi_i} \to X$ are the natural morphisms corresponding to forgetting the (H_i).

Remark 1.5. In the case (L), the subgroups $A[\pi^+]$ and $A[\pi^-]$ are Cartier duals. This comes from the compatibility between the complex conjugation and the Rosati involution. Therefore, each of these groups is totally isotropic. A flag $(H_\bullet)$ of $A[\pi^+]$ gives naturally a flag $(H_\bullet^\perp)$ of $A[\pi^-]$ with $H_i^\perp = (A[\pi^+]/H_i)^D \subset A[\pi^+]^D = A[\pi^-]$. Choosing the prime π^- instead of π^+ would have given the same definition.

Now we will explicitly describe the determinant condition for the abelian scheme A. We are still working with a prime π of F_0 above p, and assume it is of type (L). Let Σ_π be the decomposition group at π, i.e., the elements $\sigma \in \mathrm{Hom}(F_0, \overline{\mathbb{Q}}_p)$ such that σ sends π into the maximal ideal of $\overline{\mathbb{Q}}_p$. For every $\sigma \in \Sigma_\pi$, there are two embeddings σ^+ and σ^- of F into $\overline{\mathbb{Q}}_p$ above σ; the embedding

σ^+ sends π^+ into the maximal ideal of $\overline{\mathbb{Q}}_p$, and similarly for π^-. To σ we have a pair of integers (a_σ, b_σ) and the choice of the embedding σ^+ gives an order for the two elements of the pair. Let $A \to R$ be an abelian scheme over an O_K-algebra R. If we let $\omega_\pi := e^*\Omega^1_{A[\pi^\infty]}$, then we have $\omega_\pi = \omega_\pi^+ \oplus \omega_\pi^-$, where $\omega_\pi^+ := e^*\Omega^1_{A[(\pi^+)^\infty]}$. The determinant condition for A then implies that

$$\omega_\pi^+ = \bigoplus_{\sigma\in\Sigma_\pi} R^{a_\sigma}$$

with O_F acting on R^{a_σ} by σ^+. Similarly, we have

$$\omega_\pi^- = \bigoplus_{\sigma\in\Sigma_\pi} R^{b_\sigma}$$

with O_F acting on R^{b_σ} by σ^-.

Now suppose that π is of type (U). We still denote by Σ_π the decomposition group at π of F_0. For each $\sigma \in \Sigma_\pi$, there are two embeddings σ_1 and σ_2 of F above σ; the choice of σ_1 gives an order for the elements of the pair (a_σ, b_σ). Let $A \to R$ be an abelian scheme over an O_K-algebra R. If we let $\omega_\pi := e^*\Omega^1_{A[\pi^\infty]}$, then the determinant condition for A implies

$$\omega_\pi = \bigoplus_{\sigma\in\Sigma_\pi} R^{a_\sigma} \oplus R^{b_\sigma},$$

where O_F acts by σ_1 on R^{a_σ} and by σ_2 on R^{b_σ}.

1B. ***The μ-ordinary locus.*** We will describe in this section the μ-ordinary locus of the Shimura variety. Let us first introduce some notation.

Let L be an unramified extension of $\mathbb{Q}_p$ of degree f_0 and k a field of characteristic p containing the residue field of L. Let D be the Galois group of L over $\mathbb{Q}_p$; there is an isomorphism $D \simeq \mathbb{Z}/f_0\mathbb{Z}$, where 1 is identified with the Frobenius σ of L. We will write $W(k)$ for the ring of Witt vectors of k. Let $\underline{\varepsilon} = (\varepsilon_\tau)_{\tau\in D}$ be a sequence of integers equal to 0 or 1. We define a Dieudonné module $M_{\underline{\varepsilon}}$ in the following way: it is a free $W(k)$-module of rank f_0 and, if $(e_\tau)_{\tau\in D}$ is a basis of this module, then the Frobenius and Verschiebung are defined by

$$Fe_{\sigma^{-1}\tau} = p^{\varepsilon_\tau} e_\tau \quad \text{and} \quad Ve_\tau = p^{1-\varepsilon_\tau} e_{\sigma^{-1}\tau}.$$

The module $M_{\underline{\varepsilon}}$ is given an action of the ring of integers of L, namely O_L acts on $W(k)\cdot e_\tau$ by τ.

We'll call $\mathrm{BT}_{\underline{\varepsilon}}$ the p-divisible group over k corresponding to the Dieudonné module $M_{\underline{\varepsilon}}$, and $H_{\underline{\varepsilon}}$ the p-torsion of this p-divisible group.

1B1. *Linear case.* Now we come back to our Shimura variety. Consider first the case (L); we are still considering a place π of F_0 above π which splits as $\pi = \pi^+\pi^-$

in F. If k is a field containing the residue field of O_K and $x = (A, \lambda, \iota, \eta)$ is a k-point of X, then whether the abelian variety A is μ-ordinary at π will depend on the p-divisible group $A[(\pi^+)^\infty]$. Recall that this p-divisible group has an action of O_{F,π^+}, the completion of O_F at π^+; this is an unramified extension of $\mathbb{Q}_p$ of degree f. If Σ_π denotes, as before, the decomposition group of π in F_0, then there is a bijection between Σ_π and the Galois group of O_{F,π^+} and, for each $\sigma \in \Sigma_\pi$, we have a pair of integers (a_σ, b_σ). We put the elements (a_σ) in increasing order; we then have $a_1 \le a_2 \le \cdots \le a_f$. For each integer $0 \le i \le f$ we define the sequence $\underline{\varepsilon}_i = (\varepsilon_{i,j})_{1 \le j \le f}$ by $\varepsilon_{i,j} = 1$ if $j \ge i+1$ and $\varepsilon_{i,j} = 0$ otherwise. We also set by convention $a_0 = 0$ and $a_{f+1} = a + b$.

Definition 1.6. Let k be an algebraically closed field of characteristic p, and let $x = (A, \lambda, \iota, \eta)$ be a k-point of X. Then x is μ-ordinary at π if there is an isomorphism

$$A[(\pi^+)^\infty] \simeq \prod_{i=0}^{f} \mathrm{BT}_{\underline{\varepsilon}_i}^{a_{i+1}-a_i}$$

of p-divisible groups with O_{F,π^+} action.

Note that the term on the right-hand side is explicitly

$$\mathrm{BT}_{(1,\dots,1)}^{a_1} \times \mathrm{BT}_{(0,1,\dots,1)}^{a_2-a_1} \times \cdots \times \mathrm{BT}_{(0,\dots,0,1)}^{a_f-a_{f-1}} \times \mathrm{BT}_{(0,\dots,0)}^{b_f}.$$

Let X_0 denote the special fiber of X, and $X_0^{\mu\text{-}\pi\text{-ord}}$ the μ-ordinary locus at the place π. We have the following proposition, due to Wedhorn [1999, Theorem 1.6.2]:

Proposition 1.7. *The μ-ordinary locus $X_0^{\mu\text{-}\pi\text{-ord}}$ is open and dense in X_0.*

We also have the following characterization of the μ-ordinary locus:

Proposition 1.8 [Moonen 2004, Theorem 1.3.7]. *Let k be an algebraically closed field of characteristic p, and let $x = (A, \lambda, \iota, \eta)$ be a k-point of X. Then x is μ-ordinary at π if and only if there is an isomorphism*

$$A[\pi^+] \simeq \prod_{i=0}^{f} \mathrm{BT}_{\underline{\varepsilon}_i}^{a_{i+1}-a_i}[\pi^+]$$

of finite flat group schemes with O_{F,π^+} action.

Since $\mathrm{BT}_{\underline{\varepsilon}_i}$ is multiplicative for $i = 0$, étale for $i = f$ and bi-infinitesimal otherwise, we have the following criterion for the existence of the ordinary locus at π:

Proposition 1.9. *The μ-ordinary locus equals the ordinary locus (at the place π) if and only if there exists an integer a such that $a_\sigma = a$ for all $\sigma \in \Sigma_\pi$.*

This last condition is also equivalent to the fact that the local reflex field at π is equal to $\mathbb{Q}_p$ (see [Wedhorn 1999, Section 1.6] for more details).

1B2. *Unitary case.* Let us now consider the case (U). Let k be a field containing the residue field of O_K, and let $x = (A, \lambda, \iota, \eta)$ be a k-point of X. Whether the abelian variety A is μ-ordinary at π will depend on the p-divisible group $A[\pi^\infty]$. Recall that this p-divisible group has an action of $O_{F,\pi}$, the completion of O_F at π; this is an unramified extension of $\mathbb{Q}_p$ of degree $2f$. Recall that Σ_π is the decomposition group at π of F_0 and it is of cardinality f. If $\sigma \in \Sigma_\pi$, there are two embeddings σ_1 and σ_2 of F into $\overline{\mathbb{Q}}_p$, and the choice of one of the two gives elements a_σ and b_σ. We suppose that the choice is made so that $a_\sigma \leq b_\sigma$; we also order the elements in Σ_π so that the sequence (a_σ) is increasing. We then have

$$a_1 \leq a_2 \leq \cdots \leq a_f \leq b_f \leq b_{f-1} \leq \cdots \leq b_1.$$

This gives an order on the embeddings of F into $\overline{\mathbb{Q}}_p$. For each integer $0 \leq i \leq 2f$ we define the sequence $\underline{\varepsilon}_i = (\varepsilon_{i,j})_{1\leq j\leq 2f}$ by $\varepsilon_{i,j} = 1$ if $j \geq i+1$ and $\varepsilon_{i,j} = 0$ otherwise. We define a sequence $(\alpha_i)_{0\leq i\leq 2f+1}$ by $\alpha_0 = 0$, $\alpha_i = a_i$ for $1 \leq i \leq f$, $\alpha_i = b_{2f+1-i}$ for $f+1 \leq i \leq 2f$ and $\alpha_{2f+1} = a+b$.

Definition 1.10. Let k be an algebraically closed field of characteristic p, and let $x = (A, \lambda, \iota, \eta)$ be a k-point of X. Then x is μ-ordinary at π if there is an isomorphism

$$A[\pi^\infty] \simeq \prod_{i=0}^{2f} \mathrm{BT}_{\underline{\varepsilon}_i}^{\alpha_{i+1}-\alpha_i}$$

of p-divisible groups with $O_{F,\pi}$ action.

Note that the term in the right-hand side is explicitly

$$\mathrm{BT}^{a_1}_{(1,\ldots,1)} \times \mathrm{BT}^{a_2-a_1}_{(0,1,\ldots,1)} \times \cdots \times \mathrm{BT}^{a_f-a_{f-1}}_{(0,\ldots,0,1,1\ldots,1)} \\ \times \mathrm{BT}^{b_f-a_f}_{(0,\ldots,0,1\ldots,1)} \times \mathrm{BT}^{a_f-a_{f-1}}_{(0,\ldots,0,0,1\ldots,1)} \times \cdots \times \mathrm{BT}^{a_1}_{(0,\ldots,0)}.$$

(We have used the fact that $b_i - b_{i-1} = a_{i-1} - a_i$, since the quantity $a_j + b_j$ is independent of j.)

Let X_0 denote the special fiber of X, and $X_0^{\mu\text{-}\pi\text{-ord}}$ the μ-ordinary locus. We have the following proposition, due to Wedhorn [1999, Theorem 1.6.2]:

Proposition 1.11. *The μ-ordinary locus $X_0^{\mu\text{-}\pi\text{-ord}}$ is open and dense in X_0.*

We also have the following characterization of the μ-ordinary locus:

Proposition 1.12 [Moonen 2004, Theorem 1.3.7]. *Let k be an algebraically closed field of characteristic p, and let $x = (A, \lambda, \iota, \eta)$ be a k-point of X. Then x is μ-ordinary at π if and only if there is an isomorphism*

$$A[\pi] \simeq \prod_{i=0}^{2f} \mathrm{BT}_{\underline{\varepsilon}_i}^{\alpha_{i+1}-\alpha_i}[\pi]$$

of finite flat group schemes with $O_{F,\pi}$ action.

Since $\mathrm{BT}_{\underline{\varepsilon}_i}$ is multiplicative for $i = 0$, étale for $i = 2f$ and bi-infinitesimal otherwise, we have the following criterion for the existence of the ordinary locus at π:

Proposition 1.13. *The μ-ordinary locus equals the ordinary locus* (*at the place π*) *if and only if $a_\sigma = b_\sigma = \frac{1}{2}(a+b)$ for all $\sigma \in \Sigma_\pi$.*

Again, this last condition is equivalent to the fact that the local reflex field at π is equal to $\mathbb{Q}_p$.

We'll need later to work with the rigid space associated to X. Let us call this rigid space X_{rig}; it is the generic fiber of the formal completion of X along its special fiber. We refer to [Berthelot 1996] for more details on rigid spaces. We have a specialization map $\mathrm{sp} : X_{\mathrm{rig}} \to X_0$, and we'll write $X_{\mathrm{rig}}^{\mu\text{-}\pi\text{-ord}}$ for the inverse image of the μ-ordinary locus under the specialization map. We'll also write $X_{\mathrm{Iw,rig}}$ for the rigid space associated to X_{Iw}.

1C. *Canonical subgroups.*

1C1. *Degrees and partial degrees.* Before introducing the canonical subgroups on the μ-ordinary locus, we'll define the degree for a finite flat group scheme defined over a finite extension of $\mathbb{Q}_p$, and the partial degrees for these endowed with an action of the ring of integers of an unramified extension of $\mathbb{Q}_p$.

Definition 1.14. Let L be a finite extension of $\mathbb{Q}_p$ and let G be a finite flat group scheme of p-power order over O_L. Let ω_G be the conormal module along the unit section. Then the degree of G is, by definition,

$$\deg G := v(\mathrm{Fitt}_0\, \omega_G),$$

where Fitt_0 denotes the Fitting ideal, and the valuation of an ideal is defined by $v(xO_L) = v(x)$, normalized by $v(p) = 1$.

We now state some properties of this function. We refer to [Fargues 2010, Section 3] for more details.

Proposition 1.15. *The degree function has the following properties*:

- *Let G be as before. Then*, *if G^D denotes the Cartier dual of G*, *we have*

  $$\deg G^D = \mathrm{ht}\, G - \deg G.$$

  In particular, $\deg G \in [0, \mathrm{ht}\, G]$.

- *The degree function is additive*: *if we have an exact sequence*

  $$0 \to G_1 \to G_2 \to G_3 \to 0$$

  with G_i finite flat, *then* $\deg G_2 = \deg G_1 + \deg G_3$.

- *Let G and G' be two finite flat group schemes, and suppose that there exists a morphism $f : G \to G'$ which is an isomorphism in generic fiber. Then* $\deg G \le \deg G'$, *and we have equality if and only if f is an isomorphism.*

We deduce from the last property the following corollary:

Corollary 1.16. *Let G be a finite flat group scheme of p-power order defined over a finite extension of $\mathbb{Q}_p$. Suppose that H_1 and H_2 are two finite flat subgroups of G. Then*

$$\deg H_1 + \deg H_2 \le \deg(H_1 + H_2) + \deg(H_1 \cap H_2).$$

Proof. By dividing everything by $H_1 \cap H_2$, we are reduced to the case $H_1 \cap H_2 = \{0\}$. The morphism $H_1 \times H_2 \to H_1 + H_2$ is an isomorphism in generic fiber, thus by the previous proposition we get

$$\deg(H_1 \times H_2) \le \deg(H_1 + H_2).$$

But, since the degree function is additive, $\deg(H_1 \times H_2) = \deg H_1 + \deg H_2$. □

Let G be as in the previous definition and suppose now that G has an action of O_M, where M is a finite unramified extension of $\mathbb{Q}_p$. Let Σ be the set of embeddings of M into $\overline{\mathbb{Q}}_p$. Then the module ω_G has an action of O_M and has the decomposition

$$\omega_G = \bigoplus_{\sigma \in \Sigma} \omega_{G,\sigma},$$

where O_M acts on $\omega_{G,\sigma}$ by σ.

Definition 1.17. The partial degree of G is defined for all $\sigma \in \Sigma$ as

$$\deg_\sigma G = v(\mathrm{Fitt}_0\, \omega_{G,\sigma}).$$

Proposition 1.18. *The partial degree functions have the following properties*:

- *We have* $\deg G = \sum_{\sigma \in \Sigma} \deg_\sigma G$.
- *Suppose that G has height $[M : \mathbb{Q}_p]h$. If G^D denotes the Cartier dual of G, we have, for all $\sigma \in \Sigma$,*

  $$\deg_\sigma G^D = h - \deg_\sigma G.$$

  In particular, $\deg_\sigma G \in [0, h]$.
- *The partial degree functions are additive*: *if we have an exact sequence*

  $$0 \to G_1 \to G_2 \to G_3 \to 0$$

  with G_i finite flat with an action of O_M, then $\deg_\sigma G_2 = \deg_\sigma G_1 + \deg_\sigma G_3$ *for all $\sigma \in \Sigma$.*

We refer to [Bijakowski 2014, Section 1.1.5] for more details.

Remark 1.19. If we are in the situation of the third point of Proposition 1.15, i.e., if we have a morphism of finite flat group schemes $f : G \to G'$ that is an isomorphism in generic fiber and if G and G' have an action of O_M, then it is not true that the partial degrees increase. Indeed, the functions that increase are linear combinations of the partial degrees. For example, if $[M : \mathbb{Q}_p] = 2$, there are two partial degrees $\deg_1$ and $\deg_2$, and the functions that increase are $\deg_1 + p \deg_2$ and $p \deg_1 + \deg_2$. See [Bijakowski 2014, Proposition 1.1.33] for more details.

Example 1.20. Let us apply what we have said to our Shimura variety. Let $x = (A, \lambda, \iota, \eta)$ be an O_L-point of X (where L is a finite extension of $\mathbb{Q}_p$) and suppose that π is a place of F_0 above p in case (L). Then $A[\pi^+]$ has an action of $O_{F_0,\pi}$. Moreover, we have

$$\deg_\sigma A[\pi^+] = a_\sigma$$

for all $\sigma \in \Sigma_\pi$. If H is an O_F-stable subgroup of $A[\pi^+]$ of height fh, then the orthogonal $H^\perp$ is a subgroup of $A[\pi^-]$ of height $f(a+b-h)$. We have $H^\perp \simeq (A[\pi^+]/H)^D$, and thus

$$\deg_\sigma H^\perp = (a+b-h) - (a_\sigma - \deg_\sigma H) = b_\sigma - h + \deg_\sigma H$$

for all $\sigma \in \Sigma_\pi$. We see that one has the inequalities

$$\deg_\sigma H \geq h - b_\sigma \quad \text{and} \quad \deg_\sigma H^\perp \geq b_\sigma - h$$

for all $\sigma \in \Sigma_\pi$.

Example 1.21. Suppose now that π is in case (U). Then the group scheme $A[\pi]$ has an action of $O_{F,\pi}$. Recall that if $\sigma \in \Sigma_\pi$ is an embedding of F_0 into $\overline{\mathbb{Q}}_p$ above π, then there are two embeddings σ_1 and σ_2 of F extending σ. With our previous conventions, we have

$$\deg_{\sigma_1} A[\pi] = a_\sigma \quad \text{and} \quad \deg_{\sigma_2} A[\pi] = b_\sigma.$$

If H is an O_F-stable subgroup of $A[\pi]$ of height $2fh$, then the orthogonal $H^\perp$ is a subgroup of $A[\pi]$ of height $2f(a+b-h)$. We have $H^\perp \simeq (A[\pi]/H)^{D,c}$, where the superscript c means that the action of O_F on $(A[\pi]/H)^{D,c}$ is the conjugate of the natural one. This comes from the compatibility between the Rosati involution and the complex conjugation. Thus,

$$\deg_{\sigma_1} H^\perp = a_\sigma - h + \deg_{\sigma_2} H \quad \text{and} \quad \deg_{\sigma_2} H^\perp = b_\sigma - h + \deg_{\sigma_1} H$$

for all $\sigma \in \Sigma_\pi$. We see that one has the inequalities

$$\deg_{\sigma_1} H \geq h - b_\sigma \quad \text{and} \quad \deg_{\sigma_2} H \geq h - a_\sigma$$

for all $\sigma \in \Sigma_\pi$.

1C2. *Siegel variety.* Let us now recall some facts for the canonical subgroup of the Siegel variety.

Let $g \geq 1$ be an integer and $\mathcal{A}_g$ the Siegel variety. There is a universal abelian scheme A on $\mathcal{A}_g$. There is also a Hasse invariant Ha on $\mathcal{A}_g$. We quote the main result obtained by Fargues [2011] on the canonical subgroup.

Proposition 1.22. *Let A be an abelian scheme of dimension g defined over O_L (L is a finite extension of $\mathbb{Q}_p$). Suppose that the valuation w of the Hasse invariant is strictly less than $\frac{1}{2}$. Then there is a canonical subgroup $H \subset A[p]$, of height g, totally isotropic, with*

$$\deg H = g - w.$$

Let $\mathcal{A}_{g,\mathrm{rig}}$ be the rigid space associated to $\mathcal{A}_g$. Then the ordinary locus of $\mathcal{A}_{g,\mathrm{rig}}$ is defined as the locus where the associated abelian scheme is ordinary; it is also the locus where the Hasse invariant is invertible. The proposition says that, on a strict neighborhood of the ordinary locus, there exists a canonical subgroup of high degree in $A[p]$. We propose a simple reformulation of this property. We have the following observation:

Proposition 1.23. *Let A be an abelian scheme of dimension g defined over O_L (L is a finite extension of $\mathbb{Q}_p$). There exists at most one subgroup H of height g of $A[p]$ with*

$$\deg H > g - \tfrac{1}{2}.$$

Proof. Suppose not, and let H_1 and H_2 be two subgroups with $\deg H_i > g - \frac{1}{2}$. Then we have

$$2g - 1 < \deg H_1 + \deg H_2 \leq \deg(H_1 + H_2) + \deg(H_1 \cap H_2).$$

But $\deg(H_1 + H_2) \leq \deg A[p] = g$ and, since $H_1 \cap H_2$ is of height $h \leq g - 1$, we have $\deg(H_1 \cap H_2) \leq g - 1$. We get a contradiction. $\square$

This can be used to prove the existence of the canonical subgroup in the following way. Let $\mathcal{A}'_g$ be the Siegel variety parametrizing a g-dimensional abelian scheme with polarization and a subgroup H totally isotropic of height g. We have a map $f : \mathcal{A}'_g \to \mathcal{A}_g$ corresponding to forgetting H. If we let $\mathcal{A}'_{g,\mathrm{rig}}$ be the rigid space associated to $\mathcal{A}'_g$, we still have a morphism $f : \mathcal{A}'_{g,\mathrm{rig}} \to \mathcal{A}_{g,\mathrm{rig}}$. Define $X_r = \{x \in \mathcal{A}'_{g,\mathrm{rig}} : \deg H(x) \geq g - r\}$ for any rational r; it is an admissible open subset of $\mathcal{A}'_{g,\mathrm{rig}}$. Then the ordinary locus of $\mathcal{A}_{g,\mathrm{rig}}$ is $f(X_0)$, and it follows from [Bijakowski et al. 2016, Proposition 4.1.7] that $(f(X_r))_{r>0}$ forms a basis of strict neighborhoods of the ordinary locus (the map f is finite étale, and the $(X_r)_{r>0}$ are strict neighborhoods of X_0). The previous proposition shows that on $f(X_r)$ there is exactly one subgroup of height g of degree greater than or equal to $g - r$ for $r < \frac{1}{2}$; this is the canonical subgroup.

1C3. *Linear case.* Now let's get back to our Shimura varieties. Suppose we are in case (L). Then we have the following proposition:

Proposition 1.24. *Let L be a finite extension of $\mathbb{Q}_p$ and $x = (A, \lambda, \iota, \eta)$ an O_L-point of X. Let $1 \leq i \leq f$ be an integer. Then there exists at most one subgroup $H \subset A[\pi^+]$ that is stable by O_F of height $f a_i$ such that*

$$\deg H > \sum_{j=1}^{f} \min(a_j, a_i) - \tfrac{1}{2}.$$

Proof. Suppose not, and let H_1 and H_2 be two such subgroups. Let us denote by fh the height of $H_1 \cap H_2$; the height of $H_1 + H_2$ is then $f(2a_i - h)$. We have

$$\deg H_1 + \deg H_2 \leq \deg(H_1 + H_2) + \deg(H_1 \cap H_2).$$

But

$$\deg(H_1 + H_2) \leq \sum_{j=1}^{f} \min(a_j, 2a_i - h) \leq \sum_{j=1}^{i} a_j + \sum_{j=i+1}^{f} (2a_i - h)$$

and

$$\deg(H_1 \cap H_2) \leq \sum_{j=1}^{f} \min(a_j, h) \leq \sum_{j=1}^{i-1} a_j + \sum_{j=i}^{f} h.$$

We finally get

$$\deg H_1 + \deg H_2 \leq 2\sum_{j=1}^{i-1} a_j + 2\sum_{j=i+1}^{f} a_i + a_i + h \leq 2\sum_{j=1}^{f} \min(a_j, a_i) - 1$$

since $h \leq a_i - 1$, a contradiction. □

The proposition shows that there exists at most one subgroup of height $f a_i$ and of big degree. The next proposition shows that if two such subgroups exist (with different heights), then we automatically have an inclusion.

Proposition 1.25. *Let $i < j$ be two integers between 1 and f. Let $x = (A, \lambda, \iota, \eta)$ be an O_L-point of X, and suppose there exists, for $l \in \{i, j\}$, a subgroup $H_l \subset A[\pi^+]$ that is stable by O_F of height $f a_l$ such that*

$$\deg H_l > \sum_{k=1}^{f} \min(a_k, a_l) - \tfrac{1}{2}.$$

Then we have $H_i \subset H_j$.

Proof. Let fh denote the height of $H_i \cap H_j$. We have the following inequalities:

$$\deg(H_i + H_j) \leq \sum_{k=1}^{f} \min(a_k, a_i + a_j - h) \leq \sum_{k=1}^{j} a_k + \sum_{k=j+1}^{f} (a_i + a_j - h),$$

$$\deg(H_i \cap H_j) \leq \sum_{k=1}^{f} \min(a_k, h) \leq \sum_{k=1}^{i-1} a_k + \sum_{k=i}^{f} h.$$

We then get

$$\begin{aligned}\deg H_i + \deg H_j &\leq \deg(H_i + H_j) + \deg(H_i \cap H_j)\\ &\leq 2\sum_{k=1}^{i-1} a_k + \sum_{k=i}^{j}(a_k + h) + \sum_{k=j+1}^{f} (a_i + a_j).\end{aligned}$$

If we do not have the inclusion $H_i \subset H_j$, then $h \leq a_i - 1$. We then get

$$\deg H_i + \deg H_j \leq \sum_{k=1}^{f}(\min(a_k, a_i) + \min(a_k, a_j)) - (j - i + 1).$$

We get a contradiction with the hypothesis that H_i and H_j have big degrees. □

As a consequence, we directly get the existence of canonical subgroups for μ-ordinary abelian schemes. Let s be the cardinality of $\{a_1, \dots, a_f\} \cap [1, a+b-1]$, and denote by $A_1 < \cdots < A_s$ the different elements of this set.

Corollary 1.26. *Let L be a finite extension of $\mathbb{Q}_p$, and let $x = (A, \lambda, \iota, \eta)$ be an O_L-point of X. Assume that x is μ-ordinary (i.e., the special fiber of A is μ-ordinary). Then, for any integer $1 \leq k \leq s$, there exists a unique subgroup $H_k \subset A[\pi^+]$ of height fA_k, with*

$$\deg_\sigma H_k = \min(a_\sigma, A_k)$$

for all $\sigma \in \Sigma_\pi$.

Proof. We can work over $\overline{\mathbb{Q}}_p$. Since A is μ-ordinary we have, by [Moonen 2004, Proposition 2.1.9], a filtration

$$X_0 = 0 \subset X_1 \subset \cdots \subset X_{f+1} = A[(\pi^+)^\infty]$$

on $A[(\pi^+)^\infty]$ with X_i p-divisible groups such that $(X_{i+1}/X_i) \times \overline{\mathbb{F}}_p \simeq \mathrm{BT}_{\underline{\varepsilon}_i}^{a_{i+1}-a_i}$ for $0 \leq i \leq f$. Let $Y_i = X_{i+1}/X_i$ for $0 \leq i \leq f$. Then Y_i is a p-divisible group over the ring of integers of $\overline{\mathbb{Q}}_p$. The module ω_{Y_i} decomposes into $\bigoplus_{\sigma \in \Sigma_\pi} \omega_{Y_i,\sigma}$, and each $\omega_{Y_i,\sigma}$ is free over the ring of integers of $\overline{\mathbb{Q}}_p$. Recall that we have chosen an ordering for the set $\Sigma_\pi = \{\sigma_1, \dots, \sigma_f\}$. From the description of the special fiber of Y_i, one sees that ω_{Y_i,σ_j} is 0 if $j \leq i$, and is free of rank $a_{i+1} - a_i$ over the ring of integers of $\overline{\mathbb{Q}}_p$ if $j \geq i+1$. Thus, $\deg_{\sigma_j} Y_i[\pi^+]$ is 0 if $j \leq i$ and $a_{i+1} - a_i$ otherwise.

Let $H_i = X_i[\pi^+]$ for $0 \leq i \leq f$; it is a finite flat subgroup of $A[\pi^+]$ of height fa_i. Moreover, we have

$$\deg_{\sigma_j} H_i = \sum_{k=1}^{i} \deg_{\sigma_j} H_k/H_{k-1} = \sum_{k=1}^{\min(i,j)} (a_k - a_{k-1}) = a_{\min(i,j)} = \min(a_i, a_j)$$

for all i and j between 1 and f. This gives the existence of the desired subgroups. The uniqueness follows from the previous proposition. □

Remark 1.27. We also have the following description of the canonical subgroups in the special fiber. Let L be a finite extension of $\mathbb{Q}_p$, and let $x = (A, \lambda, \iota, \eta)$ be a μ-ordinary point of X defined over O_L. Let A_s denote the special fiber of A; then the Frobenius F acts on $A_s[(\pi^+)^\infty]$, and we can form the subgroups

$$C_i := (\pi^+)^{f-i} A_s[F^f, (\pi^+)^{f-i+1}]$$

for $1 \leq i \leq f$. Then we have $0 \subset C_1 \subset \cdots \subset C_f \subset A_s[\pi^+]$, and the special fiber of each of the canonical subgroups is equal to one of the C_i. More precisely, we have $H_j \times k(L) = C_{r(j)}$ for $1 \leq j \leq s$, with $k(L)$ the residue field of L and $r(j) = \min\{l : a_l = A_j\}$.

We can then define the relevant degree functions on $X_{\mathrm{Iw},\pi}$. For each integer j, define

$$d_j = \sum_{\sigma \in \Sigma_\pi} \min(a_\sigma, a_j).$$

Let $X_{\mathrm{Iw},\pi,\mathrm{rig}}$ be the associated rigid space. We define the degree function Deg : $X_{\mathrm{Iw},\pi,\mathrm{rig}} \to \prod_{j=1}^{s} [0, d_j]$ on this space by

$$\mathrm{Deg}(A, \lambda, \iota, \eta, H_\bullet) := (\deg H_{a_j})_{1 \leq j \leq s}.$$

We also define the j-th degree function by $\mathrm{Deg}_j(A, \lambda, \iota, \eta, H_\bullet) := \deg H_{a_j}$ for $1 \leq j \leq s$.

Remark 1.28. The integers s and d_j as well as the functions Deg and Deg_j depend on the place π. If the context is clear, we choose not to write the dependence on π to lighten the notation.

We then have the following description of the μ-ordinary locus.

Proposition 1.29. *The space $X_{\mathrm{rig}}^{\mu\text{-}\pi\text{-ord}} \subset X_{\mathrm{rig}}$ is exactly the image of*

$$\mathrm{Deg}^{-1}(\{d_1\} \times \cdots \times \{d_s\})$$

under the map $X_{\mathrm{Iw},\pi,\mathrm{rig}} \to X_{\mathrm{rig}}$.

Proof. If x is a μ-ordinary point, it then follows from the previous corollary that there exist subgroups H_j of height fa_j with $\deg H_j = d_j$ for all $1 \leq j \leq s$. Conversely, suppose that $(A, \lambda, \iota, \eta)$ is a point of X_{rig}, with A defined over the ring of integers of an extension L of $\mathbb{Q}_p$, and that there exist subgroups H_j of height fa_j with $\deg H_j = d_j$ for all $1 \leq j \leq s$. We want to show that A is μ-ordinary; it suffices to show that $A[\pi^+]$ has a nice description. We have $\deg_\sigma H_j \leq \min(a_\sigma, a_j)$ for all $\sigma \in \Sigma_\pi$ and, since $d_j = \sum_{\sigma \in \Sigma_\pi} \min(a_\sigma, a_j)$, this forces the last inequality to be an equality. Define $H_0 = 0$, $H_{s+1} = A[\pi^+]$, and let H_j' be a complement of H_{j-1} in H_j for all $1 \leq j \leq s+1$. This is possible if the field L is large enough. We claim that, for $1 \leq j \leq s+1$,

$$H_j \simeq H_j' \times H_{j-1}.$$

Indeed we have a morphism $H_{j-1} \to H_j/H_j'$ that is an isomorphism in generic fiber. The degree of the image of H_{j-1} in H_j/H_j' thus increases, but since the degree of H_{j-1} is maximal, it must be an equality. We deduce that $\deg H_j = \deg H_j' + \deg H_{j-1}$ and that the morphism $H_j' \times H_{j-1} \to H_j$ is an isomorphism.

We finally get

$$A[\pi^+] \simeq H_1' \times H_2' \times \cdots \times H_{s+1}'.$$

But we can explicitly describe the groups H_j'. Indeed, for all $1 \leq j \leq s+1$ and $\sigma \in \Sigma_\pi$, we have (setting $A_0 = 0$ and $A_{s+1} = a+b$)

$$\deg_\sigma H_j' = \deg_\sigma H_j - \deg_\sigma H_{j-1} = \min(a_\sigma, a_j) - \min(a_\sigma, A_{j-1}).$$

This quantity is 0 if $a_\sigma \leq A_{j-1}$ and $a_j - A_{j-1}$ if $a_\sigma \geq a_j$. Since the height of H_j' is $f(a_j - A_{j-1})$, one can see that the special fiber of H_j' is isomorphic to $\mathrm{BT}_{\mathcal{E}_{s(j)}}[\pi^+]^{A_j - A_{j-1}}$, where $s(j)$ is the number of $\sigma \in D_\pi$ with $a_\sigma \leq A_{j-1}$ (one can see this by looking at the Dieudonné module associated to the special fiber of H_j', for example). We then conclude that A is μ-ordinary by Proposition 1.8. $\square$

1C4. *Unitary case.* Suppose now we are in case (U). Then we have the following proposition:

Proposition 1.30. *Let L be a finite extension of $\mathbb{Q}_p$ and $x = (A, \lambda, \iota, \eta)$ an O_L-point of X. Let $1 \leq i \leq f$ be an integer. Then there exists at most one subgroup $H \subset A[\pi]$ that is stable by O_F of height $2fa_i$ such that*

$$\deg H > \sum_{j=1}^{2f} \min(\alpha_j, a_i) - \tfrac{1}{2}.$$

Moreover, if such a subgroup exists, it is totally isotropic.

Proof. The proof of the first part of the proposition is exactly the same as in the linear case (see Proposition 1.24). To prove that the subgroup is totally isotropic, we will use the same argument as in the proof of Proposition 1.25. We only need to get a bound for the degree of $H^{\perp}$. But we have $H^{\perp} \simeq (A[\pi]/H)^D$, so

$$\begin{aligned}\deg H^{\perp} &= 2fb_i - \deg(A[\pi]/H)\\ &= \deg H + \sum_{j=1}^{2f}(b_i - \alpha_j) > \sum_{j=1}^{2f}(b_i - \alpha_j + \min(\alpha_j, a_i)) - \tfrac{1}{2}.\end{aligned}$$

But $b_i - \alpha_j + \min(\alpha_j, a_i) = \min(b_i, a_i + b_i - \alpha_j)$ and, since $a_i + b_i$ is constant, we have $a_i + b_i - \alpha_j = \alpha_{2f+1-i}$. In conclusion, we get

$$\deg H^{\perp} > \sum_{j=1}^{2f} \min(b_i, \alpha_j) - \tfrac{1}{2}.$$

We conclude that $H \subset H^{\perp}$ by applying the proof of Proposition 1.25 directly (note that $H^{\perp}$ is of height $2fb_i$). □

The proposition shows that there exists at most one subgroup of height $2fa_i$ and of big degree. The next proposition shows that if two such subgroups exists (with different heights), then we automatically have an inclusion.

Proposition 1.31. *Let $i < j$ be two integers between 1 and f. Let $x = (A, \lambda, \iota, \eta)$ be an O_L-point of X, and suppose there exists, for $l \in \{i, j\}$, a subgroup $H_l \subset A[\pi]$ that is stable by O_F of height $2fa_l$ such that*

$$\deg H_l > \sum_{k=1}^{2f} \min(\alpha_k, a_l) - \tfrac{1}{2}.$$

Then $H_i \subset H_j$.

Proof. The proof is the same as in the linear case (see Proposition 1.25). □

As a consequence, we directly get the existence of canonical subgroups for μ-ordinary abelian schemes. Let s be the cardinality of $\{a_1, \dots, a_f\} \cap \left[1, \frac{1}{2}(a+b)\right]$, and denote by $A_1 < \cdots < A_s$ the different elements of this set.

Corollary 1.32. *Let L be a finite extension of $\mathbb{Q}_p$ and let $x = (A, \lambda, \iota, \eta)$ be an O_L-point of X. Assume that x is μ-ordinary (i.e., the special fiber of A is μ-ordinary). Then, for any integer $1 \le k \le s$, there exists a unique totally isotropic subgroup $H_k \subset A[\pi]$ of height $2fA_k$, with*

$$\deg_\sigma H_k = \min(\alpha_\sigma, A_k)$$

for all $\sigma \in \Sigma_\pi$.

Proof. The proof is similar to the linear case (see Corollary 1.26). □

Remark 1.33. We also have the following description of the canonical subgroups in the special fiber. Let L be a finite extension of $\mathbb{Q}_p$, and let $x = (A, \lambda, \iota, \eta)$ be a μ-ordinary point of X defined over O_L. Let A_s denote the special fiber of A; then the Frobenius F acts on $A_s[\pi^\infty]$, and we can form the subgroups

$$C_i := \pi^{2f-i} A_s[F^{2f}, \pi^{2f-i+1}]$$

for $1 \leq i \leq 2f$. Then we have $0 \subset C_1 \subset \cdots \subset C_{2f} \subset A_s[\pi]$, and the special fiber of each of the canonical subgroups is equal to one of the C_i. More precisely, we have $H_j \times k(L) = C_{r(j)}$ for $1 \leq j \leq s$, with $k(L)$ the residue field of L and $r(j) = \min\{l : a_l = A_j\}$. Note also that $C_i^\perp = C_{2f+1-i}$.

We can then define the relevant degree functions on $X_{\mathrm{Iw},\pi}$. For each integer k, define

$$d_k = \sum_{j=1}^{2f} \min(\alpha_j, A_k).$$

Let $X_{\mathrm{Iw},\pi,\mathrm{rig}}$ be the associated rigid space. We define the degree function $\mathrm{Deg} : X_{\mathrm{Iw},\pi,\mathrm{rig}} \to \prod_{k=1}^{s} [0, d_k]$ on this space by

$$\mathrm{Deg}(A, \lambda, \iota, \eta, H_\bullet) := (\deg H_{A_k})_{1 \leq k \leq s}.$$

We also define the k-th degree function by $\mathrm{Deg}_k(A, \lambda, \iota, \eta, H_\bullet) := \deg H_{A_k}$ for $1 \leq k \leq s$. We then have the following description of the μ-ordinary locus:

Proposition 1.34. *The space $X_{\mathrm{rig}}^{\mu\text{-ord}} \subset X_{\mathrm{rig}}$ is exactly the image of*

$$\mathrm{Deg}^{-1}(\{d_1\} \times \cdots \times \{d_s\})$$

under the map $X_{\mathrm{Iw},\pi,\mathrm{rig}} \to X_{\mathrm{rig}}$.

Proof. If x is a μ-ordinary point, it follows from the previous corollary that there exist subgroups H_k of height $2fA_k$ with $\deg H_k = d_k$ for all $1 \leq k \leq s$. Conversely, suppose that $(A, \lambda, \iota, \eta)$ is a point of X_{rig}, with A defined over the ring of integers of an extension L of $\mathbb{Q}_p$, and that there exist subgroups H_k of height $2fA_k$ with $\deg H_k = d_k$ for all $1 \leq k \leq s$. This implies that we have $\deg_\sigma H_k = \min(\alpha_\sigma, A_k)$ for all $\sigma \in D_\pi$ and $1 \leq k \leq s$. We want to show that A is μ-ordinary; it suffices to show that $A[\pi]$ has a nice description. Define $H_0 = 0$, $H_i = H_{2s+1-i}^\perp$ for $s+1 \leq i \leq 2s+1$, and let H_j' be a complement of H_{j-1} in H_j for all $1 \leq j \leq 2s+1$. This is possible if the field L is big enough. We claim that

$$H_j \simeq H_j' \times H_{j-1}.$$

Indeed we have a morphism $H_{j-1} \to H_j/H_j'$ that is an isomorphism in generic fiber. The degree of the image of H_{j-1} in H_j/H_j' thus increases, but since the degree of

H_{j-1} is maximal, it must be an equality. We deduce that $\deg H_j = \deg H'_j + \deg H_{j-1}$ and that the morphism $H'_j \times H_{j-1} \to H_j$ is an isomorphism.

We finally get

$$A[\pi] \simeq H'_1 \times H'_2 \times \cdots \times H'_{2s+1}.$$

But we can explicitly describe the groups H'_j using the same proof as Proposition 1.29. We then conclude that A is μ-ordinary by Proposition 1.12. $\square$

2. Modular forms and Hecke operators

2A. *Modular forms.* Let's now define the modular forms for the Shimura variety X. Let π be a place of F_0 above p, and suppose it is in case (L). Then we define the $O_F \otimes O_K$-module St_π by

$$\mathrm{St}_\pi := O_K^{a_\sigma} \oplus O_K^{b_\sigma},$$

where O_F acts on $O_K^{a_\sigma}$ by σ^+ and on $O_K^{b_\sigma}$ by σ^-. If π is in case (U), we define the $O_F \otimes O_K$-module St_π by

$$\mathrm{St}_\pi := O_K^{a_\sigma} \oplus O_K^{b_\sigma},$$

where O_F acts on $O_K^{a_\sigma}$ by σ_1 and on $O_K^{b_\sigma}$ by σ_2. Finally, we define the $O_F \otimes O_K$-module St by

$$\mathrm{St} = \bigoplus_\pi \mathrm{St}_\pi,$$

where π runs over the places of F_0 above p. If R is an O_K-algebra and $(A, \lambda, \iota, \eta)$ is an R-point of X, then the $R \otimes O_F$-module $e^*\Omega^1_{A/R}$ is isomorphic to $\mathrm{St} \otimes_{O_K} R$. The sheaf $\omega_A := e^*\Omega^1_{A/X}$ is then locally isomorphic to $\mathrm{St} \otimes_{O_K} \mathcal{O}_X$ (it is a locally free sheaf on $\mathcal{O}_X$).

Define

$$\mathcal{T} = \mathrm{Isom}_{O_F \otimes \mathcal{O}_X}(\mathrm{St} \otimes \mathcal{O}_X, \omega_A).$$

This is a torsor on X under the group, defined over O_K,

$$M = \prod_{\pi \in \mathcal{P}} \prod_{\sigma \in \Sigma_\pi} \mathrm{GL}_{a_\sigma} \times \mathrm{GL}_{b_\sigma},$$

where $\mathcal{P}$ is the set of primes of F_0 above π. Let B_M be the upper Borel subgroup of M, U_M its unipotent radical, and T_M its maximal torus. Let $X(T_M)$ be the character group of T_M and $X(T_M)^+$ the cone of dominant weights for B_M. If $\kappa \in X(T_M)^+$, we let $\kappa' = -w_0\kappa \in X(T_M)^+$, where w_0 is the element of highest length in the Weyl group of M relative to T_M.

Let $\phi : \mathcal{T} \to X$ be the projection morphism.

Definition 2.1. Let $\kappa \in X(T_M)^+$. The sheaf of modular forms of weight κ is $\omega^\kappa = \phi_* O_{\mathcal{T}}[\kappa']$, where $\phi_* O_{\mathcal{T}}[\kappa']$ is the subsheaf of $\phi_* O_{\mathcal{T}}$, on which $B_M = T_M U_M$ acts by κ' on T_M and trivially on U_M.

A modular form of weight κ on X with coefficients in an O_L-algebra R is thus a global section of ω^κ, so an element of $H^0(X \times_{O_K} R, \omega^\kappa)$. Using the projection $X_{\mathrm{Iw}} \to X$, we similarly define the sheaf ω^κ on X_{Iw}, as well as the modular forms on X_{Iw}.

2B. *Overconvergent modular forms.* For simplicity, we will now assume that there is only one place π of F_0 above p, that is to say that p is inert in F_0. The case with several places does not add any difficulty, and will be treated in Section 4.

We can then define the space of overconvergent modular forms. These will be sections of the sheaf of modular forms defined over a strict neighborhood of the μ-ordinary locus. Recall that we have defined in both cases a degree function

$$\mathrm{Deg} : X_{\mathrm{Iw,rig}} \to \prod_{k=1}^{s} [0, d_k].$$

Since there is only one place above p in F_0, we have $X_{\mathrm{Iw}} = X_{\mathrm{Iw},\pi}$.

We define the μ-ordinary-multiplicative locus as $\mathrm{Deg}^{-1}(\{d_1\} \times \cdots \times \{d_s\})$. By Proposition 1.29 or Proposition 1.34, this locus lies in the μ-ordinary locus.

Definition 2.2. The space of overconvergent modular forms of weight κ is defined as

$$M^\dagger := \mathrm{colim}_{\mathcal{V}}\, H^0(\mathcal{V}, \omega^\kappa),$$

where $\mathcal{V}$ runs over the strict neighborhoods of the μ-ordinary-multiplicative locus in $X_{\mathrm{Iw,rig}}$.

An overconvergent modular form is then defined over a space of the form

$$\mathrm{Deg}^{-1}([d_1 - \varepsilon, d_1] \times \cdots \times [d_s - \varepsilon, d_s])$$

for some $\varepsilon > 0$.

2C. *Hecke operators.* We now define the Hecke operators. These operators will act both on the rigid space and on the space of modular forms. We will fix the weight κ. Explicitly, κ is a collection of integers

$$\big((\kappa_{\sigma,1} \geq \cdots \geq \kappa_{\sigma,a_\sigma}),\ (\lambda_{\sigma,1} \geq \cdots \geq \lambda_{\sigma,b_\sigma})\big)_{\sigma \in \Sigma_\pi}.$$

We recall that we still assume that π is the only place of F_0 above p. To simplify the notation, we define $\kappa_\sigma := \kappa_{\sigma,a_\sigma}$ and $\lambda_\sigma := \lambda_{\sigma,b_\sigma}$.

2C1. *Linear case.* Assume that π is in case (L). Let $1 \le i \le a+b-1$ be an integer, and let C_i be the moduli space defined over K parameterizing $(A, \lambda, \iota, \eta, H_\bullet, L)$, with $(A, \lambda, \iota, \eta, H_\bullet)$ a point of X_{Iw} and $L = L_0 \oplus L_0^\perp$ a subgroup of $A[\pi]$, where L_0 is an O_F-stable subgroup of $A[\pi^+]$ with $A[\pi^+] = H_i \oplus L_0$. We have two morphisms $p_1, p_2 : C_i \to X_{\mathrm{Iw}} \times_{O_K} K$. The morphism p_1 corresponds to forgetting L, and the morphism p_2 is defined as $p_2(A, \lambda, \iota, \eta, H_\bullet, L) = (A/L, \lambda', \iota', \eta', H'_\bullet)$, with

- $H'_j = (H_j + L_0)/L_0$ if $j \le i$;
- $H'_j = ((\pi^+)^{-1}(H_j \cap L_0))/L_0$ if $j > i$.

We take the polarization λ' to be equal to $p \cdot \lambda$, which is a prime-to-p polarization. Let C_i^{an} be the analytic space associated to C_i, and define $C_{i,\mathrm{rig}} := p_1^{-1}(X_{\mathrm{Iw,rig}})$. The morphisms p_1 and p_2 give morphisms $C_{i,\mathrm{rig}} \to X_{\mathrm{Iw,rig}}$.

Definition 2.3. The i-th Hecke operator acting on the subsets of $X_{\mathrm{Iw,rig}}$ is defined by

$$U_{\pi,i}(S) = p_2(p_1^{-1}(S)).$$

This operator preserves the admissible open subsets and quasicompact admissible open subsets.

Let us denote by $p : A \to A/L$ the universal isogeny over C_i. This induces an isomorphism $p^* : \omega_{(A/L)/X} \to \omega_{A/X}$, and thus a morphism $p^*(\kappa) : p_2^*\omega^\kappa \to p_1^*\omega^\kappa$. For every admissible open $\mathcal{U}$ of $X_{\mathrm{Iw,rig}}$, we form the composed morphism

$$\widetilde{U}_{\pi,i} : H^0(U_{\pi,i}(\mathcal{U}), \omega^\kappa) \to H^0(p_1^{-1}(\mathcal{U}), p_2^*\omega^\kappa) \xrightarrow{p^*(\kappa)} H^0(p_1^{-1}(\mathcal{U}), p_1^*\omega^\kappa) \xrightarrow{\mathrm{Tr}_{p_1}} H^0(\mathcal{U}, \omega^\kappa).$$

Definition 2.4. The Hecke operator acting on modular forms is defined by $U_{\pi,i} = (1/p^{N_i})\widetilde{U}_{\pi,i}$ with

$$N_i = \sum_{\sigma \in \Sigma_\pi} \big(\min(i, a_\sigma)\min(a+b-i, b_\sigma) + \max(a_\sigma - i, 0)\kappa_\sigma + \max(i - a_\sigma, 0)\lambda_\sigma\big).$$

We will also write

$$n_i = \sum_{\sigma \in \Sigma_\pi} \min(i, a_\sigma)\min(a+b-i, b_\sigma)$$

for the constant term of N_i, which is independent of the weight.

Let us explain briefly the meaning of the normalization factor N_i. The term $\min(i, a_\sigma)\min(a+b-i, b_\sigma)$ comes from the inseparability degree of the projection p_1. The term $\max(a_\sigma - i, 0)\kappa_\sigma + \max(i - a_\sigma, 0)\lambda_\sigma$ comes from the morphism $p^*(\kappa)$. Indeed, we have the following proposition:

Proposition 2.5. *Let M be a finite extension of $\mathbb{Q}_p$, let $(A, \lambda, \iota, \eta, H_\bullet)$ be an O_M-point of X_{Iw} and let $L = L_0 \oplus L_0^\perp$ be a subgroup of $A[\pi]$, where L_0 is an O_F-stable subgroup of $A[\pi^+]$ with $A[\pi^+] = H_i \oplus L_0$ in generic fiber. Then we have, for all $\sigma \in \Sigma_\pi$,*

$$\deg_\sigma L_0 \geq a_\sigma - i \quad \textit{and} \quad \deg_\sigma L_0^\perp \geq i - a_\sigma.$$

Proof. The group $A[\pi^+]/L_0$ is of height fi and has partial degree $(a_\sigma - \deg_\sigma L_0)_\sigma$. Hence, $a_\sigma - \deg_\sigma L_0 \leq i$ and $\deg_\sigma L_0 \geq a_\sigma - i$. We get the other equality by duality (note that $b_\sigma - (a+b-i) = i - a_\sigma$). □

We have the following proposition concerning the behavior of the Hecke operator regarding the degree function:

Proposition 2.6. *Let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point of $X_{\mathrm{Iw,rig}}$, and $y \in U_{\pi,i}(x)$ corresponding to a subgroup $L \in A[\pi]$ as before. Write $y = (A/L, \lambda, \iota, \eta, H'_\bullet)$; then we have*

$$\deg H'_j \geq \deg H_j$$

for all $1 \leq j \leq a+b-1$. Moreover, we have

$$\deg H'_i = \deg A[\pi^+] - \deg L_0.$$

If $\deg H'_i = \deg H_i$, then $\deg H_i \in \mathbb{Z}$.

Proof. For all $1 \leq j \leq i$, the morphism $H_j \to H'_j$ is an isomorphism in the generic fiber. Thus, $\deg H'_j \geq \deg H_j$. If $i < j \leq a+b-1$, we have

$$\deg H'_j = \deg((\pi^+)^{-1}(H_j \cap L_0)) - \deg L_0.$$

Since $\deg((\pi^+)^{-1} H) = \deg A[\pi^+] + \deg H$ for every subgroup H of $A[\pi^+]$, we get

$$\begin{aligned} \deg H'_j &= \deg A[\pi^+] + \deg(H_j \cap L_0) - \deg L_0 \\ &= \deg(H_j + L_0) + \deg(H_j \cap L_0) - \deg L_0 \geq \deg H_j \end{aligned}$$

from the properties of the degree function.

We have $H'_i = (A[\pi^+]/L_0)$, hence the formula for the degree of H'_i. If $\deg H'_i = \deg H_i$, then we have $\deg H_i + \deg L_0 = \deg A[\pi^+]$, and $A[\pi^+] = H_i \times L_0$. Since $A[\pi^+]$ is a BT_1, so is H_i, and its degree is an integer. □

2C2. *Unitary case.* Assume now that π is in case (U). Let $1 \leq i \leq \frac{1}{2}(a+b)$ be an integer, and let C_i be the moduli space defined over K parameterizing $(A, \lambda, \iota, \eta, H_\bullet, L)$, with $(A, \lambda, \iota, \eta, H_\bullet)$ a point of X_{Iw}, where

- L is an O_F-stable, totally isotropic subgroup of $A[\pi^2]$ such that $A[\pi] = H_i \oplus L[\pi] = H_i^\perp \oplus \pi L$ if $i < \frac{1}{2}(a+b)$;

- L is an O_F-stable, totally isotropic subgroup of $A[\pi]$ such that $A[\pi] = H_i \oplus L$ if $i = \frac{1}{2}(a+b)$.

We have two morphisms $p_1,\ p_2 : C_i \to X_{\mathrm{Iw}} \times_{O_K} K$. The morphism p_1 corresponds to forgetting L, and the morphism p_2 is defined as $p_2(A, \lambda, \iota, \eta, H_\bullet, L) = (A/L, \lambda', \iota', \eta', H'_\bullet)$, with

- $H'_j = (H_j + L)/L$ if $j \leq i$;
- $H'_j = (\pi^{-1}(H_j \cap L) + L)/L$ if $i < j \leq \frac{1}{2}(a+b)$.

We take the polarization λ' to be equal to $p \cdot \lambda$, which is a prime-to-p polarization. Let C_i^{an} be the analytic space associated to C_i, and define $C_{i,\mathrm{rig}} := p_1^{-1}(X_{\mathrm{Iw,rig}})$. The morphisms p_1 and p_2 give morphisms $C_{i,\mathrm{rig}} \to X_{\mathrm{Iw,rig}}$.

Definition 2.7. The i-th Hecke operator acting on the subsets of $X_{\mathrm{Iw,rig}}$ is defined by

$$U_{\pi,i}(S) = p_2(p_1^{-1}(S)).$$

This operator preserves the admissible open subsets and quasicompact admissible open subsets.

Remark 2.8. The condition $A[\pi] = H_i^{\perp} \oplus \pi L$ is actually redundant with the condition $A[\pi] = H_i \oplus L[\pi]$. Indeed, since L is totally isotropic, we have $\pi L \subset L[\pi]^{\perp}$ (we denote by $L[\pi]^{\perp}$ the orthogonal in $A[\pi]$ of $L[\pi]$). Comparing the heights, we see that we have the equality $\pi L = L[\pi]^{\perp}$.

Let us denote by $p : A \to A/L$ the universal isogeny over C_i. This induces an isomorphism $p^* : \omega_{(A/L)/X} \to \omega_{A/X}$, and thus a morphism $p^*(\kappa) : p_2^*\omega^\kappa \to p_1^*\omega^\kappa$. For every admissible open subset $\mathcal{U}$ of $X_{\mathrm{Iw,rig}}$, we form the composed morphism

$$\begin{aligned}\widetilde{U}_{\pi,i} : H^0(U_{\pi,i}(\mathcal{U}), \omega^\kappa) &\to H^0(p_1^{-1}(\mathcal{U}), p_2^*\omega^\kappa) \\ &\xrightarrow{p^*(\kappa)} H^0(p_1^{-1}(\mathcal{U}), p_1^*\omega^\kappa) \xrightarrow{\mathrm{Tr}_{p_1}} H^0(\mathcal{U}, \omega^\kappa).\end{aligned}$$

Definition 2.9. The Hecke operator acting on modular forms is defined by $U_{\pi,i} = (1/p^{N_i})\widetilde{U}_{\pi,i}$ with

$$N_i = \sum_{\sigma \in \Sigma_\pi} \big((a+b)\min(i, a_\sigma) + \max(a_\sigma - i, 0)\kappa_\sigma + \max(b_\sigma - a_\sigma, b_\sigma - i)\lambda_\sigma\big)$$

if $i < \frac{1}{2}(a+b)$, and

$$N_{(a+b)/2} = \sum_{\sigma \in \Sigma_\pi} \big(\tfrac{1}{2}(a+b)a_\sigma + \tfrac{1}{2}(b_\sigma - a_\sigma)\lambda_\sigma\big).$$

We will also write

$$n_i = \sum_{\sigma \in \Sigma_\pi} (a+b)\min(i, a_\sigma)$$

if $i < \frac{1}{2}(a+b)$, and

$$n_{(a+b)/2} = \sum_{\sigma \in \Sigma_\pi} \tfrac{1}{2}(a+b)a_\sigma$$

for the constant term of N_i, which is independent of the weight.

Again, the reason for the normalization factor n_i comes in two parts. The term $(a+b)\min(i, a_\sigma)$ comes from the inseparability degree of the projection p_1. The term $\max(a_\sigma - i, 0)\kappa_\sigma + \max(b_\sigma - a_\sigma, b_\sigma - i)\lambda_\sigma$ comes from the morphism $p^*(\kappa)$. Indeed, we have the following proposition:

Proposition 2.10. *Let M be a finite extension of $\mathbb{Q}_p$, let $(A, \lambda, \iota, \eta, H_\bullet)$ be an O_M-point of X_{Iw} and let L be a subgroup of $A[\pi^2]$ as before. If $i < \frac{1}{2}(a+b)$, we have*

$$\deg_{\sigma_1} L \geq a_\sigma - i \quad \textit{and} \quad \deg_{\sigma_2} L \geq \max(b_\sigma - i, b_\sigma - a_\sigma).$$

If $i = \frac{1}{2}(a+b)$, we have

$$\deg_{\sigma_2} L \geq \tfrac{1}{2}(b_\sigma - a_\sigma).$$

Proof. Suppose first that $i < \frac{1}{2}(a+b)$. The subgroup L being totally isotropic, we get

$$\deg_{\sigma_2} L = \deg_{\sigma_1} L + b_\sigma - a_\sigma.$$

The group $A[\pi]/L[\pi]$ is of height $2fi$, hence we have $\deg_{\sigma_1} A[\pi]/L[\pi] \leq i$. We get

$$\deg_{\sigma_1} L \geq \deg_{\sigma_1} L[\pi] \geq a_\sigma - i.$$

We deduce that $\deg_{\sigma_2} L = \deg_{\sigma_1} L + b_\sigma - a_\sigma \geq \max(b_\sigma - i, b_\sigma - a_\sigma)$.

If $i = \frac{1}{2}(a+b)$, then L is a maximal totally isotropic subgroup of $A[\pi]$. Thus, we have

$$\deg_{\sigma_2} L = \tfrac{1}{2}(b_\sigma - a_\sigma) + \deg_{\sigma_1} L \geq \tfrac{1}{2}(b_\sigma - a_\sigma). \qquad \square$$

We have the following proposition concerning the behavior of the Hecke operator regarding the degree function:

Proposition 2.11. *Let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point of $X_{\mathrm{Iw,rig}}$, and $y \in U_{\pi,i}(x)$ corresponding to a subgroup $L \in A[\pi^2]$ as before. Write $y = (A/L, \lambda, \iota, \eta, H'_\bullet)$; then we have*

$$\deg H'_j \geq \deg H_j$$

for all $1 \leq j \leq \frac{1}{2}(a+b)$. Moreover, if $i < \frac{1}{2}(a+b)$, we have

$$\deg H'_i = 2fi - \deg(L/L[\pi]).$$

If $\deg H'_i = \deg H_i$, then $\deg H_i \in \mathbb{Z}$.

If $i = \frac{1}{2}(a+b)$, *then*

$$\deg H'_{(a+b)/2} = f(a+b) - \deg(L).$$

If $\deg H'_{(a+b)/2} = \deg H_{(a+b)/2}$, *then* $d_{(a+b)/2} - \deg H_{(a+b)/2} \in 2\mathbb{Z}$.

Proof. Suppose first that $i < \frac{1}{2}(a+b)$. For each $1 \le j \le i$, the morphism $H_j \to H'_j$ is an isomorphism in the generic fiber. Thus, $\deg H'_j \ge \deg H_j$. Suppose $i < j \le \frac{1}{2}(a+b)$. Then, observing that $\pi^{-1}(H_j \cap L) \cap L = L[\pi]$, we get

$$\begin{aligned}
\deg H'_j &= \deg(\pi^{-1}(H_j \cap L) + L) - \deg L \\
&\ge \deg(\pi^{-1}(H_j \cap L)) - \deg L[\pi] \\
&\ge \deg A[\pi] + \deg(H_j \cap L) - \deg L[\pi] \\
&= \deg(H_j + L[\pi]) + \deg(H_j \cap L[\pi]) - \deg L[\pi] \\
&\ge \deg H_j
\end{aligned}$$

from the properties of the degree function.

Let us calculate the degree of H'_i. We have

$$\begin{aligned}
\deg H'_i &= \deg(A[\pi] + L)/L \\
&= \deg(\pi^{-1} L[\pi]^{\perp})/L = \deg A[\pi] + \deg L[\pi]^{\perp} - \deg L \\
&= 2fi + \deg L[\pi] - \deg L = 2fi - \deg(L/L[\pi]).
\end{aligned}$$

If $\deg H'_i = \deg H_i$, then we have $\deg H_i + \deg L[\pi] = \deg A[\pi]$ and $A[\pi] = H_i \times L[\pi]$. Since $A[\pi]$ is a BT_1, so is H_i, and its degree is an integer.

Suppose now that $i = \frac{1}{2}(a+b)$. The same argument as before shows that we still have $\deg H'_j \ge \deg H_j$ for all $1 \le j \le \frac{1}{2}(a+b)$. Since $H'_{(a+b)/2} = A[\pi]/L$, we have

$$\deg H'_{(a+b)/2} = \deg A[\pi] - \deg L = f(a+b) - \deg L.$$

If we have the equality $\deg H'_{(a+b)/2} = \deg H_{(a+b)/2}$, then $H_{(a+b)/2}$ is a BT_1, and all its partial degrees are integers. Since $H_{(a+b)/2}$ is totally isotropic, we have the relations

$$\deg_{\sigma_2} H_{(a+b)/2} = \deg_{\sigma_1} H_{(a+b)/2} + \tfrac{1}{2}(b_\sigma - a_\sigma)$$

for all $\sigma \in \Sigma_\pi$. If we write $h_\sigma = \deg_{\sigma_1} H_{(a+b)/2}$, which is an integer under the previous assumption, we have

$$\begin{aligned}
d_{(a+b)/2} - \deg H_{(a+b)/2} &= \sum_{\sigma \in \Sigma_\pi} \big(a_\sigma + \tfrac{1}{2}(a+b) - 2h_\sigma - \tfrac{1}{2}(b_\sigma - a_\sigma)\big) \\
&= 2 \sum_{\sigma \in \Sigma_\pi} (a_\sigma - h_\sigma) \in 2\mathbb{Z}.
\end{aligned}$$

□

We will also need the following useful lemma:

Lemma 2.12. *Let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point of $X_{\mathrm{Iw,rig}}$, and let $L \in A[\pi^2]$ be a totally isotropic subgroup as before, corresponding to a point of $U_{\pi,i}(x)$, with $i < \frac{1}{2}(a+b)$. We have the inequality*

$$\deg(L/L[\pi]) \leq 2fi - \deg A[\pi] + \deg L[\pi].$$

Proof. The multiplication by π gives a morphism $L/L[\pi] \to \pi L$. But, L being totally isotropic, we have the relation $\pi L = L[\pi]^\perp$. We thus get a morphism $L/L[\pi] \to L[\pi]^\perp$, which is an isomorphism in generic fiber. Thus,

$$\deg(L/L[\pi]) \leq \deg L[\pi]^\perp = 2fi - \deg A[\pi] + \deg L[\pi]. \qquad \square$$

3. A classicality result

We will now prove a control theorem, that is to say that an overconvergent modular form is indeed classical under a certain assumption.

3A. *Decomposition of the Hecke operators.* Fix a rational $\varepsilon > 0$. We will fix rationals $\varepsilon_k \in \{\varepsilon, d_k\}$ for $1 \leq k \leq s$. Also fix an integer i between 1 and s such that $\varepsilon_i = \varepsilon$. Since we have assumed that there is only one place π of F_0 above p, we will simply let $\Sigma = \Sigma_\pi$. We define a partition of this set. First, we let, for $1 \leq k \leq s$,

$$\Sigma_k := \{\sigma \in \Sigma : a_\sigma = A_k\}.$$

We will also let $\Sigma_0 := \{\sigma \in \Sigma : a_\sigma = 0\}$ and $\Sigma_{s+1} := \{\sigma \in \Sigma : a_\sigma = a+b\}$ (of course Σ_{s+1} is always empty in case (U)). The sets $(\Sigma_k)_{0 \leq k \leq s+1}$ form a partition of Σ.

From the collection $(\varepsilon_k)_{1 \leq k \leq s}$, we define another partition of Σ. The set S_1 is defined to be

$$S_1 = \Sigma_0 \cup \bigcup_{\substack{k \neq i \\ \varepsilon_k = \varepsilon}} \Sigma_k \cup \Sigma_{s+1}.$$

The complement is the set

$$S_2 = \Sigma_i \cup \bigcup_{k, \varepsilon_k = d_k} \Sigma_k.$$

Define

$$\mathcal{U}_0 := \mathrm{Deg}^{-1}([d_1 - \varepsilon_1, d_1] \times \cdots \times [d_s - \varepsilon_s, d_s]) = \bigcap_{k=1}^{s} \mathrm{Deg}_k^{-1}[d_k - \varepsilon_k, d_k],$$

$$\mathcal{U}_1 := \bigcap_{k \neq i} \mathrm{Deg}_k^{-1}[d_k - \varepsilon_k, d_k].$$

We will define a decomposition of the Hecke operator U_{π, A_i} on subsets of $\mathcal{U}_1$. Fix a rational $\alpha > 0$, and define the integer t to be 1, except in the unitary case and if $i = \frac{1}{2}(a+b)$, where we set $t = 2$. For simplicity, we will write U_i for U_{π, A_i}. Define

$$\mathcal{U} := \mathrm{Deg}_i^{-1}([0, d_i - t(1-\alpha)]) \cap \mathcal{U}_1.$$

Theorem 3.1. *Let $N \geq 1$ and let $\beta < \varepsilon$ be a positive rational. There exists a finite ordered set M_N and a decreasing sequence of admissible open subsets $(\mathcal{U}_k(N))_{k \in M_N}$ of $\mathcal{U}$ such that, for all $k \geq 0$, we have a decomposition of the operator U_i^N on $\mathcal{U}_k(N) \backslash \mathcal{U}_{k+1}(N)$ of the form*

$$U_i^N = \left(\coprod_{j=0}^{N-1} U_i^{N-1-j} \circ T_j \right) \amalg T_N,$$

with $T_0 = U_{i,k,N}^{\mathrm{good}}$,

$$T_j = \coprod_{k_1 \in M_{N-1}, \dots, k_j \in M_{N-j}} U_{i,k_j,N}^{\mathrm{good}} U_{i,k_{j-1},k_j,N}^{\mathrm{bad}} \dots U_{i,k,k_1,N}^{\mathrm{bad}} \quad \textit{for } 0 < j < N,$$

and

$$T_N = \coprod_{k_1 \in M_{N-1}, \dots, k_{N-1} \in M_1} U_{i,k_{N-1},N}^{\mathrm{bad}} U_{i,k_{N-2},k_{N-1},N}^{\mathrm{bad}} \dots U_{i,k,k_1,N}^{\mathrm{bad}}$$

such that

- *the images of $U_{i,j,N}^{\mathrm{good}}$ for $j \in M_k$ are in $\mathrm{Deg}_i^{-1}(]d_i - t(1-\beta), d_i])$;*
- *the images of $U_{i,l,l',N}^{\mathrm{bad}}$ for $l \in M_k$ and $l' \in M_{k-1}$, and $U_{i,l,N}^{\mathrm{bad}}$ for $l \in M_1$ are in $\mathrm{Deg}_i^{-1}\big([0, d_i - t(1-\beta)]\big)$.*

The idea is that the points in $]d_i - t(1-\beta), d_i]$ are "good" (because the overconvergent modular form will be defined at these points), and the points in $\mathrm{Deg}_i^{-1}\big([0, d_i - t(1-\beta)]\big)$ are "bad". The decomposition in the theorem is then made to ensure that all the operators have their image either in good points or in bad points.

Let us describe this decomposition by looking at a point $x \in \mathcal{U}$. The set $U_i(\{x\})$ is finite and has, say, N_1 points in $\mathrm{Deg}_i^{-1}\big([0, d_i - t(1-\beta)]\big)$ and N_2 in its complement $\mathrm{Deg}_i^{-1}\big(]d_i - t(1-\beta), d_i]\big)$. Thus, one can decompose the operator U_i over x as

$$U_i = U_{i,x}^{\mathrm{good}} \amalg U_{i,x}^{\mathrm{bad}}$$

with $U_{i,x}^{\mathrm{bad}}$ corresponding to the N_1 points in $\mathrm{Deg}_i^{-1}\big([0, d_i - t(1-\beta)]\big)$ and $U_{i,x}^{\mathrm{good}}$ to the N_2 other points. This is a decomposition of operators, meaning that the set $U_i(\{x\})$ is the disjoint union of the sets $U_{i,x}^{\mathrm{good}}(\{x\})$ and $U_{i,x}^{\mathrm{bad}}(\{x\})$; moreover, the operators $U_{i,x}^{\mathrm{good}}$ and $U_{i,x}^{\mathrm{bad}}$ induce morphisms $H^0\big(U_i(\{x\}), \omega^\kappa\big) \to H^0(\{x\}, \omega^\kappa)$ for each weight κ, such that

$$U_i f = U_{i,x}^{\mathrm{good}} f + U_{i,x}^{\mathrm{bad}} f$$

for all $f \in H^0\big(U_i(\{x\}), \omega^\kappa\big)$. These morphisms are defined in the same way as the morphism U_i, with the same normalization factor. This gives the decomposition of the theorem for $N = 1$ at the point x.

To get the decomposition for $N = 2$, one needs to study the operator $U_{i,x}^{\text{bad}}$. Let $x_1, \dots, x_{N_1}$ be the elements of the set $U_{i,x}^{\text{bad}}(\{x\})$. We can thus decompose the operator $U_{i,x}^{\text{bad}}$ into

$$U_{i,x}^{\text{bad}} = \coprod_{j=1}^{N_1} U_{i,x,j}^{\text{bad}},$$

where $U_{i,x,j}^{\text{bad}}$ corresponds to the point x_j. Then, for each $1 \le j \le N_1$, we have the decomposition of the operator U_i at the point x_j as $U_{i,x_j}^{\text{good}} \amalg U_{i,x_j}^{\text{bad}}$. One then gets the decomposition of U_i^2 at the point x,

$$U_i^2 = U_i \circ U_{i,x}^{\text{good}} \amalg \coprod_{j=1}^{N_1} U_{i,x_j}^{\text{good}} \circ U_{i,x,j}^{\text{bad}} \amalg \coprod_{j=1}^{N_1} U_{i,x_j}^{\text{bad}} \circ U_{i,x,j}^{\text{bad}}.$$

Repeating this argument, one gets the desired decomposition of the operator U_i^N at the point x. Of course, this decomposition does not have a meaning on the whole open subset $\mathcal{U}$ because the morphisms used to define the different operators will not be finite (the integer N_1 depends on the point x). But, on an adequate subset, this will be the case. That is why one has to construct the subsets $(\mathcal{U}_k(N))_{k \in M_N}$. This construction and the proof of its properties have been done in [Bijakowski et al. 2016, Theorem 4.4.1].

3B. ***Analytic continuation.*** Let f be a section of the sheaf ω^κ on $\mathcal{U}_0$. We will show that f can be extended to $\mathcal{U}_1$ under a certain condition. More precisely, suppose in this section that f is an eigenform for the Hecke operator U_i with eigenvalue α_i and that

$$v(\alpha_i) + n_{A_i} < (1 - 2f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma).$$

The first step is to extend f to $\mathrm{Deg}_i^{-1}(]d_i - t, d_i]) \cap \mathcal{U}_1$. We will use the following proposition:

Proposition 3.2. *Let $0 < \gamma < 1$ be a rational. Then there exists an integer N such that*

$$U_i^N\big(\mathrm{Deg}_i^{-1}([d_i - t\gamma, d_i]) \cap \mathcal{U}_1\big) \subset \mathrm{Deg}_i^{-1}([d_i - \varepsilon, d_i]) \cap \mathcal{U}_1.$$

Proof. The key point is that on $\mathrm{Deg}_i^{-1}([d_i - t\gamma, d_i]) \cap \mathcal{U}_1$ the operator U_i strictly increases the degree function Deg_i. Since $\mathrm{Deg}_i^{-1}([d_i - t\gamma, d_i]) \cap \mathcal{U}_1$ is a quasi-compact open subset of $X_{\mathrm{Iw,rig}}$, one can then apply the argument in [Pilloni 2011, Proposition 2.5]. □

Corollary 3.3. *The overconvergent form f extends to* $\mathrm{Deg}_i^{-1}(]d_i - t, d_i]) \cap \mathcal{U}_1$.

Proof. Let $0 < \gamma < 1$ be a rational. By the previous proposition, there exists an integer N such that

$$U_i^N\big(\mathrm{Deg}_i^{-1}([d_i - t\gamma, d_i]) \cap \mathcal{U}_1\big) \subset \mathrm{Deg}_i^{-1}([d_i - \varepsilon, d_i]) \cap \mathcal{U}_1.$$

The quantity $\alpha_i^{-N} U_i^N f$ is thus defined on $\mathrm{Deg}_i^{-1}([d_i - t\gamma, d_i]) \cap \mathcal{U}_1$. This formula allows us to extend f to $\mathrm{Deg}_i^{-1}(]d_i - t, d_i]) \cap \mathcal{U}_1$. □

The second step is to define some series on $\mathcal{U} := \mathrm{Deg}_i^{-1}\big([0, d_i - t(1-\alpha)]\big) \cap \mathcal{U}_1$ for some sufficiently small rational α. We will use the decomposition of the Hecke operator U_i. First, we recall the definition of the norm of an operator. If $\mathcal{U}$ is an open subset of $X_{\mathrm{Iw,rig}}$, and $T : H^0(T(\mathcal{U}), \omega^\kappa) \to H^0(\mathcal{U}, \omega^\kappa)$ is an operator, then we define the norm of T as

$$\|T\| := \inf\big\{r > 0 : |Tf|_{\mathcal{U}} \leq r|f|_{T(\mathcal{U})} \text{ for all } f \in H^0(T(\mathcal{U}), \omega^\kappa)\big\}.$$

Theorem 3.4. *Suppose that all the operators U_i^{bad} introduced in Theorem 3.1 satisfy the relation*

$$\|\alpha_i^{-1} U_i^{\mathrm{bad}}\| < 1.$$

Then it is possible to construct sections $f_N \in H^0(\mathcal{U}, \omega^\kappa/p^{A_N})$ such that $A_N \to \infty$. Moreover, the functions f_N can be glued together with the initial form f to give an element of $H^0(\mathcal{U}_1, \omega^\kappa)$.

The construction of the series f_N and the proof of the gluing process have been done in [Bijakowski et al. 2016, Section 4.5]. Let us briefly describe the method. We take an element $\beta > 0$ and consider the open subsets $(\mathcal{U}_k(N))_{k \in M_N}$ constructed in Theorem 3.1 for this element. Let K_N be the largest element of M_N; neglecting the bad points for the operator U_i^N gives a function $g_{K_N} \in H^0(\mathcal{U}_{K_N}(N), \omega^\kappa)$ (we refer to [Bijakowski et al. 2016, Definition 4.5.2] for the precise definition of the function g_{K_N}). Now take $\beta' < \beta$, and again apply Theorem 3.1 for this other element: we get open subsets $(\mathcal{U}_k(N)')_{k \in M_N}$. We consider the subset $\mathcal{U}_{K_N - 1}(N)' \backslash \mathcal{U}_{K_N}(N)'$ and the decomposition of U_i^N on it. Neglecting the bad points gives a function $g_{K_N - 1} \in H^0(\mathcal{U}_{K_N-1}(N)' \backslash \mathcal{U}_{K_N}(N)', \omega^\kappa)$. One can then show that g_{K_N} and $g_{K_N - 1}$ can be glued together to give a function modulo p^{A_N}, where A_N is an explicit constant. One thus gets a function $f_N \in H^0(\mathcal{U}_{K_N-1}(N)', \omega^\kappa/p^{A_N})$ (see [Bijakowski et al. 2016, Proposition 4.5.6] for more details). Repeating this argument, one thus extends the function f_N to the whole $\mathcal{U}$. The hypothesis on the operators implies that the sequence $(A_N)_N$ tends to infinity. A gluing lemma ([Bijakowski et al. 2016, Proposition 4.5.7], following [Kassaei 2006]) then ensures the existence of a function $f \in H^0(\mathcal{U}, \omega^\kappa)$, which can be glued with the initial function f, thus getting a section on $\mathcal{U}_1$.

We will now prove that, under the assumption made at the beginning of Section 3B, the condition in the theorem is fulfilled, that is to say that the norm of the operators

$\alpha_i^{-1} U_i^{\text{bad}}$ is strictly less than 1. We will split the discussion between the linear and unitary cases.

3B1. *Linear case.* We recall that the integer t is equal to 1 in this case. Let M be a finite extension of $\mathbb{Q}_p$, let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point of $\mathcal{U}$ defined over O_M, and let $L = L_0 \oplus L_0^\perp$ be a subgroup of $A[\pi]$, where L_0 is an O_F-stable subgroup of $A[\pi^+]$ with $A[\pi^+] = H_{A_i} \oplus L_0$ in generic fiber. Let us write $A' = A/L$, and let H'_{A_i} be the image of H_{A_i} in A/L. Thus,

$$H'_{A_i} = A[\pi^+]/L_0.$$

We suppose that the subgroup L corresponds to a bad point, that is to say that $\deg H'_{A_i} \leq d_i - 1 + \alpha$ for a certain rational $\alpha > 0$. We write

$$\deg_\sigma L_0 = \max(a_\sigma - A_i, 0) + l_\sigma$$

for all $\sigma \in \Sigma$. As is shown by Proposition 2.5, l_σ is a positive rational for all σ. We then have, for all σ,

$$\deg_\sigma H'_{A_i} = a_\sigma - \deg_\sigma L_0 = \min(A_i, a_\sigma) - l_\sigma.$$

We deduce that $\deg H'_{A_i} = d_i - \sum_{\sigma \in \Sigma} l_\sigma$. The condition of being a bad point gives

$$\sum_{\sigma \in \Sigma} l_\sigma \geq 1 - \alpha.$$

Actually, we can control some of the l_σ. First, we prove the following technical lemma:

Lemma 3.5. *Let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point as before and let $H \subset A[\pi^+]$ be an O_F-stable subgroup. If H is of height fA_k and* $\deg H \geq d_k - \varepsilon$, *then* $\deg_\sigma H \geq \min(a_\sigma, A_k) - \varepsilon$ *for all $\sigma \in \Sigma$.*

If H is of height $f(a+b-A_k)$ and $\deg H \leq \deg A[\pi^+] - d_k + \varepsilon$, *then* $\deg_\sigma H \leq \max(a_\sigma - A_k, 0) + \varepsilon$ *for all $\sigma \in \Sigma$.*

Proof. Suppose that H is of height fA_k and $\deg H \geq d_k - \varepsilon$. If $\deg_\sigma H < \min(a_\sigma, A_k) - \varepsilon$ for some $\sigma \in \Sigma$, then

$$\deg H = \sum_{\sigma' \in \Sigma} \deg_{\sigma'} H < \min(a_\sigma, A_k) - \varepsilon + \sum_{\sigma' \neq \sigma} \min(a_{\sigma'}, A_k) = d_k - \varepsilon$$

and we get a contradiction.

If H is of height $f(a+b-A_k)$ and $\deg H \leq \deg A[\pi^+] - d_k + \varepsilon$, then we can apply the previous argument to $A[\pi^+]/H$. We get $\deg_\sigma (A[\pi^+]/H) \geq \min(a_\sigma, A_k) - \varepsilon$ for all $\sigma \in \Sigma$, and therefore $\deg_\sigma H \leq \max(a_\sigma - A_k, 0) + \varepsilon$. $\square$

Now we prove a bound for some of the l_σ.

Lemma 3.6. *If $\sigma \in S_1$, we have $l_\sigma \leq \varepsilon$.*

Proof. If $\sigma \in \Sigma_0 \cup \Sigma_{s+1}$, then $l_\sigma = 0$. If not, $\sigma \in \Sigma_k$, with $1 \leq k \leq s$, $k \neq i$, and $\varepsilon_k = \varepsilon$. Since x is a point of $\mathcal{U}$, we have $\deg H_{A_k} \geq d_k - \varepsilon$. Suppose first that $k < i$. Since H_{A_k} and L_0 are disjoint, the morphism $H_{A_k} \to H'_{A_k} := (H_{A_k} + L_0)/L_0$ is an isomorphism in the generic fiber. Thus, by the properties of the degree function, $\deg H'_{A_k} \geq \deg H_{A_k} \geq d_k - \varepsilon$. By Lemma 3.5, we get $\deg_\sigma H'_{A_k} \geq A_k - \varepsilon$. But we also have $\deg_\sigma H'_{A_k} \leq \deg_\sigma H'_{A_i} = \min(A_i, a_\sigma) - l_\sigma = A_k - l_\sigma$. In conclusion, we get $l_\sigma \leq \varepsilon$.

The case $k > i$ can be treated by duality, considering the group $A[\pi^-]$. We can also give a direct argument. Let us denote the group $L_0/(L_0 \cap H_{A_k})$ by L'_k. Since H_{A_k} and L_0 generate $A[\pi^+]$, the morphism $L'_k \to A[\pi^+]/H_{A_k}$ is an isomorphism in generic fiber, so we get $\deg L'_k \leq \deg(A[\pi^+]/H_{A_k})$. But we have

$$\deg(A[\pi^+]/H_{A_k}) = \deg A[\pi^+] - \deg H_{A_k} \leq \deg A[\pi^+] - d_k + \varepsilon.$$

From Lemma 3.5, we get $\deg_\sigma L'_k \leq \varepsilon$. Since $L_0 \cap H_{A_k}$ is of height $f(A_k - A_i)$, we have

$$l_\sigma = \deg_\sigma L_0 - (A_k - A_i) = \deg_\sigma L'_k + \deg_\sigma(L_0 \cap H_{A_k}) - (A_k - A_i) \leq \varepsilon. \quad \square$$

Remark 3.7. Actually, we can have more control on the l_σ. Indeed, suppose that there exists $\sigma \in \Sigma_k \cap S_1$. If $k < i$, then we have $l_\sigma \leq \varepsilon$ for all $\sigma \in \Sigma_j$ with $j \leq k$. If $k > i$, then we have $l_\sigma \leq \varepsilon$ for all $\sigma \in \Sigma_j$ with $j \geq k$.

Putting together all the calculations made, we get the following result:

Proposition 3.8. *We have*

$$\sum_{\sigma \in S_2} l_\sigma \geq 1 - \alpha - f\varepsilon.$$

Proof. If $\sigma \in S_1$ we have $l_\sigma \leq \varepsilon$, and we also have $\sum_{\sigma \in \Sigma} l_\sigma \geq 1 - \alpha$. Thus,

$$\sum_{\sigma \in S_2} l_\sigma \geq 1 - \alpha - \sum_{\sigma \in S_1} l_\sigma \geq 1 - \alpha - f\varepsilon. \quad \square$$

We can now prove the bound for the norm of the operator U_i^{bad}.

Proposition 3.9. *We have*

$$\|\alpha_i^{-1} U_i^{\text{bad}}\| \leq p^{v(\alpha_i) + n_{A_i} - (1 - \alpha - 2f\varepsilon)\inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma)}.$$

Proof. The term $p^{v(\alpha_i)}$ is the norm of the element α_i^{-1}. Recall also the normalization factor for the Hecke operator,

$$N_{A_i} = n_{A_i} + \sum_{\sigma \in \Sigma_\pi} \big(\max(a_\sigma - A_i, 0)\kappa_\sigma + \max(A_i - a_\sigma, 0)\lambda_\sigma\big).$$

The first term will come in the bound and the second term will be canceled out by the bounds for the partial degrees of L_0 and $L_0^{\perp}$.

Indeed, let us calculate the norm of the morphism $\omega_A^{\kappa} \to \omega_{A/L}^{\kappa}$. Let κ_1 be the weight defined by $((\kappa_\sigma, \dots, \kappa_\sigma), (\lambda_\sigma, \dots, \lambda_\sigma))_{\sigma \in \Sigma}$ and $\kappa_2 = \kappa - \kappa_1$. Thus, $\omega^{\kappa} = \omega^{\kappa_1} \otimes \omega^{\kappa_2}$ and ω^{κ_1} is a line bundle. Since κ_2 has nonnegative coefficients, the morphism $\omega_A^{\kappa_2} \to \omega_{A/L}^{\kappa_2}$ has norm less than 1. It then suffices to study the morphism $\omega_A^{\kappa_1} \to \omega_{A/L}^{\kappa_1}$. But we have

$$\omega_A^{\kappa_1} = \bigotimes_{\sigma \in \Sigma} (\det \omega_{A,\sigma}^{+})^{\kappa_\sigma} \otimes (\det \omega_{A,\sigma}^{-})^{\lambda_\sigma},$$

where $\omega_{A,\sigma}^{+}$ is the submodule of ω_A where O_F acts on σ^{+}, and similarly for $\omega_{A,\sigma}^{-}$. We recall that $L = L_0 \oplus L_0^{\perp}$. The norm of the map $\omega_A^{\kappa_1} \to \omega_{A/L}^{\kappa_1}$ is exactly p^A, with

$$A = -\sum_{\sigma \in \Sigma} (\kappa_\sigma \deg_\sigma L_0 + \lambda_\sigma \deg_\sigma L_0^{\perp}).$$

But recall that $\deg_\sigma L_0 = \max(a_\sigma - A_i, 0) + l_\sigma$ and that

$$\deg_\sigma L_0^{\perp} = A_i - a_\sigma + \deg_\sigma L_0 = \max(A_i - a_\sigma, 0) + l_\sigma.$$

Thus,

$$\begin{aligned} A &= -\sum_{\sigma \in \Sigma} \big(\max(a_\sigma - A_i, 0)\kappa_\sigma + \max(A_i - a_\sigma, 0)\lambda_\sigma + l_\sigma(\kappa_\sigma + \lambda_\sigma) \big) \\ &= -(N_{A_i} - n_{A_i}) + B \end{aligned}$$

with

$$\begin{aligned} B &= -\sum_{\sigma \in \Sigma} (l_\sigma(\kappa_\sigma + \lambda_\sigma)) \leq -\sum_{\sigma \in S_2} (l_\sigma(\kappa_\sigma + \lambda_\sigma)) \\ &\leq -\inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma) \sum_{\sigma \in S_2} l_\sigma \leq -(1 - \alpha - f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma). \end{aligned}$$

Hence the bound for the operator $\alpha_i^{-1} U_i^{\mathrm{bad}}$ follows, noting that $1 - \alpha - f\varepsilon \geq 1 - \alpha - 2f\varepsilon$. □

Since we have made the assumption

$$v(\alpha_i) + n_{A_i} < (1 - 2f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma),$$

we see that, if α is small enough, the norm of the operator $\alpha_i^{-1} U_i^{\mathrm{bad}}$ is strictly less than 1.

Remark 3.10. Actually, we only needed the weaker hypothesis

$$v(\alpha_i) + n_{A_i} < (1 - f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma),$$

but we used ours to have the same condition in both the linear and unitary cases.

3B2. *Unitary case.* We now turn to the case (U). We first assume that $i < \frac{1}{2}(a+b)$, so that the integer t is equal to 1. Let M be a finite extension of $\mathbb{Q}_p$, let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point of $\mathcal{U}$ defined over O_M, and let L be a totally isotropic O_F-stable subgroup of $A[\pi^2]$ such that $A[\pi] = H_i \oplus L[\pi]$. Let $A' = A/L$, and let H'_{A_i} be the image of H_{A_i} in A/L. Thus,

$$H'_{A_i} = (A[\pi] + L)/L.$$

We suppose that the subgroup L corresponds to a bad point, that is to say that $\deg H'_{A_i} \leq d_i - 1 + \alpha$ for a certain rational $\alpha > 0$. We write

$$\deg_{\sigma_1} L = \max(a_\sigma - A_i, 0) + l_\sigma,$$

and

$$\deg_{\sigma_2} L = b_\sigma - a_\sigma + \deg_{\sigma_1} L = \max(b_\sigma - A_i, b_\sigma - a_\sigma) + l_\sigma$$

for all $\sigma \in \Sigma$. As it is shown by Proposition 2.10, l_σ is a positive rational for all σ. We then have $\deg H'_{A_i} = 2fA_i - \deg(L/L[\pi])$; therefore,

$$\deg(L/L[\pi]) \geq 2fA_i - d_i + 1 - \alpha.$$

We also have, by Lemma 2.12, $\deg(L/L[\pi]) \leq 2fA_i - \deg A[\pi] + \deg L[\pi]$. Thus,

$$\deg L[\pi] \geq \deg A[\pi] - d_i + 1 - \alpha.$$

We conclude that

$$\deg L \geq 2fA_i - 2d_i + \deg A[\pi] + 2(1 - \alpha).$$

By an explicit computation, one checks the equality

$$2fA_i - 2d_i + \deg A[\pi] = \sum_{\sigma \in \Sigma} (\max(a_\sigma - A_i, 0) + \max(b_\sigma - A_i, b_\sigma - a_\sigma)).$$

We conclude that the condition of being a bad point gives

$$\sum_{\sigma \in \Sigma} l_\sigma \geq 1 - \alpha.$$

Actually, we can control some of the l_σ. First, we prove the following technical lemma:

Lemma 3.11. *Let $x = (A, \lambda, \iota, \eta, H_\bullet)$ be a point as before, and let $H \subset A[\pi]$ be an O_F-stable subgroup. If H is of height $2fA_k$ and* $\deg H \geq d_k - \varepsilon$, *then* $\deg_{\sigma_1} H \geq \min(a_\sigma, A_k) - \varepsilon$ *for all $\sigma \in \Sigma$.*

If H is of height $2f(a+b-A_k)$ and $\deg H \leq \deg A[\pi] - d_k + \varepsilon$, *then* $\deg_{\sigma_1} H \leq \max(a_\sigma - A_k, 0) + \varepsilon$ *for all $\sigma \in \Sigma$.*

Proof. Suppose that H is of height $2fA_k$ and $\deg H \geq d_k - \varepsilon$. If $\deg_{\sigma_1} H < \min(a_\sigma, A_k) - \varepsilon$ for some $\sigma \in \Sigma$, then

$$\begin{aligned}\deg H &= \sum_{\sigma' \in \Sigma} (\deg_{\sigma'_1} H + \deg_{\sigma'_2} H) \\ &< \min(a_\sigma, A_k) - \varepsilon + A_k + \sum_{\sigma' \neq \sigma} (\min(a_{\sigma'}, A_k) + A_k) = d_k - \varepsilon,\end{aligned}$$

and we get a contradiction.

If H is of height $2f(a+b-A_k)$ and $\deg H \leq \deg A[\pi] - d_k + \varepsilon$, then we can apply the previous argument to $A[\pi]/H$. We get $\deg_{\sigma_1}(A[\pi]/H) \geq \min(a_\sigma, A_k) - \varepsilon$ for all $\sigma \in \Sigma$, and therefore $\deg_{\sigma_1} H \leq \max(a_\sigma - A_k, 0) + \varepsilon$. □

Now we prove a bound for some of the l_σ.

Lemma 3.12. *If $\sigma \in S_1$, we have $l_\sigma \leq 2\varepsilon$.*

Proof. If $\sigma \in \Sigma_0$ then $l_\sigma = 0$. If not, $\sigma \in \Sigma_k$ with $1 \leq k \leq s$, $k \neq i$, and $\varepsilon_k = \varepsilon$. Since x is a point of $\mathcal{U}$, we have $\deg H_{A_k} \geq d_k - \varepsilon$. Suppose first that $k < i$. Since H_{A_k} and $L[\pi]$ are disjoint, the morphism $H_{A_k} \to H'_{A_k} := (H_{A_k} + L[\pi])/L[\pi]$ is an isomorphism in the generic fiber. Thus, by the properties of the degree function, $\deg H'_{A_k} \geq \deg H_{A_k} \geq d_k - \varepsilon$. By Lemma 3.11, we have $\deg_{\sigma_1} H'_{A_k} \geq A_k - \varepsilon$. But we also have $\deg_{\sigma_1} H'_{A_k} \leq \deg_{\sigma_1} A[\pi]/L[\pi] = a_\sigma - \deg_{\sigma_1} L[\pi]$, and therefore $\deg_{\sigma_1} L[\pi] \leq \varepsilon$. Next, we study $H''_{A_k} := (H'_{A_k} + L/L[\pi])/(L/L[\pi])$. Using Lemma 3.11, we get

$$A_k - \varepsilon \leq \deg_{\sigma_1} H_{A''_k} \leq A_k - \deg_{\sigma_1} L/L[\pi].$$

Therefore, $\deg_{\sigma_1}(L/L[\pi]) \leq \varepsilon$. In conclusion, we get $l_\sigma \leq 2\varepsilon$.

The case $k > i$ cannot be treated by duality in this case. Let us denote the group $L[\pi]/(L[\pi] \cap H_{A_k})$ by L'_k. Since H_{A_k} and $L[\pi]$ generate $A[\pi]$, the morphism $L'_k \to A[\pi]/H_{A_k}$ is an isomorphism in generic fiber, so we get $\deg L'_k \leq \deg(A[\pi]/H_{A_k})$. But we have

$$\deg(A[\pi]/H_{A_k}) = \deg A[\pi] - \deg H_{A_k} \leq \deg A[\pi] - d_k + \varepsilon.$$

Lemma 3.11 gives $\deg_{\sigma_1} L'_k \leq \varepsilon$. Since $L[\pi] \cap H_{A_k}$ is of height $2f(A_k - A_i)$, we have

$$\deg_{\sigma_1} L[\pi] = \deg_{\sigma_1} L'_k + \deg_{\sigma_1} L[\pi] \cap H_{A_k} \leq \varepsilon + A_k - A_i.$$

Now we study $A' := A/L[\pi]$ and set $H'_{A_k} := \big(\pi^{-1}(H_{A_k} \cap L[\pi])\big)/L[\pi]$. If H''_{A_k} is the image of the morphism $H'_{A_k} \to A'[\pi]/(L/L[\pi])$, then we get $\deg H''_{A_k} \geq d_k - \varepsilon$. Lemma 3.11 gives

$$A_k - \varepsilon \leq \deg_{\sigma_1} H''_{A_k} \leq A_k - \deg_{\sigma_1}(L/L[\pi]).$$

We deduce that $\deg_{\sigma_1} L/L[\pi] \leq \varepsilon$. Finally, we have

$$l_\sigma = \deg_{\sigma_1} L - (A_k - A_i) \leq 2\varepsilon. \qquad \square$$

Remark 3.13. Actually, we can have more control on the l_σ. Indeed, suppose that there exists $\sigma \in \Sigma_k \cap S_1$. If $k < i$, then we have $l_\sigma \leq \varepsilon$ for all $\sigma \in \Sigma_j$ with $j \leq k$. If $k > i$, then we have $l_\sigma \leq \varepsilon$ for all $\sigma \in \Sigma_j$ with $j \geq k$.

Putting together all the calculations made, we get the following result:

Proposition 3.14. *We have*

$$\sum_{\sigma \in S_2} l_\sigma \geq 1 - \alpha - 2f\varepsilon.$$

Proof. If $\sigma \in S_1$, we have $l_\sigma \leq 2\varepsilon$, and we also have $\sum_{\sigma \in \Sigma} l_\sigma \geq 1 - \alpha$. Thus,

$$\sum_{\sigma \in S_2} l_\sigma \geq 1 - \alpha - \sum_{\sigma \in S_1} l_\sigma \geq 1 - \alpha - 2f\varepsilon. \qquad \square$$

We can now prove the bound for the norm of the operator U_i^{bad}.

Proposition 3.15. *We have*

$$\|\alpha_i^{-1} U_i^{\text{bad}}\| \leq p^{v(\alpha_i) + n_{A_i} - (1-\alpha-2f\varepsilon)\inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma)}.$$

Proof. The proof is exactly the same as in the linear case. $\square$

Since we have made the assumption

$$v(\alpha_i) + n_{A_i} < (1 - 2f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma),$$

we see that, if α is small enough, the norm of the operator $\alpha_i^{-1} U_i^{\text{bad}}$ is strictly less than 1.

We now deal with the case $i = \frac{1}{2}(a+b)$; the integer t is now equal to 2. The subgroup L is now a maximal totally isotropic subgroup of $A[\pi]$. We write $\deg_{\sigma_1} L = l_\sigma$, and we have $\deg_{\sigma_2} L = \frac{1}{2}(b_\sigma - a_\sigma) + l_\sigma$ for all $\sigma \in \Sigma$. Recall that

$$\deg H'_{(a+b)/2} = f(a+b) - \deg L = \sum_{\sigma \in \Sigma} \left(a_\sigma + \tfrac{1}{2}(a+b) - 2l_\sigma\right),$$

so that

$$d_{(a+b)/2} - \deg H'_{(a+b)/2} = 2 \sum_{\sigma \in \Sigma} l_\sigma.$$

The condition of being a bad point gives

$$\sum_{\sigma \in \Sigma} l_\sigma \geq 1 - \alpha$$

for a certain $\alpha > 0$. The calculations are now exactly the same as previously, and the bound obtained in Proposition 3.15 is still valid in the case $i = \frac{1}{2}(a+b)$.

3B3. *Conclusion.* We now recall the analytic continuation result we got, putting together all the results of the past sections.

Recall that we have defined subsets

$$\mathcal{U}_0 := \mathrm{Deg}^{-1}([d_1-\varepsilon_1, d_1] \times \cdots \times [d_s - \varepsilon_s, d_s]) = \bigcap_{k=1}^{s} \mathrm{Deg}_k^{-1}[d_k - \varepsilon_k, d_k],$$

$$\mathcal{U}_1 := \bigcap_{k \neq i} \mathrm{Deg}_k^{-1}[d_k - \varepsilon_k, d_k].$$

We also have defined a partition $\Sigma = S_1 \amalg S_2$. The result of this section is thus contained in the following theorem:

Theorem 3.16. *Let κ be a weight and f a section of ω^κ on $\mathcal{U}_0$. Suppose that f is an eigenform for the Hecke operators U_i, with eigenvalues α_i, for all $1 \leq i \leq s$, and that we have the relations*

$$n_{A_i} + v(\alpha_i) < (1 - 2f\varepsilon) \inf_{\sigma \in S_2} (\kappa_\sigma + \lambda_\sigma).$$

Then f can be extended to a section of ω^κ on $\mathcal{U}_1$.

3C. ***The classicality theorem.*** We recall that we have assumed that p is inert in F_0. We have defined integers $(A_i)_{1 \leq i \leq s}$, and to each of these integers corresponds a canonical subgroup and a relevant Hecke operator. We also have a partition

$$\Sigma = \Sigma_0 \amalg \coprod_{j=1}^{s} \Sigma_j \amalg \Sigma_{s+1}.$$

We can now state the classicality result.

Theorem 3.17. *Let f be an overconvergent modular form of weight κ. Suppose that f is an eigenform for the Hecke operators U_{π, A_i}, with eigenvalues α_i, and that*

$$n_{A_i} + v(\alpha_i) < \inf_{\sigma \in \Sigma_i} (\kappa_\sigma + \lambda_\sigma)$$

for $1 \leq i \leq s$. Then f is classical.

Before giving the proof of the theorem, let us make the conditions explicit. In the case (L), the conditions become

$$v(\alpha_i) + \sum_{j=1}^{f} \min(a_j, A_i) \min(b_j, B_i) < \inf_{\sigma \in \Sigma_i} (\kappa_\sigma + \lambda_\sigma).$$

In the special case where all the (a_i) are distinct and different from 0 and $a+b$, we have $s = d$ conditions, which can be written

$$v(\alpha_\sigma) + \sum_{\sigma' \in \Sigma} \min(a_\sigma, a_{\sigma'}) \min(b_\sigma, b_{\sigma'}) < \kappa_\sigma + \lambda_\sigma,$$

where α_σ is the eigenvalue of U_{π, a_σ}.

In the case (U), they become

$$v(\alpha_i) + \sum_{j=1}^{f} (a+b)\min(a_j, A_i) < \inf_{\sigma\in\Sigma_i} (\kappa_\sigma + \lambda_\sigma)$$

if $A_i < \frac{1}{2}(a+b)$, and

$$v(\alpha_i) + \sum_{j=1}^{f} \tfrac{1}{2}(a+b)a_j < \inf_{\sigma\in\Sigma_i} (\kappa_\sigma + \lambda_\sigma)$$

if $A_i = \frac{1}{2}(a+b)$.

In the special case where all the (a_i) are distinct and different from 0 and $\frac{1}{2}(a+b)$, we have $s = d$ conditions, which can be written

$$v(\alpha_\sigma) + \sum_{\sigma'\in\Sigma} (a+b)\min(a_\sigma, a_{\sigma'}) < \kappa_\sigma + \lambda_\sigma,$$

where α_σ is the eigenvalue of U_{π,a_σ}.

Of course, in the case where the ordinary locus is nonempty, we find the same conditions as in [Bijakowski et al. 2016]. In the case (L), the condition of ordinariness is $a_\sigma = a,\ b_\sigma = b$ for some pair (a, b) and for all $\sigma \in \Sigma$. There is one relevant Hecke operator, $U_{\pi,a}$, and the classicality condition is

$$fab + v(\alpha) < \inf_{\sigma\in\Sigma} (\kappa_\sigma + \lambda_\sigma),$$

where α is the eigenvalue of $U_{\pi,a}$.

In the case (U), the condition of ordinariness is

$$a_\sigma = b_\sigma = \tfrac{1}{2}(a+b)$$

for all $\sigma \in \Sigma$. There is one relevant Hecke operator, $U_{\pi,(a+b)/2}$, and the condition is

$$\tfrac{1}{4} f(a+b)^2 + v(\alpha) < \inf_{\sigma\in\Sigma} (\kappa_\sigma + \lambda_\sigma),$$

where α is the eigenvalue of $U_{\pi,(a+b)/2}$.

Remark 3.18. Since we need to use all the Hecke operators U_{π,A_i} for the classicality result, maybe the relevant operator is a product of these ones. For example, in the linear case the operator $\prod_\sigma U_{\pi,a_\sigma}$ parametrizes complements of (a lifting of) the kernel of the f-th power of the Frobenius on the μ-ordinary locus.

3C1. *Proof of the theorem.* The overconvergent form f is a section of the sheaf ω^κ defined over

$$\mathcal{V}_0 := \bigcap_{i=1}^{s} \mathrm{Deg}_i^{-1}([d_i - \varepsilon, d_i])$$

for some $\varepsilon > 0$. Of course, ε can be taken as small as we want. Let us write $K_i = \inf_{\sigma \in \Sigma_i}(\kappa_\sigma + \lambda_\sigma)$ for all $1 \leq i \leq s$. We put the elements $(K_i)_{1 \leq i \leq s}$ in decreasing order:

$$K_{i_1} \geq K_{i_2} \geq \cdots \geq K_{i_s}.$$

We will use the analytic continuation theorem successively for the operators $U_{i_1}, \ldots, U_{i_s}$, in that order.

We first consider the operator U_{i_1} (we recall that we let $U_i := U_{\pi, A_i}$). We take all the rationals ε_k to be equal to ε. In that case, $S_2 = \Sigma_{i_1}$. We can apply the analytic continuation theorem if the condition

$$n_{A_{i_1}} + v(\alpha_{i_1}) < (1 - 2f\varepsilon) \inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma)$$

is fulfilled. But $\inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma) = K_{i_1}$ and we have, by hypothesis,

$$n_{A_{i_1}} + v(\alpha_{i_1}) < K_{i_1}.$$

If ε is small enough, then we can apply the theorem. We can thus extend f to

$$\mathcal{V}_1 := \bigcap_{i \neq i_1} \mathrm{Deg}_i^{-1}([d_i - \varepsilon, d_i]).$$

We then use the operator U_{i_2}. In this case, we take all the rationals ε_k to be equal to ε, except $\varepsilon_{i_1} = d_{i_1}$. In that case, $S_2 = \Sigma_{i_1} \cup \Sigma_{i_2}$. We can apply the analytic continuation theorem if the condition

$$n_{A_{i_2}} + v(\alpha_{i_2}) < (1 - 2f\varepsilon) \inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma)$$

is fulfilled. But $\inf_{\sigma \in S_2}(\kappa_\sigma + \lambda_\sigma) = \inf(K_{i_1}, K_{i_2}) = K_{i_2}$ and we have, by hypothesis,

$$n_{A_{i_2}} + v(\alpha_{i_2}) < K_{i_2}.$$

If ε is small enough, then we can apply the theorem. We extend f to

$$\mathcal{V}_2 := \bigcap_{i \notin \{i_1, i_2\}} \mathrm{Deg}_i^{-1}([d_i - \varepsilon, d_i]).$$

Repeating this argument, we can extend the overconvergent form f to the whole rigid variety $X_{\mathrm{Iw,rig}}$. We conclude by applying a Koecher principle and a GAGA theorem, which proves that the space

$$H^0(X_{\mathrm{Iw,rig}}, \omega^\kappa)$$

consists of classical modular forms. This will be done in the next section.

3C2. *Compactifications and Koecher's principle.* To complete the proof of the theorem, we need to prove a Koecher principle and to introduce compactifications of the Shimura variety.

Proposition 3.19. *There exists a toroidal compactification $\overline{X}_{\mathrm{Iw}}$ of X_{Iw} defined over O_K. It is a proper scheme, and the sheaf ω^κ extends to $\overline{X}_{\mathrm{Iw}}$.*

The construction of the compactification of X has been done in [Lan 2008], and the one for X_{Iw} follows from [Lan 2015, Section 3] (see also [Bijakowski et al. 2016, Section 5.1]). One can also construct the minimal compactification X^*_{Iw} of X_{Iw}. The Koecher principle states that, under a certain condition, the sections on X_{Iw} automatically extend to the toroidal compactification. The next proposition follows from [Lan 2015, Theorem 8.7] (see also [Bijakowski et al. 2016, Section 5.2]).

Proposition 3.20. *Suppose that the codimension of the boundary of X^*_{Iw} is greater than 2. Then for any O_K-algebra R the restriction map*

$$H^0(\overline{X}_{\mathrm{Iw}} \times R, \omega^\kappa) \to H^0(X_{\mathrm{Iw}} \times R, \omega^\kappa)$$

is an isomorphism.

From this, we deduce a rigid Koecher principle (under the same dimension assumption):

$$H^0(\overline{X}_{\mathrm{Iw,rig}}, \omega^\kappa) \simeq H^0(X_{\mathrm{Iw,rig}}, \omega^\kappa),$$

where $\overline{X}_{\mathrm{Iw,rig}}$ is the rigid space associated to $\overline{X}_{\mathrm{Iw}}$. Finally, since $\overline{X}_{\mathrm{Iw}}$ is proper, we have a GAGA theorem.

Proposition 3.21. *The analytification morphism*

$$H^0(\overline{X}_{\mathrm{Iw}} \times K, \omega^\kappa) \to H^0(\overline{X}_{\mathrm{Iw,rig}}, \omega^\kappa)$$

is an isomorphism.

To conclude, we need to make explicit the dimension condition. We have the following cases:

- If there exists σ such that $a_\sigma b_\sigma = 0$, then the varieties X and X_{Iw} are compact [Lan 2008, Remark 5.3.3.2].

- If this is not the case, the codimension of the boundary of X^*_{Iw} is equal to $d(a+b-1)$. Indeed, the dimension of the variety is equal to $\sum_\sigma a_\sigma b_\sigma$. From [Lan 2008, Theorem 7.2.4.1], there is a stratification of X^*_{Iw}, and any top-dimensional strata of the boundary is isomorphic to a Shimura variety with signatures $(a_\sigma - 1, b_\sigma - 1)$, hence has dimension $\sum_\sigma (a_\sigma - 1)(b_\sigma - 1)$.

  If $d > 1$ or $a + b \geq 3$, then the condition for the Koecher principle is fulfilled.

- If $d = a = b = 1$, we cannot apply the Koecher principle.

Since we have excluded the third case by Hypothesis 1.3, it follows from what has been said before that we have an isomorphism

$$H^0(\overline{X}_{\mathrm{Iw}} \times K, \omega^\kappa) \simeq H^0(X_{\mathrm{Iw,rig}}, \omega^\kappa),$$

that is to say that the space $H^0(X_{\mathrm{Iw,rig}}, \omega^\kappa)$ consists of classical modular forms. If the variety is compact, this is only the GAGA theorem, and in the noncompact case it is a combination of the rigid Koecher principle and the GAGA theorem. This concludes the proof of the theorem.

In the exceptional remaining case, the variety is essentially a modular curve. To prove the classicality theorem, one has to take the cusps into account in the series constructed. We refer to [Kassaei 2006] for more details.

4. The case with several primes above *p*

In this section, we no longer assume that there is only one prime of F_0 above p. We will define the degree functions and the overconvergent modular forms, and prove the classicality result.

Let $\mathcal{P}$ be the set of primes of F_0 above p. We write $\mathcal{P} = \{\pi_1, \dots, \pi_g\}$. Recall that we have defined the Shimura variety of Iwahori level at p

$$X_{\mathrm{Iw}} = X_{\mathrm{Iw},\pi_1} \times_X X_{\mathrm{Iw},\pi_2} \times_X \cdots \times_X X_{\mathrm{Iw},\pi_g}.$$

We will denote by $X_{\mathrm{Iw,rig}}$ the rigid space associated to the scheme X_{Iw}. Let π be an element of $\mathcal{P}$. In the previous sections, we have defined an integer s_π, integers $A_{\pi,1} < \cdots < A_{\pi,s_\pi}$, other integers $d_{\pi,1}, \dots, d_{\pi,s_\pi}$, together with a degree function

$$\mathrm{Deg}_\pi : X_{\mathrm{Iw},\pi,\mathrm{rig}} \to \prod_{i=1}^{s_\pi} [0, d_{\pi,i}].$$

We will define the degree function on $X_{\mathrm{Iw,rig}}$ by

$$\begin{aligned} \mathrm{Deg} : X_{\mathrm{Iw,rig}} &\to \prod_{\pi \in \mathcal{P}} \prod_{i=1}^{s_\pi} [0, d_{\pi,i}], \\ x &\mapsto \big(\mathrm{Deg}_\pi(p_\pi(x))\big)_{\pi \in \mathcal{P}}. \end{aligned}$$

Here p_π is the projection $X_{\mathrm{Iw,rig}} \to X_{\mathrm{Iw},\pi,\mathrm{rig}}$.

Let $X_{\mathrm{Iw}}^{\mu\text{-ord-mult}}$ be the space $\mathrm{Deg}^{-1}(\{d_{\pi,i}\})$. Let us fix a weight κ.

Definition 4.1. The space of overconvergent modular forms of weight κ is defined as

$$M^\dagger := \mathrm{colim}_{\mathcal{V}} H^0(\mathcal{V}, \omega^\kappa),$$

where $\mathcal{V}$ runs over the strict neighborhoods of $X_{\mathrm{Iw}}^{\mu\text{-ord-mult}}$ in $X_{\mathrm{Iw,rig}}$.

We have defined Hecke operators $U_{\pi,A_{\pi,1}}, \dots, U_{\pi,A_{\pi,s_\pi}}$. There is a slight ambiguity for the polarization λ'. If π is the only place above p, one can take $p \cdot \lambda$, which is a prime-to-p polarization. In general, λ' is not well defined. Let x be a totally positive element of O_F such that $v_{\pi'}(x)$ equals 1 if $\pi' = \pi$, and 0 if π' is a place of F_0 above p different from π. Then one takes for λ' the polarization $x \cdot \lambda$.

These operators act both on classical and overconvergent modular forms. We can finally state and prove the classicality theorem.

Theorem 4.2. *Let f be an overconvergent modular form. Suppose that f is an eigenform for the Hecke operators $U_{\pi,A_{\pi,i}}$, with eigenvalues $\alpha_{\pi,i}$. If we have the relations*

$$n_{\pi,A_{\pi,i}} + v(\alpha_{\pi,i}) < \inf_{\sigma \in \Sigma_{\pi,i}} (\kappa_\sigma + \lambda_\sigma)$$

for all $\pi \in \mathcal{P}$, $1 \leq i \leq s_\pi$, then f is classical.

We recall that $n_{\pi,i}$ is the constant term of the normalization factor for the Hecke operator $U_{\pi,i}$ and that $\Sigma_{\pi,i} = \{\sigma \in \Sigma_\pi : a_\sigma = A_{\pi,i}\}$.

Proof. We will use for each place π the analytic continuation theorem we got in the previous section to extend the modular form to the whole rigid variety, and deduce its classicality. Luckily, the order of the places π is not important here.

We start with an overconvergent form f. It is a section of ω^κ defined on a set of the form

$$\mathrm{Deg}^{-1}\Bigg(\prod_{\pi \in \mathcal{P}} \prod_{i=1}^{s_\pi} [d_{\pi,i} - \varepsilon, d_{\pi,i}]\Bigg).$$

First, we consider the place π_1. Using the Hecke operators and the relations satisfied by the eigenvalues, by the previous section one can extend f to

$$\mathrm{Deg}^{-1}\Bigg(\prod_{i=1}^{s_{\pi_1}} [0, d_{\pi_1,i}] \times \prod_{\pi \neq \pi_1} \prod_{i=1}^{s_\pi} [d_{\pi,i} - \varepsilon, d_{\pi,i}]\Bigg).$$

Using this argument successively with the different places, one can extend f to the whole rigid variety $X_{\mathrm{Iw,rig}}$. Using a Koecher principle (if the variety is not compact) and a GAGA theorem as in Section 3C2, one can prove that f is a classical modular form. □

Acknowledgement

The author would like to thank Benoît Stroh, Valentin Hernandez and Vincent Pilloni for helpful discussions.

References

[Berthelot 1996] P. Berthelot, "Cohomologie rigide et cohomologie à supports propres, première partie", preprint IRMAR 96-03, 1996, available at https://perso.univ-rennes1.fr/pierre.berthelot/publis/Cohomologie_Rigide_I.pdf.

[Bijakowski 2014] S. Bijakowski, *Classicité de formes modulaires surconvergentes sur une variété de Shimura*, Ph.D. thesis, Université Paris-Nord, 2014, available at https://tel.archives-ouvertes.fr/tel-01275872.

[Bijakowski et al. 2016] S. Bijakowski, V. Pilloni, and B. Stroh, "Classicité de formes modulaires surconvergentes", *Ann. Math.* **183**:3 (2016), 975–1014.

[Buzzard 2003] K. Buzzard, "Analytic continuation of overconvergent eigenforms", *J. Amer. Math. Soc.* **16**:1 (2003), 29–55. MR Zbl

[Coleman 1996] R. F. Coleman, "Classical and overconvergent modular forms", *Invent. Math.* **124**:1-3 (1996), 215–241. MR Zbl

[Fargues 2010] L. Fargues, "La filtration de Harder–Narasimhan des schémas en groupes finis et plats", *J. Reine Angew. Math.* **645** (2010), 1–39. MR Zbl

[Fargues 2011] L. Fargues, "La filtration canonique des points de torsion des groupes p-divisibles", *Ann. Sci. École Norm. Sup.* (4) **44**:6 (2011), 905–961. MR Zbl

[Johansson 2013] C. Johansson, "Classicality for small slope overconvergent automorphic forms on some compact PEL Shimura varieties of type C", *Math. Ann.* **357**:1 (2013), 51–88. MR Zbl

[Kassaei 2006] P. L. Kassaei, "A gluing lemma and overconvergent modular forms", *Duke Math. J.* **132**:3 (2006), 509–529. MR Zbl

[Kottwitz 1992] R. E. Kottwitz, "Points on some Shimura varieties over finite fields", *J. Amer. Math. Soc.* **5**:2 (1992), 373–444. MR Zbl

[Lan 2008] K.-W. Lan, *Arithmetic compactifications of PEL-type Shimura varieties*, thesis, Harvard University, 2008, available at http://search.proquest.com/docview/304599957. MR Zbl

[Lan 2015] K.-W. Lan, "Integral models of toroidal compactifications with projective cone decompositions", preprint, 2015, available at http://www.math.umn.edu/~kwlan/articles/cpt-ram-nbl.pdf.

[Moonen 2004] B. Moonen, "Serre–Tate theory for moduli spaces of PEL type", *Ann. Sci. École Norm. Sup.* (4) **37**:2 (2004), 223–269. MR Zbl

[Pilloni 2011] V. Pilloni, "Prolongement analytique sur les variétés de Siegel", *Duke Math. J.* **157**:1 (2011), 167–222. MR Zbl

[Pilloni and Stroh 2011] V. Pilloni and B. Stroh, "Surconvergence et classicité: le cas Hilbert", preprint, 2011, available at http://perso.ens-lyon.fr/vincent.pilloni/surconv_hilbert.pdf.

[Sasaki 2010] S. Sasaki, "Analytic continuation of overconvergent Hilbert eigenforms in the totally split case", *Compos. Math.* **146**:3 (2010), 541–560. MR Zbl

[Tian and Xiao 2013] Y. Tian and L. Xiao, "p-adic cohomology and classicality of overconvergent Hilbert modular forms", preprint, 2013. To appear in *Arithmétique p-adique des formes de Hilbert*, Astérisque. arXiv

[Wedhorn 1999] T. Wedhorn, "Ordinariness in good reductions of Shimura varieties of PEL-type", *Ann. Sci. École Norm. Sup.* (4) **32**:5 (1999), 575–618. MR Zbl

Communicated by Brian Conrad
Received 2015-05-29 Revised 2016-04-11 Accepted 2016-05-12

s.bijakowski@imperial.ac.uk *Department of Mathematics, Imperial College, 180 Queen's Gate, London, SW7 2AZ, United Kingdom*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.887

A note on secondary K-theory

Gonçalo Tabuada

We prove that Toën's secondary Grothendieck ring is isomorphic to the Grothendieck ring of smooth proper pretriangulated dg categories previously introduced by Bondal, Larsen, and Lunts. Along the way, we show that those short exact sequences of dg categories in which the first term is smooth proper and the second term is proper are necessarily split. As an application, we prove that the canonical map from the derived Brauer group to the secondary Grothendieck ring has the following injectivity properties: in the case of a commutative ring of characteristic zero, it distinguishes between dg Azumaya algebras associated to nontorsion cohomology classes and dg Azumaya algebras associated to torsion cohomology classes (= ordinary Azumaya algebras); in the case of a field of characteristic zero, it is injective; in the case of a field of positive characteristic p, it restricts to an injective map on the p-primary component of the Brauer group.

1. Introduction and statement of results

A *dg category* $\mathcal{A}$, over a base commutative ring k, is a category enriched over complexes of k-modules; see Section 3. Every (dg) k-algebra A gives rise naturally to a dg category with a single object. Another source of examples is provided by schemes since the category of perfect complexes of every quasicompact quasiseparated k-scheme admits a canonical dg enhancement; see [Lunts and Orlov 2010]. Following [Kontsevich 1998], a dg category $\mathcal{A}$ is called *smooth* if it is compact as a bimodule over itself and *proper* if the complexes of k-modules $\mathcal{A}(x, y)$ are compact. Examples include the finite dimensional k-algebras of finite global dimension (when k is a perfect field) and the dg categories of perfect complexes associated to smooth proper k-schemes. Following [Bondal and Kapranov 1990], a dg category $\mathcal{A}$ is called *pretriangulated* if the Yoneda functor $\mathrm{H}^0(\mathcal{A}) \to \mathcal{D}_c(\mathcal{A})$, $x \mapsto \hat{x}$, is an equivalence of categories. As explained in Section 3, every dg category $\mathcal{A}$ admits a pretriangulated "envelope" $\mathrm{perf}_{\mathrm{dg}}(\mathcal{A})$.

The author was partially supported by the NSF CAREER Award #1350472 and by the Portuguese Foundation for Science and Technology grant PEst-OE/MAT/UI0297/2014.

MSC2010: primary 14A22; secondary 14F22, 16E20, 16H05, 16K50, 18D20.

Keywords: dg category, semiorthogonal decomposition, Azumaya algebra, Brauer group, Grothendieck ring, noncommutative motives, noncommutative algebraic geometry.

Bondal, Larsen, and Lunts [Bondal et al. 2004, §5] introduced the *Grothendieck ring of smooth proper pretriangulated dg categories* $\mathcal{PT}(k)$. This ring is defined by generators and relations. The generators are the quasiequivalence classes of smooth proper pretriangulated dg categories.[1] The relations $[\mathcal{B}]=[\mathcal{A}]+[\mathcal{C}]$ arise from the dg categories $\mathcal{A}, \mathcal{C} \subseteq \mathcal{B}$ for which the triangulated subcategories $\mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C}) \subseteq \mathrm{H}^0(\mathcal{B})$ are admissible and induce a semiorthogonal decomposition $\mathrm{H}^0(\mathcal{B})=\langle \mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C})\rangle$. The multiplication law is given by $\mathcal{A} \bullet \mathcal{B} := \mathrm{perf}_{\mathrm{dg}}(\mathcal{A} \otimes^{\boldsymbol{L}} \mathcal{B})$, where $- \otimes^{\boldsymbol{L}} -$ stands for the derived tensor product of dg categories. Among other applications, Bondal, Larsen, and Lunts constructed an interesting motivic measure with values in $\mathcal{PT}(k)$.

Toën [2009; 2011, §5.4] introduced a "categorified" version of the classical Grothendieck ring named the *secondary Grothendieck ring* $K_0^{(2)}(k)$. By definition, $K_0^{(2)}(k)$ is the quotient of the free abelian group on the Morita equivalence classes of smooth proper dg categories by the relations $[\mathcal{B}]=[\mathcal{A}]+[\mathcal{C}]$ arising from short exact sequences of dg categories $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$. Thanks to [Drinfeld 2004, Proposition 1.6.3], the derived tensor product of dg categories endows $K_0^{(2)}(k)$ with a commutative ring structure. Among other applications, the ring $K_0^{(2)}(k)$ was used in the study of derived loop spaces; see [Ben-Zvi and Nadler 2012; Toën and Vezzosi 2015; 2009].

Theorem 1.1. *The rings $\mathcal{PT}(k)$ and $K_0^{(2)}(k)$ are isomorphic.*

The proof of Theorem 1.1 is based on the fact that those short exact sequences of dg categories $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$ in which $\mathcal{A}$ is smooth proper and $\mathcal{B}$ is proper are necessarily split; see Theorem 4.4. This result is of independent interest. Intuitively speaking, it shows us that the smooth proper dg categories behave as "injective" objects. In the setting of triangulated categories, this idea of "injectivity" goes back to the pioneering work [Bondal and Kapranov 1989].

2. Applications

Following [Toën 2012], a dg k-algebra A is called a *dg Azumaya algebra* if the underlying complex of k-modules is a compact generator of the derived category $\mathcal{D}(k)$ and the canonical morphism $A^{\mathrm{op}} \otimes^{\boldsymbol{L}} A \to \boldsymbol{R}\mathrm{Hom}(A, A)$ in $\mathcal{D}(k)$ is invertible. The ordinary Azumaya algebras (see [Grothendieck 1995a]) are the dg Azumaya algebras whose underlying complex is k-flat and concentrated in degree zero. For every nontorsion étale cohomology class $\alpha \in H^2_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{G}_m)$ there exists a dg Azumaya algebra A_α, representing α, which is *not* Morita equivalent to an ordinary Azumaya algebra; see [Toën 2012, page 584]. Unfortunately, the

[1] Bondal, Larsen, and Lunts worked originally with pretriangulated dg categories. In this case the classical Eilenberg's swindle argument implies that the associated Grothendieck ring is trivial. In order to obtain a nontrivial Grothendieck ring, we need to restrict ourselves to smooth proper dg categories; consult [Tabuada 2005, §7] for further details.

construction of A_α is highly inexplicit; consult [Tabuada and Van den Bergh 2014, Appendix B] for some properties of these mysterious dg algebras. In the case where k is a field, every dg Azumaya algebra is Morita equivalent to an ordinary Azumaya algebra; see [Toën 2012, Proposition 2.12].

The *derived Brauer group* $\mathrm{dBr}(k)$ *of* k is the set of Morita equivalence classes of dg Azumaya algebras. The (multiplicative) group structure is induced by the derived tensor product of dg categories and the inverse of A is given by A^{op}. Since every dg Azumaya algebra is smooth proper, we have a canonical map

$$\mathrm{dBr}(k) \to K_0^{(2)}(k). \tag{2.1}$$

By analogy with the canonical map from the Picard group to the Grothendieck ring

$$\mathrm{Pic}(k) \to K_0(k), \tag{2.2}$$

it is natural to ask[2] if (2.1) is injective. Note that, in contrast with (2.2), the canonical map (2.1) does not seem to admit a "determinant" map in the converse direction. In this note, making use of Theorem 1.1 and of the recent theory of noncommutative motives (see Section 5), we establish several injectivity properties of (2.1).

Recall that k has characteristic zero (resp. positive prime characteristic p) if the kernel of the unique ring homomorphism $\mathbb{Z} \to k$ is $\{0\}$ (resp. $p\mathbb{Z}$).

Theorem 2.3. *Let k be a noetherian*[3] *commutative ring of characteristic zero* (*resp. positive prime characteristic p*) *and A a dg Azumaya algebra which is not Morita equivalent to an ordinary Azumaya algebra. If $K_0(k)_{\mathbb{Q}} \simeq \mathbb{Q}$* (*resp. $K_0(k)_{\mathbb{F}_p} \simeq \mathbb{F}_p$*), *then the image of* $[A]$ *under the canonical map* (2.1) *is nontrivial. Moreover*, *when k is of characteristic zero* (*resp. positive prime characteristic p*), *this nontrivial image is different from the images of the ordinary Azumaya algebras* (*resp. of the ordinary Azumaya algebras whose index is not a multiple of p*).

As proved in [Gabber 1981, Theorem II.1], every torsion étale cohomology class $\alpha \in H^2_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{G}_m)_{\mathrm{tor}}$ can be represented by an ordinary Azumaya algebra A_α. Therefore, Theorem 2.3 shows us that in some cases the canonical map (2.1) distinguishes between torsion and nontorsion classes.

Example 2.4. Let k be the noetherian local ring of the singular point of the normal complex algebraic surface constructed in [Mumford 1961, page 16]. As explained in [Grothendieck 1995b, page 75], k is a local $\mathbb{C}$-algebra of dimension 2 whose étale cohomology group $H^2_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{G}_m)$ contains nontorsion classes α. Therefore,

[2] In the case where k is a field, Toën [2011, §5.4] asked if the canonical map (2.1) is nonzero. This now follows automatically from Theorems 2.5 and 2.7.

[3] As pointed out by the anonymous referee, this assumption can be removed using absolute noetherian approximation. We leave the details to the reader.

since $K_0(k) \simeq \mathbb{Z}$, Theorem 2.3 can be applied to the associated dg Azumaya algebras A_α.

Theorem 2.5. *Let k be a field of characteristic zero. In this case, the canonical map* (2.1) *is injective.*

Example 2.6. (i) When k is the field of real numbers $\mathbb{R}$, we have $\mathrm{Br}(\mathbb{R}) \simeq \mathbb{Z}/2\mathbb{Z}$.

(ii) When k is the field of p-adic numbers $\mathbb{Q}_p$, we have $\mathrm{Br}(\mathbb{Q}_p) \simeq \mathbb{Q}/\mathbb{Z}$.

Theorem 2.7. *Let k be a field of positive characteristic p and A, B two central simple k-algebras. If $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$, where* ind *stands for index, then the images of $[A]$ and $[B]$ under the canonical map* (2.1) *are different. This holds in particular when* $\mathrm{ind}(A)$ *and* $\mathrm{ind}(B)$ *are coprime*[4] *and p divides* $\mathrm{ind}(A)$ *or* $\mathrm{ind}(B)$.

Corollary 2.8. *When k is a field of positive characteristic p, the restriction of the canonical map* (2.1) *to the p-primary torsion subgroup* $\mathrm{Br}(k)\{p\}$ *is injective. Moreover, the image of* $\mathrm{Br}(k)\{p\} - 0$ *is disjoint from the image of* $\bigoplus_{q \neq p} \mathrm{Br}(k)\{q\}$.

Proof. The index and the period of a central simple algebra have the same prime factors. Therefore, the proof of the first claim follows from the divisibility relation $\mathrm{ind}(A^{\mathrm{op}} \otimes B) \mid \mathrm{ind}(A) \cdot \mathrm{ind}(B)$. The proof of the second claim is now clear. $\square$

Example 2.9. Let k be a field of characteristic $p > 0$. Given a character χ and an invertible element $b \in k^\times$, the associated cyclic algebra (χ, b) belongs to the p-primary torsion subgroup $\mathrm{Br}(k)\{p\}$. Moreover, thanks to the work of Albert (see [Gille and Szamuely 2006, Theorem 9.1.8]), every element of $\mathrm{Br}(k)\{p\}$ is of this form. Making use of Corollary 2.8, we hence conclude that the canonical map (2.1) distinguishes all these cyclic algebras. Furthermore, the image of $\mathrm{Br}(k)\{p\} - 0$ in the secondary Grothendieck ring $K_0^{(2)}(k)$ is disjoint from the image of $\bigoplus_{q \neq p} \mathrm{Br}(k)\{q\}$.

Every ring homomorphism $k \to k'$ gives rise to the following commutative square:

$$\begin{tikzcd} \mathrm{dBr}(k) \arrow[r, "(2.1)"] \arrow[d, "- \otimes_k^{\boldsymbol{L}} k'"'] & K_0^{(2)}(k) \arrow[d, "- \otimes_k^{\boldsymbol{L}} k'"] \\ \mathrm{dBr}(k') \arrow[r, "(2.1)"'] & K_0^{(2)}(k') \end{tikzcd}$$

By combining it with Theorems 2.5 and 2.7, we hence obtain the following result:

Corollary 2.10. *Let A and B be dg Azumaya k-algebras. If there exists a ring homomorphism $k \to k'$, with k' a field of characteristic zero* (*resp. positive characteristic p*) *such that* $[A \otimes_k^{\boldsymbol{L}} k'] \neq [B \otimes_k^{\boldsymbol{L}} k']$ *in* $\mathrm{Br}(k')$ (*resp.* $p \mid \mathrm{ind}((A^{\mathrm{op}} \otimes^{\boldsymbol{L}} B) \otimes_k^{\boldsymbol{L}} k')$), *then the images of $[A]$ and $[B]$ under the canonical map* (2.1) *are different.*

[4] When $\mathrm{ind}(A)$ and $\mathrm{ind}(B)$ are coprime we have $\mathrm{ind}(A^{\mathrm{op}} \otimes B) = \mathrm{ind}(A) \cdot \mathrm{ind}(B)$.

Example 2.11 (local rings). Let k be a complete local ring with residue field k' of characteristic zero (resp. positive characteristic p). As proved in [Auslander and Goldman 1961, Theorem 6.5], the assignment $A \mapsto A \otimes_k^{\boldsymbol{L}} k'$ gives rise to a group isomorphism $\mathrm{Br}(k) \simeq \mathrm{Br}(k')$. Therefore, by combining Corollary 2.10 with Theorem 2.5 (resp. Corollary 2.8), we conclude that the restriction of the canonical map (2.1) to the subgroup $\mathrm{Br}(k) \subset \mathrm{dBr}(k)$ (resp. $\mathrm{Br}(k)\{p\} \subset \mathrm{dBr}(k)$) is injective.

Example 2.12 (domains). Let k be a regular noetherian domain of characteristic zero (resp. positive prime characteristic p) with field of fractions k'. Since $H^1_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{Z}) = 0$ and all étale cohomology classes of $H^2_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{G}_m)$ are torsion (see [Grothendieck 1995b, Proposition 1.4]), Theorem II.1 of [Gabber 1981] implies that the derived Brauer group $\mathrm{dBr}(k) \simeq H^1_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{Z}) \times H^2_{\mathrm{et}}(\mathrm{Spec}(k), \mathbb{G}_m)$ agrees with $\mathrm{Br}(k)$. As proved in [Auslander and Goldman 1961, Theorem 7.2], the assignment $A \mapsto A \otimes_k^{\boldsymbol{L}} k'$ gives rise to an injective group homomorphism $\mathrm{Br}(k) \to \mathrm{Br}(k')$. Therefore, by combining Corollary 2.10 with Theorem 2.5 (resp. Corollary 2.8), we conclude that the canonical map (2.1) (resp. the restriction of (2.1) to $\mathrm{Br}(k)\{p\} \subset \mathrm{Br}(k)$) is injective.

Example 2.13 (Weyl algebras). Let F be a field of positive characteristic p. Thanks to [Revoy 1973], the classical Weyl algebra $W_n(F)$, $n \geq 1$, defined as the quotient of $F\langle x_1, \dots, x_n, \partial_1, \dots, \partial_n\rangle$ by the relations $[\partial_i, x_j] = \delta_{ij}$, can be considered as an (ordinary) Azumaya algebra over the ring of polynomials $k := F[x_1^p, \dots, x_n^p, \partial_1^p, \dots, \partial_n^p]$, $n \geq 1$. Consider the composition

$$k := F[x_1^p, \dots, x_n^p, \partial_1^p, \dots, \partial_n^p] \to F[x_1^p, \partial_1^p] \to \mathrm{Frac}(F[x_1^p, \partial_1^p]) =: k',$$

where the first homomorphism sends x_i^p, ∂_i^p, $i > 1$, to zero and $\mathrm{Frac}(F[x_1^p, \partial_1^p])$ denotes the field of fractions of the integral domain $F[x_1^p, \partial_1^p]$. As explained in [Wodzicki 2011, §4], we have $\mathrm{ind}(W_n(F) \otimes_k^{\boldsymbol{L}} k') = p$. Therefore, thanks to Corollary 2.10, we conclude that the image of $W_n(F)$ under (2.1) is nontrivial.

Example 2.14 (algebras of p-symbols). Let F be a field of positive characteristic p, $k := F[x_1^p, \partial_1^p]$ the algebra of polynomials, and $k' := \mathrm{Frac}(F[x_1^p, \partial_1^p])$ the field of fractions. Following [Wodzicki 2011, §1], given elements $a, b \in k$, let us denote by $\mathcal{S}_{ab}(k) \in {}_p\mathrm{Br}(k)$ the associated (ordinary) Azumaya k-algebra of p-symbols. For example, when $a = x_1^p$ and $b = \partial_1^p$, we have $\mathcal{S}_{ab}(k) = W_1(F)$. As proved in [Wodzicki 2011, §6], we have $\mathrm{ind}(\mathcal{S}_{ab}(k) \otimes_k^{\boldsymbol{L}} k') = p$ if and only if

$$b \neq c_0^p + c_1^p a + \cdots + c_{p-1}^p a^{p-1} - c_{p-1} \quad \text{for every } c_0 + c_1 t + \cdots + c_{p-1} t^{p-1} \in k'[t]. \tag{2.15}$$

Therefore, thanks to Corollary 2.10, we conclude that whenever a and b satisfy condition (2.15) the image of $\mathcal{S}_{ab}(k)$ under the canonical map (2.1) is nontrivial.

Remark 2.16 (stronger results). As explained in Section 6, Theorems 2.3, 2.5 and 2.7, and Corollaries 2.8 and 2.10, follow from stronger analogous results where instead of $K_0^{(2)}(k)$ we consider the Grothendieck ring of the category of noncommutative Chow motives; consult Theorems 6.5, 6.12 and 6.20, Corollary 6.22, and Remark 6.23.

3. Background on dg categories

Let $(\mathcal{C}(k), \otimes, k)$ be the symmetric monoidal category of cochain complexes of k-modules. A *dg category* $\mathcal{A}$ is a category enriched over $\mathcal{C}(k)$ and a *dg functor* $F : \mathcal{A} \to \mathcal{B}$ is a functor enriched over $\mathcal{C}(k)$; consult Keller's ICM survey [2006]. Let us denote by $\mathrm{dgcat}(k)$ the category of (small) dg categories and dg functors.

Let $\mathcal{A}$ be a dg category. The opposite dg category $\mathcal{A}^{\mathrm{op}}$ has the same objects and is defined as $\mathcal{A}^{\mathrm{op}}(x, y) := \mathcal{A}(y, x)$. The category $\mathrm{H}^0(\mathcal{A})$ has the same objects as $\mathcal{A}$ and morphisms $\mathrm{H}^0(\mathcal{A})(x, y) := H^0(\mathcal{A}(x, y))$, where $H^0(-)$ stands for the zeroth cohomology.

A *right dg $\mathcal{A}$-module* is a dg functor $M : \mathcal{A}^{\mathrm{op}} \to \mathcal{C}_{\mathrm{dg}}(k)$ with values in the dg category $\mathcal{C}_{\mathrm{dg}}(k)$ of complexes of k-modules. Given $x \in \mathcal{A}$, let us write $\hat{x}$ for the Yoneda right dg $\mathcal{A}$-module defined by $y \mapsto \mathcal{A}(x, y)$. Let $\mathcal{C}(\mathcal{A})$ be the category of right dg $\mathcal{A}$-modules. As explained in [Keller 2006, §3.2], $\mathcal{C}(\mathcal{A})$ carries a Quillen model structure whose weak equivalences (resp. fibrations) are the objectwise quasi-isomorphisms (resp. surjections). The *derived category $\mathcal{D}(\mathcal{A})$ of $\mathcal{A}$* is the associated homotopy category. Let $\mathcal{D}_c(\mathcal{A})$ be the full triangulated subcategory of compact objects. The dg structure of $\mathcal{C}_{\mathrm{dg}}(k)$ makes $\mathcal{C}(\mathcal{A})$ naturally into a dg category $\mathcal{C}_{\mathrm{dg}}(\mathcal{A})$. Let us write $\mathrm{perf}_{\mathrm{dg}}(\mathcal{A})$ for the full dg subcategory of $\mathcal{C}_{\mathrm{dg}}(\mathcal{A})$ consisting of those cofibrant right dg $\mathcal{A}$-modules which belong to $\mathcal{D}_c(\mathcal{A})$. Note that we have the Yoneda dg functor $\mathcal{A} \to \mathrm{perf}_{\mathrm{dg}}(\mathcal{A}) \subset \mathcal{C}_{\mathrm{dg}}(\mathcal{A})$, $x \mapsto \hat{x}$, and that $\mathrm{H}^0(\mathrm{perf}_{\mathrm{dg}}(\mathcal{A})) \simeq \mathcal{D}_c(\mathcal{A})$.

A dg functor $F : \mathcal{A} \to \mathcal{B}$ is called a *quasiequivalence* if the morphisms of k-modules $F(x, y) : \mathcal{A}(x, y) \to \mathcal{B}(F(x), F(y))$ are quasi-isomorphisms and the induced functor $\mathrm{H}^0(F) : \mathrm{H}^0(\mathcal{A}) \to \mathrm{H}^0(\mathcal{B})$ is an equivalence of categories. More generally, F is called a *Morita equivalence* if it induces an equivalence of derived categories $\mathcal{D}(\mathcal{A}) \to \mathcal{D}(\mathcal{B})$; see [Keller 2006, §4.6]. As proved in [Tabuada 2005, Theorem 5.3], $\mathrm{dgcat}(k)$ carries a Quillen model structure whose weak equivalences are the Morita equivalences. Let us denote by $\mathrm{Hmo}(k)$ the associated homotopy category.

The tensor product $\mathcal{A} \otimes \mathcal{B}$ of two dg categories $\mathcal{A}$ and $\mathcal{B}$ is defined as follows: the set of objects is the cartesian product and $(\mathcal{A} \otimes \mathcal{B})((x, w), (y, z)) := \mathcal{A}(x, y) \otimes \mathcal{B}(w, z)$. As explained in [Keller 2006, §4.3], this construction can be derived — and denoted by $\otimes^{\boldsymbol{L}}$ — thus giving rise to a symmetric monoidal structure on $\mathrm{Hmo}(k)$ with $\otimes$-unit the dg category k.

4. Proof of Theorem 1.1

The smooth proper dg categories can be characterized as the dualizable objects of the symmetric monoidal category $\mathrm{Hmo}(k)$; see [Cisinski and Tabuada 2012, §5]. Consequently, Kontsevich's notions of smoothness and properness are invariant under Morita equivalence.

Recall from [Keller 2006, §4.6] that a short exact sequence of dg categories is a sequence of morphisms $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$ in the homotopy category $\mathrm{Hmo}(k)$ inducing an exact sequence of triangulated categories $0 \to \mathcal{D}_c(\mathcal{A}) \to \mathcal{D}_c(\mathcal{B}) \to \mathcal{D}_c(\mathcal{C}) \to 0$ in the sense of Verdier. As proved in [Tabuada 2008, Lemma 10.3], the morphism $\mathcal{A} \to \mathcal{B}$ is isomorphic to an inclusion of dg categories $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{C}$ identifies with Drinfeld's dg quotient $\mathcal{B}/\mathcal{A}$.

Definition 4.1. A short exact sequence of dg categories $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$ is called *split* if the triangulated subcategory $\mathcal{D}_c(\mathcal{A}) \subseteq \mathcal{D}_c(\mathcal{B})$ is admissible.

Remark 4.2. In the case of a split short exact sequence of dg categories we have an induced equivalence between $\mathcal{D}_c(\mathcal{C})$ and the right orthogonal $\mathcal{D}_c(\mathcal{A})^{\perp} \subseteq \mathcal{D}_c(\mathcal{B})$. Consequently, we obtain a semiorthogonal decomposition $\mathcal{D}_c(\mathcal{B}) = \langle \mathcal{D}_c(\mathcal{A}), \mathcal{D}_c(\mathcal{C}) \rangle$.

Let us write $K_0^{(2)}(k)^s$ for the ring defined similarly to $K_0^{(2)}(k)$ but with *split* short exact sequences of dg categories instead of short exact sequences of dg categories.

Proposition 4.3. *The rings $\mathcal{PT}(k)$ and $K_0^{(2)}(k)^s$ are isomorphic.*

Proof. The assignment $\mathcal{A} \mapsto \mathcal{A}$ clearly sends quasiequivalence classes of smooth proper pretriangulated dg categories to Morita equivalence classes of smooth proper dg categories. Let $\mathcal{A}, \mathcal{C} \subseteq \mathcal{B}$ be smooth proper pretriangulated dg categories for which the triangulated subcategories $\mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C}) \subseteq \mathrm{H}^0(\mathcal{B})$ are admissible and induce a semiorthogonal decomposition $\mathrm{H}^0(\mathcal{B}) = \langle \mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C}) \rangle$. Consider the full dg subcategory $\mathcal{B}'$ of $\mathcal{B}$ consisting of those objects which belong to $\mathrm{H}^0(\mathcal{A})$ or to $\mathrm{H}^0(\mathcal{C})$. Thanks to the preceding semiorthogonal decomposition, the inclusion dg functor $\mathcal{B}' \subseteq \mathcal{B}$ is a Morita equivalence. Consider also the dg functor $\pi : \mathcal{B}' \to \mathcal{C}$ which is the identity on $\mathcal{C}$ and which sends all the remaining objects to a fixed zero object 0 of $\mathcal{C}$. Under this notation, we have the split short exact sequence of dg categories $0 \to \mathcal{A} \hookrightarrow \mathcal{B}' \xrightarrow{\pi} \mathcal{C} \to 0$. We hence conclude that the assignment $\mathcal{A} \mapsto \mathcal{A}$ gives rise to a group homomorphism $\mathcal{PT}(k) \to K_0^{(2)}(k)^s$.

As explained in [Tabuada 2005, §5], the pretriangulated dg categories can be (conceptually) characterized as the fibrant objects of the Quillen model structure on $\mathrm{dgcat}(k)$; see Section 3. Moreover, given a dg category $\mathcal{A}$, the Yoneda dg functor $\mathcal{A} \to \mathrm{perf}_{\mathrm{dg}}(\mathcal{A})$, $x \mapsto \hat{x}$, is a fibrant resolution. This implies that $\mathcal{PT}(k) \to K_0^{(2)}(k)^s$ is moreover a surjective ring homomorphism. It remains then only to show its injectivity. Given a split short exact sequence of smooth proper dg categories $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$, which we can assume pretriangulated, recall from Remark 4.2

that we have an associated semiorthogonal decomposition $\mathrm{H}^0(\mathcal{B}) = \langle \mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C})\rangle$. Let us write $\mathcal{C}'$ for the full dg subcategory of $\mathcal{B}$ consisting of those objects which belong to $\mathrm{H}^0(\mathcal{C})$. Note that $\mathcal{C}'$ is pretriangulated and quasiequivalent to $\mathcal{C}$. Note also that since the triangulated subcategory $\mathrm{H}^0(\mathcal{A}) \subseteq \mathrm{H}^0(\mathcal{B})$ is admissible, the triangulated subcategory $\mathrm{H}^0(\mathcal{C}') \simeq \mathrm{H}^0(\mathcal{A})^\perp \subseteq \mathrm{H}^0(\mathcal{B})$ is also admissible. We hence conclude that the relation $[\mathcal{B}] = [\mathcal{A}] + [\mathcal{C}] \Leftrightarrow [\mathcal{B}] = [\mathcal{A}] + [\mathcal{C}']$ holds in $\mathcal{PT}(k)$, and consequently that the surjective ring homomorphism $\mathcal{PT}(k) \to K_0^{(2)}(k)^s$ is moreover injective. $\square$

Thanks to Proposition 4.3, the proof of Theorem 1.1 now follows automatically from the following result of independent interest:

Theorem 4.4. *Let* $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$ *be a short exact sequence of dg categories. If* $\mathcal{A}$ *is smooth proper and* $\mathcal{B}$ *is proper, then the sequence is split.*

Proof. Without loss of generality, we can assume that the dg categories $\mathcal{A}$ and $\mathcal{B}$ are pretriangulated. Let us prove first that the triangulated subcategory $\mathrm{H}^0(\mathcal{A}) \subseteq \mathrm{H}^0(\mathcal{B})$ is right admissible, i.e., that the inclusion functor admits a right adjoint. Given an object $z \in \mathcal{B}$, consider the composition

$$\mathrm{H}^0(\mathcal{A})^{\mathrm{op}} \xrightarrow{\mathrm{H}^0(\mathcal{B}(-,z))} \mathrm{H}^0(\mathrm{perf}_{\mathrm{dg}}(k)) \simeq \mathcal{D}_c(k) \xrightarrow{H^0(-)} \mathrm{mod}(k) \tag{4.5}$$

with values in the category of finitely generated k-modules. Thanks to Proposition 4.8 (with $F = \mathcal{B}(-, z)$), the functor (4.5) is representable. Let us denote by x the representing object. Since the composition (4.5) is naturally isomorphic to the (contravariant) functor $\mathrm{Hom}_{\mathrm{H}^0(\mathcal{B})}(-, z) : \mathrm{H}^0(\mathcal{A})^{\mathrm{op}} \to \mathrm{mod}(k)$, we have $\mathrm{Hom}_{\mathrm{H}^0(\mathcal{A})}(y, x) \simeq \mathrm{Hom}_{\mathrm{H}^0(\mathcal{B})}(y, z)$ for every $y \in \mathcal{A}$. By taking $y = x$, we hence obtain a canonical morphism $\eta : x \to z$ and consequently a distinguished triangle $x \xrightarrow{\eta} z \to \mathrm{cone}(\eta) \to \Sigma(x)$ in the triangulated category $\mathrm{H}^0(\mathcal{B})$. The associated long exact sequences allow us then to conclude that $\mathrm{cone}(\eta)$ belongs to the right orthogonal $\mathrm{H}^0(\mathcal{A})^\perp \subseteq \mathrm{H}^0(\mathcal{B})$. This implies that the triangulated subcategory $\mathrm{H}^0(\mathcal{A}) \subseteq \mathrm{H}^0(\mathcal{B})$ is right admissible. The proof of left admissibility is similar: simply replace $\mathcal{B}(-, z)$ by the covariant dg functor $\mathcal{B}(z, -)$; see Remark 4.9. $\square$

Notation 4.6 (bimodules). Let $\mathcal{A}$ and $\mathcal{B}$ be two dg categories. A *dg* $\mathcal{A}$-$\mathcal{B}$-*bimodule* is a dg functor $\mathrm{B} : \mathcal{A} \otimes^{\boldsymbol{L}} \mathcal{B}^{\mathrm{op}} \to \mathcal{C}_{\mathrm{dg}}(k)$, i.e., a right dg $(\mathcal{A}^{\mathrm{op}} \otimes^{\boldsymbol{L}} \mathcal{B})$-module. Associated to a dg functor $F : \mathcal{A} \to \mathcal{B}$, we have the dg $\mathcal{A}$-$\mathcal{B}$-bimodule

$${}_F\mathrm{B} : \mathcal{A} \otimes^{\boldsymbol{L}} \mathcal{B}^{\mathrm{op}} \to \mathcal{C}_{\mathrm{dg}}(k), \quad (x, z) \mapsto \mathcal{B}(z, F(x)). \tag{4.7}$$

Let us write $\mathrm{rep}(\mathcal{A}, \mathcal{B})$ for the full triangulated subcategory of $\mathcal{D}(\mathcal{A}^{\mathrm{op}} \otimes^{\boldsymbol{L}} \mathcal{B})$ consisting of those dg $\mathcal{A}$-$\mathcal{B}$-bimodules B such that for every object $x \in \mathcal{A}$ the associated right dg $\mathcal{B}$-module $\mathrm{B}(x, -)$ belongs to $\mathcal{D}_c(\mathcal{B})$. Similarly, let $\mathrm{rep}_{\mathrm{dg}}(\mathcal{A}, \mathcal{B})$ be the full dg subcategory of $\mathcal{C}_{\mathrm{dg}}(\mathcal{A}^{\mathrm{op}} \otimes^{\boldsymbol{L}} \mathcal{B})$ consisting of those cofibrant right dg $\mathcal{A}$-$\mathcal{B}$-bimodules which belong to $\mathrm{rep}(\mathcal{A}, \mathcal{B})$. By construction, $\mathrm{H}^0(\mathrm{rep}_{\mathrm{dg}}(\mathcal{A}, \mathcal{B})) \simeq \mathrm{rep}(\mathcal{A}, \mathcal{B})$.

Proposition 4.8 (representability). *Let $\mathcal{A}$ be a smooth proper pretriangulated dg category and $G : \mathrm{H}^0(\mathcal{A})^{\mathrm{op}} \to \mathrm{mod}(k)$ be a (contravariant) functor with values in the category of finitely generated k-modules. Assume that there exists a dg functor $F : \mathcal{A}^{\mathrm{op}} \to \mathrm{perf}_{\mathrm{dg}}(k)$ and a natural isomorphism between G and the composition*

$$\mathrm{H}^0(\mathcal{A})^{\mathrm{op}} \xrightarrow{\mathrm{H}^0(F)} \mathrm{H}^0(\mathrm{perf}_{\mathrm{dg}}(k)) \simeq \mathcal{D}_c(k) \xrightarrow{H^0(-)} \mathrm{mod}(k).$$

Under these assumptions, the functor G is representable.

Remark 4.9. Given a smooth proper pretriangulated dg category $\mathcal{A}$, the opposite dg category $\mathcal{A}^{\mathrm{op}}$ is also smooth, proper, and pretriangulated. Therefore, Proposition 4.8 (with $\mathcal{A}$ replaced by $\mathcal{A}^{\mathrm{op}}$) is also a corepresentability result.

Proof. Following Notation 4.6, let ${}_F\mathrm{B} \in \mathrm{rep}(\mathcal{A}^{\mathrm{op}}, \mathrm{perf}_{\mathrm{dg}}(k))$ be the dg $\mathcal{A}^{\mathrm{op}}$-$\mathrm{perf}_{\mathrm{dg}}(k)$-bimodule associated to the dg functor F. Thanks to Lemma 4.11 below, there exists an object $x \in \mathcal{A}$ and an isomorphism in the triangulated category $\mathrm{rep}(\mathcal{A}^{\mathrm{op}}, \mathrm{perf}_{\mathrm{dg}}(k))$ between the dg $\mathcal{A}^{\mathrm{op}}$-$\mathrm{perf}_{\mathrm{dg}}(k)$-bimodules ${}_F\mathrm{B}$ and ${}_{\hat{x}}\mathrm{B}$. Making use of the functor

$$\mathrm{rep}(\mathcal{A}^{\mathrm{op}}, \mathrm{perf}_{\mathrm{dg}}(k)) \to \mathrm{Fun}_{\Delta}(\mathrm{H}^0(\mathcal{A})^{\mathrm{op}}, \mathcal{D}_c(k)), \quad \mathrm{B} \mapsto - \otimes^{\boldsymbol{L}}_{\mathcal{A}^{\mathrm{op}}} \mathrm{B}, \tag{4.10}$$

where $\mathrm{Fun}_{\Delta}(-,-)$ stands for the category of triangulated functors, we obtain an isomorphism between the functors $- \otimes^{\boldsymbol{L}}_{\mathcal{A}^{\mathrm{op}}} {}_F\mathrm{B} \simeq \mathrm{H}^0(F)$ and $- \otimes^{\boldsymbol{L}}_{\mathcal{A}^{\mathrm{op}}} {}_{\hat{x}}\mathrm{B} \simeq \mathrm{H}^0(\hat{x})$. By composing them with $H^0(-) : \mathcal{D}_c(k) \to \mathrm{mod}(k)$, we hence conclude that G is naturally isomorphic to the representable functor $\mathrm{Hom}_{\mathrm{H}^0(\mathcal{A})}(-, x)$. $\square$

Lemma 4.11. *Given a smooth proper pretriangulated dg category $\mathcal{A}$, the dg functor*

$$\mathcal{A} \to \mathrm{rep}_{\mathrm{dg}}(\mathcal{A}^{\mathrm{op}}, \mathrm{perf}_{\mathrm{dg}}(k)), \quad x \mapsto {}_{\hat{x}}\mathrm{B} \tag{4.12}$$

is a quasiequivalence.

Proof. As proved in [Cisinski and Tabuada 2012, Theorem 5.8], the dualizable objects of the symmetric monoidal category $\mathrm{Hmo}(k)$ are the smooth proper dg categories. Moreover, the dual of a smooth proper dg category $\mathcal{A}$ is the opposite dg category $\mathcal{A}^{\mathrm{op}}$ and the evaluation morphism is given by the dg functor

$$\mathcal{A} \otimes^{\boldsymbol{L}} \mathcal{A}^{\mathrm{op}} \to \mathrm{perf}_{\mathrm{dg}}(k), \quad (x, y) \mapsto \mathcal{A}(y, x). \tag{4.13}$$

The symmetric monoidal category $\mathrm{Hmo}(k)$ is closed; see [Keller 2006, §4.3]. Given dg categories $\mathcal{A}$ and $\mathcal{B}$, their internal Hom is given by $\mathrm{rep}_{\mathrm{dg}}(\mathcal{A}, \mathcal{B})$. Therefore, by adjunction, (4.13) corresponds to the dg functor (4.12). Thanks to the unicity of dualizable objects, we hence conclude that (4.12) is a Morita equivalence. The proof now follows from the fact that a Morita equivalence between pretriangulated dg categories is necessarily a quasiequivalence; see [Tabuada 2005, §5]. $\square$

Alternative proof of Theorem 4.4 when k is a field. Without loss of generality, we can assume that the dg categories $\mathcal{A}$ and $\mathcal{B}$ are pretriangulated. Note first that $\mathcal{B}$ is proper if and only if we have $\sum_n \dim \mathrm{Hom}_{\mathrm{H}^0(\mathcal{B})}(w, z[n]) < \infty$ for any two objects w and z. Since the dg category $\mathcal{A}$ is smooth, the triangulated category $\mathrm{H}^0(\mathcal{A})$ admits a strong generator in the sense of [Bondal and van den Bergh 2003]; see [Lunts 2010, Lemmas 3.5 and 3.6]. Using the fact that the (contravariant) functor (4.5) is cohomological and that the triangulated category $\mathrm{H}^0(\mathcal{A})$ is idempotent complete, we hence conclude from Bondal and Van den Bergh's powerful (co)representability result [2003, Theorem 1.3] that the functor $\mathrm{Hom}_{\mathrm{H}^0(\mathcal{B})}(-, z) : \mathrm{H}^0(\mathcal{A})^{\mathrm{op}} \to \mathrm{vect}(k)$ is representable. The remainder of the proof is now similar.

Remark 4.14 (Orlov's regularity). Following [Orlov 2015, Definition 3.13], a dg category $\mathcal{A}$ is called *regular* if the triangulated category $\mathcal{D}_c(\mathcal{A})$ admits a strong generator in the sense of [Bondal and van den Bergh 2003]. Examples include the dg categories of perfect complexes associated to regular separated noetherian k-schemes. Smoothness implies regularity but the converse does not hold. The preceding proof shows us that Theorem 4.4 holds more generally when $\mathcal{A}$ is a regular proper dg category.

5. Noncommutative motives

For a survey or book on noncommutative motives, we invite the reader to consult [Tabuada 2012] or [Tabuada 2015], respectively. Let $\mathcal{A}$ and $\mathcal{B}$ be two dg categories. As proved in [Tabuada 2005, Corollary 5.10], we have an identification between $\mathrm{Hom}_{\mathrm{Hmo}(k)}(\mathcal{A}, \mathcal{B})$ and the isomorphism classes of the category $\mathrm{rep}(\mathcal{A}, \mathcal{B})$, under which the composition law of $\mathrm{Hmo}(k)$ corresponds to the derived tensor product of bimodules. Since the dg $\mathcal{A}$-$\mathcal{B}$-bimodules (4.7) belong to $\mathrm{rep}(\mathcal{A}, \mathcal{B})$, we hence obtain a symmetric monoidal functor

$$\mathrm{dgcat}(k) \to \mathrm{Hmo}(k), \quad \mathcal{A} \mapsto \mathcal{A}, \quad F \mapsto {}_F\mathrm{B}. \tag{5.1}$$

The *additivization* of $\mathrm{Hmo}(k)$ is the additive category $\mathrm{Hmo}_0(k)$ with the same objects as $\mathrm{Hmo}(k)$ and with morphisms given by $\mathrm{Hom}_{\mathrm{Hmo}_0(k)}(\mathcal{A}, \mathcal{B}) := K_0 \,\mathrm{rep}(\mathcal{A}, \mathcal{B})$, where $K_0 \,\mathrm{rep}(\mathcal{A}, \mathcal{B})$ stands for the Grothendieck group of the triangulated category $\mathrm{rep}(\mathcal{A}, \mathcal{B})$. The composition law is induced by the derived tensor product of bimodules and the symmetric monoidal structure extends by bilinearity from $\mathrm{Hmo}(k)$ to $\mathrm{Hmo}_0(k)$. Note that we have a symmetric monoidal functor

$$\mathrm{Hmo}(k) \to \mathrm{Hmo}_0(k), \quad \mathcal{A} \mapsto \mathcal{A}, \quad \mathrm{B} \mapsto [\mathrm{B}]. \tag{5.2}$$

Given a commutative ring of coefficients R, the *R-linearization* of $\mathrm{Hmo}_0(k)$ is the R-linear category $\mathrm{Hmo}_0(k)_R$ obtained by tensoring the morphisms of $\mathrm{Hmo}_0(k)$ with R. Note that $\mathrm{Hmo}_0(k)_R$ inherits an R-linear symmetric monoidal structure and that we have the symmetric monoidal functor

$$\mathrm{Hmo}_0(k) \to \mathrm{Hmo}_0(k)_R, \quad \mathcal{A} \mapsto \mathcal{A}, \quad [\mathrm{B}] \mapsto [\mathrm{B}]_R. \tag{5.3}$$

Let us denote by $U(-)_R : \mathrm{dgcat}(k) \to \mathrm{Hmo}_0(k)_R$ the composition (5.3)∘(5.2)∘(5.1).

Noncommutative Chow motives. The category of *noncommutative Chow motives* $\mathrm{NChow}(k)_R$ is defined as the idempotent completion of the full subcategory of $\mathrm{Hmo}_0(k)_R$ consisting of the objects $U(\mathcal{A})_R$ with $\mathcal{A}$ a smooth proper dg category. This category is not only R-linear and idempotent complete, but also additive and rigid[5] symmetric monoidal; see [Tabuada 2012, §4]. Given dg categories $\mathcal{A}$ and $\mathcal{B}$, with $\mathcal{A}$ smooth proper, we have $\mathrm{rep}(\mathcal{A}, \mathcal{B}) \simeq \mathcal{D}_c(\mathcal{A}^{\mathrm{op}} \otimes^{\boldsymbol{L}} \mathcal{B})$. Hence, we obtain isomorphisms

$$\mathrm{Hom}_{\mathrm{NChow}(k)_R}(U(\mathcal{A})_R, U(\mathcal{B})_R) := K_0(\mathrm{rep}(\mathcal{A}, \mathcal{B}))_R \simeq K_0(\mathcal{A}^{\mathrm{op}} \otimes^{\boldsymbol{L}} \mathcal{B})_R.$$

When $R = \mathbb{Z}$, we write $\mathrm{NChow}(k)$ instead of $\mathrm{NChow}(k)_{\mathbb{Z}}$ and U instead of $U(-)_{\mathbb{Z}}$.

Noncommutative numerical motives. Given an R-linear, additive, rigid symmetric monoidal category $\mathcal{C}$, its *$\mathcal{N}$-ideal* is defined as

$$\mathcal{N}(a, b) := \{f \in \mathrm{Hom}_{\mathcal{C}}(a, b) \mid \text{for all } g \in \mathrm{Hom}_{\mathcal{C}}(b, a) \text{ we have } \mathrm{tr}(g \circ f) = 0\},$$

where $\mathrm{tr}(g \circ f)$ stands for the categorical trace of the endomorphism $g \circ f$. The category of *noncommutative numerical motives* $\mathrm{NNum}(k)_R$ is defined as the idempotent completion of the quotient of $\mathrm{NChow}(k)_R$ by the ⊗-ideal $\mathcal{N}$. By construction, this category is R-linear, additive, rigid symmetric monoidal, and idempotent complete.

Notation 5.4. In the case where k is a field, we write $\mathrm{CSA}(k)_R$ for the full subcategory of $\mathrm{NNum}(k)_R$ consisting of the objects $U(A)_R$ with A a central simple k-algebra, and $\mathrm{CSA}(k)_R^{\oplus}$ for the closure of $\mathrm{CSA}(k)_R$ under finite direct sums.

The next result is a slight variant of [Marcolli and Tabuada 2014b, Theorem 1.10].

Theorem 5.5 (semisimplicity). *Let k be a commutative ring of characteristic zero* (*resp. positive prime characteristic p*) *and R a field with the same characteristic. If $K_0(k)_{\mathbb{Q}} \simeq \mathbb{Q}$* (*resp. $K_0(k)_{\mathbb{F}_p} \simeq \mathbb{F}_p$*), *then* $\mathrm{NNum}(k)_R$ *is abelian semisimple.*

Proof. As explained in [Cisinski and Tabuada 2012, Example 8.9], Hochschild homology gives rise to an additive symmetric monoidal functor $HH : \mathrm{Hmo}_0(k) \to \mathcal{D}(k)$. The dualizable objects of the derived category $\mathcal{D}(k)$ are the compact ones. Therefore, since the symmetric monoidal subcategory $\mathrm{NChow}(k)$ of $\mathrm{Hmo}_0(k)$ is rigid and every symmetric monoidal functor preserves dualizable objects, the preceding functor restricts to an additive symmetric monoidal functor

$$\mathrm{NChow}(k) \to \mathcal{D}_c(k), \quad U(\mathcal{A}) \mapsto HH(\mathcal{A}). \tag{5.6}$$

[5] Recall that a symmetric monoidal category is called *rigid* if all its objects are dualizable.

Assume first that k is of characteristic zero. Let us write $k_{\mathbb{Q}}$ for the localization of k at the multiplicative set $\mathbb{Z}\setminus\{0\}$. Choose a maximal ideal $\mathfrak{m}$ of $k_{\mathbb{Q}}$ and consider the associated residue field $F := k_{\mathbb{Q}}/\mathfrak{m}$. By composing (5.6) with the base-change (derived) symmetric monoidal functors $\mathcal{D}_c(k) \to \mathcal{D}_c(k_{\mathbb{Q}}) \to \mathcal{D}_c(F)$, we hence obtain an induced $\mathbb{Q}$-linear symmetric monoidal functor

$$\mathrm{NChow}(k)_{\mathbb{Q}} \to \mathcal{D}_c(F), \quad U(\mathcal{A})_{\mathbb{Q}} \mapsto HH(\mathcal{A}) \otimes_k^{\boldsymbol{L}} F. \tag{5.7}$$

Since by assumption we have $\mathrm{End}_{\mathrm{NChow}(k)_{\mathbb{Q}}}(U(k)_{\mathbb{Q}}) = K_0(k)_{\mathbb{Q}} \simeq \mathbb{Q}$, we conclude from André and Kahn's general results [2005, Theorem 1a; 2002, Theorem A.2.10], applied to the functor (5.7), that the category $\mathrm{NNum}(k)_{\mathbb{Q}}$ is abelian semisimple. The proof now follows from the fact that the $\otimes$-ideal $\mathcal{N}$ is compatible with change of coefficients along the field extension $R/\mathbb{Q}$; consult [Bruguières 2000, Proposition 1.4.1] for further details.

Assume now that k is of positive prime characteristic p. Choose a maximal ideal $\mathfrak{m}$ of k and consider the associated residue field $F := k/\mathfrak{m}$. As in the characteristic zero case, we obtain an induced $\mathbb{F}_p$-linear symmetric monoidal functor

$$\mathrm{NChow}(k)_{\mathbb{F}_p} \to \mathcal{D}_c(F), \quad U(\mathcal{A})_{\mathbb{F}_p} \mapsto HH(\mathcal{A}) \otimes_k^{\boldsymbol{L}} F,$$

which allows us to conclude that the category $\mathrm{NNum}(k)_R$ is abelian semisimple. □

6. Proof of Theorems 2.3, 2.5, and 2.7

We start by studying the noncommutative Chow motives of dg Azumaya algebras. These results are of independent interest.

Proposition 6.1. *Let k be a commutative ring and A a dg Azumaya k-algebra which is not Morita equivalent to an ordinary Azumaya algebra. If k is noetherian, then we have $U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}}$ in* $\mathrm{NChow}(k)_{\mathbb{Q}}$ *and $U(A)_{\mathbb{F}_q} \not\simeq U(k)_{\mathbb{F}_q}$ in* $\mathrm{NChow}(k)_{\mathbb{F}_q}$ *for every prime number q.*

Proof. As proved in [Tabuada and Van den Bergh 2014, Theorem B.15], we have $U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}}$ in $\mathrm{NChow}(k)_{\mathbb{Q}}$. The proof that $U(A)_{\mathbb{F}_q} \not\simeq U(k)_{\mathbb{F}_q}$ in $\mathrm{NChow}(k)_{\mathbb{F}_q}$ is similar: simply further assume that q does not divides the positive integers m, n used in [loc. cit.]. □

Proposition 6.2. *Let k be a field, A and B two central simple k-algebras, and R a commutative ring of positive prime characteristic p.*

(i) *If $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$, then $U(A)_R \not\simeq U(B)_R$ in* $\mathrm{NChow}(k)_R$. *Moreover, we have* $\mathrm{Hom}_{\mathrm{NNum}(k)_R}(U(A)_R, U(B)_R) = \mathrm{Hom}_{\mathrm{NNum}(k)_R}(U(B)_R, U(A)_R) = 0$.

(ii) *If $p \nmid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$ and R is a field, then $U(A)_R \simeq U(B)_R$ in* $\mathrm{NChow}(k)_R$.

Proof. As explained in the proof of [Tabuada and Van den Bergh 2014, Proposition 2.25], we have natural identifications $\mathrm{Hom}_{\mathrm{NChow}(k)}(U(A), U(B)) \simeq \mathbb{Z}$, under which the composition law (in $\mathrm{NChow}(k)$)

$$\mathrm{Hom}(U(A), U(B)) \times \mathrm{Hom}(U(B), U(C)) \to \mathrm{Hom}(U(A), U(C))$$

corresponds to the bilinear pairing

$$\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (n, m) \mapsto n \cdot \mathrm{ind}(A^{\mathrm{op}} \otimes B) \cdot \mathrm{ind}(B^{\mathrm{op}} \otimes C) \cdot m.$$

Hence, we obtain natural identifications $\mathrm{Hom}_{\mathrm{NChow}(k)_R}(U(A)_R, U(B)_R) \simeq R$. Moreover, since $\mathrm{ind}(A^{\mathrm{op}} \otimes B) = \mathrm{ind}(B^{\mathrm{op}} \otimes A)$, the composition law (in $\mathrm{NChow}(k)_R$)

$$\mathrm{Hom}(U(A)_R, U(B)_R) \times \mathrm{Hom}(U(B)_R, U(A)_R) \to \mathrm{Hom}(U(A)_R, U(A)_R)$$

corresponds to the bilinear pairing

$$R \times R \to R, \quad (n, m) \mapsto n \cdot \mathrm{ind}(A^{\mathrm{op}} \otimes B)^2 \cdot m; \tag{6.3}$$

similarly with A and B replaced by B and A, respectively.

If $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$, then the bilinear pairing (6.3) is zero. This implies that $U(A)_R \not\simeq U(B)_R$ in $\mathrm{NChow}(k)_R$. Moreover, since the categorical trace of the zero endomorphism is zero, we conclude that all the elements of the R-modules $\mathrm{Hom}_{\mathrm{NChow}(k)_R}(U(A)_R, U(B)_R)$ and $\mathrm{Hom}_{\mathrm{NChow}(k)_R}(U(B)_R, U(A)_R)$ belong to the $\mathcal{N}$-ideal. In other words, we have $\mathrm{Hom}_{\mathrm{NNum}(k)_R}(U(A)_R, U(B)_R) = 0$ and also $\mathrm{Hom}_{\mathrm{NNum}(k)_R}(U(B)_R, U(A)_R) = 0$. This proves item (i).

If $p \nmid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$ and R is a field, then $\mathrm{ind}(A^{\mathrm{op}} \otimes B)$ is invertible in R. It follows then from the bilinear pairing (6.3) that $U(A)_R \simeq U(B)_R$ in $\mathrm{NChow}(k)_R$. This proves item (ii). □

Let $\mathcal{A}, \mathcal{C} \subseteq \mathcal{B}$ be smooth proper pretriangulated dg categories for which the triangulated subcategories $\mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C}) \subseteq \mathrm{H}^0(\mathcal{B})$ are admissible and induce a semiorthogonal decomposition $\mathrm{H}^0(\mathcal{B}) = \langle \mathrm{H}^0(\mathcal{A}), \mathrm{H}^0(\mathcal{C})\rangle$. As proved in [Tabuada 2005, Theorem 6.3], the inclusion dg functors $\mathcal{A}, \mathcal{C} \subseteq \mathcal{B}$ induce an isomorphism $U(\mathcal{A}) \oplus U(\mathcal{C}) \simeq U(\mathcal{B})$ in the additive category $\mathrm{NChow}(k)$. Consequently, if we denote by $K_0(\mathrm{NChow}(k))$ the Grothendieck ring of the symmetric monoidal additive $\mathrm{NChow}(k)$, we obtain a well-defined ring homomorphism

$$\mathcal{PT}(k) \to K_0(\mathrm{NChow}(k)), \quad [\mathcal{A}] \mapsto [U(\mathcal{A})].$$

By precomposing it with the isomorphism $K_0^{(2)}(k) \simeq \mathcal{PT}(k)$ of Theorem 1.1 and with (2.1), we hence obtain the canonical map

$$\mathrm{dBr}(k) \to K_0(\mathrm{NChow}(k)), \quad [A] \mapsto [U(A)]. \tag{6.4}$$

The proof of Theorem 2.3 now follows from the next result.

Theorem 6.5. *Let k be a noetherian commutative ring of characteristic zero (resp. positive prime characteristic p) and A a dg Azumaya algebra which is not Morita equivalent to an ordinary Azumaya algebra. If $K_0(k)_{\mathbb{Q}} \simeq \mathbb{Q}$ (resp. $K_0(k)_{\mathbb{F}_p} \simeq \mathbb{F}_p$), then the image of $[A]$ under the canonical map* (6.4) *is nontrivial. Moreover, when k is of characteristic zero (resp. positive prime characteristic p), this nontrivial image is different from the images of the ordinary Azumaya algebras (resp. of the ordinary Azumaya algebras whose index is not a multiple of p).*

Proof. Similarly to ordinary Azumaya algebras (see [Tabuada and Van den Bergh 2015, Lemma 8.10]), we have the equivalence of symmetric monoidal triangulated categories

$$\mathcal{D}_c(k) \xrightarrow{\simeq} \mathcal{D}_c(A^{\mathrm{op}} \otimes^{\boldsymbol{L}} A), \quad M \mapsto M \otimes^{\boldsymbol{L}} A,$$

where the symmetric monoidal structure on $\mathcal{D}_c(k)$ (resp. $\mathcal{D}_c(A^{\mathrm{op}} \otimes^{\boldsymbol{L}} A)$) is induced by $- \otimes^{\boldsymbol{L}} -$ (resp. $- \otimes^{\boldsymbol{L}}_A -$). Consequently, we obtain an induced ring isomorphism

$$\mathrm{End}_{\mathrm{NChow}(k)}(U(k)) \xrightarrow{\sim} \mathrm{End}_{\mathrm{NChow}(k)}(U(A)). \tag{6.6}$$

Let us prove the first claim. Assume that k is of characteristic zero; the proof of the cases where k is of positive prime characteristic p is similar. By definition of the category of noncommutative Chow motives, the left-hand side of (6.6) is given by the Grothendieck ring $K_0(k)$. Therefore, the assumption $K_0(k)_{\mathbb{Q}} \simeq \mathbb{Q}$ combined with the isomorphism (6.6) implies that $\mathrm{End}_{\mathrm{NChow}(k)_{\mathbb{Q}}}(U(k)_{\mathbb{Q}}) \simeq \mathbb{Q}$ and $\mathrm{End}_{\mathrm{NChow}(k)_{\mathbb{Q}}}(U(A)_{\mathbb{Q}}) \simeq \mathbb{Q}$. By construction of the category of noncommutative numerical motives, we have $\mathrm{End}_{\mathrm{NNum}(k)_{\mathbb{Q}}}(U(k)_{\mathbb{Q}}) \simeq \mathbb{Q}$. Using the fact that $U(A)_{\mathbb{Q}} \in \mathrm{NChow}(k)_{\mathbb{Q}}$ is a $\otimes$-invertible object and that the $\mathbb{Q}$-linear quotient functor $\mathrm{NChow}(k)_{\mathbb{Q}} \to \mathrm{NNum}(k)_{\mathbb{Q}}$ is symmetric monoidal, we hence conclude that $\mathrm{End}_{\mathrm{NNum}(k)_{\mathbb{Q}}}(U(A)_{\mathbb{Q}})$ is also isomorphic to $\mathbb{Q}$. This gives rise to the implication

$$U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}} \text{ in } \mathrm{NChow}(k)_{\mathbb{Q}} \implies U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}} \text{ in } \mathrm{NNum}(k)_{\mathbb{Q}}. \tag{6.7}$$

Note that since the quotient functor $\mathrm{NChow}(k)_{\mathbb{Q}} \to \mathrm{NNum}(k)_{\mathbb{Q}}$ is full, implication (6.7) is equivalent to the fact that every morphism $U(A)_{\mathbb{Q}} \to U(k)_{\mathbb{Q}}$ in $\mathrm{NChow}(k)_{\mathbb{Q}}$ which becomes invertible in $\mathrm{NNum}(k)_{\mathbb{Q}}$ is already invertible in $\mathrm{NChow}(k)_{\mathbb{Q}}$.

Recall from Proposition 6.1 that $U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}}$ in $\mathrm{NChow}(k)_{\mathbb{Q}}$ when k is noetherian. Making use of (6.7), we hence conclude that $U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}}$ in $\mathrm{NNum}(k)_{\mathbb{Q}}$. By definition, we have $[U(A)] = [U(k)]$ in the Grothendieck ring $K_0(\mathrm{NChow}(k))$ if and only if the following condition holds:

$$\text{there exists an } NM \in \mathrm{NChow}(k) \text{ such that } U(A) \oplus NM \simeq U(k) \oplus NM. \tag{6.8}$$

Thanks to Theorem 5.5, the category $\mathrm{NNum}(k)_{\mathbb{Q}}$ is abelian semisimple. Consequently, it satisfies the cancellation property with respect to direct sums. Therefore, if condition (6.8) holds, one would conclude that $U(A)_{\mathbb{Q}} \simeq U(k)_{\mathbb{Q}}$ in $\mathrm{NNum}(k)_{\mathbb{Q}}$, which is a contradiction. This finishes the proof of the first claim.

Let us now prove the second claim. Note first that by combining Proposition 6.1 with implication (6.7), we conclude that

$$U(A)_{\mathbb{Q}} \not\simeq U(k)_{\mathbb{Q}} \text{ in } \mathrm{NNum}(k)_{\mathbb{Q}} \quad \big(\text{resp. } U(A)_{\mathbb{F}_p} \not\simeq U(k)_{\mathbb{F}_p} \text{ in } \mathrm{NNum}(k)_{\mathbb{F}_p}\big). \tag{6.9}$$

Let B be an ordinary Azumaya k-algebra (resp. an ordinary Azumaya k-algebra whose index is not a multiple of p). In the latter case, by definition of index, we can assume without loss of generality that the rank of B is not a multiple of p. We have $[U(A)] = [U(B)]$ in the Grothendieck ring $K_0(\mathrm{NChow}(k))$ if and only if

$$\text{there exists an } NM \in \mathrm{NChow}(k) \text{ such that } U(A) \oplus NM \simeq U(B) \oplus NM. \tag{6.10}$$

Thanks to Theorem 5.5, the category $\mathrm{NNum}(k)_{\mathbb{Q}}$ (resp. $\mathrm{NNum}(k)_{\mathbb{F}_p}$) is abelian semisimple. Consequently, it satisfies the cancellation property with respect to direct sums. Therefore, if (6.10) holds, one would conclude that $U(A)_{\mathbb{Q}} \simeq U(B)_{\mathbb{Q}}$ in $\mathrm{NNum}(k)_{\mathbb{Q}}$ (resp. $U(A)_{\mathbb{F}_p} \simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NNum}(k)_{\mathbb{F}_p}$). On one hand, Corollary B.14 of [Tabuada and Van den Bergh 2014] implies that $U(B)_{\mathbb{Q}} \simeq U(k)_{\mathbb{Q}}$ in $\mathrm{NNum}(k)_{\mathbb{Q}}$. This contradicts the left-hand side of (6.9). On the other hand, since the rank of B is invertible in $\mathbb{F}_p$, the corollary implies that $U(B)_{\mathbb{F}_p} \simeq U(k)_{\mathbb{F}_p}$. This contradicts the right-hand side of (6.9). The proof of the second claim is then finished. □

Proposition 6.11. *Let k be a field and R a field of positive characteristic p. In this case, the category* $\mathrm{CSA}(k)_R^{\oplus}$ (see Notation 5.4) *is equivalent to the category of* $\mathrm{Br}(k)\{p\}$*-graded finite dimensional R-vector spaces.*

Proof. Let A be a central simple k-algebra. Similarly to the proof of Theorem 6.5, we have a ring isomorphism $\mathrm{End}_{\mathrm{CSA}(k)_R}(U(A)_R) \simeq R$.

Let A, B be central simple k-algebras such that $[A], [B] \in \mathrm{Br}(k)\{p\}$ and $[A] \neq [B]$. Since $\mathrm{ind}(A^{\mathrm{op}} \otimes B) \mid \mathrm{ind}(A^{\mathrm{op}}) \cdot \mathrm{ind}(B)$ and $[A] \neq [B]$, we have $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$. Therefore, Proposition 6.2(i) implies that $\mathrm{Hom}_{\mathrm{CSA}(k)_R}(U(A)_R, U(B)_R) = 0$ and also that $\mathrm{Hom}_{\mathrm{CSA}(k)_R}(U(B)_R, U(A)_R) = 0$.

Let A be a central simple k-algebra such that $[A] \in \bigoplus_{q \neq p} \mathrm{Br}(k)\{q\}$. Then, Proposition 6.2(ii) implies that $U(A)_R \simeq U(k)_R$ in $\mathrm{CSA}(k)_R$.

The proof now follows automatically from the combination of the above facts. □

The proof of Theorem 2.5 now follows from the next result.

Theorem 6.12. *Let k be a field of characteristic zero. In this case, the canonical map* (6.4) *is injective.*

Proof. Let A and B be two central simple k-algebras such that $[A] \neq [B]$ in $\mathrm{Br}(k)$. Recall that $\mathrm{ind}(A^{\mathrm{op}} \otimes B) = 1$ if and only if $[A] = [B]$. Therefore, let us choose a prime number p such that $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$. Thanks to Proposition 6.2(i), we have

$U(A)_{\mathbb{F}_p} \not\simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NChow}(k)_{\mathbb{F}_p}$. Consequently, similarly to implication (6.7), we have $U(A)_{\mathbb{F}_p} \not\simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NNum}(k)_{\mathbb{F}_p}$. By definition, we have $[U(A)] = [U(B)]$ in the Grothendieck ring $K_0(\mathrm{NChow}(k))$ if and only if the following condition holds:

$$\text{there exists an } NM \in \mathrm{NChow}(k) \text{ such that } U(A) \oplus NM \simeq U(B) \oplus NM. \tag{6.13}$$

Thanks to Lemma 6.17 below, if condition (6.13) holds, then there exist nonnegative integers $n, m \geq 0$ and a noncommutative numerical motive NM' such that

$$\bigoplus_{i=1}^{n+1} U(A)_{\mathbb{F}_p} \oplus \bigoplus_{j=1}^{m} U(B)_{\mathbb{F}_p} \oplus NM' \simeq \bigoplus_{i=1}^{n} U(A)_{\mathbb{F}_p} \oplus \bigoplus_{j=1}^{m+1} U(B)_{\mathbb{F}_p} \oplus NM' \tag{6.14}$$

in $\mathrm{NNum}(k)_{\mathbb{F}_p}$. Note that the composition bilinear pairing (in $\mathrm{NNum}(k)_{\mathbb{F}_p}$)

$$\mathrm{Hom}(U(A)_{\mathbb{F}_p}, NM') \times \mathrm{Hom}(NM', U(A)_{\mathbb{F}_p}) \to \mathrm{Hom}(U(A)_{\mathbb{F}_p}, U(A)_{\mathbb{F}_p}) \tag{6.15}$$

is zero; similarly for $U(B)_{\mathbb{F}_p}$. This follows from the fact that the right-hand side of (6.15) identifies with $\mathbb{F}_p$, from the fact that the category $\mathrm{NNum}(k)_{\mathbb{F}_p}$ is $\mathbb{F}_p$-linear, and from the fact that the noncommutative numerical motive NM' does not contain $U(A)_{\mathbb{F}_p}$ as a direct summand. The composition bilinear pairing (in $\mathrm{NNum}(k)_{\mathbb{F}_p}$)

$$\mathrm{Hom}(U(A)_{\mathbb{F}_p}, NM') \times \mathrm{Hom}(NM', U(B)_{\mathbb{F}_p}) \to \mathrm{Hom}(U(A)_{\mathbb{F}_p}, U(B)_{\mathbb{F}_p}) \tag{6.16}$$

is also zero; similarly with A and B replaced by B and A, respectively. This follows automatically from the fact that the right-hand side of (6.16) is zero; see Proposition 6.2(i). Now, note that the triviality of the pairings (6.15)–(6.16) implies that the isomorphism (6.14) restricts to an isomorphism

$$U(A)_{\mathbb{F}_p} \oplus \bigoplus_{i=1}^{n} U(A)_{\mathbb{F}_p} \oplus \bigoplus_{j=1}^{m} U(B)_{\mathbb{F}_p} \simeq U(B)_{\mathbb{F}_p} \oplus \bigoplus_{i=1}^{n} U(A)_{\mathbb{F}_p} \oplus \bigoplus_{j=1}^{m} U(B)_{\mathbb{F}_p}$$

in the category $\mathrm{CSA}(k)^{\oplus}_{\mathbb{F}_p} \subset \mathrm{NNum}(k)_{\mathbb{F}_p}$. Since $\mathrm{CSA}(k)^{\oplus}_{\mathbb{F}_p}$ is equivalent to the category of $\mathrm{Br}(k)\{p\}$-graded finite dimensional $\mathbb{F}_p$-vector spaces (see Proposition 6.11), it satisfies the cancellation property with respect to direct sums. Consequently, we conclude from the preceding isomorphism that $U(A)_{\mathbb{F}_p} \simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NNum}(k)_{\mathbb{F}_p}$, which is a contradiction. This finishes the proof. □

Lemma 6.17. *There exist nonnegative integers $n, m \geq 0$ and a noncommutative numerical motive $NM' \in \mathrm{NNum}(k)_{\mathbb{F}_p}$ such that*:

(i) *We have $NM_{\mathbb{F}_p} \simeq \bigoplus_{i=1}^{n} U(A)_{\mathbb{F}_p} \oplus \bigoplus_{j=1}^{m} U(B)_{\mathbb{F}_p} \oplus NM'$ in* $\mathrm{NNum}(k)_{\mathbb{F}_p}$.

(ii) *The noncommutative numerical motive NM' does not contain $U(A)_{\mathbb{F}_p}$ or $U(B)_{\mathbb{F}_p}$ as a direct summand.*

Proof. Recall that the category $\mathrm{NNum}(k)_{\mathbb{F}_p}$ is idempotent complete. Therefore, by inductively splitting the (possible) direct summands $U(A)_{\mathbb{F}_p}$ and $U(B)_{\mathbb{F}_p}$ of the noncommutative numerical motive $NM_{\mathbb{F}_p}$, we obtain an isomorphism

$$NM_{\mathbb{F}_p} \simeq U(A)_{\mathbb{F}_p} \oplus \cdots \oplus U(A)_{\mathbb{F}_p} \oplus U(B)_{\mathbb{F}_p} \oplus \cdots \oplus U(B)_{\mathbb{F}_p} \oplus NM'$$

in $\mathrm{NNum}(k)_{\mathbb{F}_p}$, with NM' satisfying condition (ii). We claim that the number of copies of $U(A)_{\mathbb{F}_p}$ and $U(B)_{\mathbb{F}_p}$ is finite; note that this concludes the proof. We will focus ourselves on the case $U(A)_{\mathbb{F}_p}$; the proof of the case $U(B)_{\mathbb{F}_p}$ is similar. Suppose that the number of copies of $U(A)_{\mathbb{F}_p}$ is infinite. Since we have natural isomorphisms

$$\mathrm{Hom}_{\mathrm{NNum}(k)_{\mathbb{F}_p}}(U(A)_{\mathbb{F}_p}, U(A)_{\mathbb{F}_p}) \simeq \mathbb{F}_p, \tag{6.18}$$

this would allow us to construct an infinite sequence $f_1, f_2, \ldots$ of vectors in the $\mathbb{F}_p$-vector space $\mathrm{Hom}_{\mathrm{NNum}(k)_{\mathbb{F}_p}}(U(A)_{\mathbb{F}_p}, NM_{\mathbb{F}_p})$, with f_i corresponding to the element $1 \in \mathbb{F}_p$ of (6.18), such that $f_1, \ldots, f_r$ is linearly independent for every positive integer r. In other words, this would allow us to conclude that the $\mathbb{F}_p$-vector space $\mathrm{Hom}_{\mathrm{NNum}(k)_{\mathbb{F}_p}}(U(A)_{\mathbb{F}_p}, NM_{\mathbb{F}_p})$ is infinite dimensional. Recall from the proof of [Bruguières 2000, Proposition 1.4.1] that the map $\mathbb{Z} \to \mathbb{F}_p$ gives rise to a surjective homomorphism

$$\mathrm{Hom}_{\mathrm{NNum}(k)}(U(A), NM) \otimes_{\mathbb{Z}} \mathbb{F}_p \twoheadrightarrow \mathrm{Hom}_{\mathrm{NNum}(k)_{\mathbb{F}_p}}(U(A)_{\mathbb{F}_p}, NM_{\mathbb{F}_p}). \tag{6.19}$$

Since, by assumption, the base field k is of characteristic zero, the abelian group $\mathrm{Hom}_{\mathrm{NNum}(k)}(U(A), NM)$ is finitely generated; see [Tabuada and Van den Bergh 2014, Theorem 1.2]. Therefore, we conclude that the right-hand side of (6.19) is a finite dimensional $\mathbb{F}_p$-vector space, which is a contradiction. This finishes the proof. □

The proof of Theorem 2.7 now follows from the next result.

Theorem 6.20. *Let k be a field of positive characteristic p and A, B two central simple k-algebras. If $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$, then the images of $[A]$ and $[B]$ under the canonical map* (6.4) *are different. This holds in particular when* $\mathrm{ind}(A)$ *and* $\mathrm{ind}(B)$ *are coprime and p divides* $\mathrm{ind}(A)$ *or* $\mathrm{ind}(B)$.

Proof. If $p \mid \mathrm{ind}(A^{\mathrm{op}} \otimes B)$, then Proposition 6.2(i) implies that $U(A)_{\mathbb{F}_p} \not\simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NChow}(k)_{\mathbb{F}_p}$. Consequently, similarly to implication (6.7), we conclude that $U(A)_{\mathbb{F}_p} \not\simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NNum}(k)_{\mathbb{F}_p}$. By definition, we have $[U(A)] = [U(B)]$ in the Grothendieck ring $K_0(\mathrm{NChow}(k))$ if and only if the following condition holds:

$$\text{there exists an } NM \in \mathrm{NChow}(k) \text{ such that } U(A) \oplus NM \simeq U(B) \oplus NM. \tag{6.21}$$

Thanks to Theorem 5.5, the category $\mathrm{NNum}(k)_{\mathbb{F}_p}$ is abelian semisimple. Consequently, it satisfies the cancellation property with respect to direct sums. Therefore, if condition (6.21) holds, one would conclude that $U(A)_{\mathbb{F}_p} \simeq U(B)_{\mathbb{F}_p}$ in $\mathrm{NNum}(k)_{\mathbb{F}_p}$, which is a contradiction. This finishes the proof. $\square$

Corollary 6.22. *When k is a field of positive characteristic p, the restriction of the canonical map* (6.4) *to the p-primary torsion subgroup* $\mathrm{Br}(k)\{p\}$ *is injective. Moreover, the image of* $\mathrm{Br}(k)\{p\} - 0$ *is disjoint from the image of* $\bigoplus_{q \neq p} \mathrm{Br}(k)\{q\}$.

Remark 6.23. As proved in [Marcolli and Tabuada 2014a, Theorem 7.1], every ring homomorphism $k \to k'$ gives rise to the following commutative square:

$$\begin{array}{ccc} \mathrm{dBr}(k) & \xrightarrow{(6.4)} & K_0(\mathrm{NChow}(k)) \\ {\scriptstyle -\otimes^{\mathbf{L}}_k k'} \downarrow & & \downarrow {\scriptstyle -\otimes^{\mathbf{L}}_k k'} \\ \mathrm{dBr}(k') & \xrightarrow[(6.4)]{} & K_0(\mathrm{NChow}(k')) \end{array}$$

Therefore, by combining it with Theorems 6.12 and 6.20, we conclude that Corollary 2.10 also holds with (2.1) replaced by (6.4).

Acknowledgements

The author is very grateful to Ofer Gabber, Maxim Kontsevich, and Michel Van den Bergh for useful discussions. He also would like to thank the anonymous referee for his/her comments as well as the Institut des Hautes Études Scientifiques (IHÉS) for its hospitality, excellent working conditions, and stimulating atmosphere.

References

[André and Kahn 2002] Y. André and B. Kahn, "Nilpotence, radicaux et structures monoïdales", *Rend. Sem. Mat. Univ. Padova* **108** (2002), 107–291. MR 1956434 Zbl 1165.18300

[André and Kahn 2005] Y. André and B. Kahn, "Erratum: "Nilpotency, radicals and monoidal structures" [Rend. Sem. Mat. Univ. Padova **108** (2002), 107–291; MR1956434]", *Rend. Sem. Mat. Univ. Padova* **113** (2005), 125–128. MR 2168984 Zbl 1167.18302

[Auslander and Goldman 1961] M. Auslander and O. Goldman, "The Brauer group of a commutative ring", *Trans. Amer. Math. Soc.* **97** (1961), 367–409. MR 0121392 Zbl 0100.26304

[Ben-Zvi and Nadler 2012] D. Ben-Zvi and D. Nadler, "Loop spaces and connections", *J. Topol.* **5**:2 (2012), 377–430. MR 2928082 Zbl 1246.14027

[Bondal and Kapranov 1989] A. I. Bondal and M. M. Kapranov, "Representable functors, Serre functors, and reconstructions", *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:6 (1989), 1183–1205, 1337. In Russian; translated in *Math. USSR-Izv.* **35**:3 (1990), 519–541. MR 1039961

[Bondal and Kapranov 1990] A. I. Bondal and M. M. Kapranov, "Framed triangulated categories", *Mat. Sb.* **181**:5 (1990), 669–683. In Russian; translated in *Math. USSR-Sb.* **70**:1 (1991), 93–107. MR 1055981

[Bondal and van den Bergh 2003] A. Bondal and M. van den Bergh, "Generators and representability of functors in commutative and noncommutative geometry", *Mosc. Math. J.* **3**:1 (2003), 1–36, 258. MR 1996800 Zbl 1135.18302

[Bondal et al. 2004] A. I. Bondal, M. Larsen, and V. A. Lunts, "Grothendieck ring of pretriangulated categories", *Int. Math. Res. Not.* **2004**:29 (2004), 1461–1495. MR 2051435 Zbl 1079.18008

[Bruguières 2000] A. Bruguières, "Tresses et structure entière sur la catégorie des représentations de SL_N quantique", *Comm. Algebra* **28**:4 (2000), 1989–2028. MR 1747368 Zbl 0951.18003

[Cisinski and Tabuada 2012] D.-C. Cisinski and G. Tabuada, "Symmetric monoidal structure on non-commutative motives", *J. K-Theory* **9**:2 (2012), 201–268. MR 2922389 Zbl 1256.19002

[Drinfeld 2004] V. Drinfeld, "DG quotients of DG categories", *J. Algebra* **272**:2 (2004), 643–691. MR 2028075 Zbl 1064.18009

[Gabber 1981] O. Gabber, "Some theorems on Azumaya algebras", pp. 129–209 in *The Brauer group* (Sem., Les Plans-sur-Bex, 1980), edited by M. Kervaire and M. Ojanguren, Lecture Notes in Math. **844**, Springer, Berlin-New York, 1981. MR 611868 Zbl 0472.14013

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics **101**, Cambridge University Press, 2006. MR 2266528 Zbl 1137.12001

[Grothendieck 1995a] A. Grothendieck, "Le groupe de Brauer, I: Algèbres d'Azumaya et interprétations diverses", exposé no. 290, 199–219 in *Séminaire Bourbaki,* 1964/65, Soc. Math. France, Paris, 1995. MR 1608798

[Grothendieck 1995b] A. Grothendieck, "Le groupe de Brauer, II: Théorie cohomologique", exposé no. 297, 287–307 in *Séminaire Bourbaki,* 1965/66, Soc. Math. France, Paris, 1995. MR 1608805

[Keller 2006] B. Keller, "On differential graded categories", pp. 151–190 in *International Congress of Mathematicians*, vol. II, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR 2275593 Zbl 1140.18008

[Kontsevich 1998] M. Kontsevich, "Triangulated categories and geometry", lecture notes, ENS, Paris, 1998, available at http://www.math.uchicago.edu/~mitya/langlands/kontsevich.ps.

[Lunts 2010] V. A. Lunts, "Categorical resolution of singularities", *J. Algebra* **323**:10 (2010), 2977–3003. MR 2609187 Zbl 1202.18006

[Lunts and Orlov 2010] V. A. Lunts and D. O. Orlov, "Uniqueness of enhancement for triangulated categories", *J. Amer. Math. Soc.* **23**:3 (2010), 853–908. MR 2629991 Zbl 1197.14014

[Marcolli and Tabuada 2014a] M. Marcolli and G. Tabuada, "Noncommutative Artin motives", *Selecta Math.* (*N.S.*) **20**:1 (2014), 315–358. MR 3147418 Zbl 1296.14002

[Marcolli and Tabuada 2014b] M. Marcolli and G. Tabuada, "Noncommutative motives, numerical equivalence, and semi-simplicity", *Amer. J. Math.* **136**:1 (2014), 59–75. MR 3163353 Zbl 1315.18020

[Mumford 1961] D. Mumford, "The topology of normal singularities of an algebraic surface and a criterion for simplicity", *Inst. Hautes Études Sci. Publ. Math.* **9** (1961), 5–22. MR 0153682 Zbl 0108.16801

[Orlov 2015] D. Orlov, "Smooth and proper noncommutative schemes and gluing of DG categories", preprint, 2015. arXiv 1402.7364

[Revoy 1973] P. Revoy, "Algèbres de Weyl en caractéristique p", *C. R. Acad. Sci. Paris Sér. A-B* **276** (1973), A225–A228. MR 0335564 Zbl 0265.16007

[Tabuada 2005] G. Tabuada, "Invariants additifs de DG-catégories", *Int. Math. Res. Not.* **2005**:53 (2005), 3309–3339. MR 2196100 Zbl 1094.18006

[Tabuada 2008] G. Tabuada, "Higher K-theory via universal invariants", *Duke Math. J.* **145**:1 (2008), 121–206. MR 2451292 Zbl 1166.18007

[Tabuada 2012] G. Tabuada, "A guided tour through the garden of noncommutative motives", pp. 259–276 in *Topics in noncommutative geometry*, edited by G. Cortiñas, Clay Math. Proc. **16**, American Mathematical Society, Providence, RI, 2012. MR 2986869 Zbl 1316.19001

[Tabuada 2015] G. Tabuada, *Noncommutative motives*, University Lecture Series **63**, American Mathematical Society, Providence, RI, 2015. MR 3379910 Zbl 06501755

[Tabuada and Van den Bergh 2014] G. Tabuada and M. Van den Bergh, "Noncommutative motives of separable algebras", preprint, 2014. arXiv 1411.7970

[Tabuada and Van den Bergh 2015] G. Tabuada and M. Van den Bergh, "Noncommutative motives of Azumaya algebras", *J. Inst. Math. Jussieu* **14**:2 (2015), 379–403. MR 3315059 Zbl 06420271

[Toën 2009] B. Toën, "Secondary K-theory", Talk at the Workshop on Topological Field Theories, Northwestern University, May 2009, available at https://www.ma.utexas.edu/users/benzvi/GRASP/lectures/NWTFT/toennw.pdf.

[Toën 2011] B. Toën, "Lectures on dg-categories", pp. 243–302 in *Topics in algebraic and topological K-theory*, edited by G. Cortiñas, Lecture Notes in Math. **2008**, Springer, Berlin, 2011. MR 2762557 Zbl 1216.18013

[Toën 2012] B. Toën, "Derived Azumaya algebras and generators for twisted derived categories", *Invent. Math.* **189**:3 (2012), 581–652. MR 2957304 Zbl 1275.14017

[Toën and Vezzosi 2009] B. Toën and G. Vezzosi, "Chern character, loop spaces and derived algebraic geometry", pp. 331–354 in *Algebraic topology*, edited by N. A. Baas et al., Abel Symp. **4**, Springer, Berlin, 2009. MR 2597742 Zbl 1177.14042

[Toën and Vezzosi 2015] B. Toën and G. Vezzosi, "Caractères de Chern, traces équivariantes et géométrie algébrique dérivée", *Selecta Math.* (*N.S.*) **21**:2 (2015), 449–554. MR 3338682 Zbl 06441726

[Wodzicki 2011] M. Wodzicki, "Algebras of p-symbols, noncommutative p-residue, and the Brauer group", pp. 283–304 in *Noncommutative geometry and global analysis*, edited by A. Connes et al., Contemp. Math. **546**, American Mathematical Society, Providence, RI, 2011. MR 2815141 Zbl 1246.58024

Communicated by Yuri Manin
Received 2015-09-21 Revised 2016-02-04 Accepted 2016-03-16

tabuada@math.mit.edu *Departamento de Matemática e Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica, Portugal*

Current address: *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States*

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.907

Nef cones of Hilbert schemes of points on surfaces

Barbara Bolognese, Jack Huizenga, Yinbang Lin, Eric Riedl, Benjamin Schmidt, Matthew Woolf and Xiaolei Zhao

Let X be a smooth projective surface of irregularity 0. The Hilbert scheme $X^{[n]}$ of n points on X parametrizes zero-dimensional subschemes of X of length n. We discuss general methods for studying the cone of ample divisors on $X^{[n]}$. We then use these techniques to compute the cone of ample divisors on $X^{[n]}$ for several surfaces where the cone was previously unknown. Our examples include families of surfaces of general type and del Pezzo surfaces of degree 1. The methods rely on Bridgeland stability and the positivity lemma of Bayer and Macrì.

A correction was posted on 16 June 2018 in an online supplement.

1. Introduction

If X is a projective variety, the cone $\mathrm{Amp}(X) \subset N^1(X)$ of ample divisors controls the various projective embeddings of X. It is one of the most important invariants of X, and carries detailed information about the geometry of X. Its closure is the *nef cone* $\mathrm{Nef}(X)$, which is dual to the Mori cone of curves (see for example [Lazarsfeld 2004]). In this paper, we will study the nef cone of the Hilbert scheme of points $X^{[n]}$, where X is a smooth projective surface over $\mathbb{C}$.

Nef divisors on Hilbert schemes of points on surfaces $X^{[n]}$ are sometimes easy to construct by classical methods. If L is an $(n-1)$-very ample line bundle

J. Huizenga was partially supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship. B. Schmidt was partially supported by NSF grant DMS-1523496 (PI Emanuele Macrì) and a Presidential Fellowship of the Ohio State University.
MSC2010: primary 14C05; secondary 14E30, 14J29, 14J60.
Keywords: Hilbert schemes, surfaces, nef cone, ample cone, birational geometry, Bridgeland stability.

on X, then for any $Z \in X^{[n]}$ we have an inclusion $H^0(L \otimes I_Z) \to H^0(L)$ which defines a morphism from $X^{[n]}$ to the Grassmannian $G(h^0(L) - n, h^0(L))$. The pullback of an ample divisor on the Grassmannian is nef on $X^{[n]}$. It is frequently possible to construct extremal nef divisors by this method. For example, this method completely computes the nef cone of $X^{[n]}$ when X is a del Pezzo surface of degree ≥ 2 or a Hirzebruch surface (see [Arcara et al. 2013; Bertram and Coskun 2013]). Unfortunately, this approach to computing the nef cone is insufficient in general. At the very least, to study nef cones of more interesting surfaces it would be necessary to study an analog of k-very ampleness for higher rank vector bundles, which is considerably more challenging than line bundles.

More recently, many nef cones have been computed by making use of Bridgeland stability conditions and the positivity lemma of Bayer and Macrì (see [Bridgeland 2007; 2008; Arcara and Bertram 2013; Bayer and Macrì 2014b] for background on these topics, which will be reviewed in Section 2). Let $\boldsymbol{v} = \operatorname{ch}(I_Z) \in K_0(X)$, where $Z \in X^{[n]}$. In the stability manifold $\operatorname{Stab}(X)$ for X there is an open *Gieseker chamber* $\mathcal{C}$ such that if $\sigma \in \mathcal{C}$ then $M_\sigma(\boldsymbol{v}) \cong X^{[n]}$, where $M_\sigma(\boldsymbol{v})$ is the moduli space of σ-semistable objects with invariants $\boldsymbol{v}$. The positivity lemma associates to any $\sigma \in \bar{\mathcal{C}}$ a nef divisor on $X^{[n]}$. Stability conditions in the boundary $\partial\mathcal{C}$ frequently give rise to extremal nef divisors. The positivity lemma also classifies the curves orthogonal to a nef divisor constructed in this way, and so gives a tool for checking extremality.

The stability manifold is rather large in general, so computation of the full Gieseker chamber can be unwieldy. We deal with this problem by focusing on a small *slice* of the stability manifold parametrized by a half-plane. Up to scale, the corresponding divisors in $N^1(X^{[n]})$ form an affine ray. The nef cone $\operatorname{Nef}(X^{[n]})$ is spanned by a codimension-1 subcone identified with $\operatorname{Nef}(X)$ and other more interesting classes which are positive on curves contracted by the Hilbert–Chow morphism. Since $\operatorname{Nef}(X^{[n]})$ is convex, we can study $\operatorname{Nef}(X^{[n]})$ by looking at positivity properties of divisors along rays in $N^1(X^{[n]})$ starting from a class in $\operatorname{Amp}(X) \subset \operatorname{Nef}(X^{[n]})$. The positivity lemma gives us an effective criterion for testing when divisors along the ray are nef.

The slices of the stability manifold that we consider are given by a pair of divisors (H, D) on X with H ample and $-D$ effective. The following is a weak version of one of our main theorems.

Theorem 1.1. *Let X be a smooth projective surface. If $n \gg 0$, then there is an extremal nef divisor on $X^{[n]}$ coming from the (H, D)-slice. It can be explicitly computed if both the intersection pairing on* $\operatorname{Pic}(X)$ *and the set of effective classes in* $\operatorname{Pic}(X)$ *are known. An orthogonal curve class is given by n points moving in a g^1_n on a curve of a particular class.*

See Section 3, especially Corollary 3.7 and Theorem 3.11, for more explicit statements. Stronger statements can also be shown under strong assumptions on $\mathrm{Pic}(X)$; for example, we study the Picard rank 1 case in detail in Section 4. Recall that if X is a surface of irregularity $q := H^1(\mathcal{O}_X) = 0$, then $N^1(X^{[n]})$ is spanned by the divisor B of nonreduced schemes and divisors $L^{[n]}$ induced by divisors $L \in \mathrm{Pic}(X)$; see Section 2A for details.

Theorem 1.2. *Let X be a smooth projective surface with* $\mathrm{Pic}\, X \cong \mathbb{Z}H$, *where H is an ample divisor. Let $a > 0$ be the smallest integer such that aH is effective. If*

$$n \geq \max\{a^2 H^2,\ p_a(aH) + 1\},$$

then $\mathrm{Nef}(X^{[n]})$ *is spanned by the divisor $H^{[n]}$ and the divisor*

$$\frac{1}{2}K_X^{[n]} + \Big(\frac{a}{2} + \frac{n}{aH^2}\Big)H^{[n]} - \frac{1}{2}B. \tag{$*$}$$

An orthogonal curve class is given by letting n points move in a g^1_n on a curve in X of class aH.

Note that in the Picard rank 1 case the divisor class $(*)$ is frequently of the form $\lambda H^{[n]} - \frac{1}{2}B$ for a noninteger number $\lambda \in \mathbb{Q}$. Any divisor constructed from an $(n-1)$-very ample line bundle will be of the form $\lambda H^{[n]} - \frac{1}{2}B$ with $\lambda \in \mathbb{Z}$, so in general the edge of the nef cone cannot be obtained from line bundles in this way.

The required lower bound on n in Theorem 1.2 can be improved in specific examples where special linear series on hyperplane sections are better understood.

Theorem 1.3. *Let X be one of the following surfaces*:

(1) *a very general hypersurface in $\mathbb{P}^3$ of degree $d \geq 4$, or*

(2) *a very general degree-d cyclic branched cover of $\mathbb{P}^2$ of general type.*

In either case, $\mathrm{Pic}(X) \cong \mathbb{Z}H$ *with H effective. Suppose $n \geq d - 1$ in the first case, and $n \geq d$ in the second case. Then* $\mathrm{Nef}(X^{[n]})$ *is spanned by $H^{[n]}$ and the divisor class* $(*)$ *with* $a = 1$.

Finally, in Section 5 we compute the nef cone of $X^{[n]}$ where X is a smooth del Pezzo surface of degree 1 and $n \geq 2$ is arbitrary. This computation was an open problem posed by Bertram and Coskun [2013]; they noted that the method of k-very ample line bundles would not be sufficient to prove the expected answer. Since X has Picard rank 9, this computation makes full use of the general methods developed in Section 3. If $C \subset X$ is a reduced, irreducible curve which admits a g^1_n, we write $C_{[n]}$ for the curve in the Hilbert scheme $X^{[n]}$ given by letting n points move in a g^1_n on C.

Theorem 1.4. *Let X be a smooth del Pezzo surface of degree* 1. *The Mori cone of curves* $\mathrm{NE}(X^{[n]})$ *is spanned by the* 240 *classes $E_{[n]}$ given by* (-1)*-curves $E \subset X$,* *the class of a curve contracted by the Hilbert–Chow morphism, and the class $F_{[n]}$, where $F \in |-K_X|$ is an anticanonical curve. The nef cone is determined by duality.*

Many previous authors have used Bridgeland stability conditions to study nef cones and wall-crossing for Hilbert schemes $X^{[n]}$ and moduli spaces of sheaves $M_H(\boldsymbol{v})$ for various classes of surfaces. For instance, the program was studied for $\mathbb{P}^2$ in [Arcara et al. 2013; Coskun and Huizenga 2016; Bertram et al. 2014; Li and Zhao 2013], Hirzebruch and del Pezzo surfaces in [Bertram and Coskun 2013], abelian surfaces in [Yanagida and Yoshioka 2014; Maciocia and Meachan 2013], K3 surfaces in [Bayer and Macrì 2014b; 2014a; Hassett and Tschinkel 2010], and Enriques surfaces in [Nuer 2014]. Our results unify several of these approaches. Additionally, nef cones were classically studied in the context of k-very ample line bundles in papers such as [Ellingsrud et al. 2001; Beltrametti and Sommese 1991; Beltrametti et al. 1989; Catanese and Göttsche 1990].

2. Preliminaries

Throughout the paper, we let X be a smooth projective surface over $\mathbb{C}$.

2A. ***Divisors and curves on $X^{[n]}$.*** For simplicity we assume that X has irregularity $q = h^1(\mathcal{O}_X) = 0$ in this subsection. By work of Fogarty [1968], the Hilbert scheme $X^{[n]}$ is a smooth projective variety of dimension $2n$ which resolves the singularities in the symmetric product $X^{(n)}$ via the Hilbert–Chow morphism $X^{[n]} \to X^{(n)}$. A line bundle L on X induces the S_n-equivariant line bundle $L^{\boxtimes n}$ on X^n which descends to a line bundle $L^{(n)}$ on the symmetric product $X^{(n)}$. The pullback of $L^{(n)}$ by the Hilbert–Chow morphism $X^{[n]} \to X^{(n)}$ defines a line bundle on $X^{[n]}$ which we will denote by $L^{[n]}$. Intuitively, if $L \cong \mathcal{O}_X(D)$ for a reduced effective divisor $D \subset X$, then $L^{[n]}$ can be represented by the divisor $D^{[n]}$ of schemes $Z \subset X$ which meet D.

Fogarty shows that

$$\operatorname{Pic}(X^{[n]}) \cong \operatorname{Pic}(X) \oplus \mathbb{Z}(B/2),$$

where $\operatorname{Pic}(X) \subset \operatorname{Pic}(X^{[n]})$ is embedded by $L \mapsto L^{[n]}$ and B is the locus of nonreduced schemes, i.e., the exceptional divisor of the Hilbert–Chow morphism [Fogarty 1973]. Tensoring by the real numbers, the Néron–Severi space $N^1(X^{[n]})$ is therefore spanned by $N^1(X)$ and B.

There are also curve classes in $X^{[n]}$ induced by curves in X. Two different constructions are immediate. Let $C \subset X$ be a reduced and irreducible curve.

(1) There is a curve $\tilde{C}_{[n]}$ in $X^{[n]}$ given by fixing $n-1$ general points of X and letting an n-th point move along C.

(2) If C admits a g^1_n, i.e., a degree-n map to $\mathbb{P}^1$, then the fibers of $C \to \mathbb{P}^1$ give a rational curve $\mathbb{P}^1 \to X^{[n]}$. We write $C_{[n]}$ for this class.

These constructions preserve intersection numbers, in the sense that if $D \subset X$ is a divisor and $C \subset X$ is a curve then

$$D^{[n]} \cdot \tilde{C}_{[n]} = D^{[n]} \cdot C_{[n]} = D \cdot C.$$

Part of the nef cone $\mathrm{Nef}(X^{[n]})$ is easily described in terms of the nef cone of X. If D is an ample divisor, then $D^{(n)}$ is ample so $D^{[n]}$ is nef. In the limit, we find that if D is nef then $D^{[n]}$ is nef. Conversely, if D is not nef then there is an irreducible curve C with $D \cdot C < 0$, so $D^{[n]} \cdot \tilde{C}_{[n]} < 0$ and $D^{[n]}$ is not nef. Under the Fogarty isomorphism,

$$\mathrm{Nef}(X^{[n]}) \cap N^1(X) = \mathrm{Nef}(X).$$

The hyperplane $N^1(X) \subset N^1(X^{[n]})$ is orthogonal to any curve contracted by the Hilbert–Chow morphism, so all the divisors in $\mathrm{Nef}(X) \subset \mathrm{Nef}(X^{[n]})$ are extremal. Since B is the exceptional locus of the Hilbert–Chow morphism, we see that any nef class must have nonpositive coefficient of B. After scaling, then, we see that computation of the cone $\mathrm{Nef}(X^{[n]})$ reduces to describing the nef classes of the form $L^{[n]} - \frac{1}{2}B$ lying outside $\mathrm{Nef}(X) \subset \mathrm{Nef}(X^{[n]})$.

2B. *Bridgeland stability conditions.* We now recall some basic definitions and properties of Bridgeland stability conditions. We fix a polarization $H \in \mathrm{Pic}(X)_{\mathbb{R}}$. For any divisor $D \in \mathrm{Pic}(X)_{\mathbb{R}}$, the twisted Chern character $\mathrm{ch}^D = e^{-D}\mathrm{ch}$ can be expanded as

$$\mathrm{ch}_0^D = \mathrm{ch}_0, \quad \mathrm{ch}_1^D = \mathrm{ch}_1 - D\,\mathrm{ch}_0, \quad \mathrm{ch}_2^D = \mathrm{ch}_2 - D \cdot \mathrm{ch}_1 + \tfrac{1}{2} D^2\,\mathrm{ch}_0\,.$$

Recall that a Bridgeland stability condition is a pair $\sigma = (Z, \mathcal{A})$ for which $Z : K_0(X) \to \mathbb{C}$ is an additive homomorphism and $\mathcal{A} \subset D^b(X)$ is the heart of a bounded t-structure. In particular, $\mathcal{A}$ is an abelian category. Moreover, Z maps any nontrivial object in $\mathcal{A}$ to the upper half-plane or the negative real line. The σ-slope function is defined by

$$\nu_\sigma = -\frac{\Re Z}{\Im Z},$$

and σ-(semi)stability of objects of $\mathcal{A}$ is defined in terms of this slope function. More technical requirements are the existence of Harder–Narasimhan filtrations and the support property. We recommend [Bridgeland 2007] for a more precise definition. The support property is well explained in Appendix A of [Bayer et al. 2014].

In the case of surfaces, Bridgeland [2008] and Arcara and Bertram [2013] showed how to construct Bridgeland stability conditions in a slice corresponding to a choice of an ample divisor $H \in \mathrm{Pic}(X)_{\mathbb{R}}$ and arbitrary twisting divisor $D \in \mathrm{Pic}(X)_{\mathbb{R}}$. The classical Mumford slope function for twisted Chern characters is defined by

$$\mu_{H,D} = \frac{H\cdot \mathrm{ch}_1^D}{H^2\,\mathrm{ch}_0^D},$$

where torsion sheaves are interpreted as having positive infinite slope. Given a real number $\beta \in \mathbb{R}$, there are two categories defined as

$$\mathcal{T}_\beta = \{E \in \mathrm{Coh}(X) : \text{any quotient } E \twoheadrightarrow G \text{ satisfies } \mu_{H,D}(G) > \beta\},$$
$$\mathcal{F}_\beta = \{E \in \mathrm{Coh}(X) : \text{any subsheaf } F \hookrightarrow E \text{ satisfies } \mu_{H,D}(F) \le \beta\}.$$

A new heart of a bounded t-structure is defined as the extension closure $\mathcal{A}_\beta := \langle \mathcal{F}_\beta[1], \mathcal{T}_\beta\rangle$. We fix an additional positive real number α and define the homomorphism as

$$Z_{\beta,\alpha} = -\,\mathrm{ch}_2^{D+\beta H} + \tfrac{1}{2}\alpha^2 H^2\,\mathrm{ch}_0^{D+\beta H} + i\,H\cdot \mathrm{ch}_1^{D+\beta H}.$$

The pair $\sigma_{\beta,\alpha} := (Z_{\beta,\alpha}, \mathcal{A}_\beta)$ is then a Bridgeland stability condition. The (H, D)-*slice* of stability conditions is the family of stability conditions

$$\{\sigma_{\beta,\alpha} : \beta, \alpha \in \mathbb{R}, \alpha > 0\}$$

parametrized by the (β, α) upper half-plane.

Definition 2.1. Fix a set of invariants $\boldsymbol{v} \in K_0(X)$.

(1) Let $\boldsymbol{w} \in K_0(X)$ be a vector such that $\boldsymbol{v}$ and $\boldsymbol{w}$ do not have the same $\sigma_{\beta,\alpha}$-slope everywhere in the (H, D)-slice. The *numerical wall* for $\boldsymbol{v}$ given by $\boldsymbol{w}$ is the set of points (β, α) where $\boldsymbol{v}$ and $\boldsymbol{w}$ have the same $\sigma_{\beta,\alpha}$-slope.

(2) A numerical wall for $\boldsymbol{v}$ given by a vector $\boldsymbol{w}$ as above is a *wall* (or *actual wall*) if there is a point (β, α) on the wall and an exact sequence $0 \to F \to E \to G \to 0$ in $\mathcal{A}_\beta$, where $\mathrm{ch}\, F = \boldsymbol{w}$, $\mathrm{ch}\, E = \boldsymbol{v}$, and F, E, and G are $\sigma_{\beta,\alpha}$-semistable objects (of the same $\sigma_{\beta,\alpha}$-slope).

We write $K_{\mathrm{num}}(X)$ for the numerical Grothendieck group of classes in $K_0(X)$ modulo numerical equivalence. Note that numerical walls for $\boldsymbol{v} \in K_0(X)$ only depend on the numerical class of $\boldsymbol{v}$, while actual walls a priori depend on $c_1(\boldsymbol{v}) \in \mathrm{Pic}(X)$. The structure of walls in a slice is heavily restricted by Bertram's nested wall theorem. This was first observed for Picard rank 1 with $D = 0$, but the proof immediately generalizes by replacing ch by ch^D everywhere.

Theorem 2.2 [Maciocia 2014]. *Let $\boldsymbol{v} \in K_0(X)$.*

(1) *Numerical walls for $\boldsymbol{v}$ can either be semicircles with center on the β-axis or the unique vertical line given by $\beta = \mu_{H,D}(\boldsymbol{v})$. Moreover, the apex of each semicircle lies on the hyperbola $\Re Z_{\beta,\alpha}(\boldsymbol{v}) = 0$.*

(2) *Numerical walls for $\boldsymbol{v}$ are disjoint, and the semicircular walls on either side of the vertical wall are nested.*

(3) *If W_1 and W_2 are two semicircular numerical walls left of the vertical wall with centers $(s_{W_1}, 0)$ and $(s_{W_2}, 0)$, then W_2 is nested inside W_1 if and only if $s_{W_1} < s_{W_2}$.*

(4) *Suppose $0 \to F \to E \to G \to 0$ is an exact sequence destabilizing an object E with $\mathrm{ch}(E) = \boldsymbol{v}$ at a point (β, α) on a numerical wall W, in the sense that all three objects have the same $\sigma_{\beta,\alpha}$-slope and this is an exact sequence in $\mathcal{A}_\beta$. Then it is an exact sequence of objects in $\mathcal{A}_{\beta'}$ with the same $\sigma_{\beta',\alpha'}$-slope for all $(\beta', \alpha') \in W$. That is, E is destabilized along the entire wall.*

2C. ***Slope and discriminant.*** The explicit geometry of walls is frequently best understood in terms of slopes and discriminants; the formulas presented here previously appeared in [Coskun and Huizenga 2014] in the context of $\mathbb{P}^2$. When the rank is nonzero, we define

$$\Delta_{H,D} = \tfrac{1}{2}\mu_{H,D}^2 - \frac{\mathrm{ch}_2^D}{H^2\,\mathrm{ch}_0^D}.$$

The Bogomolov inequality gives $\Delta_{H,D}(E) \geq 0$ whenever E is an (H, D)-twisted Gieseker semistable sheaf. Observe that $\Delta_{H,D+\beta H} = \Delta_{H,D}$ for every $\beta \in \mathbb{R}$. A straightforward calculation shows that for vectors of nonzero rank the slope function for the stability condition $\sigma_{\beta,\alpha}$ in the (H, D)-slice is given by

$$\nu_{\sigma_{\beta,\alpha}} = \frac{(\mu_{H,D} - \beta)^2 - \alpha^2 - 2\Delta_{H,D}}{(\mu_{H,D} - \beta)}. \tag{1}$$

Suppose $\boldsymbol{v}$, $\boldsymbol{w}$ are two classes with positive rank, and let their slopes and discriminants be $\mu_{H,D}$, $\Delta_{H,D}$ and $\mu'_{H,D}$, $\Delta'_{H,D}$, respectively. The numerical wall W in the (H, D)-slice where $\boldsymbol{v}$ and $\boldsymbol{w}$ have the same slope is computed as follows.

- If $\mu_{H,D} = \mu'_{H,D}$ and $\Delta_{H,D} = \Delta'_{H,D}$, then $\boldsymbol{v}$ and $\boldsymbol{w}$ have the same slope everywhere in the slice, so there is no numerical wall.
- If $\mu_{H,D} = \mu'_{H,D}$ and $\Delta_{H,D} \neq \Delta'_{H,D}$, then W is the vertical wall $\beta = \mu_{H,D}$.
- If $\mu_{H,D} \neq \mu'_{H,D}$, then (1) implies W is the semicircle with center $(s_W, 0)$ and radius ρ_W, where

$$s_W = \tfrac{1}{2}(\mu_{H,D} + \mu'_{H,D}) - \frac{\Delta_{H,D} - \Delta'_{H,D}}{\mu_{H,D} - \mu'_{H,D}}, \tag{2}$$

$$\rho_W^2 = (s_W - \mu_{H,D})^2 - 2\Delta_{H,D}, \tag{3}$$

provided that the expression defining ρ_W^2 is positive; if it is negative then the wall is empty.

Notice that if $\Delta_{H,D}(\boldsymbol{v}) \geq 0$ then numerical walls for $\boldsymbol{v}$ left of the vertical wall accumulate at the point

$$\left(\mu_{H,D}(\boldsymbol{v}) - \sqrt{2\Delta_{H,D}(\boldsymbol{v})}, 0\right) \tag{4}$$

as their radii go to 0.

2D. *Nef divisors and the positivity lemma.* In this section, we describe the positivity lemma of Bayer and Macrì. Let $\sigma = (Z, \mathcal{A})$ be a stability condition on X, $\boldsymbol{v} \in K_{\text{num}}(X)$ and S a proper algebraic space of finite type over $\mathbb{C}$. Let $\mathcal{E} \in D^b(X \times S)$ be a flat family of σ-semistable objects of class $\boldsymbol{v}$, i.e., for every $\mathbb{C}$-point $p \in S$, the derived restriction $\mathcal{E}|_{\pi_S^{-1}(\{p\})}$ is σ-semistable of class $\boldsymbol{v}$. Then Bayer and Macrì define a numerical divisor class $D_{\sigma,\mathcal{E}} \in N^1(S)$ on the space S by assigning its intersection with any projective integral curve $C \subset S$:

$$D_{\sigma,\mathcal{E}} \cdot C = \Im\left(-\frac{Z((p_X)_*\mathcal{E}|_{C\times X})}{Z(\boldsymbol{v})}\right).$$

The positivity lemma shows that this divisor inherits positivity properties from the homomorphism Z, and classifies the curve classes orthogonal to the divisor. Recall that two σ-semistable objects are *S-equivalent* with respect to σ if their sets of Jordan–Hölder factors are the same.

Theorem 2.3 (positivity lemma [Bayer and Macrì 2014b, Lemma 3.3]). *The divisor $D_{\sigma,\mathcal{E}} \in N^1(S)$ is nef. Moreover, if $C \subset S$ is a projective integral curve then $D_{\sigma,\mathcal{E}} \cdot C = 0$ if and only if two general objects parametrized by C are S-equivalent with respect to σ.*

Our primary use of the positivity lemma is to attempt to construct extremal nef divisors on Hilbert schemes of points. Thus it is important to recover Hilbert schemes of points as Bridgeland moduli spaces. Recall that a torsion-free coherent sheaf E is *(H, D)-twisted Gieseker semistable* if for every $F \subset E$ we have

$$\frac{\chi\big(F \otimes \mathcal{O}_X(mH - D)\big)}{\operatorname{rk}(F)} \leq \frac{\chi\big(E \otimes \mathcal{O}_X(mH - D)\big)}{\operatorname{rk}(E)}$$

for all $m \gg 0$, where the Euler characteristic is computed formally via Riemann–Roch; see [Matsuki and Wentworth 1997]. For any class $\boldsymbol{v} \in K_0(X)$, there are projective moduli spaces $M_{H,D}(\boldsymbol{v})$ of S-equivalence classes of (H, D)-twisted Gieseker semistable sheaves with class $\boldsymbol{v}$. If $\boldsymbol{v} = (1, 0, -n)$ is the Chern character of an ideal sheaf of n points then $M_{H,D}(\boldsymbol{v}) = X^{[n]}$. Note that if the irregularity of X is nonzero, then it is crucial to fix the determinant.

Fix an (H, D)-slice in the stability manifold, and fix a vector $\boldsymbol{v} \in K_0(X)$ with positive rank. If β lies to the left of the vertical wall $\beta = \mu_{H,D}(\boldsymbol{v})$ for $\boldsymbol{v}$, then for $\alpha \gg 0$ the moduli space coincides with a twisted Gieseker moduli space.

Proposition 2.4 (the large volume limit [Bridgeland 2008; Maciocia 2014]). *Fix divisors (H, D) giving a slice in* $\mathrm{Stab}(X)$. *Let $\boldsymbol{v} \in K_0(X)$ be a vector with positive rank, and let $\beta \in \mathbb{R}$ be such that $\mu_{H,D}(\boldsymbol{v}) > \beta$. If $E \in \mathcal{A}_\beta$ has* $\mathrm{ch}(E) = \boldsymbol{v}$ *then E is $\sigma_{\beta,\alpha}$-semistable for all $\alpha \gg 0$ if and only if E is an $\big(H, D - \frac{1}{2}K_X\big)$-twisted Gieseker semistable sheaf.*

Moreover, in the quadrant of the (H, D)-slice left of the vertical wall there is a largest semicircular wall for $\boldsymbol{v}$, called the Gieseker wall. For all (β, α) between this wall and the vertical wall, the moduli space $M_{\sigma_{\beta,\alpha}}(\boldsymbol{v})$ coincides with the moduli space $M_{H,D-K_X/2}(\boldsymbol{v})$ of $\big(H, D - \frac{1}{2}K_X\big)$-twisted Gieseker semistable sheaves.

We use these results as follows. Let $\boldsymbol{v} = (1, 0, -n) \in K_0(X)$ be the vector for the Hilbert scheme $X^{[n]}$, and let σ_+ be a stability condition in the (H, D)-slice lying above the Gieseker wall, so that $M_{\sigma_+}(\boldsymbol{v}) \cong X^{[n]}$. Let $\mathcal{E}/(X \times X^{[n]})$ be the universal ideal sheaf, and let σ_0 be a stability condition on the Gieseker wall. By the definition of the Gieseker wall, $\mathcal{E}$ is a family of σ_0-semistable objects, so there is an induced nef divisor $D_{\sigma_0,\mathcal{E}}$ on $X^{[n]}$. Furthermore, curves orthogonal to $D_{\sigma_0,\mathcal{E}}$ are understood in terms of destabilizing sequences along the wall, so it is possible to test for extremality.

3. Gieseker walls and the nef cone

Fix an ample divisor $H \in \mathrm{Pic}(X)$ with $H^2 = d$ and an *antieffective* divisor D. In this section we study the nef divisor arising from the Gieseker wall (i.e., the largest wall where some ideal sheaf is destabilized) in the slice of the stability manifold given by the pair (H, D). We first compute the Gieseker wall, and then investigate when the corresponding nef divisor is in fact extremal.

3A. ***Bounding higher rank walls.*** The main difficulty in computing extremal rays of the nef cone is to show that a destabilizing subobject along the Gieseker wall is a line bundle, and not some higher rank sheaf. We first prove a lemma which generalizes Proposition 8.3 of [Coskun and Huizenga 2016] from $X = \mathbb{P}^2$ to an arbitrary surface. We prove the result in slightly more generality than we will need here as we expect it to be useful in future work.

Lemma 3.1. *Let σ_0 be a stability condition in the (H, D)-slice, and suppose*

$$0 \to F \to E \to G \to 0$$

is an exact sequence of σ_0-semistable objects of the same σ_0-slope, where E is an (H, D)-twisted Gieseker semistable torsion-free sheaf. If the map $F \to E$ of sheaves is not injective, then the radius ρ_W of the wall W defined by this sequence satisfies

$$\rho_W^2 \leq \frac{(\min\{\mathrm{rk}(F)-1, \mathrm{rk}(E)\})^2}{2\,\mathrm{rk}(F)} \Delta_{H,D}(E).$$

Proof. The proof is similar to the proof in [Coskun and Huizenga 2016] given in the case of $\mathbb{P}^2$; we present it for completeness. The object F is a torsion-free sheaf by the standard cohomology sequence and the fact that the heart of the t-structure in the slice we are working in consists of objects which only have nonzero cohomology sheaves in degrees 0 and -1. The exact sequence along W gives an exact sequence

$$0 \to K \to F \to E \to C \to 0$$

of sheaves of ranks k, f, e, c, respectively. By assumption, $k, f, e > 0$. Let $(s_W, 0)$ be the center of W. As F is in the category $\mathcal{T}_\beta$ whenever (β, α) is on W, we find that $\mu_{H,D}(F) \geq s_W + \rho_W$, so

$$\begin{aligned} df(s_W + \rho_W) \leq df\mu_{H,D}(F) = \mathrm{ch}_1^D(F)\cdot H &= (\mathrm{ch}_1^D(K) + \mathrm{ch}_1^D(E) - \mathrm{ch}_1^D(C))\cdot H \\ &= dk\mu_{H,D}(K) + de\mu_{H,D}(E) - \mathrm{ch}_1^D(C)\cdot H. \end{aligned}$$

Similarly, $K \in \mathcal{F}_\beta$ along W, so $\mu_{H,D}(K) \leq s_W - \rho_W$ and

$$df(s_W + \rho_W) \leq dk(s_W - \rho_W) + de\mu_{H,D}(E) - \mathrm{ch}_1^D(C)\cdot H,$$

which gives

$$d(k+f)\rho_W \leq d(k-f)s_W + de\mu_{H,D}(E) - \mathrm{ch}_1^D(C)\cdot H. \tag{5}$$

We now wish to eliminate the term $\mathrm{ch}_1^D(C)\cdot H$ in inequality (5). If C is either 0 or torsion, then $\mathrm{ch}_1^D(C)\cdot H \geq 0$ and $-e = k - f$, and we deduce

$$(k+f)\rho_W \leq (k-f)(s_W - \mu_{H,D}(E)). \tag{6}$$

Suppose instead that C is not torsion. Since C is a quotient of the semistable sheaf E, we have $\mu_{H,D}(C) \geq \mu_{H,D}(E)$, so $\mathrm{ch}_1^D(C)\cdot H = dc\mu_{H,D}(C) \geq dc\mu_{H,D}(E)$. As $k - f = c - e$, we find that inequality (6) also holds in this case.

Both sides of inequality (6) are positive, so squaring both sides gives

$$(k+f)^2\rho_W^2 \leq (k-f)^2(s_W - \mu_{H,D}(E))^2.$$

Formula (3) for ρ_W^2 shows this is equivalent to

$$(k+f)^2\rho_W^2 \leq (k-f)^2(\rho_W^2 + 2\Delta_{H,D}(E)),$$

from which we obtain

$$\rho_W^2 \leq \frac{(k-f)^2}{2kf}\Delta_{H,D}(E).$$

Since $k = f - e + c$, we see that $k \geq \max\{1, f - e\}$. By taking derivatives in k, we see that $(k-f)^2/(2kf)$ is decreasing for $k + f > 0$, and so the maximum possible value of the right-hand side must occur when $k = \max\{1, f - e\}$. The denominator will be at least $2f$ in this case, and the numerator is $\min\{(f-1)^2, e^2\}$. The result follows. □

For our present work we will only need the next consequence of Lemma 3.1, which follows immediately from computing $\Delta_{H,D}(I_Z)$.

Corollary 3.2. *With the hypotheses of Lemma 3.1, if E is an ideal sheaf $I_Z \in X^{[n]}$ and F has rank at least 2, then the radius of the corresponding wall satisfies*

$$\rho_W^2 \le \frac{2nd + (H\cdot D)^2 - dD^2}{8d^2} := \varrho_{H,D,n}.$$

The number $\varrho_{H,D,n}$ therefore bounds the squares of the radii of higher rank walls for $X^{[n]}$.

3B. *Rank-1 walls and critical divisors.* In the cases where we compute the Gieseker wall, the ideal sheaf that is destabilized along the wall will be destabilized by a rank-1 subobject. We first compute the numerical walls given by rank-1 subobjects.

Lemma 3.3. *Consider a rank-1 torsion-free sheaf $F = I_{Z'}(-L)$, where Z' is a zero-dimensional scheme of length w and L is an effective divisor. In the (H, D)-slice, the numerical wall W for $X^{[n]}$ where F has the same slope as an ideal I_Z of n points has center $(s_W, 0)$ given by*

$$s_W = -\frac{2(n-w) + L^2 + 2(D\cdot L)}{2(H\cdot L)}.$$

Proof. This is an immediate consequence of (2) for the center of a wall. □

Recalling that walls for $X^{[n]}$ left of the vertical wall get larger as their centers decrease, we deduce the following consequence.

Lemma 3.4. *If the Gieseker wall in the (H, D)-slice is given by a rank-1 subobject, then it is a line bundle $\mathcal{O}_X(-L)$ for some effective divisor L.*

Proof. Suppose some $I_Z \in X^{[n]}$ is destabilized along the Gieseker wall W by a sheaf of the form $I_{Z'}(-L)$, where Z' is a nonempty zero-dimensional scheme and L is effective. By Lemma 3.3, the numerical wall W' given by $\mathcal{O}_X(-L)$ is strictly larger than W. Since $\mathcal{O}_X(-L)$ has the same $\mu_{H,D}$-slope as $I_{Z'}(-L)$, which is in the categories along W, we find that $\mathcal{O}_X(-L)$ is in at least some of the categories along W'. But then W' is an actual wall, since any ideal sheaf I_Z where Z lies on a curve $C \in |L|$ is destabilized along it. This contradicts the fact that W is the Gieseker wall. □

Less trivially, there is a further minimality condition automatically satisfied by a line bundle $\mathcal{O}_X(-L)$ which gives the Gieseker wall. We define the set of *critical effective divisors* with respect to H and D by

$$\mathrm{CrDiv}(H, D) = \{-D\} \cup \{L \in \mathrm{Pic}(X) \text{ effective} : H\cdot L < H\cdot(-D)\}.$$

By [Hartshorne 1977, Exercise V.1.11], the set $\mathrm{CrDiv}(H, D)/\sim$ of critical divisors modulo numerical equivalence is finite. Therefore, the set of numerical walls for $X^{[n]}$ given by line bundles $\mathcal{O}_X(-L)$ with $L \in \mathrm{CrDiv}(H, D)$ is also finite. Note that the inequality $H \cdot L < H \cdot (-D)$ is equivalent to the inequality $\mu_{H,D}(\mathcal{O}_X(-L)) > 0$. The next proposition demonstrates the importance of critical divisors.

Proposition 3.5. *Assume $2n > D^2$, and suppose the subobject giving the Gieseker wall for $X^{[n]}$ in the (H, D)-slice is a line bundle. Then the Gieseker wall is computed by $\mathcal{O}_X(-L)$, where $L \in \mathrm{CrDiv}(H, D)$ is chosen so that the numerical wall given by $\mathcal{O}_X(-L)$ is as large as possible.*

Proof. First, consider the numerical wall W given by $\mathcal{O}_X(D)$. By Lemma 3.3, the center $(s_W, 0)$ has

$$s_W = \frac{2n - D^2}{2(H \cdot D)} < 0, \tag{7}$$

since $2n > D^2$ and D is antieffective. Since $\mu_{H,D}(\mathcal{O}_X(D)) = \Delta_{H,D}(\mathcal{O}_X(D)) = 0$, formula (3) for the radius of W gives $\rho_W^2 = s_W^2$. In particular, W is nonempty, and $\mathcal{O}_X(D)$ lies in at least some of the categories along W. Since D is antieffective, there are exact sequences of the form

$$0 \to \mathcal{O}_X(D) \to I_Z \to I_{Z \subset C} \to 0,$$

where $C \in |-D|$ and $Z \subset C$ is a collection of n points. If no actual wall is larger than W, it follows that W is an actual wall and it is the Gieseker wall.

Suppose the Gieseker wall is larger than W and computed by a line bundle $\mathcal{O}_X(-L)$ with L effective. Since W passes through the origin in the (β, α)-plane, $\mathcal{O}_X(-L)$ must lie in the category $\mathcal{T}_0$. Therefore $\mu_{H,D}(\mathcal{O}_X(-L)) > 0$, and $L \in \mathrm{CrDiv}(H, D)$.

Conversely, suppose $L \in \mathrm{CrDiv}(H, D)$ is chosen to maximize the wall W' given by $\mathcal{O}_X(-L)$. Then no actual wall is larger than W'. Since $s_W < 0$ and $\mu_{H,D}(\mathcal{O}_X(-L)) \geq 0$, we find that $\mathcal{O}_X(-L)$ is in at least some of the categories along W, and hence in at least some of the categories along W'. We conclude that W' is an actual wall, and therefore that it is the Gieseker wall. □

Combining Corollary 3.2 and Proposition 3.5 gives our primary tool to compute the Gieseker wall.

Theorem 3.6. *Assume $2n > D^2$, and let $L \in \mathrm{CrDiv}(H, D)$ be a critical divisor such that the wall for $X^{[n]}$ given by $\mathcal{O}_X(-L)$ is as large as possible. If this wall has radius ρ satisfying $\rho^2 \geq \varrho_{H,D,n}$, then it is the Gieseker wall.*

Conversely, if the Gieseker wall has radius satisfying $\rho^2 \geq \varrho_{H,D,n}$ then it is obtained in this way.

While the theorem is our sharpest result, it is useful to lose some generality to get a more explicit version. Since $-D \in \mathrm{CrDiv}(H, D)$, if the wall given by $\mathcal{O}_X(D)$ satisfies $\rho^2 \geq \varrho_{H,D,n}$ then the Gieseker wall is computed by Theorem 3.6. This allows us to compute the Gieseker wall so long as n is large enough, depending only on the intersection numbers of H and D.

Corollary 3.7. *Let*

$$\eta_{H,D} := \frac{(H \cdot D)^2 + dD^2}{2d}.$$

If $n \geq \eta_{H,D}$ then the Gieseker wall is the largest wall given by a critical divisor.

Furthermore, if $n > \eta_{H,D}$ then every I_Z destabilized along the Gieseker wall fits into an exact sequence

$$0 \to \mathcal{O}_X(-C) \to I_Z \to I_{Z \subset C} \to 0$$

for some curve $C \in |L|$, where L is a critical divisor computing the Gieseker wall. If the critical divisor computing the Gieseker wall is unique, then $\mathcal{O}_X(-C)$ and $I_{Z\subset C}$ are the Jordan–Hölder factors of any I_Z destabilized along the Gieseker wall.

Proof. Observe that the inequality $n \geq \eta_{H,D}$ automatically implies the inequality $2n > D^2$ needed to apply Theorem 3.6.

Let W be the wall for $X^{[n]}$ in the (H, D)-slice corresponding to $\mathcal{O}_X(D)$. The center $(s_W, 0)$ of W was computed in (7), and $\rho_W^2 = s_W^2$. We find that $\rho_W^2 \geq \varrho_{H,D,n}$ holds when $n \geq \eta_{H,D}$, with strict inequality when $n > \eta_{H,D}$.

When $n > \eta_{H,D}$ there can be no higher-rank destabilizing subobject of an I_Z destabilized along the Gieseker wall, so there is an exact sequence as claimed. Furthermore, if there is only one critical divisor computing the wall, then there is a unique destabilizing subobject along the wall, so the Jordan–Hölder filtration has length two. □

3C. ***Classes of divisors.*** In this subsection we give an elementary computation of the class of the divisor corresponding to a wall in a given slice of the stability manifold. Similar results have been obtained by Liu [2015], but the result is critical to our discussion so we include the proof. See [Bayer and Macrì 2014b, §4] for more details on the definitions and results we use here.

Throughout this subsection, let $\boldsymbol{v} \in K_0(X)$ be a vector such that the moduli space $M_{H,D}(\boldsymbol{v})$ of (H, D)-Gieseker semistable sheaves admits a (quasi-)universal family $\mathcal{E}$ which is unique up to equivalence (Hilbert schemes $X^{[n]}$ are examples of such spaces). We also let $\sigma = (Z, \mathcal{A})$ be a stability condition in the closure of the Gieseker chamber for $\boldsymbol{v}$ in the (H, D)-slice. Then there is a well-defined corresponding divisor $D_\sigma \in N^1(M_{H,D-K_X/2}(\boldsymbol{v}))$ which is independent of the choice of $\mathcal{E}$.

Let $(\boldsymbol{v}, \boldsymbol{w}) = \chi(\boldsymbol{v}\cdot\boldsymbol{w})$ be the Euler pairing on $K_{\mathrm{num}}(X)_{\mathbb{R}}$, and write $\boldsymbol{v}^{\perp} \subset K_{\mathrm{num}}(X)_{\mathbb{R}}$ for the orthogonal complement with respect to this pairing. The correspondence between stability conditions and divisor classes is understood in terms of the Donaldson homomorphism

$$\lambda : \boldsymbol{v}^{\perp} \to N^1(M_{H,D-K_X/2}(\boldsymbol{v})).$$

Since the Euler pairing is nondegenerate, there is a unique vector $\boldsymbol{w}_\sigma \in \boldsymbol{v}^{\perp}$ such that

$$\Im\Big(-\frac{Z(\boldsymbol{w}')}{Z(\boldsymbol{v})}\Big) = (\boldsymbol{w}', \boldsymbol{w}_\sigma),$$

for all $\boldsymbol{w}' \in K_{\mathrm{num}}(X)_{\mathbb{R}}$. Bayer and Macrì show that $D_\sigma = \lambda(\boldsymbol{w}_\sigma)$. In what follows, we write vectors in $K_{\mathrm{num}}(X)_{\mathbb{R}}$ as $(\mathrm{ch}_0, \mathrm{ch}_1, \mathrm{ch}_2)$.

Proposition 3.8. *With the above assumptions, suppose σ lies on a numerical wall W in the (H, D)-slice with center $(s_W, 0)$. Then $\boldsymbol{w}_\sigma$ is a multiple of*

$$\big(-1, -\tfrac{1}{2}K_X + s_W H + D, m\big) \in \boldsymbol{v}^{\perp},$$

where m is determined by the requirement $\boldsymbol{w}_\sigma \in \boldsymbol{v}^{\perp}$.

In particular, if X has irregularity 0 and $\boldsymbol{v} = (1, 0, -n)$ is the vector for $X^{[n]}$, then the divisor D_σ is a multiple of

$$\tfrac{1}{2}K_X^{[n]} - s_W H^{[n]} - D^{[n]} - \tfrac{1}{2}B.$$

Remark 3.9. Suppose X has irregularity 0. Up to scale, the divisors induced by stability conditions in the (H, D)-slice give a ray in $N^1(X^{[n]})$ emanating from the class $H^{[n]} \in \mathrm{Nef}(X) \subset \mathrm{Nef}(X^{[n]})$. The particular ray is determined by the choice of the twisting divisor D.

Proof of Proposition 3.8. Since σ is in the (H, D)-slice, write $\sigma = \sigma_{\beta,\alpha}$ and $(Z, \mathcal{A}) = (Z_{\beta,\alpha}, \mathcal{A}_\beta)$ for short. Put $z = -1/Z(\boldsymbol{v}) = u + iv$. We evaluate the identity

$$\Im(zZ(\boldsymbol{w}')) = (\boldsymbol{w}', \boldsymbol{w}_\sigma)$$

defining $\boldsymbol{w}_\sigma$ on various classes $\boldsymbol{w}'$ to compute $\boldsymbol{w}_\sigma$.

Write the Chern character $\boldsymbol{w}_\sigma = (r, C, d)$. Then

$$-v = \Im(zZ(0, 0, 1)) = ((0, 0, 1), \boldsymbol{w}_\sigma) = r,$$

so $r = -v$. Next, for any curve class C',

$$\begin{aligned}(u + \beta v)(C' \cdot H) + v(C' \cdot D) &= \Im(zZ(0, C', 0)) \\ &= ((0, C', 0), \boldsymbol{w}_\sigma) = \chi(0, -vC', C' \cdot C).\end{aligned}$$

By Riemann–Roch and adjunction,

$$\chi(0,-vC',C'\cdot C) = -v\Big(\Big(-\frac{1}{v}(C'\cdot C)+\frac{1}{2}(C')^2\Big)-\frac{1}{2}(C')^2-\frac{1}{2}(C'\cdot K_X)\Big) = C'\cdot C+\frac{v}{2}(K_X\cdot C'),$$

so

$$C'\cdot C = (u+\beta v)(C'\cdot H)+v(C'\cdot D)-\frac{v}{2}(C'\cdot K_X)$$

for every class C'. Thus for any class C' with $C'\cdot H=0$, we have $C'\cdot C = v(C'\cdot D)-\frac{1}{2}v(C'\cdot K_X)$; it follows that there is some number a with

$$C = -\frac{v}{2}K_X + aH + vD.$$

Considering $C=H$ shows that $a=u+\beta v$. Therefore,

$$\boldsymbol{w}_\sigma = \Big(-v, -\frac{v}{2}K_X+(u+\beta v)H, m\Big),$$

where m is chosen such that $\boldsymbol{w}_\sigma \in \boldsymbol{v}^\perp$.

Finally, a straightforward calculation shows that

$$\frac{u}{v}+\beta = \nu_\sigma(\boldsymbol{v})+\beta = s_W$$

holds for all (β,α) along W. The follow-up statement for Hilbert schemes follows by computing the Donaldson homomorphism. □

3D. *Dual curves.* Suppose D_{σ_0} is the nef divisor corresponding to the Gieseker wall for $X^{[n]}$ in the (H,D)-slice. Showing that D_{σ_0} is an extremal nef divisor amounts to showing that there is some curve $\gamma\subset X^{[n]}$ with $D_{\sigma_0}\cdot\gamma=0$. By the positivity lemma, this happens when γ parametrizes objects of $X^{[n]}$ which are generically S-equivalent with respect to σ_0.

In every case where we computed the Gieseker wall, the wall can be given by a destabilizing subobject which is a line bundle $\mathcal{O}_X(-C)$ with C an effective curve. If Z is a length-n subscheme of C, then there is a destabilizing sequence

$$0\to\mathcal{O}_X(-C)\to I_Z\to I_{Z\subset C}\to 0.$$

If $\operatorname{ext}^1(I_{Z\subset C},\mathcal{O}_X(-C))\ge 2$, then curves of objects of $X^{[n]}$ which are generically S-equivalent with respect to σ_0 are obtained by varying the extension class. We obtain the following general result.

Lemma 3.10. *Suppose the Gieseker wall for $X^{[n]}$ in the (H,D)-slice is computed by the subobject $\mathcal{O}_X(-C)$, where C is an effective curve class of arithmetic genus $p_a(C)$. If $n\ge p_a(C)+1$, then the corresponding nef divisor D_{σ_0} is extremal.*

Proof. Bilinearity of the Euler characteristic $\chi(\cdot,\cdot)$ and Serre duality show that

$$\chi(I_{Z\subset C}, \mathcal{O}_X(-C)) = p_a(C) - 1 - n.$$

Therefore, once $n \geq p_a(C)+1$ we will have $\chi(I_{Z\subset C}, \mathcal{O}_X(-C)) \leq -2$, and curves orthogonal to D_{σ_0} can be constructed by varying the extension class. □

Combining Lemma 3.10 with our previous results on the computation of the Gieseker wall gives us the following asymptotic result.

Theorem 3.11. *Fix a slice* (H, D) *for* $\operatorname{Stab}(X)$. *There is some* $L \in \operatorname{CrDiv}(H, D)$ *such that for all* $n \gg 0$ *the Gieseker wall is computed by* $\mathcal{O}_X(-L)$. *Furthermore, the corresponding nef divisor is extremal.*

Proof. Recall that the set $\operatorname{CrDiv}(H, D)/\!\sim$ of critical divisors modulo numerical equivalence is finite; say $\{L_1, \dots, L_m\}$ is a set of representatives. For $1 \leq i \leq m$, let $(s_i(n), 0)$ be the center of the wall $\mathcal{O}_X(-L_i)$ for $X^{[n]}$. Then $s_i(n)$ is a linear function of n by Lemma 3.3, so there is some i with $s_i(n) \leq s_j(n)$ for all $1 \leq j \leq m$ and $n \gg 0$. Then by Corollary 3.7 the Gieseker wall is given by $\mathcal{O}_X(-L_i)$. Again increasing n if necessary, the divisor D_{σ_0} corresponding to the Gieseker wall is extremal by Lemma 3.10. □

Remark 3.12. The requirement $n \geq p_a(C)+1$ in Lemma 3.10 is not typically sharp. For example, if $|C|$ contains a smooth curve we may as well assume C is smooth. Then $I_{Z\subset C}$ is a line bundle on C, and

$$\operatorname{Ext}^1(I_{Z\subset C}, \mathcal{O}_X(-C)) \cong H^0(\mathcal{O}_C(Z)).$$

Thus a g^1_n on C gives a curve which is orthogonal to D_{σ_0}. The following fact from Brill–Noether theory therefore provides curves on $X^{[n]}$ for smaller values on n.

Lemma 3.13 [Arbarello et al. 1985]. *If C is smooth of genus g, then it has a g^1_n for any* $n \geq \lceil (g+2)/2 \rceil$.

For specific surfaces, some curves in $|C|$ may have highly special linear series giving better constructions of curves on $X^{[n]}$.

4. Picard rank 1 examples

For the rest of the paper, we will apply the methods of Section 3 to compute $\operatorname{Nef}(X^{[n]})$ for several interesting surfaces X. These applications form the heart of the paper.

4A. ***Picard rank 1 in general.*** Suppose $\operatorname{Pic}(X) \cong \mathbb{Z}H$ for some ample divisor H. If we choose $D = -aH$, where $a > 0$ is the smallest positive integer such that aH is effective, then $\operatorname{CrDiv}(H, D) = \{-D\}$.

Lemma 4.1. *Suppose* $\mathrm{Pic}(X) = \mathbb{Z}H$ *and* aH *is the minimal effective class. If* $n \geq (aH)^2 = a^2d$, *then the Gieseker wall for* $X^{[n]}$ *is the wall given by* $\mathcal{O}_X(-aH)$.

Proof. Apply Corollary 3.7 with $D = -aH$. □

Note that when $n > a^2d$, additional information about the Jordan–Hölder filtration can be obtained as in Corollary 3.7. We use formula (7) to see that the wall W given by $\mathcal{O}_X(-aH)$ has center $(s_W, 0)$ with

$$s_W = \frac{a}{2} - \frac{n}{ad}.$$

Combining Lemmas 4.1 and 3.10 with Proposition 3.8, we have proved the following general result.

Theorem 4.2. *Suppose* $\mathrm{Pic}\, X \cong \mathbb{Z}H$ *and* aH *is the minimal effective class. If* $n \geq a^2d$ *then the divisor*

$$\frac{1}{2}K_X^{[n]} + \Big(\frac{a}{2} + \frac{n}{ad}\Big)H^{[n]} - \frac{1}{2}B \tag{8}$$

is nef. Additionally, if $n \geq p_a(aH) + 1$ *then this divisor is extremal, so* $\mathrm{Nef}(X^{[n]})$ *is spanned by this divisor and* $H^{[n]}$. *An orthogonal curve is given by letting* n *points move in a* g^1_n *on a curve of class* aH.

Remark 4.3. If $\mathrm{Pic}(X) = \mathbb{Z}H$ and H is already effective, then a different argument computes the Gieseker wall so long as $2n > d$, improving the bound in Lemma 4.1. However, fine information about the Jordan–Hölder filtration of a destabilized ideal sheaf is not obtained. In fact, if $n \leq d$ then the destabilizing behavior can be complicated. For instance, a scheme Z contained in the complete intersection of two curves of class H will admit an interesting map from $\mathcal{O}_X(-H)^{\oplus 2}$.

Proposition 4.4. *Suppose* $\mathrm{Pic}\, X = \mathbb{Z}H$ *and* H *is effective. If* $2n > d$, *then the Gieseker wall for* $X^{[n]}$ *in the* $(H, -H)$*-slice is the wall given by* $\mathcal{O}_X(-H)$. *Thus the divisor* (8) *with* $a = 1$ *is nef.*

Proof. Let W be the numerical wall given by $\mathcal{O}_X(-H)$. If no actual wall is larger than W then, by the proof of Proposition 3.5, W is an actual wall, and hence the Gieseker wall. If there is a destabilizing sequence

$$0 \to F \to I_Z \to G \to 0$$

giving a wall W' larger than W, then $F, G \in \mathcal{A}_0$ since W passes through the origin in the (β, α)-plane. Fix $\alpha > 0$ such that $(0, \alpha)$ lies on W'. We have

$$H \cdot \mathrm{ch}_1^{-H}(F) = \Im Z_{0,\alpha}(F) \geq 0 \quad \text{and} \quad H \cdot \mathrm{ch}_1^{-H}(G) = \Im Z_{0,\alpha}(G) \geq 0.$$

Since d is the smallest intersection number of H with an integral divisor and

$$d = \Im Z_{0,\alpha}(I_Z) = \Im Z_{0,\alpha}(F) + \Im Z_{0,\alpha}(G),$$

we conclude that either $\Im Z_{0,\alpha}(F) = 0$ or $\Im Z_{0,\alpha}(G) = 0$. Thus either F or G has infinite $\sigma_{0,\alpha}$-slope, contradicting the fact that $(0, \alpha)$ is on W'. □

We now further relax the lower bound on n needed to guarantee the existence of orthogonal curve classes in special cases.

4B. ***Surfaces in*** $\mathbb{P}^3$**.** By the Noether–Lefschetz theorem, a very general surface $X \subset \mathbb{P}^3$ of degree $d \geq 4$ is smooth of Picard rank 1 and irregularity 0. Let H be the hyperplane class and put $D = -H$. We have $K_X = (d-4)H$, so Proposition 4.4 shows that if $2n > d$ then the divisor

$$\left(\frac{d}{2} - \frac{3}{2} + \frac{n}{d}\right)H^{[n]} - \frac{1}{2}B$$

is nef. If C is any smooth hyperplane section then the projection from a point on C gives a map of degree $d-1$ to $\mathbb{P}^1$, so C carries a g^1_n for any $n \geq d-1$. We have proved the following result.

Proposition 4.5. *Let X be a smooth degree-d hypersurface in $\mathbb{P}^3$ with Picard rank* 1*. The divisor*

$$\left(\frac{d}{2} - \frac{3}{2} + \frac{n}{d}\right)H^{[n]} - \frac{1}{2}B$$

on $X^{[n]}$ is nef if $2n > d$. If $n \geq d-1$, then it is extremal, and together with $H^{[n]}$ it spans $\mathrm{Nef}(X^{[n]})$.

Remark 4.6. The behavior of $\mathrm{Nef}(X^{[n]})$ for smaller n in Proposition 4.5 is more mysterious. Even the cases $d = 5$ and $n = 2, 3$ are interesting.

Remark 4.7. The case $d = 4$ of Proposition 4.5 recovers a special case of [Bayer and Macrì 2014b, Proposition 10.3] for K3 surfaces. The case $d = 1$ recovers the computation of the nef cone of $\mathbb{P}^{2[n]}$ [Arcara et al. 2013].

4C. ***Branched covers of*** $\mathbb{P}^2$**.** Next we consider cyclic branched covers of $\mathbb{P}^2$. Let X be a very general cyclic degree-d cover of $\mathbb{P}^2$, branched along a degree-e curve. Note that this means that d necessarily divides e. We can view these covers as hypersurfaces in a weighted projective space, which gives us a Noether–Lefschetz type theorem: $\mathrm{Pic}\, X = \mathbb{Z}H$, generated by the pullback H of the hyperplane class on $\mathbb{P}^2$, provided that X has positive geometric genus. The canonical bundle of X is

$$K_X = -3H + e\left(\frac{d-1}{d}\right)H = \left(\frac{e(d-1)}{d} - 3\right)H.$$

Then X will have positive geometric genus if $e \geq 3d/(d-1)$.

Setting $D = -H$, we see that if $2n > d$ then the divisor class

$$\left(\frac{e(d-1)}{2d} - 1 + \frac{n}{d}\right)H^{[n]} - \frac{1}{2}B$$

is nef by Proposition 4.4. The preimage of a line is a curve of class H, and it carries a g_d^1 given by the map to $\mathbb{P}^2$. Therefore, the above divisor is extremal once $n \geq d$.

Proposition 4.8. *Let X be a very general degree-d cyclic cover of $\mathbb{P}^2$ ramified along a degree-e curve, where d divides e and $e \geq 3d/(d-1)$. The divisor*

$$\left(\frac{e(d-1)}{2d} - 1 + \frac{n}{d}\right) H^{[n]} - \frac{1}{2} B$$

on $X^{[n]}$ is nef if $2n > d$. For $n \geq d$, this class is extremal, and together with $H^{[n]}$ it spans $\mathrm{Nef}(X^{[n]})$.

5. Del Pezzo surfaces of degree 1

Bertram and Coskun [2013] studied the birational geometry of $X^{[n]}$ when X is a minimal rational surface or a del Pezzo surface. In particular, they completely computed the nef cones of all these Hilbert schemes except in the case of a del Pezzo surface of degree 1. The constructions they gave were classical: they produced nef divisors from k-very ample line bundles, and dual curves by letting collections of points move in linear pencils on special curves.

In this section, we will compute the nef cone of $X^{[n]}$, where X is a smooth del Pezzo surface of degree 1. Then $X \cong \mathrm{Bl}_{p_1,\dots,p_8} \mathbb{P}^2$ for distinct points $p_1, \dots, p_8$ with the property that $-K_X$ is ample (see [Manin 1974, Theorem 24.4] or [Beauville 1996, Exercise V.21.1]). This application exhibits the full strength of the methods of Section 3.

5A. ***Notation and statement of results.*** Let H be the class of a line and let the 8 exceptional divisors over the p_i be $E_1, \dots, E_8$, so $\mathrm{Pic}(X) \cong \mathbb{Z}H \oplus \mathbb{Z}E_1 \oplus \cdots \oplus \mathbb{Z}E_8$ and $K_X = -3H + \sum_i E_i$. Recall that a (-1)-*curve* on X is a smooth rational curve of self-intersection -1. It is simplest to describe the dual cone of effective curves. We recommend reviewing Section 2A for notation.

Theorem 5.1. *The cone of curves* $\mathrm{NE}(X^{[n]})$ *is spanned by all the classes $E_{[n]}$ given by (-1)-curves $E \subset X$, the class of a curve contracted by the Hilbert–Chow morphism, and the class $F_{[n]}$, where $F \in |-K_X|$ is an anticanonical curve.*

The 240 (-1)-curves E on X are well-known. The possible classes are

$$(0; 1), \quad (1; 1^2), \quad (2; 1^5), \quad (3; 2, 1^6), \quad (4; 2^3, 1^5), \quad (5; 2^6, 1^2), \quad (6; 3, 2^7),$$

where, e.g., $(4; 2^3, 1^5)$ denotes any class equivalent to

$$4H - 2E_1 - 2E_2 - 2E_3 - E_4 - E_5 - E_6 - E_7 - E_8$$

under the natural action of S_8 on $\mathrm{Pic}(X)$. The cone of curves $\mathrm{NE}(X)$ is spanned by the classes of the (-1)-curves. The Weyl group action on $\mathrm{Pic}(X)$ acts transitively on (-1)-curve classes. It also acts transitively on systems of 8 pairwise disjoint (-1)-curves; dually, it acts transitively on the extremal rays of the nef cone $\mathrm{Nef}(X)$. We refer the reader to [Manin 1974, §26] for details.

Consider the divisor class $(n-1)(-K_X)^{[n]} - \frac{1}{2}B$. If E is any (-1)-curve on X, then $-K_X \cdot E = 1$, so

$$E_{[n]} \cdot \big((n-1)(-K_X)^{[n]} - \tfrac{1}{2}B\big) = (n-1)(-K_X \cdot E) - (n-1) = 0.$$

Let $\Lambda \subset N^1(X^{[n]})$ be the cone spanned by divisors which are nonnegative on all classes $E_{[n]}$ and curves contracted by the Hilbert–Chow morphism. It follows that $\Lambda \supset \mathrm{Nef}(X^{[n]})$ is spanned by $\mathrm{Nef}(X) \subset \mathrm{Nef}(X^{[n]})$ and the single additional class $(n-1)(-K_X)^{[n]} - \frac{1}{2}B$.

However, $\mathrm{Nef}(X^{[n]}) \subset \Lambda$ is a *proper* subcone. Indeed, if $F \in |-K_X|$ is an anticanonical curve then by Riemann–Hurwitz $F_{[n]} \cdot B = 2n$, so $F_{[n]} \cdot \big((n-1)(-K_X)^{[n]} - \frac{1}{2}B\big) = -1$. Let $\Lambda' \subset \Lambda$ be the subcone of $F_{[n]}$-nonnegative divisors. Taking duals, we see that Theorem 5.1 is equivalent to the next result.

Theorem 5.2. *The nef cone of $X^{[n]}$ is Λ'.*

To prove Theorem 5.2, we must show that all the extremal rays of Λ' are actually nef. Suppose $N \in \mathrm{Nef}(X)$ spans an extremal ray of $\mathrm{Nef}(X)$. Then the cone spanned by $N^{[n]}$ and $(n-1)(-K_X)^{[n]} - \frac{1}{2}B$ contains a single ray of $F_{[n]}$-orthogonal divisors, and this ray is an extremal ray of Λ'. Conversely, due to our description of the cone Λ, the extremal rays of Λ' which are not in $\mathrm{Nef}(X)$ are all obtained in this way.

5B. ***Choosing a slice.*** More concretely, making use of the Weyl group action we may as well assume our extremal nef class $N \in \mathrm{Nef}(X)$ is $H - E_1$. The corresponding $F_{[n]}$-orthogonal ray described in the previous paragraph is spanned by

$$(n-1)(-K_X)^{[n]} + \tfrac{1}{2}(H^{[n]} - E_1^{[n]}) - \tfrac{1}{2}B; \tag{9}$$

our job is to show that this class is nef. We will prove this by exhibiting this divisor as the nef divisor on $X^{[n]}$ corresponding to the Gieseker wall for a suitable choice of slice of $\mathrm{Stab}(X)$.

To apply the methods of Section 3, it is convenient to choose our polarization to be

$$P = \big(n - \tfrac{3}{2}\big)(-K_X) + \tfrac{1}{2}(H - E_1)$$

(which depends on n!) and our antieffective class to be $D = K_X$. Observe that P is ample since it is the sum of an ample and a nef class. If we show that the Gieseker wall W in the (P, K_X)-slice has center $(s_W, 0) = (-1, 0)$, then Proposition 3.8 implies the divisor class (9) is nef.

5C. ***Critical divisors.*** Our plan is to apply Corollary 3.7 to compute the Gieseker wall in the (P, K_X)-slice. We must first identify the set $\mathrm{CrDiv}(P, K_X)$ of critical divisors.

Lemma 5.3. *If $n > 2$, then the set $\mathrm{CrDiv}(P, K_X)$ consists of $-K_X$ and the classes L of (-1)-curves on X with $L\cdot(H-E_1) \le 1$.*

When $n = 2$, the above classes are still critical. Additionally, the class $H - E_1$ is critical, as is any sum of two (-1)-curves L_1, L_2 with $L_i\cdot(H-E_1) = 0$.

Proof. Write $2P = A + N$, where $A = (2n-3)(-K_X)$ is ample and $N = H - E_1$ is nef. Then $A\cdot(-K_X) = 2n-3$ and $N\cdot(-K_X) = 2$, so an effective curve class $L \ne -K_X$ is in $\mathrm{CrDiv}(P, K_X)$ if and only if $L\cdot(2P) < 2n-1$.

First suppose $n > 2$, and let $L \in \mathrm{CrDiv}(P, K_X)$. If $L\cdot(-K_X) \ge 2$, then $L\cdot(2P) \ge 4n-6 > 2n-1$, so L is not critical. Therefore, $L\cdot(-K_X) = 1$. Thus any curve of class L is reduced and irreducible. By the Hodge index theorem,

$$L^2 = L^2\cdot(-K_X)^2 \le (L\cdot(-K_X))^2 = 1,$$

with equality if and only if $L = -K_X$. If the inequality is strict, then by adjunction we must have $L^2 = -1$ and L is a (-1)-curve. Since $L\cdot(2P) < 2n-1$, we further have $L\cdot N \le 1$.

Suppose instead that $n = 2$ and $L \in \mathrm{CrDiv}(P, K_X)$. The cases $L\cdot(2P) \le 1$ and $L\cdot(2P) \ge 3$ follow as in the previous case. The only other possibility is that $L\cdot(-K_X) = 2$ and $L\cdot N = 0$. Since $L\cdot N = 0$, the curve L is a sum of curves in fibers of the projection $X \to \mathbb{P}^1$ given by $|N|$. This easily implies the result. □

The next application of Corollary 3.7 completes the proof of Theorems 5.1 and 5.2.

Proposition 5.4. *The Gieseker wall for $X^{[n]}$ in the (P, K_X)-slice has center $(-1, 0)$, and is given by the subobject $\mathcal{O}_X(K_X)$. It coincides with the wall given by $\mathcal{O}_X(-L)$, where L is any (-1)-curve with $L\cdot(H-E_1) = 0$.*

Proof. By (7), the center of the wall for $\mathcal{O}_X(K_X)$ is $(s_W, 0)$ with

$$s_W = \frac{2n - K_X^2}{(2P)\cdot K_X} = -1.$$

A straightforward computation shows $\eta_{P,K_X} < n$ for all $n \ge 2$. Therefore, by Corollary 3.7, the Gieseker wall is computed by a critical divisor.

We only need to verify that no other critical divisor gives a larger wall. Let $L \in \mathrm{CrDiv}(P, K_X)$. By Lemma 3.3, the center of the wall given by $\mathcal{O}_X(-L)$ lies at the point $(s_L, 0)$, where

$$s_L = -\frac{2n + L^2 + 2(K_X\cdot L)}{(2P)\cdot L}.$$

If L is a (-1)-curve, then

$$s_L = -\frac{2n-3}{(2P)\cdot L} = -\frac{2n-3}{2n-3+L\cdot(H-E_1)} \geq -1,$$

with equality if and only if $L\cdot(H-E_1)=0$. This proves the result if $n>2$.

To complete the proof when $n=2$, we only need to consider the additional critical classes mentioned in Lemma 5.3. For every such $L \in \mathrm{CrDiv}(P, K_X)$ we have $L\cdot K_X = -2$ and $L^2 \leq 0$. Thus $s_L \geq 0$ for every such divisor. □

Acknowledgements

This work was initiated at the 2015 Algebraic Geometry Bootcamp preceding the Algebraic Geometry Summer Research Institute organized by the AMS and the University of Utah. We would like to thank the organizers of both programs for providing the wonderful environment where this collaboration could happen. Additionally, we would like to thank Arend Bayer, Izzet Coskun, and Emanuele Macrì for many valuable discussions on Bridgeland stability.

References

[Arbarello et al. 1985] E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves*, vol. I, Grundlehren der Math. Wissenschaften **267**, Springer, New York, 1985. MR 770932 Zbl 0559.14017

[Arcara and Bertram 2013] D. Arcara and A. Bertram, "Bridgeland-stable moduli spaces for K-trivial surfaces", *J. Eur. Math. Soc.* (*JEMS*) **15**:1 (2013), 1–38. MR 2998828 Zbl 1259.14014

[Arcara et al. 2013] D. Arcara, A. Bertram, I. Coskun, and J. Huizenga, "The minimal model program for the Hilbert scheme of points on $\mathbb{P}^2$ and Bridgeland stability", *Adv. Math.* **235** (2013), 580–626. MR 3010070 Zbl 1267.14023

[Bayer and Macrì 2014a] A. Bayer and E. Macrì, "MMP for moduli of sheaves on K3s via wall-crossing: nef and movable cones, Lagrangian fibrations", *Invent. Math.* **198**:3 (2014), 505–590. MR 3279532 Zbl 1308.14011

[Bayer and Macrì 2014b] A. Bayer and E. Macrì, "Projectivity and birational geometry of Bridgeland moduli spaces", *J. Amer. Math. Soc.* **27**:3 (2014), 707–752. MR 3194493 Zbl 1314.14020

[Bayer et al. 2014] A. Bayer, E. Macrì, and P. Stellari, "The Space of Stability Conditions on Abelian Threefolds, and on some Calabi–Yau Threefolds", preprint, 2014. arXiv 1410.1585

[Beauville 1996] A. Beauville, *Complex algebraic surfaces*, 2nd ed., London Mathematical Society Student Texts **34**, Cambridge University Press, 1996. MR 1406314 Zbl 0849.14014

[Beltrametti and Sommese 1991] M. Beltrametti and A. J. Sommese, "Zero cycles and k-th order embeddings of smooth projective surfaces", pp. 33–48 in *Problems in the theory of surfaces and their classification* (Cortona, 1988), edited by F. Catanese et al., Sympos. Math. **XXXII**, Academic Press, London, 1991. MR 1273371 Zbl 0827.14029

[Beltrametti et al. 1989] M. Beltrametti, P. Francia, and A. J. Sommese, "On Reider's method and higher order embeddings", *Duke Math. J.* **58**:2 (1989), 425–439. MR 1016428 Zbl 0702.14010

[Bertram and Coskun 2013] A. Bertram and I. Coskun, "The birational geometry of the Hilbert scheme of points on surfaces", pp. 15–55 in *Birational geometry, rational curves, and arithmetic*, edited by F. Bogomolov et al., Springer, New York, 2013. MR 3114922 Zbl 1273.14032

[Bertram et al. 2014] A. Bertram, C. Martinez, and J. Wang, "The birational geometry of moduli spaces of sheaves on the projective plane", *Geom. Dedicata* **173** (2014), 37–64. MR 3275289 Zbl 1314.14087

[Bridgeland 2007] T. Bridgeland, "Stability conditions on triangulated categories", *Ann. of Math.* (2) **166**:2 (2007), 317–345. MR 2373143 Zbl 1137.18008

[Bridgeland 2008] T. Bridgeland, "Stability conditions on $K3$ surfaces", *Duke Math. J.* **141**:2 (2008), 241–291. MR 2376815 Zbl 1138.14022

[Catanese and Göttsche 1990] F. Catanese and L. Göttsche, "d-very-ample line bundles and embeddings of Hilbert schemes of 0-cycles", *Manuscripta Math.* **68**:3 (1990), 337–341. MR 1065935 Zbl 0729.14006

[Coskun and Huizenga 2014] I. Coskun and J. Huizenga, "Interpolation, Bridgeland stability and monomial schemes in the plane", *J. Math. Pures Appl.* (9) **102**:5 (2014), 930–971. MR 3271294 Zbl 06361756

[Coskun and Huizenga 2016] I. Coskun and J. Huizenga, "The ample cone of moduli spaces of sheaves on the plane", *Algebr. Geom.* **3**:1 (2016), 106–136. MR 3455422

[Ellingsrud et al. 2001] G. Ellingsrud, L. Göttsche, and M. Lehn, "On the cobordism class of the Hilbert scheme of a surface", *J. Algebraic Geom.* **10**:1 (2001), 81–100. MR 1795551 Zbl 0976.14002

[Fogarty 1968] J. Fogarty, "Algebraic families on an algebraic surface", *Amer. J. Math* **90** (1968), 511–521. MR 0237496 Zbl 0176.18401

[Fogarty 1973] J. Fogarty, "Algebraic families on an algebraic surface, II: The Picard scheme of the punctual Hilbert scheme", *Amer. J. Math.* **95** (1973), 660–687. MR 0335512 Zbl 0299.14020

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 0463157 Zbl 0367.14001

[Hassett and Tschinkel 2010] B. Hassett and Y. Tschinkel, "Intersection numbers of extremal rays on holomorphic symplectic varieties", *Asian J. Math.* **14**:3 (2010), 303–322. MR 2755719 Zbl 1216.14012

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting: line bundles and linear series*, Ergebnisse der Mathematik (3) **48**, Springer, Berlin, 2004. MR 2095471 Zbl 1093.14501

[Li and Zhao 2013] C. Li and X. Zhao, "The MMP for deformations of Hilbert schemes of points on the projective plane", preprint, 2013. arXiv 1312.1748

[Liu 2015] W. Liu, "Bayer–Macrì decomposition on Bridgeland moduli spaces over surfaces", preprint, 2015. arXiv 1501.06397

[Maciocia 2014] A. Maciocia, "Computing the walls associated to Bridgeland stability conditions on projective surfaces", *Asian J. Math.* **18**:2 (2014), 263–279. MR 3217637 Zbl 1307.14022

[Maciocia and Meachan 2013] A. Maciocia and C. Meachan, "Rank 1 Bridgeland stable moduli spaces on a principally polarized abelian surface", *Int. Math. Res. Not.* **2013**:9 (2013), 2054–2077. MR 3053413 Zbl 1312.14036

[Manin 1974] Y. I. Manin, *Cubic forms: algebra, geometry, arithmetic*, North-Holland, Amsterdam, 1974. MR 0460349 Zbl 0277.14014

[Matsuki and Wentworth 1997] K. Matsuki and R. Wentworth, "Mumford–Thaddeus principle on the moduli space of vector bundles on an algebraic surface", *Internat. J. Math.* **8**:1 (1997), 97–148. MR 1433203 Zbl 0879.14002

[Nuer 2014] H. Nuer, "Projectivity and Birational Geometry of Bridgeland Moduli spaces on an Enriques Surface", preprint, 2014. arXiv 1406.0908

[Yanagida and Yoshioka 2014] S. Yanagida and K. Yoshioka, "Bridgeland's stabilities on abelian surfaces", *Math. Z.* **276**:1-2 (2014), 571–610. MR 3150219 Zbl 1292.14012

Communicated by János Kollár
Received 2015-11-01 Revised 2016-03-12 Accepted 2016-05-07

bolognese.b@husky.neu.edu *Department of Mathematics, Northeastern University, Boston, MA 02115, United States*

huizenga@psu.edu *Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, United States*

lin.yinb@husky.neu.edu *Department of Mathematics, Northeastern University, Boston, MA 02115, United States*

ebriedl@uic.edu *Department of Mathematics, Statistics, and CS, University of Illinois, Chicago, IL 60607, United States*

schmidt.707@osu.edu *Department of Mathematics, The Ohio State University, Columbus, OH 43210, United States*

mwoolf@ed.ac.uk *Department of Mathematics, Statistics, and CS, University of Illinois, Chicago, IL 60607, United States*

x.zhao@neu.edu *Department of Mathematics, Northeastern University, Boston, MA 02115, United States*

ALGEBRA AND NUMBER THEORY 10:4 (2016)
dx.doi.org/10.2140/ant.2016.10.931

Interpolation for restricted tangent bundles of general curves

Eric Larson

Let $q_1, q_2, \dots, q_n \in \mathbb{P}^r$ be a general collection of points, and $(C, p_1, p_2, \dots, p_n)$ a general marked curve of genus g. We determine when there exists a nondegenerate degree-d map $f : C \to \mathbb{P}^r$ such that $f(p_i) = q_i$ for all i. This is a consequence of our main theorem, which states that the restricted tangent bundle $f^*T_{\mathbb{P}^r}$ of a general curve of genus g, equipped with a general degree-d map f to $\mathbb{P}^r$, satisfies the property of *interpolation*, i.e., that for a general effective divisor D of any degree on C, either $H^0(f^*T_{\mathbb{P}^r}(-D)) = 0$ or $H^1(f^*T_{\mathbb{P}^r}(-D)) = 0$. We also prove an analogous theorem for the twist $f^*T_{\mathbb{P}^r}(-1)$.

1. Introduction

The goal of this paper is to answer the following basic question about incidence conditions for curves.

Question 1.1. Fix a general marked curve $(C, p_1, p_2, \dots, p_n)$ of genus g, and n general points $q_1, q_2, \dots, q_n \in \mathbb{P}^r$. When does there exist a nondegenerate degree-d map $f : C \to \mathbb{P}^r$ so that $f(p_i) = q_i$ for all i?

(An analogous question when $(C, p_1, p_2, \dots, p_n)$ is allowed to vary in moduli was recently answered in [Atanasov et al. 2016] for curves with nonspecial hyperplane section.)

In order for there to exist any nondegenerate degree-d maps $f : C \to \mathbb{P}^r$, the Brill–Noether theorem [Griffiths and Harris 1980] states that the *Brill–Noether number*

$$\rho(d, g, r) := (r+1)d - rg - r(r+1)$$

must be nonnegative; so we assume this is the case for the remainder of the paper. In this case, writing $\mathrm{Map}_d(C, \mathbb{P}^r)$ for the space of nondegenerate degree-d maps $C \to \mathbb{P}^r$, the Brill–Noether theorem additionally gives

$$\dim \mathrm{Map}_d(C, \mathbb{P}^r) = \rho(d, g, r) + \dim \mathrm{Aut}\, \mathbb{P}^r = (r+1)d - rg + r.$$

MSC2010: 14H99.
Keywords: restricted tangent bundle, interpolation.

Thus, the answer to our main question can only be positive when

$$(r+1)d - rg + r - rn \geq 0. \tag{1}$$

Our main theorem will imply (as an immediate consequence) that, conversely, the answer to our main question is positive when the above inequality holds. To state the main theorem, we first need to make a definition.

Definition 1.2. We say that a vector bundle $\mathcal{E}$ on a curve C *satisfies interpolation* if it is nonspecial (i.e., $H^1(\mathcal{E}) = 0$), and for a general effective divisor D of any degree $d \geq 0$, either

$$H^0(\mathcal{E}(-D)) = 0 \quad \text{or} \quad H^1(\mathcal{E}(-D)) = 0.$$

For C reducible, we require that the above holds for an effective divisor D which is general in *some* (not every) component of $\operatorname{Sym}^d C$ (for each $d \geq 0$).

With this notation, we state our main result.

Theorem 1.3. *Let C be a general curve of genus g, equipped with a general nondegenerate map $f : C \to \mathbb{P}^r$ of degree d. Then $f^* T_{\mathbb{P}^r}$ satisfies interpolation.*

To see that this theorem implies a positive answer to the main question subject to the inequality (1), we first note that by basic deformation theory, the map $\operatorname{Map}_d(C, \mathbb{P}^r) \to (\mathbb{P}^r)^n$, defined via $f \mapsto (f(p_i))_{i=1}^n$, is smooth at f provided that $H^1(f^* T_{\mathbb{P}^r}(-p_1 - \cdots - p_n)) = 0$. Note that the left-hand side of (1) equals $\chi(f^* T_{\mathbb{P}^r})$. Consequently, it suffices to show, for C a general curve of genus g, equipped with a general nondegenerate map $f : C \to \mathbb{P}^r$, that $H^1(f^* T_{\mathbb{P}^r}(-p_1 - \cdots - p_n))$ vanishes whenever $\chi(f^* T_{\mathbb{P}^r}(-p_1 - \cdots - p_n)) \geq 0$. But this is immediate provided that $f^* T_{\mathbb{P}^r}$ satisfies interpolation. We have thus shown the following corollary of our main theorem.

Corollary 1.4. *Fix a general marked curve $(C, p_1, p_2, \ldots, p_n)$ of genus g, and n general points $q_1, q_2, \ldots, q_n \in \mathbb{P}^r$. There exists a nondegenerate degree-d map $f : C \to \mathbb{P}^r$ such that $f(p_i) = q_i$ for all i if and only if $\rho(d, g, r) \geq 0$ and*

$$(r+1)d - rg + r - rn \geq 0.$$

Additionally, we prove a similar theorem for the twist of the tangent bundle; this answers a similar question, where d of the n points are constrained to lie on a hyperplane.

Theorem 1.5. *Let C be a general curve of genus g, equipped with a general nondegenerate map $f : C \to \mathbb{P}^r$ of degree d. Then $f^* T_{\mathbb{P}^r}(-1)$ satisfies interpolation if and only if*

$$d - rg - 1 \geq 0.$$

To see the "only if" direction, first note that since C is nondegenerate, the restriction map $H^0(T_{\mathbb{P}^r}(-1)) \to H^0(f^*T_{\mathbb{P}^r}(-1))$ is injective, so $\dim H^0(f^*T_{\mathbb{P}^r}(-1)) \geq \dim H^0(T_{\mathbb{P}^r}(-1)) = r+1$. If $d-rg-1<0$, then $\chi(f^*T_{\mathbb{P}^r}(-1)) = d-rg+r < r+1$; consequently $H^1(f^*T_{\mathbb{P}^r}(-1)) \neq 0$, so $f^*T_{\mathbb{P}^r}(-1)$ does not satisfy interpolation. For the remainder of the paper, we will therefore assume $d-rg-1 \geq 0$ in Theorem 1.5. As before, Theorem 1.5 gives the "if" direction of the next result.

Corollary 1.6. *Fix a general marked curve $(C, p_1, p_2, \dots, p_n)$ of genus g, and n points $q_1, q_2, \dots, q_n \in \mathbb{P}^r$, which are general subject to the constraint that $q_1, q_2, \dots, q_d$ lie on a hyperplane H (for $d \leq n$). There exists a nondegenerate degree-d map $f : C \to \mathbb{P}^r$ such that $f(p_i) = q_i$ for all i if and only if $\rho(d, g, r) \geq 0$ and*

$$(r+1)d - rg + r - rn \geq 0 \quad \textit{and} \quad d - rg - 1 \geq 0.$$

For the "only if" direction, we first compute the dimension of the space of degree-d maps $\mathrm{Map}_d\big((C, p_1 \cup p_2 \cup \dots \cup p_d), (\mathbb{P}^r, H)\big)$, from C to $\mathbb{P}^r$, which send $p_1 \cup p_2 \cup \dots \cup p_d$ to H. Choose coordinates $[x_1 : x_2 : \dots : x_{r+1}]$ on $\mathbb{P}^r$ so H is given by $x_{r+1} = 0$. Then any such map is given by $r+1$ sections $[s_1 : s_2 : \dots : s_r : 1]$ of $\mathcal{O}_C(p_1 + p_2 + \dots + p_d)$, where we write $s_{r+1} = 1 \in H^0(\mathcal{O}_C) \subset H^0(\mathcal{O}_C(p_1 + \dots + p_d))$ for the constant section. By taking the sections $s_1, s_2, \dots, s_r$, we have identified $\mathrm{Map}_d\big((C, p_1 \cup p_2 \cup \dots \cup p_d), (\mathbb{P}^r, H)\big)$ as an open subset of $H^0(\mathcal{O}_C(p_1 + \dots + p_d))^r$, so its dimension is $r(d+1-g)$. Therefore, in order for $f \mapsto (f(p_i))_{i=1}^n$ to dominate $H^d \times (\mathbb{P}^r)^{n-d}$, we must also have

$$r(d+1-g) \geq (r-1)d + r(n-d) \quad \Longleftrightarrow \quad (r+1)d - rg + r - rn \geq 0.$$

In addition, $f \mapsto (f(p_i))_{i=1}^d$ must dominate H^d. But the fibers of this map have a free action of the $(r+1)$-dimensional group $\mathrm{Stab}_H(\mathrm{Aut}\,\mathbb{P}^r)$; we must therefore also have

$$r(d+1-g) \geq (r-1)d + (r+1) \quad \Longleftrightarrow \quad d - rg - 1 \geq 0.$$

Remark 1.7. For $k \geq 2$, the bundles $f^*T_{\mathbb{P}^r}(-k)$ almost never satisfy interpolation. Namely, in the setting of Theorem 1.3, the bundle $f^*T_{\mathbb{P}^r}(-2)$ satisfies interpolation if and only if either $(g, r) = (0, 1)$ or $(d, g, r) = (2, 0, 2)$; and $f^*T_{\mathbb{P}^r}(-k)$ never satisfies interpolation for $k \geq 3$.

For $r = 1$, this can be seen by observing that $f^*T_{\mathbb{P}^1}(-k) \simeq \mathcal{O}_C(2-k)$ is a line bundle, so it satisfies interpolation if and only if it is nonspecial (see Proposition 4.7 of [Atanasov et al. 2016]); as $k \geq 2$, this only happens if $k = 2$ and C is rational ($g = 0$).

In general, if $f^*T_{\mathbb{P}^r}(-k)$ satisfies interpolation,

$$\chi(f^*T_{\mathbb{P}^r}(-k)) = (r+1)d - rg + r - krd \geq 0.$$

When $r \geq 2$ and $k \geq 2$, this is only satisfied for $(d, g, r, k) = (2, 0, 2, 2)$; conversely, one may easily check that $f^*T_{\mathbb{P}^2}(-2)$ satisfies interpolation for $f : \mathbb{P}^1 \to \mathbb{P}^2$ a general degree-2 map (for instance by combining Theorem 1.3 with Proposition 4.12 of [Atanasov et al. 2016]).

The remainder of the paper will be devoted to the proof of Theorem 1.3 and Theorem 1.5 using inductive degeneration. In Section 2, we begin by explaining how to use degeneration to approach the main theorems, and how to work with the condition of interpolation on reducible curves. Then in Section 3, we prove the main theorems for rational curves by inductively degenerating the curve C to a union $D \cup L$, where L is a 1-secant line to D. Finally in Section 4, we prove the main theorems for arbitrary genus by degenerating the curve C to a union $D \cup R$, where R is a rational normal curve meeting D at $1 \leq s \leq r + 2$ points.

2. Preliminaries

Lemma 2.1. *The locus of $f : C \to \mathbb{P}^r$ in $\overline{M}_g(\mathbb{P}^r, d)$ for which $f^*T_{\mathbb{P}^r}$ satisfies interpolation is open, and likewise for $f^*T_{\mathbb{P}^r}(-1)$. Moreover, every component of these loci dominates $\overline{M}_g$.*

In particular, to prove Theorem 1.3 (respectively, Theorem 1.5) for curves of degree d and genus g, it suffices to exhibit one, possibly singular, nondegenerate $f : C \to \mathbb{P}^r$ of degree d and genus g, for which $f^*T_{\mathbb{P}^r}$ (respectively, $f^*T_{\mathbb{P}^r}(-1)$) satisfies interpolation.

Proof. Since the vanishing of cohomology groups is an open condition, it follows that interpolation is an open condition as well. (For a more careful proof, see Theorem 5.8 of [Atanasov 2014].)

If $f^*T_{\mathbb{P}^r}$ (respectively, $f^*T_{\mathbb{P}^r}(-1)$) satisfies interpolation, then in particular $H^1(f^*T_{\mathbb{P}^r}) = 0$ (respectively, $H^1(f^*T_{\mathbb{P}^r}(-1)) = 0$). Since $H^1(f^*T_{\mathbb{P}^r}(-1)) = 0$ implies $H^1(f^*T_{\mathbb{P}^r}) = 0$, we know either way that $H^1(f^*T_{\mathbb{P}^r}) = 0$. This completes the proof as the obstruction to smoothness of $\overline{M}_g(\mathbb{P}^r, d) \to \overline{M}_g$ lies in $H^1(f^*T_{\mathbb{P}^r})$. $\square$

In order to work with reducible curves, it will be helpful to introduce the following somewhat more general variant on interpolation.

Definition 2.2. Let $\mathcal{E}$ be a rank n vector bundle over a curve C. We say that a subspace of sections $V \subseteq H^0(\mathcal{E})$ satisfies interpolation if $\mathcal{E}$ is nonspecial and, for every $d \geq 0$, there exists an effective Cartier divisor D of degree d such that

$$\dim\bigl(V \cap H^0(\mathcal{E}(-D))\bigr) = \max\{0, \dim V - dn\}.$$

We now state for the reader the elementary properties of interpolation that we shall need.

Lemma 2.3 [Atanasov et al. 2016, Proposition 4.5]. *A vector bundle $\mathcal{E}$ satisfies interpolation if and only if its full space of sections $V = H^0(\mathcal{E}) \subseteq H^0(\mathcal{E})$ satisfies interpolation.*

Proof sketch. The result follows by examining the long exact sequence

$$0 \to H^0(\mathcal{E}(-D)) \to H^0(\mathcal{E}) \to H^0(\mathcal{E}|_D) \to H^1(\mathcal{E}(-D)) \to H^1(\mathcal{E}) = 0.$$

A complete proof is given at the location cited; see also Definition 4.1 in the same reference. □

Lemma 2.4 [Atanasov et al. 2016, Proposition 8.1]. *Let $\mathcal{E}$ be a vector bundle on a reducible curve $X \cup Y$, and D be an effective divisor on X disjoint from $X \cap Y$. Assume that*

$$H^0(\mathcal{E}|_X(-D - X \cap Y)) = 0.$$

Let

$$\mathrm{ev}_X : H^0(\mathcal{E}|_X) \to H^0(\mathcal{E}|_{X\cap Y}), \quad \mathrm{ev}_Y : H^0(\mathcal{E}|_Y) \to H^0(\mathcal{E}|_{X\cap Y})$$

denote the natural evaluation morphisms. Then $\mathcal{E}$ satisfies interpolation provided

$$V = \mathrm{ev}_Y^{-1}\big(\mathrm{ev}_X\big(H^0(\mathcal{E}|_X(-D))\big)\big) \subseteq H^0(\mathcal{E}|_Y)$$

satisfies interpolation and has dimension $\chi(\mathcal{E}|_Y) + \chi(\mathcal{E}|_X(-D - X \cap Y))$.

Proof sketch. As $H^0(\mathcal{E}|_X(-D - X \cap Y)) = 0$, restriction to Y gives an isomorphism $H^0(\mathcal{E}(-D)) \simeq V$. Also, the final dimension statement implies $H^1(\mathcal{E}(-D)) = 0$. Therefore $\mathcal{E}(-D)$, and hence $\mathcal{E}$, satisfies interpolation. □

Lemma 2.5 [Atanasov 2014, Theorem 8.1 and Section 3]. *Let $\mathcal{E}$ be a vector bundle on an irreducible curve C, and $p \in C_{\mathrm{sm}}$ be a general point. If $\mathcal{E}$ satisfies interpolation, and $\Lambda \subseteq \mathcal{E}|_p$ is a general subspace of any dimension, then*

$$\{\sigma \in H^0(\mathcal{E}) : \sigma|_p \in \Lambda\} \subseteq H^0(\mathcal{E})$$

satisfies interpolation and has dimension $\max\{0, \chi(\mathcal{E}) - \operatorname{codim} \Lambda\}$.

Proof sketch. Let $n = \operatorname{rk} \mathcal{E}$, and D be a general effective divisor of any degree $d \geq 0$. Since $\mathcal{E}$ satisfies interpolation,

$$\dim H^0(\mathcal{E}(-D)) = \max\{0, \chi(\mathcal{E}) - dn\},$$
$$\dim H^0(\mathcal{E}(-D - p)) = \max\{0, \chi(\mathcal{E}) - dn - n\}.$$

These inequalities imply the restriction map $H^0(\mathcal{E}(-D)) \to \mathcal{E}|_p$ is either injective or surjective, which (since Λ is general) in turn implies the composition $H^0(\mathcal{E}(-D)) \to \mathcal{E}|_p \to \mathcal{E}|_p/\Lambda$ is either injective or surjective. Together with the first of the above equalities, this implies

$$\dim\{\sigma \in H^0(\mathcal{E}(-D)) : \sigma|_p \in \Lambda\} = \max\{0, \chi(\mathcal{E}) - \operatorname{codim} \Lambda - dn\}.$$ □

3. Rational curves

In this section, we prove Theorem 1.3 in the case $g = 0$.

Proposition 3.1. *Let $L \subseteq \mathbb{P}^r$ be a line. Then $T_{\mathbb{P}^r}|_L \simeq \mathcal{O}_L(2) \oplus \mathcal{O}_L(1)^{r-1}$, where the $\mathcal{O}_L(2)$ summand comes from the inclusion $T_L \hookrightarrow T_{\mathbb{P}^r}|_L$; in particular, $T_{\mathbb{P}^r}|_L$ and $T_{\mathbb{P}^r}|_L(-1)$ satisfy interpolation.*

Proof. The first assertion follows from the exact sequence

$$0 \to T_L \simeq \mathcal{O}_L(2) \to T_{\mathbb{P}^r}|_L \to N_{L/\mathbb{P}^r} \simeq \mathcal{O}_L(1)^{r-1} \to 0.$$

(We have $N_{L/\mathbb{P}^r} \simeq \mathcal{O}_L(1)^{r-1}$ since L is the complete intersection of $r-1$ hyperplanes.)

The second assertion follows from the first by inspection (or alternatively using Proposition 3.11 of [Atanasov 2014]). □

Proposition 3.2. *Let $f : \mathbb{P}^1 \to \mathbb{P}^r$ be a general degree-d map (allowed to be degenerate if $d < r$). Then $f^*T_{\mathbb{P}^r}(-1)$, and consequently $f^*T_{\mathbb{P}^r}$, satisfies interpolation.*

Proof. To show $f^*T_{\mathbb{P}^r}(-1)$ satisfies interpolation, we argue by induction on the degree d of f; the base case $d = 1$ is given by Proposition 3.1. For the inductive step $d \geq 2$, we degenerate $f : \mathbb{P}^1 \to \mathbb{P}^r$ to a map $g : D \cup_p L \to \mathbb{P}^r$ from a two-component reducible rational curve; write $g_D = g|_D$ and $g_L = g|_L$. Assume that $\deg g_D = d-1$ and $\deg g_L = 1$. By Lemma 2.1, it suffices to show $g^*T_{\mathbb{P}^r}(-1)$ satisfies interpolation. Write

$$\mathrm{ev}_D : H^0(g_D^* T_{\mathbb{P}^r}(-1)) \to H^0(T_{\mathbb{P}^r}(-1)|_p),$$
$$\mathrm{ev}_L : H^0(g_L^* T_{\mathbb{P}^r}(-1)) \to H^0(T_{\mathbb{P}^r}(-1)|_p)$$

for the natural evaluation morphisms. Pick a point $x \in L \setminus p$. Then by Lemma 2.4, it suffices to show that

$$V = \mathrm{ev}_D^{-1}\big(\mathrm{ev}_L\big(H^0(g_L^* T_{\mathbb{P}^r}(-1)(-x))\big)\big) \subseteq H^0(g_D^* T_{\mathbb{P}^r}(-1))$$

satisfies interpolation and has dimension

$$\chi(g_D^* T_{\mathbb{P}^r}(-1)) + \chi(g_L^* T_{\mathbb{P}^r}(-1)(-x-p)) = \chi(g_D^* T_{\mathbb{P}^r}(-1)) - (r-1).$$

Using the description of $g_L^* T_{\mathbb{P}^r}$ from Proposition 3.1,

$$V = \{\sigma \in H^0(g_D^* T_{\mathbb{P}^r}(-1)) : \sigma|_p \in T_L(-1)|_p\}.$$

Since $T_L(-1)|_p \subseteq T_{\mathbb{P}^r}(-1)|_p$ is a general subspace of codimension $r-1$, Lemma 2.5 implies V satisfies interpolation and has dimension $\max\{0, \chi(g_D^* T_{\mathbb{P}^r}(-1)) - (r-1)\}$. It thus suffices to note that

$$\chi(g_D^* T_{\mathbb{P}^r}(-1)) - (r-1) = d + 1 \geq 0.$$

By inspection (or alternatively using Proposition 4.11 of [Atanasov et al. 2016]), interpolation for $f^*T_{\mathbb{P}^r}(-1)$ implies interpolation for $f^*T_{\mathbb{P}^r}$. □

4. Curves of higher genus

Lemma 4.1. *Let $\mathcal{E}$ be a vector bundle of rank n on a reducible nodal curve $X \cup Y$, such that $X \cap Y$ is a general collection of k points on X (relative to $\mathcal{E}|_X$). Suppose that $\mathcal{E}|_X$ and $\mathcal{E}|_Y$ satisfy interpolation. If $\chi(\mathcal{E}|_X) \equiv 0 \mod n$ and $\chi(\mathcal{E}|_X) \geq nk$, then $\mathcal{E}$ satisfies interpolation.*

Proof. Let D be a general divisor on X (in particular disjoint from $X \cap Y$) of degree $(\chi(\mathcal{E}|_X) - nk)/n$. Write

$$\mathrm{ev}_X : H^0(\mathcal{E}|_X) \to H^0(\mathcal{E}|_{X\cap Y}), \quad \mathrm{ev}_Y : H^0(\mathcal{E}|_Y) \to H^0(\mathcal{E}|_{X\cap Y})$$

for the natural evaluation morphisms, and let

$$V = \mathrm{ev}_Y^{-1}\big(\mathrm{ev}_X\big(H^0(\mathcal{E}|_X(-D))\big)\big) \subseteq H^0(\mathcal{E}|_Y).$$

Since $\mathcal{E}|_X$ satisfies interpolation and $\chi(\mathcal{E}|_X(-D)) = nk$, we conclude ev_X is an isomorphism when restricted to $H^0(\mathcal{E}|_X(-D))$. In particular,

$$H^0(\mathcal{E}|_X(-D - X \cap Y)) = 0,$$

and $V = H^0(\mathcal{E}|_Y)$ is the full space of sections. Because $\mathcal{E}|_Y$ satisfies interpolation, we conclude V satisfies interpolation and has dimension

$$\chi(\mathcal{E}|_Y) = \chi(\mathcal{E}|_Y) + \chi(\mathcal{E}|_X(-D - X \cap Y)).$$

This implies, via Lemma 2.4, that $\mathcal{E}$ satisfies interpolation. □

Proof of Theorem 1.3. We argue by induction on g; the base case $g = 0$ is given by Proposition 3.2. For the inductive step $g \geq 1$, we let

$$(s, g', d') = \begin{cases} (r+2, g-r-1, d-r) & \text{if } g \geq r+1, \\ (g+1, 0, d-r) & \text{otherwise.} \end{cases}$$

By construction, $g' \geq 0$ and $1 \leq s \leq r+2$. Moreover, since $\rho(d, g, r) \geq 0$, we have either $\rho(d', g', r) \geq 0$, or $g' = 0$ and $d' \geq 1$; and in the second case $d' \geq s - 1$.

We now degenerate $f : C \to \mathbb{P}^r$ to a map $g : D \cup_\Gamma \mathbb{P}^1 \to \mathbb{P}^r$ from a two-component reducible curve; write $g_D = g|_D$ and $g_{\mathbb{P}^1} = g|_{\mathbb{P}^1}$. By the above, we may take D to be a general curve of genus g', and Γ a collection of s points general on both D and $\mathbb{P}^1$. We also take g_D and $g_{\mathbb{P}^1}$ to be general maps of degrees d' and r respectively (composing with an automorphism of $\mathbb{P}^r$ so that $g_D(\Gamma) = g_{\mathbb{P}^1}(\Gamma)$ — which exists since $\operatorname{Aut}\mathbb{P}^r$ acts $(r+2)$-transitively on points in linear general position; note that $s \le r+2$, and $g_{\mathbb{P}^1}$ is nondegenerate, while g_D spans at least a $\mathbb{P}^{\min(r,s-1)}$). By Lemma 2.1, it suffices to show $g^* T_{\mathbb{P}^r}$ satisfies interpolation.

By induction (and direct application of Proposition 3.2 in the case $g' = 0$ and $d' \ge 1$), we know $g_D^* T_{\mathbb{P}^r}$ and $g_{\mathbb{P}^1}^* T_{\mathbb{P}^r}$ satisfy interpolation. Moreover, $\chi(g_{\mathbb{P}^1}^* T_{\mathbb{P}^r}) = r(r+2)$ is a multiple of r which is at least rs. Lemma 4.1 thus yields the desired conclusion. □

Proof of Theorem 1.5. We argue by induction on g; the base case $g = 0$ is given by Proposition 3.2. For the inductive step $g \ge 1$, we write $g' = g - 1 \ge 0$ and $d' = d - r$. Since $d - rg - 1 \ge 0$, we have $d' - rg' - 1 \ge 0$. This in turn implies either $\rho(d', g', r) \ge 0$, or $g' = 0$ and $d' \ge 1$.

We now degenerate $f : C \to \mathbb{P}^r$ to a map $g : D \cup_\Gamma \mathbb{P}^1 \to \mathbb{P}^r$ from a two-component reducible curve; write $g_D = g|_D$ and $g_{\mathbb{P}^1} = g|_{\mathbb{P}^1}$. By the above, we may take D to be a general curve of genus g', and Γ a collection of 2 points general on both D and $\mathbb{P}^1$. We also take g_D and $g_{\mathbb{P}^1}$ to be general maps of degrees d' and r respectively (composing with an automorphism of $\mathbb{P}^r$ so that $g_D(\Gamma) = g_{\mathbb{P}^1}(\Gamma)$ — which exists since $\operatorname{Aut}\mathbb{P}^r$ acts 2-transitively). By Lemma 2.1, it suffices to show $g^* T_{\mathbb{P}^r}(-1)$ satisfies interpolation.

By induction (and direct application of Proposition 3.2 in the case $g' = 0$ and $d' \ge 1$), we know $g_D^* T_{\mathbb{P}^r}(-1)$ and $g_{\mathbb{P}^1}^* T_{\mathbb{P}^r}(-1)$ satisfy interpolation. Moreover, $\chi(g_{\mathbb{P}^1}^* T_{\mathbb{P}^r}(-1)) = 2r$. Lemma 4.1 thus yields the desired conclusion. □

Acknowledgements

The author would like to thank Joe Harris for his guidance throughout this research, as well as Atanas Atanasov and David Yang for helpful conversations. The author would also like to acknowledge the generous support both of the Fannie and John Hertz Foundation, and of the Department of Defense (NDSEG fellowship).

References

[Atanasov 2014] A. Atanasov, "Interpolation and vector bundles on curves", preprint, 2014. arXiv

[Atanasov et al. 2016] A. Atanasov, E. Larson, and D. Yang, "Interpolation for normal bundles of general curves", preprint, 2016. arXiv

[Griffiths and Harris 1980] P. Griffiths and J. Harris, "On the variety of special linear systems on a general algebraic curve", *Duke Math. J.* **47**:1 (1980), 233–272. MR Zbl

Communicated by Ravi Vakil
Received 2015-11-30 Accepted 2016-05-16

elarson3@gmail.com *Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, United States*

mathematical sciences publishers msp

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 10 No. 4 2016

1937-0652(2016)10:4;1-L