Algebra & Number Theory

Volume 10
2016
No. 5

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 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.939

Conjugacy classes of special automorphisms of the affine spaces

Jérémy Blanc

In the group of polynomial automorphisms of the plane, the conjugacy class of an element is closed if and only if the element is diagonalisable. In this article, we show that this does not hold for the group of special automorphisms, giving a first step in the direction of showing that this group is not simple, as an infinite-dimensional algebraic group.

1. Introduction

In this article, k will always denote an algebraically closed field. The conjugacy classes of the algebraic groups $\mathrm{GL}(n, \mathrm{k})$ and $\mathrm{SL}(n, \mathrm{k})$ are well known. In particular, the following observation is classical:

An element is diagonalisable if and only if its conjugacy class is Zariski-closed.

As observed in [Furter and Maubach 2010], the same holds for the group $\mathrm{Aut}(\mathbb{A}^2_{\mathbb{C}})$ of complex polynomial automorphisms of the affine plane. Here, the topology corresponds to the topology of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ induced by families parametrised by algebraic varieties A, called morphisms $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ and corresponding to elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}[A]})$ (see Section 2A).

In fact, there is one easy direction in the result of [Furter and Maubach 2010], which corresponds to showing that if the conjugacy class is closed, then the element is diagonalisable. This works over any algebraically closed field k and follows from the following observation: If $f \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ is an element that fixes the origin, the conjugation of f by

$$(x_1, \ldots, x_n) \longmapsto \left(\frac{1}{t}x_1, \ldots, \frac{1}{t}x_n\right)$$

The author gratefully acknowledges support by the Swiss National Science Foundation Grant “Birational Geometry” PP00P2_153026 /1 and by the French National Research Agency Grant “BirPol”, ANR-11-JS01-004-01.

MSC2010: primary 14R10; secondary 14R20.

Keywords: plane automorphisms, conjugacy classes, topology.

yields an element of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}(t)})$ whose value at $t=0$ is the linear part of f, which is an element of $\mathrm{GL}(n,\mathrm{k})$. Moreover elements of $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ which do not have fixed points are easy to handle (these are conjugate to $(x_1,x_2)\mapsto(x_1+1,ax_2+P(x_2))$ for some polynomial $P\in\mathrm{k}[x_2]$ and $a\in\mathrm{k}^*$).

In this article, we focus on the closed normal subgroup $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ of elements of Jacobian 1. We will show that the conjugacy classes of the two groups $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ and $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ have a very different behaviour.

We say that an element $f\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ is *dynamically regular* if the extensions of f and f^{-1} to $\mathbb{P}^n_{\mathrm{k}}$ have disjoint indeterminacy loci (see Section 2B). We will also say that f is *algebraic* if $\{\deg(f^n)\}_{n\in\mathbb{N}}$ is bounded. The dynamically regular elements are never algebraic, and in dimension 2, nonalgebraic elements are conjugate to dynamically regular elements (see Remark 4.3).

The first result that we obtain is to show that there is no degeneration of conjugates of dynamically regular elements in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, contrary to the case of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$.

Theorem 1.1. *Let $f\in\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ be a dynamically regular element.*

(1) *If $\alpha\in\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}((t))})$ is such that $\alpha f\alpha^{-1}$ has a value at $t=0$, then this value is conjugate to f by $\alpha(0)$ in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ (in particular α is defined at $t=0$).*

(2) *For each integer d, the set $\{gfg^{-1}\mid g\in\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}}),\ \deg(g)\le d\}$ is closed.*

(3) *If $n=2$, the conjugacy class of f in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ is closed.*

(4) *If k is uncountable, the following holds: for each morphism $A\to\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, where A is an algebraic variety, the preimage of the conjugacy class of f contains the closure of each locally closed subset $B\subset A$ that it contains.*

Remark 1.2. The proof of this result is given in Section 3. As we will show, part (1) implies the others.

In the notation of [Furter and Kraft $\geq$ 2016], assertion (4) can be reinterpreted by saying that the conjugacy class $C(f)$ of a dynamically regular element f is *weakly closed* in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$.

In dimension 2, an easy consequence of Theorem 1.1 and of the Jung–van der Kulk theorem is the fact the conjugacy class of nonalgebraic elements of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is in fact closed, contrary to the case of $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$. In fact, we can be much more precise: we describe in Section 4 the conjugacy classes of elements in $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$, and decide which ones are closed. In particular, we obtain the following complete description.

Theorem 1.3. *Let $f\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$. Then:*

(1) *If f is diagonalisable, its conjugacy class is closed.*

(2) *If f is algebraic but not diagonalisable, its conjugacy class is not closed. More precisely, there exists an element $F \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}[t]})$ such that for each $t \neq 0$, $F(t) \in \mathrm{SAut}(\mathbb{A}^2_K)$ is conjugate to f, and $F(0) \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is diagonalisable.*

(3) *If f is not algebraic, its conjugacy class is closed.*

Remark 1.4. Since the set $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})_{\mathrm{alg}}$ of algebraic elements of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is closed (Corollary 4.4), we can decompose $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ as an infinite union of disjoint closed sets, one being $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})_{\mathrm{alg}}$ and the others being conjugacy classes of nonalgebraic elements. The group $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is however irreducible, by the simple observation made above.

Remark 1.5. Note that these results show that the group $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is more rigid than the group $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$, in the sense that there are less possible degenerations of conjugates.

One can check, using the conjugations around fixed points as above, that every normal closed subgroup of $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ is either trivial, $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ or $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$. The interesting question is then to know whether $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ contains nontrivial closed normal subgroups (by [Furter and Lamy 2010], it contains many nontrivial normal subgroups which contain only nonalgebraic elements, and the identity).

The fact that the conjugacy classes of nonalgebraic elements are closed suggests that $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ could contain nontrivial closed normal subgroups. This text can then be viewed as a first step towards the study of the simplicity of $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, viewed as an infinite-dimensional algebraic group (ind-group). In [Shafarevich 1966], it is claimed that this one is simple, but the proof contains serious gaps.

2. Preliminaries

As we said, in the sequel k will always be an algebraically closed field. We will sometimes also work on a general field (most of the time with an extension of k), and will denote it by K.

In this section, we introduce the terminology and give some basic results (most of them classical, maybe in alternate formulations) on the topology of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ (Section 2A), the relation between the iterations of a map and the indeterminacy sets at infinity (Section 2B) and the families of automorphisms parametrised by formal series (Section 2C), that we will need in Sections 3 and 4.

2A. *Topology on* $\mathbf{Aut}(\mathbb{A}^n_{\mathrm{k}})$.

Notation 2.1. Let R be any commutative unitary ring.

(1) We denote by $\mathrm{End}(\mathbb{A}^n_R)$ the set of algebraic endomorphisms of $\mathbb{A}^n_R$. An element $f \in \mathrm{End}(\mathbb{A}^n_R)$ is given by

$$f : (x_1, \dots, x_n) \longmapsto (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))$$

for some polynomials $f_1, \dots, f_n \in R[x_1, \dots, x_n]$. The degree of f is by definition the maximal degree of the f_i, and we will use the notation

$$f = (f_1, \dots, f_n).$$

This corresponds to a natural bijection $\mathrm{End}(\mathbb{A}^n_R) \to (R[x_1, \dots, x_n])^n$.

(2) The group $\mathrm{Aut}(\mathbb{A}^n_R)$ is equal to the group of automorphisms of $\mathbb{A}^n_R$, i.e., to the elements of $\mathrm{End}(\mathbb{A}^n_R)$ that admit an inverse in this set.

(3) For each $f \in \mathrm{End}(\mathbb{A}^n_R)$, we denote by

$$\mathrm{Jac}(f) = \det\left(\frac{\partial f_i}{x_j}\right)^n_{i,j=1} \in R[x_1, \dots, x_n]$$

the Jacobian of f, and denote by $\mathrm{SAut}(\mathbb{A}^n_R)$, the normal subgroup of $\mathrm{Aut}(\mathbb{A}^n_R)$ given by $\{f \in \mathrm{Aut}(\mathbb{A}^n_R) \mid \mathrm{Jac}(f) = 1\}$.

(4) We denote by $\mathrm{End}(\mathbb{A}^n_R)_{\le d}$ and $\mathrm{Aut}(\mathbb{A}^n_R)_{\le d}$ the subsets of $\mathrm{End}(\mathbb{A}^n_R)$ and $\mathrm{Aut}(\mathbb{A}^n_R)$ respectively, given by elements of degree $\le d$.

Example 2.2. For each $p_1 \in R[x_1]$, $p_2 \in R[x_1, x_2], \dots, p_{n-1} \in R[x_1, \dots, x_{n-1}]$ and $a_1, \dots, a_n \in R^*$ the element

$$(a_1x_1, a_2x_2 + p_1, a_3x_3 + p_2, \dots, a_nx_n + p_{n-1})$$

belongs to $\mathrm{Aut}(\mathbb{A}^n_R)$. Such elements are usually called *triangular*, or *de Jonquières*.

Remark 2.3. Suppose that R is a field K. Extending the scalars to an algebraically closed field, we observe that the Jacobian matrix of every element of $\mathrm{Aut}(\mathbb{A}^n_K)$ is invertible everywhere, so $\mathrm{Jac}(f) \in K^*$. In particular,

$$\mathrm{Aut}(\mathbb{A}^n_K) \subset \{f \in \mathrm{End}(\mathbb{A}^n_K) \mid \mathrm{Jac}(f) \in K^*\},$$

and the equality, when K is of characteristic zero, is the classical Jacobian conjecture, open for any $n \ge 2$.

If Z is an algebraic variety defined over k, where k is algebraically closed as before, there is a natural way to endow the group $\mathrm{Bir}(Z)$ of birational transformations of Z with a topology (see for example [Demazure 1970; Serre 2010; Blanc 2010; Blanc and Furter 2013]). When restricted to the subgroup $\mathrm{Aut}(Z)$ of automorphisms, we obtain the following:

Definition 2.4. Let A, Z be two algebraic varieties defined over k. We say that a *morphism* $f : A \to \mathrm{Aut}(Z)$ is a map given by an A-automorphism of $A \times Z$.

The *Zariski topology* on $\mathrm{Aut}(Z)$ is defined as follows: a set $F \subset \mathrm{Aut}(Z)$ is closed if and only if $f^{-1}(F) \subset A$ is closed for any algebraic variety A and any morphism $f : A \to \mathrm{Aut}(Z)$.

Remark 2.5. When the group $\mathrm{Aut}(Z)$ has a natural structure of an algebraic group (for example when $Z = \mathbb{P}^n$), the topology defined above agrees with the classical topology of the algebraic group, and morphisms $A \to \mathrm{Aut}(Z)$ correspond to morphisms of algebraic varieties.

However, in general the group $\mathrm{Aut}(Z)$ is too big to be an algebraic variety, for instance for $Z = \mathbb{A}^n$, $n \geq 2$ (see Example 2.2).

We will observe (in Lemma 2.7 below) that this topology, when restricted to $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, gives the one introduced by Shafarevich in [Shafarevich 1966], i.e., the inductive limit topology given by the inclusion of affine algebraic varieties

$$\mathrm{Aff}(\mathbb{A}^n_{\mathrm{k}}) = \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq 1} \subset \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq 2} \subset \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq 3} \subset \cdots$$

and corresponds in fact to an infinite-dimensional algebraic group. In order to do this, we recall how one obtains natural structures of affine varieties for $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ and $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$.

Lemma 2.6. *Let us fix some integers $d, n \geq 1$, and see $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ as an affine space, via the bijection $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}}) \to (\mathrm{k}[x_1, \ldots, x_n]_{\leq d})^n$. Then, the following hold*:

(1) $J_{\mathrm{k}^*} = \{f \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \mid \mathrm{Jac}(f) \in \mathrm{k}^*\}$ *is locally closed in* $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$, *and inherits from it the structure of an affine variety.*

(2) $J_1 = \{f \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \mid \mathrm{Jac}(f) = 1\}$ *is closed in* $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$.

(3) $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ *is a closed subset of* J_{k^*}.

(4) $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ *is a closed subset of* J_1.

Proof. The Jacobian being a morphism $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \to \mathrm{k}[x_1, \ldots, x_n]_{\leq n(d-1)}$, the sets

$$J_k = \{f \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \mid \mathrm{Jac}(f) \in \mathrm{k}\} \text{ and } J_1$$

are closed in $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ and are thus affine algebraic varieties. Since $J_0 = \{f \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \mid \mathrm{Jac}(f) = 0\}$ is given by one equation in J_{k}, the set $J_{\mathrm{k}^*} = J_{\mathrm{k}} \setminus J_0$ is affine (and locally closed in $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$). This yields assertions (1) and (2).

In order to show (3) and (4), we define W_d to be the set of nonzero $(n+1)$-tuples $(h_0, \ldots, h_n)$ of homogeneous polynomials $h_i \in \mathrm{k}[x_0, \ldots, x_n]$ of degree d, where $h_0 = \mu x_0^d$, $\mu \in \mathrm{k}$, up to linear equivalence: $(h_0, \ldots, h_n) \sim (\lambda h_0, \ldots, \lambda h_n)$ for any $\lambda \in \mathrm{k}^*$. The equivalence class of $(h_0, \ldots, h_n)$ will be denoted by $[h_0 : \cdots : h_n]$. Since the set of homogeneous polynomials of degree d in $n+1$ variables is a k-vector space, this gives to W_d a canonical projective space structure. We then denote by $B_d \subset W_d$ the hyperplane given by $h_0 = 0$ and obtain a canonical isomorphism of affine spaces

$$\begin{aligned} W_d \setminus B_d &\xrightarrow{\sim} \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}, \\ [x_0^d : h_1 : \cdots : h_n] &\longmapsto (h_1(1, x_1, \ldots, x_n), \ldots, h_n(1, x_1, \ldots, x_n)). \end{aligned}$$

We denote by $Y \subseteq W_{d^{n-1}} \times (W_d \setminus B_d)$ the set consisting of elements (g, f), such that $h := (g_0(f_0, \dots, f_n), \dots, g_n(f_0, \dots, f_n))$ is a multiple (maybe 0) of the identity, i.e., $h_i x_j = h_j x_i$ for all i, j. The description of Y shows that it is closed in $W_{d^{n-1}} \times (W_d \setminus B_d)$. Since $W_{d^{n-1}}$ is a complete variety, the projection $p_2 \colon W_{d^{n-1}} \times (W_d \setminus B_d) \to (W_d \setminus B_d)$ is a Zariski-closed morphism, so $p_2(Y)$ is closed in $(W_d \setminus B_d) \simeq \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$.

In order to show (3) and (4), we only need to show that $p_2(Y) \cap J_{\mathrm{k}^*} = \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$, which implies that $p_2(Y) \cap J_1 = \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$. We then show both inclusions.

(i) If $f \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$, there exists $g \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d^{n-1}}$ such that $g \circ f = \mathrm{id}$ ([Bass et al. 1982, Theorem 1.5, page 292]). In consequence, we obtain $(g, f) \in Y$, so $f \in p_2(Y)$. The fact that f belongs to J_{k^*} is given by Remark 2.3, so $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \subseteq p_2(Y) \cap J_{\mathrm{k}^*}$.

(ii) Let $(g, f) \in Y$, with $f \in J_{\mathrm{k}^*}$. By definition of Y, the element $(g_0(f_0, \dots, f_n)$, $\dots$, $g_n(f_0, \dots, f_n))$ is a multiple of the identity. There exists some j such that $g_j \neq 0$. The fact that $f \in J_{\mathrm{k}^*}$ implies that $f \colon \mathbb{A}^n_{\mathrm{k}} \to \mathbb{A}^n_{\mathrm{k}}$ is dominant. Hence, $g_j(f_0(1, x_0, \dots, x_n), \dots, f_n(1, x_0, \dots, x_n))$ is not equal to zero. This implies that $g \in \mathrm{End}(\mathbb{A}^{n-1})_{\leq d^{n-1}}$ and that $g \circ f = \mathrm{id}$. Hence, $f \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$. This yields $p_2(Y) \cap J_{\mathrm{k}^*} \subseteq \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$. □

Lemma 2.7. *Let us fix some integers $d, n \geq 1$, and see $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ and $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ as affine varieties, via their inclusion in $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \simeq (\mathrm{k}[x_1, \dots, x_n]_{\leq d})^n$ (see Lemma 2.6). Then, the following hold:*

(1) *Morphisms $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ correspond to elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}[A]})$.*
(2) *For each k-algebraic variety A, the morphisms $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ that have image in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ correspond to morphisms of algebraic varieties $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$.*
(3) *The set $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ is closed in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, and the restriction of the topology of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ on it yields the topology of its algebraic variety structure.*
(4) *A subset of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ is closed if and only if its intersection with $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ is closed, for each integer $d \geq 1$.*

Moreover, everything works the same replacing Aut *with* SAut.

Proof. We do the proof with Aut, the same proof works replacing Aut with SAut.

The map

$$\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \times \mathbb{A}^n_{\mathrm{k}} \longrightarrow \mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d} \times \mathbb{A}^n_{\mathrm{k}}, \quad (f, x) \longmapsto (f, f(x))$$

is a morphism of algebraic varieties, so every morphism of algebraic varieties $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ yields an A-automorphism of $A \times \mathbb{A}^n_{\mathrm{k}}$, and thus a morphism $A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ with image in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$.

Conversely, let $f \colon A \to \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ be a morphism. By definition, this is given by an A-automorphism of $A \times \mathbb{A}^n_{\mathrm{k}}$. Composing with the projection $A \times \mathbb{A}^n_{\mathrm{k}} \to \mathbb{A}^n_{\mathrm{k}}$ and then with the function x_i on $\mathbb{A}^n_{\mathrm{k}}$, we obtain an element $f_i \in \mathrm{k}[A\times\mathbb{A}^n_{\mathrm{k}}]=\mathrm{k}[A][x_1,\dots,x_n]$. Hence, $(f_1,\dots,f_n)$ is an element of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}[A]})$. Such an element yields a morphism of algebraic varieties $A\to\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\le m}$ for some $m\in\mathbb{N}$, having image in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$. This yields (1) and (2). Assertion (3) also follows, after observing that the preimage of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le i}$ is a closed subset of A, for each $i\ge 0$.

It remains to show (4). Let $Y\subset\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ be any subset. If Y is closed in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, then $Y\cap\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le d}$ is closed (in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ or $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le d}$) for each d, by (3). Suppose conversely that $Y\cap\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le d}$ is closed for each d. In order to show that Y is closed in $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, we take any morphism $A\to\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, and show that the preimage of Y is closed. As we observed before, we can see this morphism as a morphism of algebraic varieties $A\to\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le m}$ for some m. Since $Y\cap\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le m}$ is closed, the preimage of Y is closed. □

2B. ***Dynamical properties and the behaviour at infinity.*** In the sequel, we fix a canonical open embedding $\mathbb{A}^n_{\mathrm{k}}\to\mathbb{P}^n_{\mathrm{k}}$ given by

$$(x_1,\dots,x_n)\longmapsto[1:x_1:\cdots:x_n]$$

and denote by $H_\infty\subset\mathbb{P}^n_{\mathrm{k}}$ the complement $H_\infty=\mathbb{P}^n_{\mathrm{k}}\setminus\mathbb{A}^n_{\mathrm{k}}$, which is a hyperplane. Once this is fixed, we have a canonical extension of any element of $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})$ to a rational map $\mathbb{P}^n_{\mathrm{k}}\dashrightarrow\mathbb{P}^n_{\mathrm{k}}$. In particular, this yields inclusions

$$\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})\longrightarrow\mathrm{Bir}(\mathbb{A}^n_{\mathrm{k}})\xrightarrow{\simeq}\mathrm{Bir}(\mathbb{P}^n_{\mathrm{k}}).$$

The automorphisms of degree 1 (affine automorphisms) correspond to those which extend to elements of $\mathrm{Aut}(\mathbb{P}^n_{\mathrm{k}})$. The others extend to birational maps which are not automorphisms, and we can associate to them two sets, which are classical objects in dynamics (see for example [Sibony 1999; Guedj and Sibony 2002; Bisi 2008]):

Definition 2.8. For each $f\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ of degree ≥ 2, we define $I_f\subset\mathbb{P}^n_{\mathrm{k}}$ to be the indeterminacy locus of the extension of f to $\mathbb{P}^n_{\mathrm{k}}$, and $X_f\subset\mathbb{P}^n_{\mathrm{k}}$ to be the image of $H_\infty\backslash(I_f)$.

In order to see the sets I_f, X_f explicitly, we use the following definition.

Definition 2.9. If $f=(f_1,\dots,f_n)\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ is an element of degree d and a_i is the homogeneous part of f_i of degree d, for $i=1,\dots,n$, we say that $(a_1,\dots,a_n)\in\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})$ is the *highest homogeneous part of* f.

Remark 2.10. If $(a_1,\dots,a_n)$ is the highest homogeneous part of $f\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, the set $I_f\subset\mathbb{P}^n_{\mathrm{k}}$ is given by

$$x_0=0,\quad a_1=\cdots=a_n=0,$$

and is a proper closed subset of the hyperplane at infinity H_∞ (given by $x_0=0$).

Moreover, the set X_f is equal to

$$\{[0:a_1(x_1,\dots,x_n):\dots:a_n(x_1,\dots,x_n)]\mid[0:x_1:\dots:x_n]\in H_\infty\backslash I_f\}.$$

Remark 2.11 (regular terminology). In dynamics, the elements $f\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ such that $I_f\cap I_{f^{-1}}=\varnothing$ are called "regular" and elements such that $X_f\cap I_f=\varnothing$ are called "weakly regular" in [Guedj and Sibony 2002] or "quasiregular" in [Bisi 2008]. The use of this terminology in algebraic geometry can be confusing, because of the common use of the word regular for maps (in fact all elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ are morphisms, hence biregular). This is why we will say "dynamically regular" if $I_f\cap I_{f^{-1}}=\varnothing$; we will not use the words "quasiregular" or "weakly regular".

Remark 2.12. Fixing the degree d, we obtain an algebraic variety $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_d$. It follows from the definition that the set of dynamically regular elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_d$ is an open subset (in general not dense). However, the set of dynamically regular elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})_{\le d}$ (and thus of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$) is not open in general. For $n=2$, this can be seen by taking for example

$$f=(-x_2,x_1+x_2^2)\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}}),\qquad \alpha=(x_1+tx_2^2,x_2)\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}[t]}),$$

and considering the family $\beta=\alpha f\alpha^{-1}\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}[t]})$. Then, $\beta(t)$ is not dynamically regular for $t\in\mathrm{k}^*$ but $\beta(0)=f$ is dynamically regular.

Remark 2.13. If $n=2$, we find $X_f=I_{f^{-1}}$. Indeed, the extension of an element $f\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ of degree >1 is an element of $\mathrm{Bir}(\mathbb{P}^2_{\mathrm{k}})$ which contracts the line at infinity onto one point X_f. The inverse of this birational map is then defined at any other point of H_∞, so we find $X_f=I_{f^{-1}}$.

We now show that the degree of a composition is determined by the sets I_f and X_f defined above. The following results are classical in the world of dynamics, we recall the easy proofs for self-containedness.

Lemma 2.14. *Let $f,g\in\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ be of degree ≥ 2. Then, the following are equivalent*:

(1) $\deg(gf)=\deg(g)\cdot\deg(f)$.

(2) *The set X_f is not contained in I_g.*

Moreover, if both conditions hold, then $X_{gf}\subset X_g$.

Proof. Denoting by $(a_1,\dots,a_n)$ and $(b_1,\dots,b_n)$ the highest homogeneous parts of f and g respectively, the equality $\deg(gf)=\deg(g)\cdot\deg(f)$ is equivalent to the fact that one of the polynomials

$$b_1(a_1,\dots,a_n),\dots,b_n(a_1,\dots,a_n)$$

is not equal to zero.

By definition, the sets $I_g, X_f \subset H_\infty$ are given respectively by

$$I_g = \{[0 : y_1 : \cdots : y_n] \in H_\infty \mid b_1(y_1, \ldots, y_n) = \cdots = b_n(y_1, \ldots, y_n) = 0\},$$
$$X_f = \{[0 : a_1(y_1, \ldots, y_n) : \cdots : a_n(y_1, \ldots, y_n)] \mid [0 : y_1 : \cdots : y_n] \in H_\infty \backslash I_f\}.$$

If $X_f \not\subset I_g$, then there exists a point $[0 : y_1 : \cdots : y_n] \in H_\infty \backslash I_f$ such that $[0 : a_1(y_1, \ldots, y_n) : \cdots : a_n(y_1, \ldots, y_n)] \notin I_g$, which corresponds to the existence of an index $i \in \{1, \ldots, n\}$ such that $b_i(a_1(y_1, \ldots, y_n), \ldots, a_n(y_1, \ldots, y_n)) \neq 0$. In particular, the polynomial $b_i(a_1, \ldots, a_n)$ is not zero, so $\deg(gf) = \deg(g)\deg(f)$.

Conversely, if $b_i(a_1, \ldots, a_n)$ is a nonzero polynomial, the open subset $U_i \subset H_\infty$ corresponding to the nonvanishing of this polynomial is nonempty. Intersecting this open set with $H_\infty \backslash I_f$ yields a nonempty open subset of points in H_∞ which have image in X_f and not in I_g.

Now that the equivalence between (1) and (2) is shown, we show that these imply that $X_{gf} \subset X_g$. Since $\deg(gf) = \deg(g)\deg(f)$, the homogeneous part of highest degree of gf is $(b_1(a_1, \ldots, a_n), \ldots, b_n(a_1, \ldots, a_n))$. This implies that the points of X_{gf} are in the image by g (or more precisely of its extension to $\mathbb{P}^n_{\mathrm{k}}$) of the set $X_f \setminus I_g$, and thus lie in X_g. □

Corollary 2.15. *Let $f, g \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ be of degree ≥ 2. Then, the following hold:*

(1) $X_f \subset I_{f^{-1}}$.

(2) *If* $I_g \cap I_{f^{-1}} = \varnothing$, *then* $\deg(gf) = \deg(g) \cdot \deg(f)$.

(3) $X_f \not\subset I_f \Leftrightarrow \deg(f^2) = \deg(f)^2$.

Proof. Since $\deg(f^{-1}f) < \deg(f^{-1})\deg(f)$, part (1) follows from Lemma 2.14. If $I_g \cap I_{f^{-1}} = \varnothing$, then $X_f \not\subset I_g$ by (1), so the equality $\deg(gf) = \deg(g)\deg(f)$ follows again from Lemma 2.14. Part (3) corresponds to Lemma 2.14, in the case $f = g$. □

Corollary 2.16. *If $f \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$ is an element such that $X_f \cap I_f = \varnothing$, then*

$$\deg(f^m) = \deg(f)^m \quad \textit{and} \quad X_{f^m} \subset X_f$$

for each $m \geq 1$.

Proof. We prove the result by induction on m, the case $m = 1$ being obvious. For $m \geq 2$, we use the facts that $X_f \cap I_f = \varnothing$ and $X_{f^{m-1}} \subset X_f$, which imply that $X_{f^{m-1}} \not\subset I_f$. Applying Lemma 2.14 to the composition $f \circ f^{m-1}$, we obtain that $\deg(f^m) = \deg(f^{m-1})\deg(f)$, which is equal to $\deg(f)^m$ by the induction hypothesis, and also that $X_{f^m} \subset X_f$. □

Restricting to dimension 2, we obtain the following result.

Corollary 2.17. *Let $f \in \mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ be of degree ≥ 1. The following are equivalent:*

(1) $I_f \neq I_{f^{-1}}$;

(2) $I_f \cap I_{f^{-1}} = \varnothing$;

(3) $\deg(f^2) = \deg(f)^2$;

(4) $\deg(f^m) = \deg(f)^m$ *for each* $m \geq 1$.

Proof. As we are in dimension 2, we have $X_f = I_{f^{-1}}$ and $X_{f^{-1}} = I_f$, which are two points of H_∞, which can be distinct or not (see Remark 2.13).

The equality $\deg(f^2) = \deg(f)^2$ is equivalent to $X_f \not\subset I_f$ (Corollary 2.15), which is here equivalent to $I_{f^{-1}} \neq I_f$ or $I_{f^{-1}} \cap I_f = \varnothing$. Hence, (1), (2), and (3) are equivalent, and of course implied by (4). It remains to see that (2) corresponds to $X_f \cap I_f = \varnothing$, which implies that $\deg(f^m) = \deg(f)^m$ for each $m \geq 1$ by Corollary 2.16. □

Remark 2.18. The most interesting implication of this corollary is (3) ⇒ (4), i.e., that $\deg(f^2) = \deg(f)^2$ implies that $\deg(f^m) = \deg(f)^m$ for $m \geq 1$, a fact already observed by Jean-Philippe Furter [1999] (at least when $\mathrm{char}(\mathrm{k}) = 0$).

Looking at the proof, this result has no reason to be true in dimension $n \geq 3$, and is in fact false on the easiest nontrivial example, that we describe now.

Example 2.19. Let $f = (x_1 + x_2^2, x_2 + x_3^2, x_3) \in \mathrm{SAut}(\mathbb{A}^3_{\mathrm{k}})$, which has highest homogeneous part $a = (x_2^2, x_3^2, 0) \in \mathrm{End}(\mathbb{A}^3_{\mathrm{k}})$. Then,

$$I_f = [0:1:0:0], \qquad X_f = \{[0:x_1:x_2:0] \mid [x_1:x_2] \in \mathbb{P}^1\}.$$

Since $X_f \not\subset I_f$, we have $\deg(f^2) = \deg(f)^2 = 4$ and $X_{f^2} \subset X_f$. More precisely, the homogeneous part of f^2 is $a^2 = (x_3^4, 0, 0)$, so

$$I_{f^2} = I_f = X_{f^2} = [0:1:0:0].$$

In particular, $X_{f^2} \subset I_f$, so $\deg(f^3) < \deg(f^2) \cdot \deg(f)$.

In fact, one easily checks with the formulas that $\deg(f^n) \leq 4$ for each n, and that $\deg(f^n) = 4$ for each $n \geq 2$ if $\mathrm{char}(\mathrm{k}) = 0$. Indeed

$$f^n = \bigl(x_1 + n x_2^2 + n(n-1) x_2 x_3^2 + \bigl(\textstyle\sum_{i=1}^{n-1} i^2\bigr) x_3^4,\ x_2 + n x_3^2,\ x_3\bigr),$$

for each $n \geq 0$.

2C. *Families of automorphisms and valuations.* In the sequel, we will study families of elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, which correspond to elements of $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}((t))})$.

It is then natural to use the valuation

$$\nu\colon \mathrm{k}((t))[x_1, \dots, x_n] \to \mathbb{Z} \cup \{-\infty\}$$

associated to t. We define precisely this valuation here, as we will use it often afterwards.

Definition 2.20. Every element $f \in \mathrm{k}((t))[x_1, \dots, x_n] \setminus \{0\}$ can be written as

$$f = \sum_{k=m}^{\infty} a_k t^k$$

where $m \in \mathbb{Z}$, $a_i \in \mathrm{k}[x_1, \dots, x_n]$ for $i \geq m$, and $a_m \neq 0$. We then define $\nu(f) = m$. Choosing $\nu(0) = -\infty$, we obtain a valuation

$$\nu \colon \mathrm{k}((t))[x_1, \dots, x_n] \longrightarrow \mathbb{Z} \cup \{-\infty\}.$$

(1) If $\nu(f) = 0$, we define $f(0) = a_0$.

(2) If $\nu(f) > 0$, we define $f(0) = 0$.

(3) If $\nu(f) < 0$, we say that f *has a pole at* $t = 0$ and that $f(0)$ is not defined.

Definition 2.21. Let $f = (f_1, \dots, f_n) \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}((t))})$. We define

$$\nu(f) = \min\{\nu(f_i) \mid i = 1, \dots, n\}.$$

(1) If $\nu(f) \geq 0$, we say that f is *defined at the origin* (or *has a value*), and define

$$f(0) = (f_1(0), \dots, f_n(0)),$$

to be its value, which is an element of $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})$.

(2) If $\nu(f) < 0$, we say that f *has a pole at* $t = 0$, and say that $f(0)$ is not defined.

Remark 2.22. In the above definition, it is possible that the element $f(0)$ is defined, but does not belong to $\mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$. This is for example the case when one of the components becomes 0, or for $f = (tx_1 + x_2, x_2, \dots, x_n)$. This phenomenon is however impossible for $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, as the following result shows.

Lemma 2.23. *Let $\alpha \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}((t))})$ be an element which has no pole at $t = 0$.*

Then, α^{-1} has no pole at $t = 0$, and replacing t with 0 yields two automorphisms $\beta, \gamma \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$,

$$\beta = \alpha(0), \quad \gamma = \alpha^{-1}(0),$$

such that $\beta\gamma = \mathrm{id}$.

Proof. Since α has no pole at $t = 0$, it sends $(0, \dots, 0) \in \mathbb{A}^n_{\mathrm{k}((t))}$ onto an element having coordinates in $\mathrm{k}[[t]]$. We can thus replace α by its composition with a translation and assume that α (and thus α^{-1}) fixes the origin. Its linear part is then equal to an element of $\mathrm{SL}(n, \mathrm{k}[[t]])$, whose inverse also belongs to $\mathrm{SL}(n, \mathrm{k}[[t]])$. Replacing with the composition by this inverse, we can assume that α has a trivial linear part. We denote by $\mathfrak{m}$ the ideal of $\mathrm{k}((t))[x_1, \dots, x_n]$ generated by the x_i, and can then write

$$\alpha = (f_1, \dots, f_n)$$

for some $f_1, \dots, f_n \in \mathrm{k}[\![t]\!][x_1, \dots, x_n]$, $f_i \equiv x_i \pmod{\mathfrak{m}^2}$. We write then

$$\alpha^{-1} = \left(x_1 + \sum_{i=2}^{d} g_{i,1}, \dots, x_n + \sum_{i=2}^{d_2} g_{i,n} \right)$$

where the $g_{i,j} \in \mathrm{k}((t))[x_1, \dots, x_n]$ are homogeneous polynomials of degree i and d_2 is the degree of g. Assume for contradiction that one of the $g_{k,j}$ does not belong to $\mathrm{k}[\![t]\!][x_1, \dots, x_n]$, and choose k to be minimal for this. Then, the j-th coordinate of $\alpha^{-1} \circ \alpha$ is equal to

$$x_j = f_j + \sum_{i=2}^{d_2} g_{i,j}(f_1, \dots, f_n),$$

which implies that $\sum_{i=k}^{d} g_{i,j}(f_1, \dots, f_n) \in \mathrm{k}[\![t]\!][x_1, \dots, x_n]$. But the part of degree k of this sum is in fact equal to $g_{i,j}(x_1, \dots, x_n)$, which does not belong to $\mathrm{k}[\![t]\!][x_1, \dots, x_n]$.

We have proved that α^{-1} does not have any pole at the origin. We then denote by d_1, d_2 the degrees of α and α^{-1} (which are the maximal degree of their components), and denote by End_{d_i} the set of endomorphisms of $\mathbb{A}^n$ of degree $\leq d_i$, which is naturally isomorphic to an affine space. Observe that (α, α^{-1}) corresponds to a $\mathrm{k}((t))$-point of the algebraic variety

$$\{(f, g) \in \mathrm{End}_{d_1} \times \mathrm{End}_{d_1} \mid f \circ g = \mathrm{id}\}.$$

Since neither α neither α^{-1} has a pole at $t = 0$, all coefficients are defined at the origin. Replacing t with 0 gives then the result. □

Remark 2.24. The result of Lemma 2.23 can also be obtained from the fact that each $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ is closed in $\mathrm{End}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ (Lemma 2.6).

We will apply a classical valuative result (Lemma 2.25 below), and recall the argument of the proof, given in [Furter 2009, §1.2] (the version that we need here is slightly more general, but the proof is analogous).

Lemma 2.25. *Let Y, Z be two quasiprojective* k*-algebraic varieties, let $\varphi\colon Y \to Z$ be a morphism and let $z \in Z$ be a (closed) point of Z. The following assertions are equivalent:*

(1) *The point z belongs to the closure $\overline{\varphi(Y)}$ of the image.*

(2) *There is an irreducible* k*-curve Γ, a smooth closed point $p \in \Gamma$ and a rational map $\iota\colon \Gamma \dashrightarrow Y$ such that $\varphi \circ \iota\colon \Gamma \dashrightarrow Z$ is defined at p and sends it onto z.*

(3) *There is a $\mathrm{k}((t))$-point $y \in Y(\mathrm{k}((t)))$ such that $\varphi(y) \in Z(\mathrm{k}((t)))$ has no pole at $t = 0$ and $\varphi(y)(0) = z$.*

Proof. (1) $\Rightarrow$ (2): Replacing Y with one of its irreducible components, we can assume that Y is irreducible. Putting Z into a projective variety, we can assume that Z is projective. Then, replacing Z with $\overline{\varphi(Y)}$ we can assume that φ is dominant. Since $\varphi(Y)$ is an irreducible constructible subset of Z, it contains a dense open subset U of Z. If Z is a point, then $\varphi(Y) = Z$, in which case (2) is obvious, so we can assume that $\dim Z \geq 1$, and can take a closed irreducible curve $C \subset Z$ containing z and meeting U (through every two points passes at least one irreducible closed curve, see [Mumford 1970, §II.6, Lemma, page 56]).

We then denote by $Y' \subset Y$ an irreducible component of $\varphi^{-1}(C)$ such that the restriction of φ yields a dominant morphism $Y' \to C$. We choose two points $p_1, p_2 \in Y'$ such that $\varphi(p_1) \neq \varphi(p_2)$ and take an irreducible curve $\Gamma_0 \subset Y'$ passing through p_1, p_2 (using again the lemma from [Mumford 1970]). The restriction of φ to Γ_0 yields a dominant morphism $\Gamma_0 \to C$. We can take an open embedding $\nu \colon \Gamma_0 \to \overline{\Gamma}_0$, where $\overline{\Gamma}_0$ is an irreducible projective, and denote by $\eta \colon \Gamma \to \overline{\Gamma}_0$ the normalisation. Then, $\varphi \circ \nu^{-1} \circ \eta$ yields a rational map $\Gamma \dashrightarrow C$, which is a surjective morphism. It remains to choose for $\iota \colon \Gamma \dashrightarrow \Gamma_0 \subset Y' \subset Y$ the rational map $\nu^{-1} \circ \eta$.

(2) $\Rightarrow$ (3): The rational maps $\Gamma \dashrightarrow Y \to Z$ correspond to field homomorphisms $\mathrm{k}(Z) \to \mathrm{k}(Y) \to \mathrm{k}(\Gamma)$, sending the local ring $\mathcal{O}_{z,Z}$ to $\mathcal{O}_{p,\Gamma}$.

Denote by $\widehat{\mathcal{O}}_{p,\Gamma}$ the completion of $\mathcal{O}_{p,\Gamma}$, with respect to its maximal ideal. Because $\mathcal{O}_{p,\Gamma}$ is a Noetherian regular local ring of dimension 1 with residue field k, its completion is a complete Noetherian regular local ring with the same properties, and by the Cohen theorem, it must be isomorphic to a ring of formal power series. The dimension being 1, one has a k-isomorphism $\widehat{\mathcal{O}}_{p,\Gamma} \simeq \mathrm{k}[[t]]$, which induces a field homomorphism $\mathrm{k}(\Gamma) \to \mathrm{k}((t))$.

The composition $\mathrm{k}(Y) \to \mathrm{k}(\Gamma) \to \mathrm{k}((t))$ corresponds to the $\mathrm{k}((t))$-point y that we want, and its image corresponds to the composition $\mathrm{k}(Z) \to \mathrm{k}(Y) \to \mathrm{k}(\Gamma) \to \mathrm{k}((t))$.

(3) $\Rightarrow$ (1): View Z as a locally closed subset of $\mathbb{P}^n$ and take a polynomial equation $F \in \mathrm{k}[x_0, \ldots, x_n]$ that vanishes on $\varphi(Y)$. Since $\varphi(y)$ is a $\mathrm{k}((t))$-point of $\varphi(Y)$, we have $F(\varphi(y)) = 0$. Replacing t with 0 we obtain $F(z) = 0$. This shows that z belongs to the closure of Z. $\square$

Corollary 2.26. *Let $f \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, let $d \geq 1$ be an integer and let Y be the* k-*algebraic variety* $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$. *The following assertions are equivalent*:

(1) *The set $\{gfg^{-1} \mid g \in Y\}$ is closed in* $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$.

(2) *If Γ is an irreducible* k-*curve Γ and $\iota \colon \Gamma \dashrightarrow Y$ is a rational map such that $\varphi \circ \iota \colon \Gamma \dashrightarrow Z$ is defined at a smooth point $p \in \Gamma$, the image of p belongs to $\{gfg^{-1} \mid g \in Y\}$.*

(3) *If $\varphi \in Y(\mathrm{k}((t)))$ is an element that has poles at $t = 0$ and $h = \varphi f \varphi^{-1}$ has no poles at $t = 0$, then $h(0)$ belongs to $\{gfg^{-1} \mid g \in Y\}$.*

Proof. The degree of the inverse of an element $g \in Y$ is at most d^{n-1} (see [Bass et al. 1982, Theorem 1.5, page 292]). Hence, the map $Y \to \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ that sends g onto gfg^{-1} corresponds to a morphism of algebraic varieties $\varphi\colon Y \to \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq m}$ for some m. The result follows then from Lemma 2.25 applied to φ, and Lemma 2.7. □

3. Conjugacy classes of dynamically regular automorphisms of $\mathbb{A}^n_{\mathbf{k}}$

3A. ***Image at infinity of elements of* $\mathbf{Aut}(\mathbb{A}^n_{\mathbf{k}((t))})$.** In Section 2B, we explained how the set $X_f \subset H_\infty$ is defined, for each element $f \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}})$, by extending the map to $\mathbb{P}^n_{\mathrm{k}}$ and looking at the image of the hyperplane H_∞ at infinity.

We now associate similarly a subset $X_\alpha \subset H_\infty$ to an element $\alpha \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}((t))})$ which has a pole at $t = 0$, by taking the limit of the image of $\alpha(t)$, when t goes towards 0. The formal definition of X_α is the following:

Definition 3.1. Let $\alpha \in \mathrm{Aut}(\mathbb{A}^n_{\mathrm{k}((t))})$ be an element of valuation $\nu(\alpha) = -m < 0$, that we write

$$\alpha = \left(\frac{1}{t^m}\alpha_1, \ldots, \frac{1}{t^m}\alpha_n\right),$$

where $\alpha_1, \ldots, \alpha_n \in \mathrm{k}[\![t]\!][x_1, \ldots, x_n]$ are such that

$$\tilde{\alpha} = (\alpha_1(0), \ldots, \alpha_n(0)) = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n) \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}}) \setminus \{0\}.$$

We then define

$$X_\alpha = \{[0 : \tilde{\alpha}_1(y_1, \ldots, y_n) : \cdots : \tilde{\alpha}_n(y_1, \ldots, y_n)] \mid (y_1, \ldots, y_n) \in \mathbb{A}^n_{\mathrm{k}} \setminus \tilde{\alpha}^{-1}(\{0\})\} \subset H_\infty.$$

This definition can be geometrically understood:

Remark 3.2. In the above definition, α is not defined at $t = 0$, but extending α to $\mathbb{P}^n$ we obtain the element of $\mathrm{Bir}(\mathbb{P}^n_{\mathrm{k}((t))})$ given by

$$[x_0 : \cdots : x_n] \dashrightarrow [t^m x_0^d : F_1(x_0, \ldots, x_n, t) : \cdots : F_n(x_0, \ldots, x_n, t)],$$

where d is the degree of α and each $F_i(x_0, \ldots, x_n, t) \in \mathrm{k}[\![t]\!][x_0, \ldots, x_n]$ is the homogenisation of α_i. This corresponds to a family of rational maps of $\mathbb{P}^n_{\mathrm{k}}$ parametrised by t, which has a value at $t = 0$, corresponding to

$$[x_0 : \cdots : x_n] \dashrightarrow [0 : F_1(x_0, \ldots, x_n, 0) : \cdots : F_n(x_0, \ldots, x_n, 0)].$$

The set $X_\alpha \subset H_\infty$ is then the image of this map by points of $\mathbb{A}^n_{\mathrm{k}}$ which are well-defined under this map.

Writing such a point as $[1 : x_1 : \cdots : x_n]$, its image by the extension of α is

$$\begin{aligned}[t^m : F_1(1, \ldots, x_n, t) : \cdots : F_n(1, \ldots, x_n, t)]& \\ = [t^m : \alpha_1(x_1, \ldots, x_n, t) : \cdots : \alpha_n(x_1, \ldots, x_n, t)]&\end{aligned}$$

and corresponds to a curve in $\mathbb{P}^n$ whose point, when $t=0$, belongs to X_α.

The set X_α corresponds then to the limit, viewed in $\mathbb{P}^n_{\mathrm{k}}$, of the image of points of $\mathbb{A}^n_{\mathrm{k}}$ under $\alpha(t)$ when t goes towards 0.

Proposition 3.3. *Let $f \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ be an element of degree $d > 1$, and let $\alpha \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}((t))})$ be an element that has a pole at $t=0$. If*

$$X_\alpha \not\subset I_f,$$

then $\alpha^{-1} f \alpha$ has a pole at $t=0$.

Proof. We write

$$f = (f_1, \dots, f_n)$$

for some $f_1, \dots, f_n \in \mathrm{k}[x_1, \dots, x_n]$ and denote by $(a_1, \dots, a_n)$ the highest homogeneous part of f (which is of degree d). In particular, the indeterminacy locus $I_f \subset \mathbb{P}^n_{\mathrm{k}}$ of the extension of f to $\mathbb{P}^n_{\mathrm{k}}$ is given by

$$x_0 = 0, \quad a_1 = \cdots = a_n = 0.$$

Because α has a pole at $t=0$, we have $\nu(\alpha) = -m < 0$ and can write $\alpha = (t^{-m}\alpha_1, \dots, t^{-m}\alpha_n)$, where $\alpha_1, \dots, \alpha_n \in \mathrm{k}[\![t]\!][x_1, \dots, x_n]$ are such that

$$\tilde{\alpha} = (\alpha_1(0), \dots, \alpha_n(0)) = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_n) \in \mathrm{End}(\mathbb{A}^n_{\mathrm{k}}) \setminus \{0\},$$

and so the subset $X_\alpha \subset H_\infty$ is described by

$$X_\alpha = \{[0 : \tilde{\alpha}_1(y_1, \dots, y_n) : \cdots : \tilde{\alpha}_n(y_1, \dots, y_n)] \mid (y_1, \dots, y_n) \in \mathbb{A}^n_{\mathrm{k}} \setminus \tilde{\alpha}^{-1}(\{0\})\}.$$

The fact that X_α is not included in I_f yields a point $(y_1, \dots, y_n) \in \mathbb{A}^n_{\mathrm{k}} \setminus \tilde{\alpha}^{-1}(\{0\})$ and an integer $i \in \{1, \dots, n\}$ such that

$$a_i(\tilde{\alpha}_1(y_1, \dots, y_n), \dots, \tilde{\alpha}_n(y_1, \dots, y_n)) \neq 0.$$

In particular, we have

$$a_i(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n) = a_i(\alpha_1(0), \dots, \alpha_n(0)) \in \mathrm{k}[x_1, \dots, x_n] \setminus \{0\}.$$

Thus $a_i(t^{-m}\alpha_1, \dots, t^{-m}\alpha_n) = t^{-md} a_i(\alpha_1, \dots, \alpha_n)$ has valuation $-md$, and hence $f\alpha$ has also valuation $-md$.

It remains to show that this implies that $\beta = \alpha^{-1} f \alpha$ has a pole at $t=0$. Indeed, if β had no pole, we would have $\nu(\alpha\beta) \geq \nu(\alpha) = -m$, which is impossible since $f\alpha = \alpha\beta$ and $\nu(f\alpha) = -md$. $\square$

Corollary 3.4. *Let $f \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ be a dynamically regular element.*

(1) *If $\alpha \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}((t))})$ is such that $g = \alpha^{-1} f \alpha$ has no pole at $t=0$, then α and α^{-1} have no pole at $t=0$. In particular, $g(0)$ is an element of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ that is conjugate to f.*

(2) *For each* $d \geq 1$, *the set* $\{gfg^{-1} \mid g \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}\}$ *is closed in* $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$.

Proof. (1) We suppose that α has a pole at $t=0$, which is equivalent to the fact that α^{-1} has a pole at $t=0$ (Lemma 2.23), and show that $\alpha^{-1}f\alpha$ has a pole at $t=0$. If $X_\alpha \not\subset I_f$, this is given by Proposition 3.3. Otherwise, we have $X_\alpha \not\subset I_{f^{-1}}$, (because $I_f \cap I_{f^{-1}} = \varnothing$ by hypothesis) and apply the proposition to f^{-1}. This implies that $\alpha f^{-1}\alpha^{-1}$ has a pole at $t=0$. Hence, $\alpha f \alpha^{-1}$ has also a pole at $t=0$ by Lemma 2.23.

(2) This follows from (1) and Corollary 2.26. □

We obtain thus the following two results.

Proposition 3.5. *Let* $f \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ *be a dynamically regular element having the following property*: *there exists a function* $\tau : \mathbb{N} \to \mathbb{N}$ *such that for each conjugate* $g \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ *of* f, *there exists* $h \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ *of degree* $\leq \tau(\deg(g))$ *such that* $g = hfh^{-1}$.

Then, *the conjugacy class of* f *in* $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ *is closed.*

Proof. Let us denote by $C \subset \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ the conjugacy class of f in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$. Note that C is closed (in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$) if and only if

$$C_d = \{g \in C \mid \deg(g) \leq d\}$$

is closed in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ for each $d \in \mathbb{N}$.

By Corollary 3.4, the set

$$C'_d = \{hfh^{-1} \mid h \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}\}$$

is closed in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ for each d.

By hypothesis, we have $C_d \subset C'_{\tau(d)}$ for each $d \in \mathbb{N}$, which implies that

$$C_d = C'_{\tau(d)} \cap \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$$

is closed in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}$ for each d. □

The additional hypothesis of Proposition 3.5 is fulfilled for all dynamically regular elements of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$, so we obtain that the conjugacy classes of all dynamically regular elements of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ are closed:

Proposition 3.6. *Let* $f \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ *be a dynamically regular element. Then*:

(1) *If* $g \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ *is conjugate to* f, *there exists* $h \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$, *such that*

$$g = hfh^{-1} \quad \textit{and} \quad \deg(h)^2 \leq \deg(g).$$

(2) *The conjugacy class of* f *in* $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ *is closed.*

Remark 3.7. The bound of (1) also exists for $\mathrm{Bir}(\mathbb{A}^2_{\mathrm{k}})$, but is much higher [Blanc and Cantat 2016].

Proof. Proposition 3.5 yields (1) $\Rightarrow$ (2), so we only need to prove (1).

Recall that, for each $g \in \mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ of degree greater than one, $I_g = X_{g^{-1}}$ consists of one point (Remark 2.13).

Let $g = hfh^{-1}$ be an element of degree d. Replacing h with hf^l, $l \in \mathbb{Z}$, we can assume that

$$\deg(h) \leq \deg(hf^l), \quad \text{for each } l \in \mathbb{Z}.$$

We can also assume that $\deg(h) \geq 2$. This implies, since $\deg(h) < \deg(hf)\deg(f^{-1})$, that $I_f = X_{f^{-1}} = I_{hf}$ (Lemma 2.14). Similarly, we have $\deg(h) < \deg(hf^{-1})\deg(f)$, so $I_{f^{-1}} = X_f = I_{hf^{-1}}$.

Because the two points $I_f, I_{f^{-1}} \in H_\infty$ are distinct, we have $I_h \neq I_f$ or $I_h \neq I_{f^{-1}}$. If $I_h \neq I_f = I_{hf}$, we have $\deg(hfh^{-1}) = \deg(hf)\deg(h^{-1}) \geq \deg(h)\deg(h^{-1})$. If $I_h \neq I_{f^{-1}} = I_{hf^{-1}}$, we have $\deg(hf^{-1}h^{-1}) = \deg(hf^{-1})\deg(h^{-1}) \geq \deg(h)\deg(h^{-1})$.

The degree of an element of $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$ and its inverse being the same, we find $\deg(g) \geq \deg(h)^2$. □

Proof of Theorem 1.1. Part (1) and (2) correspond to the statements of Corollary 3.4. Part (3) is provided by Proposition 3.6.

It remains to show (4). For each $d \in \mathbb{N}$, the set

$$C_d = \{gfg^{-1} \mid g \in \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})_{\leq d}\}$$

is closed in $\mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$, and the conjugacy class of f is the infinite union $C = \bigcup_d C_d$. Let A be an algebraic variety, $F\colon A \to \mathrm{SAut}(\mathbb{A}^n_{\mathrm{k}})$ be a morphism, and let $B \subset A$ be a locally closed subset which is contained in $F^{-1}(C)$. Writing $B_d = F^{-1}(C_d)$, the set B_d is closed in B for each d and $B = \bigcup_d B_d$. Since k is uncountable and $B_d \subset B_{d+1}$ for each d, we obtain $B = B_m$ for some integer m. Hence, B is contained in $F^{-1}(C_m)$, which is closed in A, so the closure of B in A is also contained in $F^{-1}(C_m)$, and thus in $F^{-1}(C)$. □

4. Conjugacy classes of elements in $\mathrm{SAut}(\mathbb{A}^2_K)$

4A. ***Overview of the Jung–van der Kulk theorem and its applications.*** For each field K, the Jung–van der Kulk theorem [Jung 1942; van der Kulk 1953] asserts that the group $\mathrm{Aut}(\mathbb{A}^2_K)$ is generated by the groups

$$\mathrm{Aff}(\mathbb{A}^2_K) = \left\{(ax_1 + bx_2 + e, cx_1 + dx_2 + f) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, K),\ e, f \in K\right\},$$

$$\mathrm{J}(\mathbb{A}^2_K) = \{(ax_1 + P(x_2), bx_2 + c) \mid a, b \in K^*,\ c \in K,\ P \in K[x_2]\}.$$

Multiplying the decomposition of an element of $\mathrm{SAut}(\mathbb{A}^2_K)$ with homotheties we can assume that each one has determinant 1. This implies that the group $\mathrm{SAut}(\mathbb{A}^2_K)$ is generated by the groups

$$\mathrm{SAff}(\mathbb{A}^2_K) = \left\{(ax_1+bx_2+e, cx_1+dx_2+f) \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,K),\ e,f\in K\right\},$$
$$\mathrm{SJ}(\mathbb{A}^2_K) = \{(ax_1+P(x_2), a^{-1}x_2+c) \mid a\in K^*,\ c\in K,\ P\in K[x_2]\}.$$

The Jung–van der Kulk theorem implies that we have an amalgamated product structure on $\mathrm{SAut}(\mathbb{A}^2_K)$. It also yields the following classical result. We recall here the simple proof.

Lemma 4.1. *Every element $f\in \mathrm{SAut}(\mathbb{A}^2_K)$ is conjugate either to an element of $\mathrm{SJ}(\mathbb{A}^2_K)$ or to an element of the form*

$$f = a_m j_m \cdots a_1 j_1,$$

where $m\geq 1$ and each $a_i\in \mathrm{SAff}(\mathbb{A}^2_K)\setminus \mathrm{SJ}(\mathbb{A}^2_K)$ and each $j_i\in \mathrm{SJ}(\mathbb{A}^2_K)\setminus \mathrm{SAff}(\mathbb{A}^2_K)$.

Proof. We write f as a product of elements of $A=\mathrm{SAff}(\mathbb{A}^2_K)$ and $J=\mathrm{SJ}(\mathbb{A}^2_K)$:

$$f = a_m j_m a_{m-1}\cdots a_1 j_1 a_0,$$

where each $a_i\in A$ and each $j_i\in J$. By merging elements we can moreover assume that $j_i\notin A$ for $i=1,\dots,m$, and that $a_i\notin J$ for $i=1,\dots,m-1$.

Suppose first that $m=0$, which implies that $f=a_0\in A$, so it extends to an element of $\mathrm{Aut}(\mathbb{P}^2_K)$. The action at infinity has a fixed point, and conjugating by an element of $\mathrm{SL}(2,K)$, we can assume that this point corresponds to $[0:1:0]$, which implies that f preserves the pencil of lines through the point, so $f\in J$.

Suppose now that $m\geq 1$, which implies that $j_1\in J\setminus A$. We conjugate f by a_0 and assume that $a_0=\mathrm{id}$. If $m=1$ and $a_m=a_1\in J$, then $f\in J$. If $m\geq 2$ and $a_m\in J$, we conjugate by j_1 and decrease m by 1. This reduces to the case $m\geq 1$, $a_1=\mathrm{id}$, $a_m\notin J$. □

Remark 4.2. Elements of the form $f=a_m j_m\cdots a_1 j_1$, where $m\geq 1$ and each $a_i\in \mathrm{SAff}(\mathbb{A}^2_K)\setminus \mathrm{SJ}(\mathbb{A}^2_K)$ and each $j_i\in \mathrm{SJ}(\mathbb{A}^2_K)\setminus \mathrm{SAff}(\mathbb{A}^2_K)$ are usually called *Hénon automorphisms*. Note that each element j_i satisfies $X_{j_i}=I_{j_i}=[0:1:0]$, and that each a_i extends to an automorphism of $\mathbb{P}^2_K$ that moves $[0:1:0]$. By Lemma 2.14, this implies that

$$\deg(f)=\prod_{i=1}^m \deg(j_i),\qquad I_f=[0:1:0],\qquad X_f=I_{f^{-1}}=a_m([0:1:0]).$$

In particular, $I_f\cap I_{f^{-1}}=\varnothing$, $\deg(f)>1$ and $\deg(f^m)=\deg(f)^m$ for each $m\geq 1$ (Corollary 2.17). Moreover, the conjugacy class of f in $\mathrm{SAut}(\mathbb{A}^2_{\overline{K}})$ is closed, where $\overline{K}$ is the algebraic closure of K (Proposition 3.6).

Remark 4.3. Elements of $\mathrm{SJ}(\mathbb{A}^2_K)$ are algebraic since they are contained in an algebraic subgroup of $\mathrm{Aut}(\mathbb{A}^2_K)$ (or equivalently have a bounded degree sequence). Hénon automorphisms are not algebraic as their degree sequence is not bounded. Hence, Lemma 4.1 corresponds to saying that nonalgebraic elements are conjugate to Hénon automorphisms, and algebraic automorphisms to elements of $\mathrm{SJ}(\mathbb{A}^2_K)$.

The following corollary follows from Lemma 4.1, and also holds (with the same proof) for $\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$. This was already observed in [Furter 1999].

Corollary 4.4. (1) *An element $f \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is algebraic if and only if*

$$\deg(f^2) \leq \deg(f).$$

(2) *The subset $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})_{\mathrm{alg}}$ of algebraic elements of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is closed.*

Proof. (1)(a) If f is algebraic it is conjugate to an element $j \in \mathrm{SJ}(\mathbb{A}^2_K)$ (Lemma 4.1, or more precisely Remark 4.3). We can thus write $f = \alpha^{-1} j \alpha$ where $\alpha \in \mathrm{SAut}(\mathbb{A}^2_K)$. Then $\alpha = a_1 j_1 \cdots$, where each $a_i \in \mathrm{SAff}(\mathbb{A}^2_K) \setminus \mathrm{SJ}(\mathbb{A}^2_K)$ and each $j_i \in \mathrm{SJ}(\mathbb{A}^2_K)$ (if α starts with an element of $\mathrm{SJ}(\mathbb{A}^2_K) \setminus \mathrm{SAff}(\mathbb{A}^2_K)$ we can merge this element with j). If $j \notin \mathrm{SAff}(\mathbb{A}^2_K)$, then $\deg(f) = \prod \deg(j_i)^2 \deg(j) = \deg(\alpha^{-1}) \deg(j) \deg(\alpha)$ (see Remark 4.2) and thus

$$\deg(f^2) = \deg(\alpha^{-1} j^2 \alpha) \leq \deg(\alpha^{-1}) \deg(j^2) \deg(\alpha) \leq \deg(f),$$

since $\deg(j^2) \leq \deg(j)$. If $j \in \mathrm{SAff}(\mathbb{A}^2_K) \cap \mathrm{SJ}(\mathbb{A}^2_K)$, then we can replace j with $a_1^{-1} j a_1$ and continue like this to obtain either the previous case or $f = \alpha^{-1} a \alpha$, where $a \in \mathrm{SAff}(\mathbb{A}^2_K)$, $\alpha = j_1 a_1 \cdots$, each $a_i \in \mathrm{SAff}(\mathbb{A}^2_K) \setminus \mathrm{SJ}(\mathbb{A}^2_K)$ and each $j_i \in \mathrm{SJ}(\mathbb{A}^2_K) \setminus \mathrm{SAff}(\mathbb{A}^2_K)$. We obtain similarly $\deg(f^2) \leq \deg(f)$.

(1)(b) If f is not algebraic, then f is conjugate to a Hénon map $h = a_m j_m \cdots a_1 j_1$, where each $a_i \in \mathrm{SAff}(\mathbb{A}^2_K) \setminus \mathrm{SJ}(\mathbb{A}^2_K)$ and each $j_i \in \mathrm{SJ}(\mathbb{A}^2_K) \setminus \mathrm{SAff}(\mathbb{A}^2_K)$ (Lemma 4.1, or more precisely Remark 4.3). Writing $f = \alpha^{-1} h \alpha$, where $\alpha \in \mathrm{SAut}(\mathbb{A}^2_K)$, we can write as before α as a product of elements of $\mathrm{SAff}(\mathbb{A}^2_K)$ and $\mathrm{SJ}(\mathbb{A}^2_K)$, replace h if α starts with an element $j_1' \in \mathrm{SJ}(\mathbb{A}^2_K)$ such that $j_1 j_1' \in \mathrm{SAff}(\mathbb{A}^2_K)$ or if it starts with an element $a_1' \in \mathrm{SAff}(\mathbb{A}^2_K)$ such that $(a_1')^{-1} \alpha_m \in \mathrm{SJ}(\mathbb{A}^2_K)$ and finish with a simplified writing where we can directly read that $\deg(f^2) > \deg(f)$.

(2) It follows from (1) that the set of algebraic elements of $\mathrm{SAut}(\mathbb{A}^2_K)_{\leq d}$ is closed for each $d \geq 1$, which implies the second assertion. □

4B. ***Conjugacy classes of elements of*** $\mathbf{SJ}(\mathbb{A}^2_K)$**.** It is easy to decide whether two Hénon automorphisms are conjugate, using the amalgamated product structure. The writing of such an element is unique up to cycling permutation of the elements and up to inserting elements of $\mathrm{SJ}(\mathbb{A}^2_K) \cap \mathrm{SAff}(\mathbb{A}^2_K)$. The classification of conjugacy classes of $\mathrm{SAut}(\mathbb{A}^2_K)$ thus reduces to the study of elements of $\mathrm{SJ}(\mathbb{A}^2_K)$.

Lemma 4.5. *Let $f \in \mathrm{SJ}(\mathbb{A}^2_K)$ be some element. After conjugation in this group, the element belongs to one of the following families:*

$$\begin{array}{lll}
\text{(i)} & (ax_1, a^{-1}x_2), & a \in K \setminus \{0,1\}, \\
\text{(ii)} & (x_1 + P(x_2), x_2), & P \in K[x_2], \\
\text{(iii)} & (\zeta x_1 + x_2^{m-1}P(x_2^m), \zeta^{-1}x_2), & \zeta \in K \setminus \{0,1\}, \quad P \in K[x_2] \setminus \{0\}, \\
 & & \zeta \textit{ a primitive } m\textit{-th root of unity}, \\
\text{(iv)} & \begin{cases} (x_1, x_2 + 1), \\ (x_1 + x_2^{p-1}P(x_2^p), x_2 + 1), \end{cases} & \begin{array}{l} \mathrm{char}(K) = 0, \\ P \in K[x_2], \quad \mathrm{char}(K) = p. \end{array}
\end{array}$$

Proof. An element $f \in \mathrm{SJ}(\mathbb{A}^2_K)$ is equal to $(ax_1 + P(x_2), a^{-1}x_2 + c)$ for some $a \in K^*$, $c \in K$, $P \in K[x_2]$.

If $a \neq 1$, we conjugate by $(x_1, x_2 + ac/(1-a))$ and can assume that $c = 0$. Conjugating by $(x_1 - \lambda x_2^n, x_2)$ yields $(ax_1 + P(x_2) + \lambda(a - a^{-n})x_2^n, a^{-1}x_2)$. We can then kill the coefficient of degree n of P, if $a \neq a^{-n}$. This gives (i) or (iii).

If $a = 1$ and $c = 0$, we get (ii). If $a = 1$ and $c \neq 0$, we conjugate by $(cx_1, c^{-1}x_2)$ and can assume that $c = 1$. Conjugating $(x_1 + P(x_2), x_2 + 1)$ by $(x_1 - \lambda x_2^{n+1}, x_2)$ yields $(x_1 + P(x_2) + \lambda(x_2^{n+1} - (x_2+1)^{n+1}), x_2 + 1)$. We can thus kill the coefficient of degree n if $n+1$ is not a multiple of $\mathrm{char}(K)$. This gives (iv). □

The most complicated case corresponds to family (iv), in the case of a field of characteristic $p > 0$. In order to describe the conjugacy classes in this family, we will need the following lemma.

Lemma 4.6. *Let K be a field of characteristic $p > 0$. Let $V \subset K[x]$ be the subvector space given by*

$$V = \{x^{p-1}P(x^p) \mid P \in K[x]\},$$

and let δ, N be the following K-linear maps

$$\begin{aligned}
\delta\colon K[x] &\longrightarrow K[x], \quad F(x) \longmapsto F(x+1) - F(x), \\
N\colon K[x] &\longrightarrow K[x], \quad F(x) \longmapsto F(x) + F(x+1) + \cdots + F(x+p-1).
\end{aligned}$$

Then, $\mathrm{Im}(\delta) = \mathrm{Ker}(N)$ *and* $K[x] = V \oplus \mathrm{Im}(\delta) = V \oplus \mathrm{Ker}(N)$.

Proof. (1) We first show that $V + \mathrm{Im}(\delta) = K[x]$. This is the same argument as in Lemma 4.5(iv). We take any polynomial $P(x) = \sum_{j=1}^{l} a_j x^j \in K[x]$. If $P \notin V$, we can define m to be the biggest integer such that $a_m \neq 0$ and $m + 1 \not\equiv 0 \pmod{p}$. Replacing P with $\tilde{P}(x) = P(x) - \delta(a_m x^{m+1}/(m+1))$, we kill the coefficient of x^m. Continuing this process until all coefficients a_i with $i + 1 \not\equiv 0 \pmod{p}$ are zero, we obtain an element of V.

(2) We now show that $\mathrm{Ker}(N) \cap V = \{0\}$. Let $F(x) = \sum_{i=1}^{n} a_i x^{ip-1}$ be an element of V, and assume that $a_n \neq 0$. Since

$$N(F(x)) = \sum_{i=1}^{n} a_i \Bigl(\sum_{k=0}^{p-1} (x+k)^{ip-1} \Bigr),$$

the coefficient of x^{np-p} of $N(F(x))$ is equal to

$$a_n \binom{np-1}{np-p} \sum_{k=0}^{p-1} k^{p-1} = a_n \binom{np-1}{np-p} (p-1) \in K^*.$$

This shows that $F \notin \mathrm{Ker}(N)$. Hence $V \cap \mathrm{Ker}(N) = \{0\}$.

(3) The inclusion $\mathrm{Im}(\delta) \subset \mathrm{Ker}(N)$ follows from the definition of δ and N, so we obtain $V \cap \mathrm{Im}(\delta) = \{0\}$ by (2). Moreover, since $\mathrm{Ker}(N)$ is contained in $\mathrm{Im}(\delta) \oplus V$ (by (1)) and $\mathrm{Ker}(N) \cap V = \{0\}$, we have $\mathrm{Im}(\delta) = \mathrm{Ker}(N)$. □

Remark 4.7. The equality $\mathrm{Ker}(N) \cap V = \{0\}$ of Lemma 4.6 corresponds to the fact that an element of the form

$$(x_1 + x_2^{p-1} P(x_2^p), x_2 + 1),$$

where $P \in K[x_2]$, $\mathrm{char}(K) = p$ (family (iv)) is of order p if and only if $P = 0$. Indeed,

$$(x_1 + Q(x_2), x_2 + 1)^p = (x_1 + Q(x_2) + \cdots + Q(x_2 + p - 1), x_2),$$

for each $Q \in K[x_2]$.

We will also need the following lemma, which is a direct consequence of the amalgamated product structure of $\mathrm{SAut}(\mathbb{A}_K^2)$.

Lemma 4.8. *Let $f \in \mathrm{SJ}(\mathbb{A}_K^2)$, which is not conjugate to $(ax_1, a^{-1}x_2)$, $a \in K^*$ or to (x_1+1, x_2) or (x_1, x_2+1) in the group $\mathrm{SJ}(\mathbb{A}_K^2)$. Then, f is conjugate to an element of $g \in \mathrm{SJ}(\mathbb{A}_K^2)$ in the group $\mathrm{SAut}(\mathbb{A}_K^2)$ if and only if it is conjugate to g in $\mathrm{SJ}(\mathbb{A}_K^2)$.*

Proof. We suppose that $g = \alpha f \alpha^{-1} \in \mathrm{SJ}(\mathbb{A}_K^2)$, for some $\alpha \in \mathrm{SAut}(\mathbb{A}_K^2) \setminus \mathrm{SJ}(\mathbb{A}_K^2)$. We can write α as

$$\alpha = j_{m+1} a_m j_m \cdots j_2 a_1 j_1,$$

where $m \geq 1$, each $j_i \in \mathrm{SJ}(\mathbb{A}_K^2)$, each $a_i \in \mathrm{SAff}(\mathbb{A}_K^2)$, and such that $a_i \notin \mathrm{SJ}(\mathbb{A}_K^2)$ for $i = 1, \ldots, m$, $j_i \notin \mathrm{SAff}(\mathbb{A}_K^2)$ for $i = 2, \ldots, m-1$. Since $\alpha f \alpha^{-1} \in \mathrm{SJ}(\mathbb{A}_K^2)$, the element $j_1 f j_1^{-1}$ belongs to $\mathrm{SAff}(\mathbb{A}_K^2) \cap \mathrm{SJ}(\mathbb{A}_K^2)$, and the same holds for the element $a_1 j_1 f j_1^{-1} a_1^{-1}$.

The fact that $f_1 = j_1 f j_1^{-1}$ and $f_2 = a_1 j_1 f j_1^{-1} a_1^{-1}$ belong to $\mathrm{SAff}(\mathbb{A}_K^2) \cap \mathrm{SJ}(\mathbb{A}_K^2)$ corresponds to the fact that both extend to automorphisms of $\mathbb{P}_K^2$ that fix the point $[0 : 1 : 0]$ at infinity. However, $a_1 \in \mathrm{SAff}(\mathbb{A}_K^2) \setminus \mathrm{SJ}(\mathbb{A}_K^2)$ so it extends to an automorphism of $\mathbb{P}^2_K$, whose action at the infinity moves the point $[0:1:0]$. This implies that f_1 fixes at least two points at infinity. Conjugating by an element of $\mathrm{SAff}(\mathbb{A}^2_K)\cap\mathrm{SJ}(\mathbb{A}^2_K)$ we can assume that these points are $[0:1:0]$ and $[0:0:1]$. This implies that $f_1=(ax_1+b,a^{-1}x_2+c)$, for some $(a,b,c)\in K^*\times K^2$. It remains to see that f_1 is conjugate to $(ax_1,a^{-1}x_2)$, (x_1+1,x_2) or (x_1,x_2+1), in the group $\mathrm{SJ}(\mathbb{A}^2_K)$. This is a straight-forward calculation, already done in the proof of Lemma 4.5. □

Proposition 4.9 (Conjugacy classes of algebraic elements in $\mathrm{SAut}(\mathbb{A}^2_K)$). *Every algebraic element of* $\mathrm{SAut}(\mathbb{A}^2_K)$ *is conjugate to one of the families* (i)–(iv) *of Lemma 4.5.*

Apart from (x_1,x_2+1), *which is conjugate to* (x_1+1,x_2), *no two elements of distinct families are conjugate in* $\mathrm{SAut}(\mathbb{A}^2_K)$. *Moreover, the conjugacy classes in* $\mathrm{SAut}(\mathbb{A}^2_K)$ *within the families are as follows*:

(i) $(ax_1,a^{-1}x_1)\sim(bx_1,b^{-1}x_2)\Longleftrightarrow a=b^{\pm1}$.

(ii) $(x_1+P(x_2),x_2)\sim(x_1+Q(x_2),x_2)\Longleftrightarrow Q(x_2)=aP(ax_2+b)$ *for some* $(a,b)\in K^*\times K$.

(iii) $(\zeta x_1+x_2^{m-1}P(x_2^m),\zeta^{-1}x_2)$ *is conjugate to* $(\zeta x_1+x_2^{m-1}a^mP(ax_2^m),\zeta^{-1}x_2)$ *for each* $a\in K^*$, *but not to any other element of family* (iii).

(iv) *Two elements*

$$f=(x_1+x_2^{p-1}P(x_2^p),x_2+1),\quad g=(x_1+x_2^{p-1}Q(x_2^p),x_2+1),$$

where $P,Q\in K[x_2]$, *are conjugate if and only their* p*-th powers*,

$$f^p=(x_1+\tilde{P}(x_2),x_2),\quad g^p=(x_1+\tilde{Q}(x_2),x_2),$$

where $\tilde{P},\tilde{Q}\in K[x_2]$, *satisfy* $\tilde{Q}(x_2)=\tilde{P}(x_2+c)$ *for some* $c\in K$.

Proof. The first assertion follows from Lemma 4.1 (see Remark 4.3). It remains to describe the conjugacy classes in the families (i)–(iv). We first observe that if $\alpha\in\mathrm{SL}(2,K)$ is conjugate by $\nu\in\mathrm{SAut}(\mathbb{A}^2_K)$ to an element $\beta\in\mathrm{SAut}(\mathbb{A}^2_K)$ that fixes the origin, the derivative of the equation $\alpha\nu=\nu\beta$ at the origin yields $\alpha D_\nu(0)=D_\nu(0)D_\beta(0)$, so the derivative of β at the origin is conjugate to α in $\mathrm{SL}(2,K)$. This shows that the families (i) and (ii) are disjoint, and gives the conjugacy classes in (i).

We now observe that two elements $f=(x_1+P(x_2),x_2)$ and $(x_1+Q(x_2),x_2)$ are conjugate in $\mathrm{SAut}(\mathbb{A}^2_K)$ if and only if they are conjugate in $\mathrm{SJ}(\mathbb{A}^2_K)$. This is given by Lemma 4.8, if f or g is not conjugate to (x_1+1,x_2) or (x_1,x_2) in $\mathrm{SJ}(\mathbb{A}^2_K)$, and is trivial in the remaining cases (when both f and g are conjugate to one of these two elements). It remains to observe that the conjugation of $f=(x_1+P(x_2),x_2)$ by

$(ax_1 + R(x_2), a^{-1}(x_2 - b))$ yields $(x_1 + aP(ax_2 + b), x_2)$ to obtain the conjugacy classes in (ii).

An element $f = (\zeta x_1 + x_2^{m-1} P(x_2^m), \zeta^{-1} x_2)$ in family (iii) satisfies

$$f^m = (x_1 + m x_2^{m-1} P(x_2^m), x_2) \neq (x_1, x_2).$$

Since f^m belongs to family (ii), f^m (and thus f) is not conjugate to an element of family (i). Moreover, f is not conjugate to an element of family (ii) because $\zeta^{-1} \in K \setminus \{0, 1\}$ is an eigenvalue of the action of f^* on $K[x_1, x_2]$. By Lemma 4.8, f is conjugate to another element g of family (iii) in $\mathrm{SAut}(\mathbb{A}_K^2)$ if and only if this conjugation holds in $\mathrm{SJ}(\mathbb{A}_K^2)$. Looking at the second coordinate, we have to conjugate by $\alpha = (ax_1 + R(x_2), a^{-1} x_2)$. This implies that $g = (\zeta x_1 + x_2^{m-1} Q(x_2^m), \zeta^{-1} x_2)$ for some polynomial Q. Looking at $g^m = \alpha f^m \alpha^{-1} = (x_1 + m a^m x_2^{m-1} P((ax_2)^m), x_2)$, we obtain $Q(x_2^m) = a^m P((ax_2)^m)$. This yields a necessary condition on Q, which is also sufficient, by choosing $R = 0$.

It remains to study the last class (iv). The element $(x_1, x_2 + 1)$ is conjugate to $(x_1 + 1, x_2)$ by $(x_2, -x_1)$. If $\mathrm{char}(K) = 0$, there is no other element in the family. So we assume that $\mathrm{char}(K) = p > 0$, and consider $f = (x_1 + x_2^{p-1} P(x_2^p), x_2 + 1)$, for some polynomial $P \in K[x_2] \setminus \{0\}$. Because the action of f^* on $K[x_1, x_2]$ has only eigenvalues equal to 1, f is not conjugate to any element of family (i) or (iii). Moreover, the fact that $P \neq 0$ implies that f^p is not the identity (see Remark 4.7). Hence, f is not conjugate to an element of family (ii).

We then take $g = (x_1 + x_2^{p-1} Q(x_2^p), x_2 + 1)$, for some $Q \in K[x_2]$, write

$$f^p = (x_1 + \tilde{P}(x_2), x_2), \quad g^p = (x_1 + \tilde{Q}(x_2), x_2),$$

for some $\tilde{P}, \tilde{Q} \in K[x_2]$, and show that f and g are conjugate if and only if $\tilde{Q}(x_2) = \tilde{P}(x_2 + c)$ for some $c \in K$. This will conclude the proof.

Suppose first that f is conjugate to g. Lemma 4.8 yields the existence of $\alpha \in \mathrm{SJ}(\mathbb{A}_K^2)$ such that $\alpha f \alpha^{-1} = g$. Looking at the second coordinate, we see that $\alpha = (x_1 + R(x_2), x_2 - c)$ for some $R \in K[x_2]$, $c \in K$. Then, $\alpha f^p \alpha^{-1} = (x_1 + \tilde{P}(x_2 + c), x_2)$, which implies that $\tilde{Q}(x_2) = \tilde{P}(x_2 + c)$ as we wanted.

Conversely, suppose that $\tilde{Q}(x_2) = \tilde{P}(x_2 + c)$ for some $c \in K$. We conjugate then g with $(x_1, x_2 + c)$; this changes maybe the polynomial Q, and replaces $\tilde{Q}$ with $\tilde{P}$. Hence, we can assume that $f^p = g^p$. This corresponds to the equality

$$N(x_2^{p-1} P(x_2^p)) = N(x_2^{p-1} Q(x_2^p)),$$

where $N \colon K[x_2] \to K[x_2]$ is the K-linear map that sends $f(x_2)$ to $f(x_2) + f(x_2+1) + \cdots + f(x_2 + p - 1)$. The element $x_2^{p-1} Q(x_2^p) - x_2^{p-1} P(x_2^p)$ is an element of the kernel of N, and is thus equal to $R(x_2 + 1) - R(x_2)$ for some polynomial $R \in K[x_2]$ (Lemma 4.6). This corresponds to the fact that f is conjugate to g by $\alpha = (x_1 + R(x_2), x_2)$. □

4C. ***Degenerations between the four families of conjugacy classes.*** In the remaining part of the article, we show that every algebraic element of $\mathrm{SAut}(\mathbb{A}^2_K)$ admits a degeneration of conjugates towards a diagonalisable element.

4C1. *From family* (ii) *and* (iii) *to family* (i). These are the easiest degenerations. Taking any $f \in \mathrm{SAut}(\mathbb{A}^2_K)$ that belongs to family (ii) or (iii), and conjugating by diagonal elements of $\mathrm{SAut}(\mathbb{A}^2_{K(t)})$ we obtain elements of $\mathrm{SAut}(\mathbb{A}^2_{K[t]})$ which yield families of conjugates of f that converge towards an element of family (i). Indeed, we have

$$(tx_1, t^{-1}x_2)(x_1 + P(x_2), x_2)(t^{-1}x_1, tx_2) = (x_1 + tP(x_2), x_2),$$
$$(tx_1, t^{-1}x_2)f(t^{-1}x_1, tx_2) = (\zeta x_1 + t^m x_2^{m-1}P(tx_2^m), \zeta^{-1}x_2).$$

In particular, there is a diagonalisable element in the closure of the conjugacy class of any element of families (ii) and (iii), and thus of any algebraic element of $\mathrm{SAut}(\mathbb{A}^2_K)$, if $\mathrm{char}(K) = 0$.

4C2. *Family* (iv). The last family, when $\mathrm{char}(K) = p > 0$, is the most interesting. The elements of this family are of the form $(x_1 + Q(x_2), x_2 + 1)$, for some polynomial $Q \in K[x_2]$. Diagonal conjugations by $(tx_1, t^{-1}x_2)$ or $(t^{-1}x_1, tx_2)$ yield elements of $\mathrm{SAut}(\mathbb{A}^2_{K[t^{\pm 1}]})$ which do not have a value at $t = 0$.

In the following proposition, we provide two explicit degenerations, typical to positive characteristic, to either $(x_1, x_2 + 1)$ or the identity.

Proposition 4.10. *Assume that* $\mathrm{char}(K) = p > 0$ *and let*

$$f = (x_1 + Q(x_2), x_2 + 1)$$

for some polynomial $Q \in K[x_2]$.

Then, there are two elements $F_1, F_2 \in \mathrm{SAut}(\mathbb{A}^2_{K[t]})$, *which are conjugate to* f *in* $\mathrm{SAut}(\mathbb{A}^2_{K[t^{\pm 1}]})$, *and have the following properties*:

(1) *For each* $t \in K \setminus \{0\}$, $F_1(t), F_2(t) \in \mathrm{SAut}(\mathbb{A}^2_K)$ *are conjugate to* f *in* $\mathrm{SAut}(\mathbb{A}^2_K)$.

(2) $F_1(0) = (x_1, x_2 + 1)$, $\quad F_2(0) = (x_1, x_2)$.

Proof. If $Q = 0$, we can choose $F_1(t) = f$ and $F_2(t)$ to be the conjugation of f by $(t^{-1}x_1, tx_2) \in \mathrm{SAut}(\mathbb{A}^2_{K[t^{\pm 1}]})$.

We can thus assume that $Q \neq 0$, denote by $d \geq 0$ the degree of Q and by $\mu \in K^*$ the coefficient of degree d. We then choose some integer $a \geq 1$ such that $q = p^a > d$, write $\lambda = 1/\mu^q \in K^*$, and define

$$\alpha = \left(\frac{x_1}{t^d}, t^d x_2 + \lambda x_1^q + \frac{1}{t}\right) \in \mathrm{SAut}(\mathbb{A}^2_{K[t,t^{-1}]}).$$

A direct computation yields

$$\alpha^{-1} f\alpha$$
$$=\left(x_1+t^d Q\left(t^d x_2+\lambda x_1^q+\frac{1}{t}\right), x_2+\frac{1}{t^d}+\frac{\lambda x_1^q}{t^d}-\frac{\lambda}{t^d}\left(x_1+t^d Q\left(t^d x_2+\lambda x_1^q+\frac{1}{t}\right)\right)^q\right).$$

The definition of d and μ implies that we can write

$$Q\left(t^d x_2+\lambda x_1^q+\frac{1}{t}\right)=\frac{\mu}{t^d}+\frac{P}{t^{d-1}},$$

for some $P \in K[x_1, x_2, t]$. This yields (remembering that $\lambda\mu^q = 1$), the following equality

$$\alpha^{-1} f\alpha=\left(x_1+\mu+tP, x_2+\frac{1}{t^d}-\frac{\lambda}{t^d}(\mu+tP)^q\right)=\left(x_1+\mu+tP, x_2-\lambda t^{q-d}P^q\right).$$

Because $q > d$, the value of this element of $\mathrm{SAut}(\mathbb{A}^2_{K[t]})$ at $t = 0$ is $(x_1 + \mu, x_2)$. Conjugating $\alpha^{-1} f\alpha$ with $(-\mu x_2, \mu^{-1}x_1)$ yields thus F_1.

To get F_2, we recall that $(tx_1, t^{-1}x_2)(x_1 + \mu, x_2)(t^{-1}x_1, tx_2) = (x_1 + t\mu, x_2)$. We are then tempted to use $(tx_1, t^{-1}x_2)\alpha^{-1} f\alpha(t^{-1}x_1, tx_2)$, but this element has in general no value at $t = 0$. We then choose $m > 0$ and define β to be

$$\beta=\left(\frac{x_1}{t^{md}}, t^{md}x_2+\lambda x_1^q+\frac{1}{t^m}\right)\in \mathrm{SAut}(\mathbb{A}^2_{K[t,t^{-1}]}).$$

Since β is obtained from α by replacing t with t^m, we obtain

$$\beta^{-1} f\beta=\left(x_1+\mu+t^m P(x_1, x_2, t^m), x_2-\lambda t^{m(q-d)}P(x_1, x_2, t^m)^q\right).$$

We can now define $F_2(t) = (tx_1, t^{-1}x_2)\beta^{-1} f\beta(t^{-1}x_1, tx_2)$, which is equal to

$$F_2(t)=\left(x_1+t\mu+t^{m+1}P(t^{-1}x_1, tx_2, t^m), x_2-\lambda t^{m(q-d)-1}P(t^{-1}x_1, tx_2, t^m)^q\right).$$

Choosing m big enough, F_2 is defined at $t = 0$ and $F_2(t) = (x_1, x_2)$. □

4D. ***The diagonalisable elements.*** In order to finish the proof of Theorem 1.3 it remains to show that the conjugacy classes of diagonalisable elements is closed. This was shown in [Furter and Maubach 2010], for the group $\mathrm{Aut}(\mathbb{A}^2_{\mathbb{C}})$, but with transcendental methods. In the case of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$, we can however give a simple proof, that works for any algebraically closed field k.

The proof uses the following result, which follows from the amalgamated structure product.

Lemma 4.11. *Let K be any field, let $\mu \in K^*$, $f = (\mu x_1, \mu^{-1}x_2) \in \mathrm{SAut}(\mathbb{A}^2_K)$, and $g \in \mathrm{SAut}(\mathbb{A}^2_K)$ be a conjugate of f in this group. Then, there exists $\alpha \in \mathrm{SAut}(\mathbb{A}^2_K)$ such that*

$$g=\alpha^{-1} f\alpha \quad \textit{and} \quad \deg\alpha \le \deg g.$$

Proof. We write $g=\alpha^{-1}f\alpha$ for some $\alpha\in\mathrm{SAut}(\mathbb{A}^2_K)$, and write α as a product of elements of $\mathrm{SAff}(\mathbb{A}^2_K)$ and $\mathrm{SJ}(\mathbb{A}^2_K)$, in a reduced way (no two consecutive elements belong to the same group). Note that f is conjugate to f^{-1} by $(x_2,-x_1)$, so we can easily exchange f with f^{-1} if needed.

If α starts with an element $a\in\mathrm{SAff}(\mathbb{A}^2_K)$, then $a'=a^{-1}ja\in\mathrm{SAff}(\mathbb{A}^2_K)$. If a' does not belong to $\mathrm{SJ}(\mathbb{A}^2_K)$, then $\deg(g)=\deg(\alpha)\deg(\alpha^{-1})=\deg(\alpha)^2$ (the degree is the product of the Jonquières elements that occur, see Remark 4.2). If a' belongs to $\mathrm{SJ}(\mathbb{A}^2_K)$, then it is conjugate to f or f^{-1} in $\mathrm{SJ}(\mathbb{A}^2_K)$ (Proposition 4.9), and we can thus replace a with an element of $\mathrm{SJ}(\mathbb{A}^2_K)$ and either decrease the length of α or it reduces to the case $\alpha\in\mathrm{SJ}(\mathbb{A}^2_K)$.

Suppose now that α starts with an element $j\in\mathrm{SJ}(\mathbb{A}^2_K)$. Replacing j with ρj, where ρ is diagonal, we can assume that $j=(x_1+P(x_2),x_2+c)$, for some $P\in K[x_2]$ and $c\in K$, we thus obtain

$$j'=j^{-1}fj=\big(\mu x_1+P(x_2)\mu-P(x_2+c(\mu^{-1}-1)),\,\mu^{-1}x_2+c(\mu^{-1}-1)\big)\in\mathrm{SJ}(\mathbb{A}^2_K).$$

We have therefore $\deg(P(x_2))\geq P(x_2)\mu-P(x_2+c(\mu^{-1}-1))$, and if equality does not hold we can kill the highest coefficient of P without changing $j'=j^{-1}fj$. We reduce then to the case where $\deg j'=\deg j$. If this degree is 1, then $j\in\mathrm{SAff}(\mathbb{A}^2_K)$ and we can reduce the length of α, or obtain $\deg(\alpha)=1$. The remaining case is when $\deg j'=\deg j$. The result is then obtained if $\alpha=j$. Otherwise, $\alpha=ja_1j_1\cdots$ is a reduced writing, as well as $g=\cdots j_1^{-1}a_1^{-1}j'a_1j_1\cdots$. We again find $\deg g\geq\deg\alpha$. □

Proposition 4.12. *Let $\mu\in\mathrm{k}^*$, $f=(\mu x_1,\mu^{-1}x_2)\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ and $d\geq1$. The set*

$$C(f)=\{gfg^{-1}\mid g\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})\}=\{gfg^{-1}\mid g\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})\}$$

is closed in $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$.

Proof. The equality between the two sets is obvious: if $g\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$, we have $g'=g\tau\in\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ for some diagonal $\tau\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})$, so $gfg^{-1}=g'fg'^{-1}$. We can moreover assume that $\mu\neq1$, otherwise the result is trivial. We fix some integer $d\geq1$, write $Z=\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})_{\leq d}$ and we will show that $C(f)\cap Z$ is closed in Z (this gives the result by Lemma 2.7).

We consider the morphism of algebraic varieties $Z\to\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}})_{\leq d^2}$ given by $g\mapsto gfg^{-1}$ and denote by $Y\subset\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})_{\leq d}$ the preimage of Z. This gives rise to a morphism $\varphi\colon Y\to Z$, whose image is equal to $C(f)\cap Z$ by Lemma 4.11. By Corollary 2.26, it suffices to take an irreducible k-curve Γ, a rational map $\iota\colon\Gamma\dashrightarrow Y$ such that $\hat\varphi=\varphi\circ\iota\colon\Gamma\dashrightarrow Z$ is defined at a smooth point $p\in\Gamma$, and to show that the image $g=\hat\varphi(p)\in Z$ belongs to $C(f)$.

Note that ι corresponds to an element $\alpha\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}(\Gamma)})$, and that $\hat\varphi$ corresponds to the element $\alpha f\alpha^{-1}\in\mathrm{Aut}(\mathbb{A}^2_{\mathrm{k}(\Gamma)})$, which is defined at p and whose value at p

corresponds to g. Since the set of algebraic elements is closed (Corollary 4.4), the element g is conjugate to an element of $\mathrm{SJ}(\mathbb{A}^2_{\mathrm{k}})$ (Remark 4.3). Looking at the actions on $\mathrm{k}(\Gamma)[x_1, x_2]$, we find that $f^*(x_1) = \mu x_1$ so $(\alpha f \alpha^{-1})^*(v) = \mu v$, where $v = (\alpha^{-1})^*(x_1) \in \mathrm{k}(\Gamma)[x_1, x_2]$. Replacing v with λv, where $\lambda \in \mathrm{k}(\Gamma)$, does not change the equation $(\alpha f \alpha^{-1})^*(v) = \mu v$. We can therefore assume that $\alpha \in \mathcal{O}_p(\Gamma)[x_1, x_2]$ and that its value at p is not zero. This shows that μ is also an eigenvalue of g^*. By Lemma 4.5, g is conjugate in $\mathrm{SJ}(\mathbb{A}^2_{\mathrm{k}})$ to $(\nu x_1, \nu^{-1} x_2)$, for some $\nu \in \mathrm{k} \setminus \{0, 1\}$, or to $(\zeta x_1 + x_2^{m-1} P(x_2^m), \zeta^{-1} x_2)$ for some primitive m-th root of unity ζ, and $P \in \mathrm{k}[x_2] \setminus \{0\}$, and μ belongs to the subgroup of $(\mathrm{k}^*, \cdot)$ generated by ν or by ζ, respectively. Let us observe that the second case is not possible. Indeed, otherwise we would have $f^m = \mathrm{id}$, so $(\alpha f \alpha^{-1})^m = \mathrm{id}$ and thus $g^m = \mathrm{id}$, which is not the case since $P \neq 0$. We can thus assume, after composing α with an element of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$, that $g = (\nu x_1, \nu^{-1} x_2)$ for some $\nu \in \mathrm{k} \setminus \{0, 1\}$.

Denote by $U \subset \Gamma$ the dense open subset where ι is defined and where Γ is smooth. Replacing Γ with $U \cup \{p\}$, we can assume that $U = \Gamma \setminus \{p\}$ (if $p \in U$, the result is obvious), and thus that $\hat{\varphi}$ is a morphism $\hat{\varphi} \colon \Gamma \to Z$. Denote by $\kappa \colon \Gamma \dashrightarrow \mathbb{A}^2_{\mathrm{k}}$ the rational map given by $\kappa(u) = \iota(u)(0, 0)$, when $u \in U$ (the image of the origin of $\mathbb{A}^2_{\mathrm{k}}$).

We now prove that κ is defined at p. To do this, we define $F \subset \Gamma \times \mathbb{A}^2_{\mathrm{k}}$ to be the closed set $F = \{(c, (x_1, x_2)) \mid \hat{\varphi}(c)(x_1, x_2) = (x_1, x_2)\}$. Since F is given by only two equations, each irreducible component of F has dimension at least 1. The intersection of F with $u \times \mathbb{A}^2_{\mathrm{k}}$ yields one point, for each $u \in \Gamma$, hence the projection to Γ yields an isomorphism $\pi \colon F \to \Gamma$ (because Γ is smooth). The rational map $\Gamma \dashrightarrow F$ given by $u \mapsto (u, \kappa(u))$ corresponds to the identity, and is thus defined at p.

The map $(x_1, x_2) \mapsto (x_1, x_2) + \kappa(u)$ corresponds therefore to an element of $\mathrm{Aut}(\mathbb{A}^2_{\mathcal{O}(\Gamma)})$, and replacing α with its composition by this one, we can assume that α fixes the origin. We obtain thus that the derivative of $\alpha f \alpha^{-1}$ at the origin is conjugate to f by an element of $\mathrm{SL}(2, \mathrm{k})$, for each $u \in U$, so its eigenvalues are μ and μ^{-1}, for any point of U. At the point p, we obtain $g = (\nu x_1, \nu^{-1} x_2)$, which implies that $\nu = \mu^{\pm 1}$, so that g and f are conjugate in $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ as we wanted. □

Proof of Theorem 1.3. By Lemma 4.1 and Remarks 4.2 and 4.3, an algebraic element of $\mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is conjugate to an element of $\mathrm{SJ}(\mathbb{A}^2_{\mathrm{k}})$ and a nonalgebraic element is conjugate to a Hénon automorphism, which is dynamically regular. In the latter case, the conjugacy class is closed by Theorem 1.1; this gives (3).

By Proposition 4.12, if $f \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is diagonalisable, its conjugacy class is closed.

If $f \in \mathrm{SAut}(\mathbb{A}^2_{\mathrm{k}})$ is algebraic but not diagonalisable, it is conjugate to an element of the families (ii), (iii) or (iv) of Lemma 4.5. The existence of the degeneration stated in (2) is given in Section 4C1 for families (ii), (iii) and in Proposition 4.10 for family (iv). □

Acknowledgements

The author thanks Jean-Philippe Furter and Hanspeter Kraft for interesting discussions during the preparation of this article and for giving him access to the preprint [Furter and Kraft ≥ 2016]. Thanks also to the referee for his precious remarks and corrections which improved the article.

References

[Bass et al. 1982] H. Bass, E. H. Connell, and D. Wright, "The Jacobian conjecture: reduction of degree and formal expansion of the inverse", *Bull. Amer. Math. Soc.* (*N.S.*) **7**:2 (1982), 287–330. MR 663785 Zbl 0539.13012

[Bisi 2008] C. Bisi, "On commuting polynomial automorphisms of $\mathbb{C}^k$, $k \geq 3$", *Math. Z.* **258**:4 (2008), 875–891. MR 2369061 Zbl 1161.32006

[Blanc 2010] J. Blanc, "Groupes de Cremona, connexité et simplicité", *Ann. Sci. Éc. Norm. Supér.* (4) **43**:2 (2010), 357–364. MR 2662668 Zbl 1193.14017

[Blanc and Cantat 2016] J. Blanc and S. Cantat, "Dynamical degrees of birational transformations of projective surfaces", *J. Amer. Math. Soc.* **29**:2 (2016), 415–471. MR 3454379

[Blanc and Furter 2013] J. Blanc and J.-P. Furter, "Topologies and structures of the Cremona groups", *Ann. of Math.* (2) **178**:3 (2013), 1173–1198. MR 3092478 Zbl 1298.14020

[Demazure 1970] M. Demazure, "Sous-groupes algébriques de rang maximum du groupe de Cremona", *Ann. Sci. École Norm. Sup.* (4) **3** (1970), 507–588. MR 0284446 Zbl 0223.14009

[Furter 1999] J.-P. Furter, "On the degree of iterates of automorphisms of the affine plane", *Manuscripta Math.* **98**:2 (1999), 183–193. MR 1667603 Zbl 0944.32017

[Furter 2009] J.-P. Furter, "Plane polynomial automorphisms of fixed multidegree", *Math. Ann.* **343**:4 (2009), 901–920. MR 2471605 Zbl 1187.14064

[Furter and Kraft ≥ 2016] J.-P. Furter and H. Kraft, "On the geometry of the automorphism group of affine n-space", in preparation.

[Furter and Lamy 2010] J.-P. Furter and S. Lamy, "Normal subgroup generated by a plane polynomial automorphism", *Transform. Groups* **15**:3 (2010), 577–610. MR 2718938 Zbl 1226.14080

[Furter and Maubach 2010] J.-P. Furter and S. Maubach, "A characterization of semisimple plane polynomial automorphisms", *J. Pure Appl. Algebra* **214**:5 (2010), 574–583. MR 2577663 Zbl 1196.14056

[Guedj and Sibony 2002] V. Guedj and N. Sibony, "Dynamics of polynomial automorphisms of $\mathbb{C}^k$", *Ark. Mat.* **40**:2 (2002), 207–243. MR 1948064 Zbl 1034.37025

[Jung 1942] H. W. E. Jung, "Über ganze birationale Transformationen der Ebene", *J. Reine Angew. Math.* **184** (1942), 161–174. MR 0008915 Zbl 0027.08503

[van der Kulk 1953] W. van der Kulk, "On polynomial rings in two variables", *Nieuw Arch. Wiskunde* (3) **1** (1953), 33–41. MR 0054574 Zbl 0050.26002

[Mumford 1970] D. Mumford, *Abelian varieties*, Tata Institute of Fundamental Research Studies in Mathematics **5**, Oxford University Press, 1970. MR 0282985 Zbl 0223.14022

[Serre 2010] J.-P. Serre, "Le groupe de Cremona et ses sous-groupes finis", pp. 75–100 in *Séminaire Bourbaki* 2008/2009 (Exposé 1000), Astérisque **332**, 2010. MR 2648675 Zbl 1257.14012

[Shafarevich 1966] I. R. Shafarevich, "On some infinite-dimensional groups", *Rend. Mat. e Appl.* (5) **25**:1-2 (1966), 208–212. MR 0485898 Zbl 0149.39003

[Sibony 1999] N. Sibony, "Dynamique des applications rationnelles de $\mathbb{P}^k$", pp. ix–x, xi–xii, 97–185 in *Dynamique et géométrie complexes* (Lyon, 1997), Panor. Synthèses **8**, Soc. Math. France, Paris, 1999. MR 1760844 Zbl 1020.37026

Communicated by Hubert Flenner
Received 2015-01-02 Revised 2015-11-04 Accepted 2016-03-04

jeremy.blanc@unibas.ch *Mathematisches Institut, Universität Basel, Spiegelgasse 1, CH-4051 Basel, Switzerland*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.969

Inversion of adjunction for rational and Du Bois pairs

Sándor J. Kovács and Karl Schwede

We prove several results about the behavior of Du Bois singularities and Du Bois pairs in families. Some of these generalize existing statements about Du Bois singularities to the pair setting while others are new even in the nonpair setting. We also prove a new inversion of adjunction result for Du Bois and rational pairs. In the nonpair setting this asserts that if a family over a smooth base has a special fiber X_0 with Du Bois singularities and the general fiber has rational singularities, then the total space has rational singularities near X_0.

1. Introduction

Rational singularities have been the gold standard for "mild" singularities in algebraic geometry for several decades. Whenever a new class of varieties with singularities is discovered, the first question usually asked is whether or not the new varieties have rational singularities. A key reason for this is that varieties with rational singularities behave cohomologically as if they were smooth. However, for many purposes rational singularities are not broad enough. For instance, nodes are not rational singularities, and more generally, singularities appearing on stable varieties, that is, mild degenerations of smooth ones that are necessary to consider in order to compactify moduli spaces, are not always rational. The class of Du Bois (or DB) singularities is slightly more inclusive than rational singularities. Du Bois singularities behave cohomologically as if they had simple normal crossing singularities (i.e., a higher-dimensional version of nodes).

Recently, Kollár [2013] and Kovács [2011a] introduced the notions of rational and Du Bois pairs (X, D) for a normal variety X and a reduced divisor $D \subseteq X$. These notions are philosophically distinct from the singularities considered typically in the minimal model program since (X, D) having rational (respectively Du

Kovács was supported in part by NSF grant DMS-0856185 and DMS-1301888 and the Craig McKibben and Sarah Merner Endowed Professorship in Mathematics at the University of Washington. Schwede was supported in part by NSF grant DMS-1064485, NSF FRG Grant DMS-1501115, NSF CAREER grant DMS-1501102 and a Sloan Fellowship.
MSC2010: primary 14J17; secondary 14E99, 14J10, 14D06, 14B05.
Keywords: Du Bois singularities, rational singularities, inversion of adjunction, vanishing theorems.

Bois) singularities does not generally imply that the ambient space X has rational (respectively Du Bois) singularities [Kollár 2013, Remark 2.81(2)] (respectively Examples 2.10, 2.14, and [Graf and Kovács 2014]). Instead the singularities of (X, D) measure the connection between the singularities of X and D (a notion obviously connected with problems related to inversion of adjunction). Furthermore, like Du Bois singularities, if (X, Z) is a Du Bois pair then the ideal sheaf of Z satisfies various Kodaira-type vanishing theorems, an observation which we hope will be useful in the future.

Even though a priori rational and Du Bois singularities are not part of the class one usually associates with the minimal model program, these singularities play important roles in both the minimal model program and moduli theory via the fact that (semi)log canonical singularities are Du Bois [Kovács et al. 2010; Kollár and Kovács 2010]. In addition, Du Bois singularities have played important roles in various other contexts recently. They are arguably the largest class of singularities for which we know that Kodaira vanishing holds [Patakfalvi 2015], they appear in proofs of extension and other vanishing theorems [Greb et al. 2011], positivity theorems [Schumacher 2012], categorical resolutions [Lunts 2012], log canonical compactifications [Hacon and Xu 2013] and many other results more directly related to the minimal model program.

It is now a basic tenet of the minimal model program that the right way to study singularities is via pairs; see [Kollár 1997; 2013]. This allows for more freedom in applications and makes inductive arguments easier. The same is true for rational and Du Bois singularities. The introduction of Du Bois pairs streamlined some existing proofs (see [Kollár 2013, Chapter 6]) and extended the realm of applications.

In this paper we extend several recent results on Du Bois singularities to the context of Du Bois pairs, notably the recent results on deformations of Du Bois singularities found in [Kovács and Schwede 2011a] and the requisite injectivity theorem, a result of Kollár and Kovács on the behavior of depth in Du Bois families, and the characterization of Cohen–Macaulay Du Bois singularities of [Kovács et al. 2010].

Furthermore, we prove a new inversion of adjunction statement for rational and Du Bois pairs. This statement is new even in the nonpair setting. Roughly speaking, in the nonpair setting it says that if $f : X \to B$ is a family over a smooth base such that the general fiber has rational singularities and the special fiber has Du Bois singularities, then X has rational singularities in a neighborhood of the special fiber. See Theorem E below for the general statement.

We state each of these theorems below. We begin with the deformation statement.

Theorem A (Theorem 4.2). *Let X be a reduced scheme essentially of finite type over $\mathbb{C}$, $Z \subseteq X$ a reduced subscheme and H a reduced effective Cartier divisor on X that does not contain any component of Z. If $(H, Z \cap H)$ is a Du Bois pair, then (X, Z) is a Du Bois pair near H.*

Just as in the nonpair setting, to prove this we first show an injectivity theorem.

Theorem B (Theorem 3.2). *Let X be a reduced scheme over $\mathbb{C}$ and $Z \subseteq X$ a reduced subscheme. Then the natural map*

$$\Phi^j : \mathcal{E}xt^j_{\mathscr{O}_X}(\underline{\Omega}^0_{X,Z}, \omega^\bullet_X) \hookrightarrow \mathcal{E}xt^j_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^\bullet_X)$$

is injective for every $j \in \mathbb{Z}$.

Here $\mathcal{E}xt^j_{\mathscr{O}_X}(\_, \omega^\bullet_X)$ is shorthand to denote $\boldsymbol{h}^j(\mathcal{R}\mathcal{H}om_{\mathscr{O}_X}(\_, \omega^\bullet_X))$.

We also generalize some of the results of [Kollár and Kovács 2010] for families to the context of Du Bois pairs.

Theorem C (Corollary 5.6). *Let $f : (X, Z) \to B$ be a flat projective family with $\mathscr{O}_Z, \mathscr{I}_Z$ flat over B as well. Assume that all the fiber pairs (X_b, Z_b) are Du Bois. Assume also that B is connected and the generic fibers $(\mathscr{I}_Z)_{\mathrm{gen}}$ are Cohen–Macaulay. Then all the fibers $(\mathscr{I}_Z)_b$ are Cohen–Macaulay.*

We have a multiplier ideal/module like characterization of Du Bois pairs.

Theorem D (Theorem 6.3). *Let X be a normal variety and $Z \subseteq X$ a divisor. Further, let $\pi : \widetilde{X} \to X$ be a log resolution of (X, Z) with $E = \pi^{-1}(Z)_{\mathrm{red}} \vee \mathrm{exc}(\pi)$. If $\mathscr{I}_Z$ is Cohen–Macaulay then (X, Z) is Du Bois if and only if*

$$\pi_*\omega_{\widetilde{X}}(E) \simeq \omega_X(Z).$$

All of the results above are used in the proof of our inversion of adjunction result.

Theorem E (Theorem 7.1). *Let $f : X \to B$ be a flat projective geometrically integral family over a smooth connected base B with $\dim B \geq 1$, $H = f^{-1}(0)$ the special fiber, and D a reduced codimension-1 subscheme of X which is flat over B. Assume that $(H, D|_H)$ is a Du Bois pair and that $(X \setminus H, D \setminus H)$ is a rational pair. Then (X, D) is a rational pair.*

This last result is *new* even in the case $D = 0$; see Corollary 7.8. In the special case when $X \setminus H$ is smooth and $D = 0$, Theorem E follows from [Schwede 2007, Theorem 5.1].

Statements similar to Theorem E have been proved in many related situations. For instance, assume $D = 0$, $X \setminus H$ is canonical and H is semi–log canonical. Then it follows from inversion of adjunction that X has canonical singularities, see [Kollár and Shepherd-Barron 1988, Theorem 5.1; Karu 2000, Theorem 2.5; Kawakita 2007]. A nonexhaustive list of some other related results includes [Fedder and Watanabe 1989, Proposition 2.13; Schwede 2009; Hacon 2014; Erickson 2014].

2. Definitions and basic properties

2A. *Rational pairs.* First we recall the notion of rational pairs defined by Kollár and Kovács, as described in [Kollár 2013, Chapter 2]. Note that a similar notion was defined by Schwede and Takagi [2008]. The two notions are closely related, but different. Their relationship is similar to how dlt singularities compare to klt singularities. Here we will discuss the former notion which in the dlt versus klt analogy corresponds to dlt.

In this subsection we work over an algebraically closed field k, although in the rest of the paper we restrict to working over the complex numbers.

Definition 2.1. Let X be a normal variety and $D \subseteq X$ an integral Weil divisor on X. A *log resolution* $(Y, D_Y) \xrightarrow{\pi} (X, D)$ is a resolution of singularities such that D_Y is the strict transform of D, and such that $(D_Y)_{\mathrm{red}} \cup \operatorname{exc}(\pi)$ is a simple normal crossing divisor.

Definition 2.2. A *reduced pair* (X, D) consists of a normal variety X and a reduced divisor D on X. For the definition of an *snc pair*, the *strata* of an snc pair and other normal crossing conditions please refer to [Kollár 2013, Definition 1.7].

One frequently wants log resolutions that do not blow up unnecessary centers. One good way to achieve this is with a thrifty resolution.

Definition 2.3 (thrifty resolution [Kollár 2013, Definition 2.79]). Let (X, D) be a reduced pair. A *thrifty resolution* of (X, D) is a resolution $\pi : Y \to X$ such that:

(a) $D_Y = \pi_*^{-1} D$ is a simple normal crossing divisor.

(b) π is an isomorphism over the generic point of every stratum of the snc locus of (X, D) and π is an isomorphism at the generic point of every stratum of (Y, D_Y).

Item (b) can also be replaced by:

(b′) The exceptional set E of π does not contain any stratum of (Y, D_Y) and $\pi(E)$ does not contain any stratum of the simple normal crossing locus of (X, D).

We can now define rational pairs [Kollár 2013, Section 2.5].

Definition 2.4 (rational pairs). A reduced pair (X, D) is called a *rational pair* if there exists a thrifty resolution $\pi : (Y, D_Y) \to (X, D)$ such that:

(i) $\mathscr{O}_X(-D) \simeq \pi_*\mathscr{O}_Y(-D_Y)$.

(ii) $\mathcal{R}^i\pi_*\mathscr{O}_Y(-D_Y) = 0$ for all $i > 0$.

(iii) $\mathcal{R}^i\pi_*\omega_Y(D_Y) = 0$ for all $i > 0$.

If (X, D) is a rational pair, and is in characteristic zero, then every thrifty resolution satisfies the properties (i), (ii), (iii) above [Kollár 2013, Corollary 2.86]. Even better though, property (iii) always holds in characteristic zero, as we point out below, whether or not (X, D) is a rational pair.

Theorem 2.5. *Assume that* $\operatorname{char} k = 0$, *and let* (X, D) *be a reduced pair and* $\pi : (Y, D_Y) \to (X, D)$ *a thrifty resolution. Then* $\mathcal{R}^i \pi_* \omega_Y(D_Y) = 0$ *for all* $i > 0$.

Proof. This follows from [Kollár 2013, Theorem 10.39]. □

Alternatively, one can prove Theorem 2.5 directly:

Claim 2.6. *Let* $\pi : E \to D$ *be a proper birational map between reduced equidimensional* $\mathbb{C}$*-schemes of finite type such that* E *is a simple normal crossing divisor in some smooth ambient space. Assume that* π *is birational onto its image when restricted to every strata of* E *(in particular, also each irreducible component of* E*). Then* $\mathcal{R}^i \pi_* \omega_E = 0$ *for* $i > 0$.

Proof. We proceed by induction on $\dim D$ and the number of irreducible components of E, and note that the base case is simply Grauert–Riemenschneider vanishing [Grauert and Riemenschneider 1970]. Write $E = E_0 \cup E'$, where E_0 is an irreducible component of E and E' denotes the remaining irreducible components. We have a short exact sequence

$$0 \to \mathcal{O}_E \to \mathcal{O}_{E_0} \oplus \mathcal{O}_{E'} \to \mathcal{O}_{E_0 \cap E'} \to 0.$$

Dualizing we obtain

$$0 \to \omega_{E_0} \oplus \omega_{E'} \to \omega_E \to \omega_{E_0 \cap E'} \to 0.$$

The intersection $E_0 \cap E'$ is a simple normal crossing divisor in the smooth ambient space E_0. It is also a union of strata of E and hence π is still birational when restricted to each strata of $E' \cap E_0$. Applying $\mathcal{R}^i \pi_*$ and the inductive hypothesis to E_0, E' and $E_0 \cap E'$ proves the claim. □

Alternative proof of Theorem 2.5. Push forward the short exact sequence

$$0 \to \omega_Y \to \omega_Y(D_Y) \to \omega_{D_Y} \to 0$$

via π and apply the claim to ω_{D_Y} and ω_Y. Note that the thrifty resolution hypothesis guarantees that π is birational when restricted to any strata of D_Y (since it is an isomorphism at the generic point of each strata). □

This gives us the following criterion.

Proposition 2.7. *Let* (X, D) *be a reduced pair and* $\pi : (Y, D_Y) \to (X, D)$ *a thrifty resolution. Then* (X, D) *is a rational pair if and only if*

$$\mathcal{R}\mathcal{H}om^{\bullet}_{\mathcal{O}_X}(\mathcal{O}_X(-D), \omega^{\bullet}_X) \simeq \mathcal{R}\pi_* \omega_Y(D_Y)[\dim X] \simeq \pi_* \omega_Y(D_Y)[\dim X]$$

for some thrifty resolution. Furthermore, in characteristic zero the second isomorphism is automatic.

Proof. Observe that the conditions (i) and (ii) of Definition 2.4 are equivalent to the isomorphism $\mathcal{R}\pi_*\mathscr{O}_Y(-D_Y) \simeq \mathscr{O}_X(-D)$. Applying Grothendieck duality and condition (iii) to this isomorphism yields the statement. The characteristic zero statement is simply Theorem 2.5. □

2B. ***Notation.*** Throughout the rest of this paper, all schemes will be assumed to be Noetherian separated schemes and essentially[1] of finite type over $\mathbb{C}$. Given divisors $D = \sum a_i D_i$ and $D' = \sum b_i D_i$ on a normal variety (possibly allowing a_i, b_j to be zero), we define

$$D \vee D' = \sum \max(a_i, b_i) D_i \quad \text{and} \quad D \wedge D' = \sum \min(a_i, b_i) D_i.$$

Of course, if D and D' have no common components then $D \vee D' = D + D'$ and $D \wedge D' = 0$. On a scheme X essentially of finite type over $\mathbb{C}$, we use $\boldsymbol{D}(\_) = \mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\_, \omega_X^{\bullet})$ to denote the Grothendieck duality functor.

2C. ***Du Bois pairs.*** The notion of Du Bois singularities is becoming more and more part of basic knowledge in higher-dimensional geometry. In particular, for the notion of the Deligne–Du Bois complex of a scheme of finite type over $\mathbb{C}$ and its degree zero associated graded complex, denoted by $\underline{\Omega}^0_X$, we refer the reader to [Kollár 2013, Section 6.1].

In contrast, the notion of Du Bois pairs is relatively new and so here we discuss some of its basic properties.

Given a (possibly nonreduced) subscheme $Z \subseteq X$ one has an induced map in $D^b_{\mathrm{coh}}(X)$,

$$\underline{\Omega}^0_X \to \underline{\Omega}^0_Z,$$

noting that by definition $\underline{\Omega}^0_Z = \underline{\Omega}^0_{Z_{\mathrm{red}}}$. Then $\underline{\Omega}^0_{X,Z}$ is defined to be the object in the derived category making the following an exact triangle:

$$\underline{\Omega}^0_{X,Z} \to \underline{\Omega}^0_X \to \underline{\Omega}^0_Z \xrightarrow{+1} . \tag{2.7.1}$$

If $\mathscr{I}_Z$ is the ideal sheaf of Z, then it is easy to see that there is a natural map $\mathscr{I}_Z \to \underline{\Omega}^0_{X,Z}$ [Kovács 2011a, Section 3.D].

Definition 2.8 [Kovács 2011a, Definition 3.13]. We say that (X, Z) is a *Du Bois pair* (or simply a *DB pair*) if the above map $\mathscr{I}_Z \to \underline{\Omega}^0_{X,Z}$ is a quasi-isomorphism.

In the original definition of a Du Bois pair in [Kovács 2011a] it was assumed that Z is reduced. As it turns out, this is not a necessary hypothesis.

[1] that is, a localization of a finite type scheme

Lemma 2.9. *If (X, Z) is a Du Bois pair and X is reduced, then Z is reduced.*

Proof. Note that $\underline{\Omega}^0_Z = \underline{\Omega}^0_{Z_{\mathrm{red}}}$ and so $\underline{\Omega}^0_{X,Z} \simeq \underline{\Omega}^0_{X,Z_{\mathrm{red}}}$. On the other hand, we also have an exact sequence,

$$\begin{tikzcd}
0 \arrow[r] & \boldsymbol{h}^0(\underline{\Omega}^0_{X,Z}) \arrow[r] \arrow[d, dotted, "\simeq"'] & \boldsymbol{h}^0(\underline{\Omega}^0_X) \arrow[r] \arrow[d, "\simeq"'] & \boldsymbol{h}^0(\underline{\Omega}^0_Z) \arrow[d, "\simeq"'] \\
0 \arrow[r] & \mathscr{I}_{\mathrm{Im}(Z^{\mathrm{SN}}_{\mathrm{red}}) \subseteq X^{\mathrm{SN}}} \arrow[r] & \mathscr{O}_{X^{\mathrm{SN}}} \arrow[r] & \mathscr{O}_{Z^{\mathrm{SN}}_{\mathrm{red}}}
\end{tikzcd}$$

where X^{SN}, $Z^{\mathrm{SN}}_{\mathrm{red}}$ are the seminormalizations of X and Z_{red} respectively, and the right two isomorphisms come from [Saito 2000]. Note that the scheme-theoretic image of $Z^{\mathrm{SN}}_{\mathrm{red}}$ in X^{SN} is reduced. The fact that the left-most vertical map is an isomorphism implies that $\boldsymbol{h}^0(\underline{\Omega}^0_{X,Z})$ is a radical ideal in $\mathscr{O}_{X^{\mathrm{SN}}}$. Since (X, Z) is Du Bois, we see that $\boldsymbol{h}^0(\underline{\Omega}^0_{X,Z}) = \mathscr{I}_{Z\subseteq X}$ and hence $\mathscr{I}_{Z\subseteq X}$ is radical in $\mathscr{O}_{X^{\mathrm{SN}}}$ and hence also in $\mathscr{O}_X = \mathscr{O}_{X^{\mathrm{red}}}$ as desired. □

Frequently we will take the Grothendieck dual of $\underline{\Omega}^0_{X,Z}$. Hence, following the notation of [Kovács and Schwede 2011a], we will write

$$\underline{\omega}^{\bullet}_{X,Z} := \mathscr{R}\mathscr{H}om^{\bullet}_{\mathscr{O}_X}(\underline{\Omega}^0_{X,Z}, \omega^{\bullet}_X). \tag{2.9.1}$$

The reader is referred to [Kollár 2013, Section 6.1] for basic properties of Du Bois pairs. As mentioned in the introduction, this notion of pairs is somewhat different in flavor from the definition of (X, Z) being log canonical or log terminal. Being a Du Bois pair is more a statement about the relationship between X and Z, not an absolute statement about the singularities of X or Z separately. In particular, Examples 2.10 and 2.14 show that (X, Z) being Du Bois *does not* imply that X is Du Bois.

Example 2.10 (Du Bois pair whose ambient space is not Du Bois). Let R denote the pullback of the diagram

$$\begin{tikzcd}
k[x] \simeq k[x,y]/\langle y \rangle & k[x,y] \arrow[l] \\
k[x^2, x^3] \arrow[u] & R \arrow[l, dotted] \arrow[u, dotted]
\end{tikzcd}$$

where the nondotted arrows are induced in the obvious ways. It is easy to see that $R = k[x^2, x^3, y, yx]$. By construction $X = \operatorname{Spec} R$ is not Du Bois since it is not seminormal. However, we claim that the pair $(\operatorname{Spec} R, V(\langle y, yx \rangle_R))$ is Du Bois.

Consider the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & \langle y, yx\rangle_R \arrow[r] \arrow[d, "\alpha"] & R \arrow[r] \arrow[d, "\beta"] & k[x^2, x^3] \arrow[r] \arrow[d, "\gamma"] & 0 \\
0 \arrow[r] & \langle y\rangle_{k[x,y]} \arrow[r] & k[x, y] \arrow[r] & k[x] \arrow[r] & 0
\end{tikzcd}$$

The maps labeled β and γ are the seminormalizations but α is an isomorphism. On the other hand, we know that $\underline{\Omega}^0_X = \underline{\Omega}^0_{X^{\mathrm{sn}}}$ in general since they have the same hyperresolution. Therefore, up to harmless identification of modules with sheaves on an affine scheme, we see $\underline{\Omega}^0_{\operatorname{Spec} R} \simeq k[x, y]$ and $\underline{\Omega}^0_{\operatorname{Spec} k[x^2,x^3]} \simeq k[x]$ and so

$$\langle y, yx\rangle_R = \langle y\rangle_{k[x,y]} \simeq \underline{\Omega}^0_{\operatorname{Spec} R, V(\langle y, yx\rangle_R)}.$$

This proves that $(\operatorname{Spec} R, V(\langle y, yx\rangle_R))$ is Du Bois and completes the example.

Next we will give an example of a normal Du Bois pair whose ambient space is not Du Bois. To this end we will use a criterion for a cone being a Du Bois pair. In order to do that we need to recall a definition [Kollár 2013, III.3.8].

Let X be a projective scheme and $\mathscr{L}$ an ample line bundle on X. We will need the *spectrum of the section ring of* $\mathscr{L}$,

$$C_a(X, \mathscr{L}) := \operatorname{Spec}_k \bigoplus_{p\geq 0} H^0(X, \mathscr{L}^p),$$

which is also called the (generalized) *ample cone* over X with conormal bundle $\mathscr{L}$. If no confusion is likely, in particular when $\mathscr{L}$ is fixed, we will use the shorthand of $CX := C_a(X, \mathscr{L})$. Notice that for a subscheme $Z \subseteq X$ there is a natural map $\iota\colon C_a(Z, \mathscr{L}|_Z) \to C_a(X, \mathscr{L})$ which is a closed embedding away from the vertex $P \in CX$. By a slight abuse of notation we will also use (CX, CZ) to denote the pair $\big(C_a(X, \mathscr{L}), \iota(C_a(Z, \mathscr{L}|_Z))\big)$.

Now we are ready to state the needed Du Bois criterion.

Proposition 2.11 ([Graf and Kovács 2014], cf. [Ma 2015]). *Let X be a smooth projective variety, $Z \subset X$ an snc divisor (possibly the empty set), and $\mathscr{L}$ an ample line bundle on X. Then (CX, CZ) is a Du Bois pair if and only if*

$$H^i(X, \mathscr{L}^m(-Z)) = 0$$

for all $i, m > 0$.

Proof. If $Z = \varnothing$, this follows from [Ma 2015, Theorem 4.4]. The general case works similarly. For a direct proof see [Graf and Kovács 2014, Theorem 2.5]. □

While the above is sufficient for our purposes, we also obtained independently a slightly different statement using similar methods.

Lemma 2.12 (Du Bois pairs for graded rings). *Let X be a projective variety with Du Bois singularities, $\mathscr{L}$ an ample line bundle and D a reduced connected divisor on X. Assume that D also has only Du Bois singularities. Form the corresponding section ring $S = \bigoplus_{i \geq 0} \Gamma(X, \mathscr{L}^i)$ and $I = \bigoplus_{i \geq 0} \Gamma(X, \mathscr{O}_X(-D) \otimes \mathscr{L}^i)$. Fix $\mathfrak{m} = S_+$ to be the irrelevant ideal. Set $Y = CX = \operatorname{Spec} S$ and $Z = CD = \operatorname{Spec}(S/I)$. If*

$$H^1(X, \mathscr{O}_X(-D) \otimes \mathscr{L}^i) = 0 \tag{2.12.1}$$

for $i \geq 0$ so that $S/I \simeq \bigoplus_{i \geq 0} \Gamma(D, \mathscr{L}^i|_D)$, then for all $i \geq 1$ we have

$$\boldsymbol{h}^i(\underline{\Omega}^0_{Y,Z}) \simeq [H^{i+1}_{\mathfrak{m}}(I)]_{>0}.$$

Again under hypothesis (2.12.1), we see immediately that (Y, Z) is Du Bois if and only if $[H^i_{\mathfrak{m}}(I)]_{>0} = 0$ for every $i > 0$.

Proof. First observe that both Y and Z are seminormal since they are saturated section rings over seminormal schemes. L. Ma [2015, Equation (4.4.4) in the proof of Theorem 4.4] showed that

$$\boldsymbol{h}^i(\underline{\Omega}^0_Y) \simeq [H^{i+1}_{\mathfrak{m}}(S)]_{>0} \tag{2.12.2}$$

for $i > 0$. Likewise $\boldsymbol{h}^i(\underline{\Omega}^0_Z) = [H^{i+1}_{\mathfrak{m}}(S/I)]_{>0}$ for $i > 0$. Now we analyze $\boldsymbol{h}^{i+j}(\mathcal{R}\Gamma_{\mathfrak{m}}(\underline{\Omega}^0_Y))$ via a spectral sequence. Since Y is Du Bois outside of the origin $V(\mathfrak{m})$, we see that $\boldsymbol{h}^j(\underline{\Omega}^0_Y)$ is supported only at the origin for $j > 0$. It follows that the E_2-page of the spectral sequence

$$H^i_{\mathfrak{m}}(\boldsymbol{h}^j(\underline{\Omega}^0_Y)) \Rightarrow \boldsymbol{h}^{i+j}(\mathcal{R}\Gamma_{\mathfrak{m}}(\underline{\Omega}^0_Y))$$

looks like

$$\begin{tikzcd}
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\boldsymbol{h}^3(\underline{\Omega}^0_Y) & 0 & 0 & 0 & 0 & \cdots \\
\boldsymbol{h}^2(\underline{\Omega}^0_Y) & 0 & 0 & 0 & 0 & \cdots \\
\boldsymbol{h}^1(\underline{\Omega}^0_Y) \arrow[drr] & 0 & 0 & 0 & 0 & \cdots \\
0 & H^1_{\mathfrak{m}}(S) & H^2_{\mathfrak{m}}(S) & H^3_{\mathfrak{m}}(S) & H^4_{\mathfrak{m}}(S) & \cdots
\end{tikzcd}$$

Here we are using the fact that S is seminormal, and so $\boldsymbol{h}^0(\underline{\Omega}^0_Y) = S$. It is not difficult to see that the unique nonzero map of the $(i-1)$-st page of this spectral sequence induces the isomorphism of (2.12.2), and so those unique nonzero maps are injective. Thus the spectral sequence contains the data of a long exact sequence

$$0 \to H^1_{\mathfrak{m}}(S) \twoheadrightarrow \mathbb{H}^1_{\mathfrak{m}}(\underline{\Omega}^0_Y) \to \boldsymbol{h}^1(\underline{\Omega}^0_Y) \hookrightarrow H^2_{\mathfrak{m}}(S) \twoheadrightarrow \mathbb{H}^2_{\mathfrak{m}}(\underline{\Omega}^0_Y) \to \boldsymbol{h}^2(\underline{\Omega}^0_Y) \hookrightarrow H^3_{\mathfrak{m}}(S) \twoheadrightarrow \cdots .$$

Hence $\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_Y) = [H^i_{\mathfrak{m}}(S)]_{\leq 0}$ for $i \geq 2$ and $\mathbb{H}^1_{\mathfrak{m}}(\underline{\Omega}^0_Y) \simeq H^1_{\mathfrak{m}}(S)$. Likewise $\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_Z) = [H^i_{\mathfrak{m}}(S/I)]_{\leq 0}$ for $i \geq 2$ and $\mathbb{H}^1_{\mathfrak{m}}(\underline{\Omega}^0_Z) \simeq H^1_{\mathfrak{m}}(S/I)$. Furthermore, since Y and Z are seminormal we see that $\boldsymbol{h}^0(\underline{\Omega}^0_{X,Z}) = I$ and so the same spectral sequence argument implies that we have a long exact sequence

$$0 \to H^1_{\mathfrak{m}}(I) \twoheadrightarrow \mathbb{H}^1_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z}) \to \boldsymbol{h}^1(\underline{\Omega}^0_{Y,Z}) \to H^2_{\mathfrak{m}}(I) \twoheadrightarrow \mathbb{H}^2_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z}) \to \boldsymbol{h}^2(\underline{\Omega}^0_{Y,Z}) \to H^3_{\mathfrak{m}}(I) \twoheadrightarrow \cdots .$$

We still have the labeled surjectivities by the Matlis dual of Theorem 3.2, which we will prove later (we assume it for now). Thus it is enough to see that the maps above make the identification $\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z}) = [H^i_{\mathfrak{m}}(I)]_{\leq 0}$ for $i \geq 2$.

We consider the diagram with distinguished triangles as rows

$$\begin{tikzcd}
I \arrow[r] \arrow[d] & S \arrow[r] \arrow[d] & S/I \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{Y,Z} \arrow[r] & \underline{\Omega}^0_Y \arrow[r] & \underline{\Omega}^0_Z \arrow[r, "+1"] & {}
\end{tikzcd}$$

We will apply the functor $\mathcal{R}\Gamma_{\mathfrak{m}}(\_)$ and take cohomology $i \geq 1$ to obtain

$$\begin{tikzcd}
H^i_{\mathfrak{m}}(S) \arrow[r] \arrow[d, "\alpha"'] & H^i_{\mathfrak{m}}(S/I) \arrow[r] \arrow[d, "\beta"'] & H^{i+1}_{\mathfrak{m}}(I) \arrow[r] \arrow[d, "\gamma"'] & H^{i+1}_{\mathfrak{m}}(S) \arrow[r] \arrow[d, "\delta"'] & H^{i+1}_{\mathfrak{m}}(S/I) \arrow[d, "\epsilon"'] \\
\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_Y) \arrow[r] \arrow[d, equal] & \mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_Z) \arrow[r] \arrow[d, equal] & \mathbb{H}^{i+1}_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z}) \arrow[r] & \mathbb{H}^{i+1}_{\mathfrak{m}}(\underline{\Omega}^0_Y) \arrow[r] \arrow[d, equal] & \mathbb{H}^{i+1}_{\mathfrak{m}}(\underline{\Omega}^0_Z) \arrow[d, equal] \\
{[H^i_{\mathfrak{m}}(S)]_{\leq 0}} & {[H^i_{\mathfrak{m}}(S/I)]_{\leq 0}} & & {[H^{i+1}_{\mathfrak{m}}(S)]_{\leq 0}} & {[H^{i+1}_{\mathfrak{m}}(S/I)]_{\leq 0}}
\end{tikzcd}$$

Note that γ is the map we already identified as surjective above. It is easy to see that the vertical maps α, β, δ, and ϵ are the projections and so $[\alpha]_{\leq 0}$, $[\beta]_{\leq 0}$, $[\delta]_{\leq 0}$, and $[\epsilon]_{\leq 0}$ are isomorphisms. Thus $[\gamma]_{\leq 0}$ is also an isomorphism. But from the second row we see that $\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z})$ is of nonpositive degree so that $\mathbb{H}^i_{\mathfrak{m}}(\underline{\Omega}^0_{Y,Z}) = [H^i_{\mathfrak{m}}(I)]_{\leq 0}$ for $i \geq 2$. □

Remark 2.13. It would be natural to try to prove a common generalization of the (independently obtained) Proposition 2.11 and Lemma 2.12.

Example 2.14 (a normal Du Bois pair whose ambient space is not Du Bois). Let W be an arbitrary smooth canonically polarized variety, that is, W is smooth and projective and ω_W is ample. Further let $n > 1$, and set $X = W \times \mathbb{P}^n$ and $\mathscr{L} = \pi_1^*\omega_W \otimes \pi_2^*\mathscr{O}_{\mathbb{P}^n}(1)$. Finally, let $K \subseteq \mathbb{P}^n$ be a smooth hypersurface of degree

$n+1$, that is, $\mathscr{O}_{\mathbb{P}^n}(K) \simeq \omega_{\mathbb{P}^n}^{-1}$, and let $Z = W \times K$. We claim that, using the above notation, (CX, CZ) is a Du Bois pair, while CX itself is not. Note also that by construction CX is normal.

Consider $H^1(X, \mathscr{O}_X(-Z) \otimes \mathscr{L}^j)$ for $j \geq 0$ and observe that

$$\mathscr{O}_X(-Z) \otimes \mathscr{L}^j = \pi_2^* \mathscr{O}_{\mathbb{P}^n}(-n-1) \otimes \pi_1^* \omega_W^j \otimes \pi_2^* \mathscr{O}_{\mathbb{P}^n}(j) = \pi_1^* \omega_W^j \otimes \pi_2^* \mathscr{O}_{\mathbb{P}^n}(j-n-1).$$

Now $H^1(\mathbb{P}^n, \mathscr{O}_{\mathbb{P}^n}(j-n-1)) = 0$ for all $j \geq 0$ and $H^0(\mathbb{P}^n, \mathscr{O}_{\mathbb{P}^n}(j-n-1)) = 0$ for $j \leq n$. But if $j > n \geq 1$, then $H^1(W, \omega_W^j) = 0$ by Kodaira vanishing and so it follows by the Künneth formula that $H^1(X, \mathscr{O}_X(-Z) \otimes \mathscr{L}^j) = 0$ for all $j \geq 0$, so the hypotheses of Lemma 2.12 are satisfied.

Let $r = \dim W$ and consider $H^r(X, \mathscr{L})$. By the Künneth formula

$$H^r(X, \mathscr{L}) \supseteq H^r(W, \omega_W) \otimes H^0(\mathbb{P}^n, \mathscr{O}_{\mathbb{P}^n}(1)) \neq 0,$$

and hence by Proposition 2.11 CX is *not* Du Bois.

On the other hand we have that

$$\mathscr{L}(-Z) \simeq \pi_1^* \omega_W \otimes \pi_2^* \mathscr{O}_{\mathbb{P}^n}(1-n-1) \simeq \omega_X \otimes \pi_2^* \mathscr{O}_{\mathbb{P}^n}(1). \tag{2.14.1}$$

Now observe that $H^q(\mathbb{P}^n, \mathscr{O}_{\mathbb{P}^n}(1-n-1)) = 0$ for *all* $q \geq 0$, so, again by the Künneth formula, it follows that $H^i(X, \mathscr{L}(-Z)) = 0$ for all $i > 0$.

In order to conclude that (CX, CZ) is a Du Bois pair we need that

$$H^i(X, \mathscr{L}^m(-Z)) = 0 \quad \text{for all } i, m > 0.$$

We just showed that $H^i(X, \mathscr{L}(-Z)) = 0$ for all $i > 0$, which handles the $m = 1$ case. If $m > 1$ then $\mathscr{M} := \mathscr{L}^{m-1} \otimes \pi_2^* \mathscr{O}_{\mathbb{P}^n}(1)$ is ample on X and by (2.14.1) and Kodaira vanishing we have that

$$H^i(X, \mathscr{L}^m(-Z)) = H^i(X, \mathscr{L}(-Z) \otimes \mathscr{L}^{m-1}) \simeq H^i(X, \omega_X \otimes \mathscr{M}) = 0,$$

and hence it follows from Proposition 2.11 that (CX, CZ) is indeed a Du Bois pair.

We will find the following lemma useful; cf. [Esnault 1990; Schwede 2007].

Lemma 2.15. *Assume that (X, Z) is a pair with $Z \subseteq X$ reduced schemes. Assume further that $X \subseteq Y$ where Y is smooth. Let $\pi : \widetilde{Y} \to Y$ be a log resolution of both X and Z in Y and set $\overline{X}$ and $\overline{Z}$ to be the reduced preimages of X and Z in $\widetilde{Y}$ respectively. Then*

$$\underline{\Omega}^0_{X,Z} \simeq \mathcal{R}\pi_* \mathscr{I}_{\overline{Z} \subseteq \overline{X}},$$

where $\mathscr{I}_{\overline{Z} \subseteq \overline{X}}$ is the ideal of $\overline{Z}$ in $\overline{X}$.

Proof. Consider the diagram

$$\begin{tikzcd}
\underline{\Omega}^0_{X,Z} \arrow[r] \arrow[d, "\alpha"] & \underline{\Omega}^0_X \arrow[r] \arrow[d, "\beta"] & \underline{\Omega}^0_Z \arrow[r, "+1"] \arrow[d, "\gamma"] & {} \\
\mathcal{R}\pi_*\underline{\Omega}^0_{\overline{X},\overline{Z}} \arrow[r] \arrow[d, equal] & \mathcal{R}\pi_*\underline{\Omega}^0_{\overline{X}} \arrow[r] \arrow[d, equal] & \mathcal{R}\pi_*\underline{\Omega}^0_{\overline{Z}} \arrow[r, "+1"] \arrow[d, equal] & {} \\
\mathcal{R}\pi_*\mathscr{I}_{\overline{Z}\subseteq\overline{X}} \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_{\overline{X}} \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_{\overline{Z}} \arrow[r, "+1"] & {}
\end{tikzcd}$$

The vertical arrows β and γ are quasi-isomorphisms by [Kovács and Schwede 2011b, Theorem 6.4] (also see [Schwede 2007, Theorem 4.3]) since $\overline{X}$ and $\overline{Z}$ are snc and hence Du Bois. The second row of equalities also follows since $\overline{X}$ and $\overline{Z}$ are Du Bois. Then α is a quasi-isomorphism as well and hence the lemma follows. □

There are some situations when a pair being Du Bois implies that the ambient space is also Du Bois. It is proved in [Graf and Kovács 2014] that this happens if X is Gorenstein, but that X being $\mathbb{Q}$-Gorenstein is not sufficient. Another simple situation in which this holds is the following.

Lemma 2.16. *Let X be a reduced $\mathbb{C}$-scheme essentially of finite type and H a Cartier divisor. If (X, H) is a Du Bois pair then X (and hence H) is also Du Bois.*

Proof. The statement is local and so we may assume that $X = \operatorname{Spec} R$ is affine. We know that $\mathscr{O}_X(-H) \to \underline{\Omega}^0_{X,H}$ is a quasi-isomorphism and thus so is $\mathscr{O}_X \to \underline{\Omega}^0_{X,H} \otimes \mathscr{O}_X(H)$. We will show that this map factors through $\mathscr{O}_X \to \underline{\Omega}^0_X$, which will complete the proof by [Kovács 1999, Theorem 2.3].

Embed $X \subseteq Y$ as a smooth scheme and let $\pi : \widetilde{Y} \to Y$ be a simultaneous log resolution of (Y, X) and (Y, H) with $\overline{X}$, $\overline{H}$ the reduced total transforms of X and H respectively. Then $\underline{\Omega}^0_{X,H} = \mathcal{R}\pi_*\mathscr{I}_{\overline{H}\subseteq\overline{X}}$ by Lemma 2.15. Fix X' to be the components of $\overline{X}$ which are not also components of $\overline{H}$ and we see that $\mathscr{I}_{\overline{H}\subseteq\overline{X}} \simeq \mathscr{O}_{X'}(-\overline{H}|_{X'})$. Thus

$$\underline{\Omega}^0_{X,H} \otimes \mathscr{O}_X(H) \simeq \mathcal{R}\pi_*\mathscr{O}_{X'}((\pi^*H - \overline{H})|_{X'}).$$

Since $\pi^*H - \overline{H}$ is effective, we obtain a map

$$\underline{\Omega}^0_X \simeq \mathcal{R}\pi_*\mathscr{O}_{\overline{X}} \to \mathcal{R}\pi_*\mathscr{O}_{X'} \to \mathcal{R}\pi_*\mathscr{O}_{X'}((\pi^*H - \overline{H})|_{X'}) \simeq \underline{\Omega}^0_{X,H} \otimes \mathscr{O}_X(H).$$

This map obviously factors the quasi-isomorphism $\mathscr{O}_X \to \underline{\Omega}^0_{X,H} \otimes \mathscr{O}_X(H)$ and hence the proof is complete. □

We recall properties of $\underline{\Omega}^0_{X,Z}$ that we will need later.

Lemma 2.17. *Let X be a scheme over $\mathbb{C}$ with $Z \subseteq X$ a closed subscheme and $j : U = X \setminus Z \hookrightarrow X$ the complement of Z. Then:*

(a) *If in addition X is proper, then $H^i(X, \mathscr{I}_Z) \to \mathbb{H}^i(X, \underline{\Omega}^0_{X,Z})$ is surjective for all $i \in \mathbb{Z}$* [Kovács 2011a, Corollary 4.2; Kollár 2013, Theorem 6.22].

(b) *If H is a general member of a basepoint-free linear system, then $\underline{\Omega}^0_{X,Z} \otimes \mathscr{O}_H \simeq \underline{\Omega}^0_{H,H\cap Z}$* [Kovács 2011a, Proposition 3.18; Kollár 2013, Theorem 6.5(6)].

(c) *If $X = U \cup V$ is a decomposition into closed subschemes and $Z \subseteq X$ is another closed subscheme, then we have a distinguished triangle*

$$\underline{\Omega}^0_{U\cup V,Z} \to \underline{\Omega}^0_{U,Z\cap U} \oplus \underline{\Omega}^0_{V,Z\cap V} \to \underline{\Omega}^0_{U\cap V,Z\cap U\cap V} \xrightarrow{+1}$$

(cf. [Kollár 2013, Theorem 6.5(11)]).

(d) *Let $X = U \cup V$ be a decomposition of X into closed subschemes. Then*

$$\underline{\Omega}^0_{U\cup V,V} \simeq \underline{\Omega}^0_{U,U\cap V}$$

(cf. [Kovács 2011a, Proposition 3.19; Kollár 2013, Theorem 6.17]).

Proof. Parts (a) and (b) follow from the references in their statements. For (c), the included reference only states the triangle in the case that $Z = \varnothing$. However, our more general version follows easily from the diagram

$$\begin{tikzcd}
\underline{\Omega}^0_{U\cup V,Z} \arrow[r, dotted] \arrow[d] & \underline{\Omega}^0_{U,Z\cap U} \oplus \underline{\Omega}^0_{V,Z\cap V} \arrow[r, dotted] \arrow[d] & \underline{\Omega}^0_{U\cap V,Z\cap U\cap V} \arrow[r, dotted, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{U\cup V} \arrow[r] \arrow[d] & \underline{\Omega}^0_{U} \oplus \underline{\Omega}^0_{V} \arrow[r] \arrow[d] & \underline{\Omega}^0_{U\cap V} \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{Z} \arrow[r] \arrow[d, "+1"'] & \underline{\Omega}^0_{Z\cap U} \oplus \underline{\Omega}^0_{Z\cap V} \arrow[r] \arrow[d, "+1"'] & \underline{\Omega}^0_{Z\cap U\cap V} \arrow[r, "+1"] \arrow[d, "+1"'] & {} \\
{} & {} & {} &
\end{tikzcd}$$

and the 9-lemma in triangulated categories [Kovács 2013, B.1].

For (d), consider the distinguished triangle

$$\underline{\Omega}^0_{U\cup V} \longrightarrow \underline{\Omega}^0_{U} \oplus \underline{\Omega}^0_{V} \longrightarrow \underline{\Omega}^0_{U\cap V} \xrightarrow{+1}$$

of part (c) with $Z = \varnothing$. Then [Kollár and Kovács 2010, Lemma 2.1] implies that the left vertical arrow of the following diagram is an isomorphism:

$$\begin{tikzcd}
\underline{\Omega}^0_{U\cup V,V} \arrow[r] \arrow[d, "\simeq"] & \underline{\Omega}^0_{U\cup V} \arrow[r] \arrow[d] & \underline{\Omega}^0_{V} \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{U,U\cap V} \arrow[r] & \underline{\Omega}^0_{U} \arrow[r] & \underline{\Omega}^0_{U\cap V} \arrow[r, "+1"] & {}
\end{tikzcd}$$

For more details see the proofs in the references and replace $\underline{\Omega}^{\times}$ with $\underline{\Omega}^0$. $\square$

The next lemma constructs a natural exact triangle for Du Bois pairs.

Lemma 2.18. *Let X be a scheme and $W, Z \subseteq X$ subschemes. Then there is a distinguished triangle*

$$\underline{\Omega}^0_{X,W\cup Z} \to \underline{\Omega}^0_{X,Z} \to \underline{\Omega}^0_{W,Z\cap W} \xrightarrow{+1} .$$

In particular, because there is also a short exact sequence

$$0 \to \mathscr{I}_{W\cup Z\subseteq X} \to \mathscr{I}_{Z\subseteq X} \to \mathscr{I}_{Z\cap W\subseteq W} \to 0,$$

if any two of $\{(X, W\cup Z), (X, Z), (W, Z\cap W)\}$ are Du Bois, so is the third.

Proof. We begin with a diagram of distinguished triangles as columns and rows (see [Kollár 2013, Theorem 6.5.11; Kovács 2013, Theorem B1]):

$$\begin{tikzcd}
\underline{\Omega}^0_{X,W\cup Z} \arrow[r, dotted] \arrow[d] & \underline{\Omega}^0_{X,Z} \oplus \underline{\Omega}^0_{X,W} \arrow[r, dotted, "-"] \arrow[d] & \underline{\Omega}^0_{X,Z\cap W} \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{X} \arrow[r, "\mathrm{id}\oplus\mathrm{id}"] \arrow[d] & \underline{\Omega}^0_{X} \oplus \underline{\Omega}^0_{X} \arrow[r, "-"] \arrow[d] & \underline{\Omega}^0_{X} \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_{W\cup Z} \arrow[r] & \underline{\Omega}^0_{Z} \oplus \underline{\Omega}^0_{W} \arrow[r, "-"] & \underline{\Omega}^0_{Z\cap W} \arrow[r, "+1"] & {}
\end{tikzcd}$$

The horizontal maps in the second column of this diagram are each obtained by subtracting the canonical maps on each factor of the direct sum, hence the minus signs. The octahedral axiom implies that there exists a diagram of distinguished triangles,

$$\begin{tikzcd}
\underline{\Omega}^0_{X,W\cup Z} \arrow[r] \arrow[d] & \underline{\Omega}^0_{X,Z} \arrow[r] \arrow[d] & K^{\bullet} \arrow[r, "+1"] \arrow[d, "\sim"] & {} \\
\underline{\Omega}^0_{X,W} \arrow[r] & \underline{\Omega}^0_{X,Z\cap W} \arrow[r] & K^{\bullet} \arrow[r, "+1"] & {}
\end{tikzcd}$$

We need to identify $K^{\bullet}$. Notice that the bottom row also fits into another diagram of distinguished triangles (see [Kovács 2013, Theorem B1]):

$$\begin{tikzcd}
\underline{\Omega}^0_{X,W} \arrow[r] \arrow[d] & \underline{\Omega}^0_{X,Z\cap W} \arrow[r] \arrow[d] & K^\bullet \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_X \arrow[r, "\sim"] \arrow[d] & \underline{\Omega}^0_X \arrow[r] \arrow[d] & 0 \arrow[r, "+1"] \arrow[d] & {} \\
\underline{\Omega}^0_W \arrow[r] \arrow[d, "+1"'] & \underline{\Omega}^0_{Z\cap W} \arrow[r] \arrow[d, "+1"'] & \underline{\Omega}^0_{W,Z\cap W}[1] \arrow[r, "+1"] \arrow[d, "+1"'] & {} \\
{} & {} & {} &
\end{tikzcd}$$

Hence $K^\bullet \simeq \underline{\Omega}^0_{W,Z\cap W}$ and the lemma follows. $\square$

Finally, note that being Du Bois is a direct generalization of being rational for pairs (see also Kollár 2013, Corollary 6.25).

Theorem 2.19 [Kovács 2011a, Corollary 5.6]. *If* (X, D) *is a rational pair then* (X, D) *is also a Du Bois pair.*

3. An injectivity theorem

A key ingredient of the proof that Du Bois singularities are deformation invariant was an injectivity theorem [Kovács and Schwede 2011a, Theorem 3.3]. In this section, we generalize that result to the context of pairs.

Lemma 3.1 (cf. [Kovács and Schwede 2011a, Lemma 3.1]). *Let* X *be a reduced scheme*, $Z \subseteq X$ *a reduced subscheme and* $\mathscr{L}$ *a semiample line bundle. Let* $s \in \mathscr{L}^n$ *be a general global section for some* $n \gg 0$ *and take the* n*-th root of this section* (as in [Kollár and Mori 1998, Definition 2.50]):

$$\eta : Y = \mathbf{Spec} \bigoplus_{i=0}^{n-1} \mathscr{L}^{-i} \to X.$$

Set $W = \eta^{-1}(Z)$ (*with the induced scheme structure*). *Note that the restriction satisfies* $\eta|_W : W = \mathbf{Spec} \bigoplus_{i=0}^{n-1} \mathscr{L}^{-i}|_Z \to Z$. *Then as before*, *writing* $\eta_* = \mathcal{R}\eta_*$,

$$\eta_* \underline{\Omega}^0_{Y,W} \simeq \underline{\Omega}^0_{X,Z} \otimes \eta_* \mathscr{O}_Y \simeq \bigoplus_{i=0}^{n-1} (\underline{\Omega}^0_{X,Z} \otimes \mathscr{L}^{-i}),$$

and this direct sum is compatible with the decomposition $\eta_* \mathscr{O}_Y = \bigoplus_{i=0}^{n-1} \mathscr{L}^{-i}$.

Proof. Although not explicitly stated, it is easy to see that [Kovács and Schwede 2011a, Lemma 3.1] is functorial in that it is compatible with the map $Z \to X$. Then by applying Lemma 2.17(b), the result follows from the diagram

$$\begin{tikzcd}
\eta_*\underline{\Omega}^0_{Y,W} \arrow[r] & \eta_*\underline{\Omega}^0_Y \arrow[rr] & & \eta_*\underline{\Omega}^0_W \arrow[r, "+1"] & {} \\
\underline{\Omega}^0_{X,Z}\otimes\eta_*\mathscr{O}_Y \arrow[r] \arrow[u, dotted] & \underline{\Omega}^0_X\otimes\eta_*\mathscr{O}_Y \arrow[r] \arrow[u, "\simeq"'] & \underline{\Omega}^0_Z\otimes\eta_*\mathscr{O}_Y \simeq & \underline{\Omega}^0_Z\otimes_{\mathscr{O}_Z}\eta_*\mathscr{O}_Z \arrow[r, "+1"] \arrow[u, "\simeq"'] & {}
\end{tikzcd}$$
□

Setting $\underline{\omega}^\bullet_{X,Z}=\mathcal{R}\mathcal{H}om^\bullet_{\mathscr{O}_X}(\underline{\Omega}^0_{X,Z},\omega^\bullet_X)$ as in (2.9.1), we easily obtain the following.

Theorem 3.2. *Let X be a reduced scheme over $\mathbb{C}$ and $Z\subseteq X$ a reduced subscheme. Then the natural map*

$$\Phi^j: \boldsymbol{h}^j(\underline{\omega}^\bullet_{X,Z})\hookrightarrow \boldsymbol{h}^j(\mathcal{R}\mathcal{H}om_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^\bullet_X))$$

is injective for every $j\in\mathbb{Z}$.

Proof. The proof is essentially the same as in [Kovács and Schwede 2011a, Theorem 3.3] so we only sketch it briefly. First, since the question is local and compatible with restricting to an open subset, we may assume that X is projective with ample line bundle $\mathscr{L}$. It follows from taking a cyclic cover with respect to a general section of $\mathscr{L}^n$, for $n\gg 0$, and applying Lemmas 2.17(a) and 3.1 that

$$H^j\Bigl(X,\mathscr{I}_Z\otimes\bigoplus_{i=0}^{n-1}\mathscr{L}^{-i}\Bigr)\to\mathbb{H}^j\Bigl(X,\underline{\Omega}^0_{X,Z}\otimes\bigoplus_{i=0}^{n-1}\mathscr{L}^{-i}\Bigr)$$

surjects for all $j\geq 0$. Therefore $H^j(X,\mathscr{I}_Z\otimes\mathscr{L}^{-i})\to\mathbb{H}^j(X,\underline{\Omega}^0_{X,Z}\otimes\mathscr{L}^{-i})$ surjects for all $i,j\geq 0$.

By an application of Serre–Grothendieck duality we obtain an injection

$$\mathbb{H}^j(X,\underline{\omega}^\bullet_{X,Z}\otimes\mathscr{L}^i)\hookrightarrow\mathbb{H}^j(X,\mathcal{R}\mathcal{H}om^\bullet_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^\bullet_X)\otimes\mathscr{L}^i) \tag{3.2.1}$$

for all $i,j\geq 0$. But for $i\gg 0$, by Serre vanishing, we obtain that

$$H^0(X,\boldsymbol{h}^j(\underline{\omega}^\bullet_{X,Z})\otimes\mathscr{L}^i)\hookrightarrow H^0(X,\boldsymbol{h}^j(\mathcal{R}\mathcal{H}om^\bullet_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^\bullet_X))\otimes\mathscr{L}^i) \tag{3.2.2}$$

is injective as well (since the spectral sequence computing (3.2.1) degenerates). On the other hand, if $\boldsymbol{h}^j(\underline{\omega}^\bullet_{X,Z})\to\boldsymbol{h}^j(\mathcal{R}\mathcal{H}om^\bullet_{\mathscr{O}_X}(\mathscr{I},\omega^\bullet_X))$ is not injective, then for some $i\gg 0$ neither is (3.2.2). □

4. Deformation of Du Bois pairs

In [Kovács and Schwede 2011a, Corollary 4.2], we showed the following result: Let $f: X\to B$ be a flat proper family over a smooth curve B with a fiber X_0, $0\in B$, having Du Bois singularities. Then there is an open neighborhood $0\in U\subseteq B$ such that the fibers X_u have Du Bois singularities for $u\in U$. In this section, we generalize this result to Du Bois pairs. We mimic our previous approach as much as possible.

First we need a lemma, which is presumably well known but for which we know no reference.

Lemma 4.1. *Let X be a reduced scheme and $Z \subseteq X$ a reduced subscheme with ideal sheaf $\mathscr{I}_Z$. Further, let $H \subseteq X$ be an effective Cartier divisor with ideal sheaf $\mathscr{I}_H$ such that H does not contain any irreducible components of either X or Z. Then*

$$\mathscr{I}_H \cap \mathscr{I}_Z = \mathscr{I}_H \cdot \mathscr{I}_Z.$$

Proof. This is left as an exercise to the reader. Earlier versions of this paper, which are available on the arXiv, also contain a detailed proof. □

Now we prove that if a special fiber supports a Du Bois pair, so does the total space near that fiber. Recall that effective Cartier divisors on a possibly nonnormal scheme are simply subschemes locally defined by a single non-zero-divisor near every point.

Theorem 4.2. *Let X be a reduced scheme essentially of finite type over $\mathbb{C}$, $Z \subseteq X$ a reduced subscheme and H a reduced effective Cartier divisor on X that does not contain any component of Z. If $(H, Z \cap H)$ is a Du Bois pair, then (X, Z) is a Du Bois pair near H. It then follows* (from Lemma 2.18) *that $(X, Z \cup H)$ is Du Bois near H.*

Proof. We follow very closely the proofs of [Kovács 2000, Theorem 3.2] and [Kovács and Schwede 2011a, Theorem 4.1], which are based on [Elkik 1978]. Choose a closed point $\mathfrak{q}$ of X contained within H. It is sufficient to prove that (X, Z) is Du Bois at $\mathfrak{q}$. Let R denote the stalk $\mathscr{O}_{X,\mathfrak{q}}$ and replace X by $\operatorname{Spec} R$. Choose $f \in R$ to denote a defining equation of H in R. Consider the following diagram, whose rows are distinguished triangles in $D^b_{\mathrm{coh}}(X)$:

$$\begin{tikzcd}
I_Z \arrow[r, "\times f"] \arrow[d] & I_Z \arrow[r] \arrow[d] & I_Z/(f \cdot I_Z) \arrow[r, "+1"] \arrow[d, "\rho"] \arrow[dd, bend left, "\simeq"] & {} \\
\underline{\Omega}^0_{X,Z} \arrow[r, "\times f"'] & \underline{\Omega}^0_{X,Z} \arrow[r] & A^\bullet \arrow[r, "+1"] \arrow[d, dotted, "\tau"] & {} \\
 & & \underline{\Omega}^0_{H,Z\cap H} &
\end{tikzcd} \tag{4.2.1}$$

where $A^\bullet$ is the term completing the second row to a distinguished triangle. We claim we have a map τ as above such that $\tau \circ \rho$ is a quasi-isomorphism. Certainly we have a diagram with distinguished triangles for rows and columns

$$
\begin{tikzcd}
\underline{\Omega}^0_{X,Z} \arrow[r, "\times f"] \arrow[dd] & \underline{\Omega}^0_{X,Z} \arrow[r] \arrow[dd] & A^\bullet \arrow[r, "+1"] \arrow[dd] \arrow[drr, dotted, "\tau"] & {} & \\
 & & & & \underline{\Omega}^0_{H,Z\cap H} \arrow[dd] \\
\underline{\Omega}^0_X \arrow[r, "\times f"] \arrow[dd] & \underline{\Omega}^0_X \arrow[r] \arrow[dd] & B^\bullet \arrow[r, "+1"] \arrow[dd] \arrow[drr, dotted, "\kappa"] & {} & \\
 & & & & \underline{\Omega}^0_H \arrow[dd] \\
\underline{\Omega}^0_Z \arrow[r, "\times f"] \arrow[d, "+1"'] & \underline{\Omega}^0_Z \arrow[r] \arrow[d, "+1"'] & C^\bullet \arrow[r, "+1"] \arrow[d, "+1"'] \arrow[drr, dotted, "\mu"'] & {} & \\
{} & {} & {} & & \underline{\Omega}^0_{Z\cap H} \arrow[d, "+1"] \\
 & & & & {}
\end{tikzcd}
$$

and the existence of τ follows immediately from the existence of κ and μ, whose existence follows from the proof of [Kovács and Schwede 2011a, Theorem 4.1]. Note that the assumptions imply that $H|_Z = H \cap Z$ is a Cartier divisor on Z, so we may indeed use [Kovács and Schwede 2011a, Theorem 4.1] for both X and Z. Since $I_Z/(f \cdot I_Z) = I_Z/((f) \cap I_Z)$ by Lemma 4.1 and because $(H, Z \cap H)$ is a Du Bois pair, we see $\tau \circ \rho$ is an isomorphism as claimed.

Next we apply the Grothendieck duality functor $\boldsymbol{D}(\_) = \mathcal{R}\operatorname{Hom}^\bullet_R(\_, \omega^\bullet_R)$ to (4.2.1) and take cohomology, using $\boldsymbol{k}^i(\_)$ as shorthand to denote $\boldsymbol{h}^i(\boldsymbol{D}(\_))$:

$$
\begin{tikzcd}
\cdots & \boldsymbol{k}^i(I_Z) \arrow[l] & \boldsymbol{k}^i(I_Z) \arrow[l, "\times f"'] & \boldsymbol{k}^i(I_Z/(f\cdot I_Z)) \arrow[l, "\delta_i"'] & \boldsymbol{k}^{i-1}(I_Z) \arrow[l, "\alpha_i"'] & \boldsymbol{k}^{i-1}(I_Z) \arrow[l, "\times f"'] & \cdots \arrow[l] \\
\cdots & \boldsymbol{h}^i(\underline{\omega}^\bullet_{X,Z}) \arrow[l] \arrow[u, hook, "\Phi^i"'] & \boldsymbol{h}^i(\underline{\omega}^\bullet_{X,Z}) \arrow[l, "\times f"] \arrow[u, hook, "\Phi^i"'] & \boldsymbol{k}^i(A^\bullet) \arrow[l] \arrow[u, two heads, "\gamma_i"'] & \boldsymbol{h}^{i-1}(\underline{\omega}^\bullet_{X,Z}) \arrow[l, "\beta_i"] \arrow[u, hook, "\Phi^{i-1}"'] & \boldsymbol{h}^{i-1}(\underline{\omega}^\bullet_{X,Z}) \arrow[l, "\times f"] \arrow[u, hook, "\Phi^{i-1}"'] & \cdots \arrow[l]
\end{tikzcd}
$$

where the $\Phi^\bullet$ are injective by Theorem 3.2 and γ_i, which was obtained from ρ, is surjective since $\tau \circ \rho$ is an isomorphism.

The proof now follows exactly as for the main theorem of [Kovács and Schwede 2011a], or dually of [Kovács 2000, Theorem 3.2]. Fix $z \in \boldsymbol{h}^{i-1}(\boldsymbol{D}(I_Z))$. Pick $w \in \boldsymbol{h}^i(\boldsymbol{D}(A^\bullet))$ such that $\alpha_i(z) = \gamma_i(w)$. Since $\delta_i(\alpha_i(z)) = 0$ and Φ^i is injective, it follows that there exists a $u \in \boldsymbol{h}^{i-1}(\underline{\omega}^\bullet_{X,Z})$ such that $\beta_i(u) = w$. Therefore, $\alpha_i(\Phi^{i-1}(u)) = \alpha_i(z)$ and so

$$z - \Phi^{i-1}(u) \in f \cdot \boldsymbol{h}^{i-1}(\boldsymbol{D}(I_Z)). \tag{4.2.2}$$

Now, fix E_{i-1} to be the cokernel of Φ^{i-1} and set $\bar{z} \in E_{i-1}$ to be the image of z. Equation (4.2.2) then guarantees that $\bar{z} \in f \cdot E_{i-1}$. The multiplication map $E_{i-1} \xrightarrow{\times f} E_{i-1}$ is then surjective and so Nakayama's lemma guarantees that Φ^{i-1} is also surjective. Therefore $\underline{\omega}^{\bullet}_{X,Z} \to \boldsymbol{D}(I_Z)$ is a quasi-isomorphism, which implies that (X, Z) is a Du Bois pair. $\square$

Corollary 4.3. *Let $f : X \to B$ be a flat proper family of varieties over a smooth one-dimensional scheme, B being essentially of finite type over $\mathbb{C}$ (for instance, a smooth curve). Further, let $Z \subseteq X$ be a subscheme such that no component of Z is contained in any component of any fiber of f and $b \in B$ a closed point such that (X_b, Z_b) is a Du Bois pair. Then there exists a neighborhood $b \in U \subseteq B$ such that*

(a) *(X, Z) is Du Bois over U, and*

(b) *the fibers (X_u, Z_u) are Du Bois for all $u \in U$.*

Proof. The non-Du Bois locus T of (X, Z) is closed, and since f is proper, $f(T)$ is also closed. Hence (a) follows from Theorem 4.2 and by replacing B with an open set, we may assume that (X, Z) is Du Bois. Then the Bertini-type theorem Lemma 2.17(b) implies that (b) follows after possibly shrinking U. $\square$

Corollary 4.4. *Let $f : X \to B$ be a flat proper family of varieties over a smooth scheme B essentially of finite type over $\mathbb{C}$. Further let $Z \subseteq X$ be a subscheme which is also flat over B and $b \in B$ a closed point such that (X_b, Z_b) is a Du Bois pair. Then there exists a neighborhood $U \subseteq B$, $b \in U$, such that (X, Z) is Du Bois over U.*

Proof. We may assume that B is affine and let $d = \dim B$. We first show that (X, Z) itself is Du Bois in a neighborhood of (X_b, Z_b). Let $H_1, \dots, H_d$ be general smooth subschemes going through b whose local defining equations generate the maximal ideal of b (i.e., locally analytically they are coordinate hyperplanes). The pair $(X_b, Z_b) = (X_{H_1 \cap H_2 \cap \dots \cap H_d}, Z_{H_1 \cap H_2 \cap \dots \cap H_d})$ is Du Bois by assumption, hence since $X_b = X_{H_1 \cap \dots \cap H_d}$ is a hypersurface in $X_{H_2 \cap \dots \cap H_d}$ it follows that the pair $(X_{H_2 \cap \dots \cap H_d}, Z_{H_2 \cap \dots \cap H_d})$ is Du Bois in a neighborhood of X_b, by Corollary 4.3. Let W_1 denote the non-Du Bois locus of $(X_{H_2 \cap \dots \cap H_d}, Z_{H_2 \cap \dots \cap H_d})$. Since W_1 is closed and f is proper, we see that $f(W_1)$ is closed in $H_2 \cap \dots \cap H_d$ and doesn't contain b. Shrinking B if necessary, we may assume that W_1 is empty. Next observe that $H_2 \cap \dots \cap H_d$ is a hypersurface in $H_3 \cap \dots \cap H_d$ and so again we see that $(X_{H_3 \cap \dots \cap H_d}, Z_{H_3 \cap \dots \cap H_d})$ is Du Bois in a neighborhood of $X_{H_2 \cap \dots \cap H_d}$ by Corollary 4.3. Set W_2 to be the non-Du Bois locus of $(X_{H_3 \cap \dots \cap H_d}, Z_{H_3 \cap \dots \cap H_d})$ and note that $f(W_2)$ does not intersect $H_2 \cap \dots \cap H_d$. We shrink B again if necessary so that $W_2 = \varnothing$. Iterating this procedure proves the statement. $\square$

In order to extend Corollary 4.3(b) to families over arbitrary-dimensional bases we need the following lemma.

Lemma 4.5. *Let $f : X \to B$ be a flat proper family of varieties over a scheme B essentially of finite type over $\mathbb{C}$. Further, let $Z \subseteq X$ be a subscheme which is also* *flat over B and assume that (X, Z) is a Du Bois pair. Then*

$$V = \{b \in B \mid (X_b, Z_b) \text{ is a Du Bois pair}\}$$

is a constructible set in B. Furthermore, if B is smooth, then V is open.

Proof. We use induction on the dimension of B.

Let $\pi : B' \to B$ be a resolution of singularities and consider the base change $f' : X' = X_{B'} \to B'$, assumed to be a flat proper family over B' and with $Z' = Z_{B'} \subseteq X'$ a subscheme that is flat over B'. Notice that all the fibers of $f' : X' \to B'$ appear as fibers of $f : X \to B$ (up to harmless field extension), so $b' \in V' = \pi^{-1}(V) \subseteq B'$ if and only if the fiber $(X'_{b'}, Z'_{b'})$ is a Du Bois pair. It follows from Corollary 4.4 that by replacing B' with an open subset we may assume that (X', Z') is a Du Bois pair. It also follows that it is enough to prove the statement over a smooth irreducible base. Indeed, that implies that V' is open in B' and hence $V = f(V')$ is constructible.

To simplify notation we will replace B with B' and assume that B is smooth and irreducible, but use the inductive hypothesis without these additional assumptions.

The Bertini-type statement Lemma 2.17(b) implies that, if $V \neq \varnothing$, there is a dense open subset $U \subseteq B$ contained in V. The case $\dim B = 1$ follows immediately via the fact that in a curve any set containing a dense open set is itself open.

In general, it follows that $\dim(B \setminus U) < \dim B$ so by induction $V \setminus U$ is a constructible set in $B \setminus U$ and hence V is constructible in B. In the case of a smooth base Corollary 4.4 implies that V is stable under generalization and since we have just proved that it is constructible it follows that it is open. □

Corollary 4.6. *Let $f : X \to B$ be a flat proper family of varieties over a smooth scheme B essentially of finite type over $\mathbb{C}$. Further let $Z \subseteq X$ be a subscheme which is also flat over B and $b \in B$ a closed point such that (X_b, Z_b) is a Du Bois pair. Then there exists a neighborhood $U \subseteq B$, $b \in U$, such that (X_u, Z_u) is a Du Bois pair for all $u \in U$.*

Proof. Observe that the non-Du Bois locus W of (X, Z) is closed in X and since f is proper, $f(W)$ is also closed in B. Note that $f(W)$ does not contain b so it also does not contain the generic point of B. Hence by replacing B by a neighborhood $U \subseteq B$ of $b \in B$, we may assume that (X, Z) is Du Bois. Then the statement follows from Lemma 4.5. □

Remark 4.7. One can recover special cases of inversion of adjunction for log canonicity [Kawakita 2007] easily from Theorem 4.2. For instance, let $(X, D + H)$ be a pair with K_X, D and H Cartier and assume that $(H, D|_H)$ is slc or equivalently Du Bois [Kollár 2013]. Then $(X, D + H)$ is Du Bois or equivalently lc by Theorem 4.2.

5. Generalizing the Kollár–Kovács result to pairs

We recall Theorem 7.12 of [Kollár and Kovács 2010]. Let $f : X \to B$ be a flat projective family of varieties with Du Bois singularities. Then if B is connected and the general fiber is Cohen–Macaulay, then all the fibers are Cohen–Macaulay.

We would like to generalize this to the context of Du Bois pairs, at least in the case when Z is a divisor. We recommend the reader have a copy of [Kollár and Kovács 2010] available when reading this section as we refer to a number of lemmas therein. We begin by generalizing a result of Du Bois and Jarraud to pairs; cf. [Du Bois and Jarraud 1974; Du Bois 1981, théorème 4.6].

Theorem 5.1. *Let $f : X \to B$ be a flat proper morphism between schemes of finite type over $\mathbb{C}$. Assume that B is smooth and let $Z \subseteq X$ be a subscheme that is flat over B. Further assume that the geometric fibers $(X_b, Z_b) \to b$ are Du Bois. Then for all i, $\mathcal{R}^i f_* \mathcal{I}_Z$ is locally free of finite rank and compatible with base change*; *in other words $(\mathcal{R}^i f_* \mathcal{I}_Z)_T \simeq \mathcal{R}^i f_* \mathcal{I}_{Z_T}$ for any morphism $T \to B$.*

Proof. For some $b \in B$, let $\mathfrak{m}$ be the maximal ideal of $\mathcal{O}_{B,b}$ and $S = S_n = \operatorname{Spec} \mathcal{O}_{B,b}/\mathfrak{m}^{n+1}$ for $n \in \mathbb{N}$. Further, let $\mathcal{I}_{Z_b}$ and $\mathcal{I}_{Z_S}$ denote the ideal sheaves of Z_b in X_b and Z_S in X_S, respectively. Consider the commutative diagram

$$\begin{tikzcd}
H^i_c((X_S \setminus Z_S)^{\mathrm{an}}, \mathbb{C}) \arrow[dd, "\lambda"'] \arrow[dr, "\alpha"] & & H^i(X_S, \mathcal{I}_{Z_S}) \arrow[dl, "\beta"'] \arrow[dd, "\nu"] \\
& H^i(X_S^{\mathrm{an}}, \mathcal{I}_{Z_S}) \arrow[dd, "\mu"] & \\
H^i_c((X_b \setminus Z_b)^{\mathrm{an}}, \mathbb{C}) \arrow[dr, "\gamma"'] & & H^i(X_b, \mathcal{I}_{Z_b}) \arrow[dl, "\delta"] \\
& H^i(X_b^{\mathrm{an}}, \mathcal{I}_{Z_b}) &
\end{tikzcd}$$

Observe that λ is an isomorphism since X_S and X_b have the same support. By [Kovács 2011a, Theorem 4.1], cf. [Kollár 2013, Theorem 6.8], γ is surjective, so $\gamma \circ \lambda = \mu \circ \alpha$ is surjective and hence μ is surjective. By Serre's GAGA principle [Serre 1956], β and δ are isomorphisms and hence ν is surjective. Finally, the statement follows by cohomology and base change [Grothendieck 1963, §7.7]. □

Next we prove the analogue of the main flatness and base change result of Kollár and Kovács [2010, Theorem 7.9] for Du Bois pairs.

Theorem 5.2. *Let $f : X \to B$ be a flat projective morphism between schemes of finite type over $\mathbb{C}$, and assume that B is smooth. Let $Z \subseteq X$ be a closed subscheme* *that is flat over B and $\mathscr{L}$ a relatively ample line bundle on X. Assume (X, Z) is Du Bois. Then*:

(a) *The sheaves $\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))$ are flat over B for all i.*

(b) *The sheaves $f_*\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f)) \otimes \mathscr{L}^q\big)$ are locally free and compatible with arbitrary base change for all $i > 0$ and $q \gg 0$.*

(c) *For any base change $\vartheta : T \to B$ and for all $i > 0$,*

$$\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))\big)_T \simeq \boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{X_T}}(\mathscr{I}_{Z_T}, \omega^{\bullet}_{f_T})).$$

Proof. We follow the proof of [Kollár and Kovács 2010, Theorem 7.9]. We may assume that $B = \operatorname{Spec} R$ is affine and hence that $\mathscr{L}^m$ is globally generated for $m \gg 0$. For such an $m \gg 0$, choose a general section $\sigma \in H^0(X, \mathscr{L}^m)$ and consider the cyclic cover induced by σ:

$$\mathscr{A} = \bigoplus_{j=0}^{m-1} \mathscr{L}^{-j} \simeq \bigoplus_{j=0}^{m-1} \mathscr{L}^{-j} t^j \big/ (t^m - \sigma).$$

Set $h : Y = \operatorname{Spec}_X \mathscr{A} \to X$, and $Z_Y = h^{-1}Z$ with the induced reduced scheme structure. Then the geometric fibers of the composition $(Y, Z_Y) \to B$ are also Du Bois by [Kollár 2013, Corollary 6.21]. Note that by construction $\mathscr{I}_{Z_Y} = \bigoplus_{j=0}^{m-1} \mathscr{I}_Z \otimes \mathscr{L}^{-j}$. Hence $\mathcal{R}^i h_* \mathscr{I}_{Z_Y}$ is locally free of finite rank and compatible with arbitrary base change by Theorem 5.1. It follows that the summands of these modules, the $\mathcal{R}^i f_*(\mathscr{I}_Z \otimes \mathscr{L}^{-j})$, are also locally free and compatible with base change. Since we may choose m arbitrarily large, this holds for all $j \in \mathbb{N}$. It follows immediately that $\mathcal{H}om_{\mathscr{O}_B}(\mathcal{R}^i f_*(\mathscr{I}_Z \otimes \mathscr{L}^{-j}), \mathscr{O}_B)$ is also locally free and compatible with base change.

By Grothendieck duality and [Kollár and Kovács 2010, Lemma 7.3] (see the proof of Lemma 7.2 in that paper follows that

$$\begin{aligned}\mathcal{H}om_{\mathscr{O}_B}(\mathcal{R}^i f_*(\mathscr{I}_Z \otimes \mathscr{L}^{-q}), \mathscr{O}_B) &\simeq f_* \boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f \otimes \mathscr{L}^q)) \\ &\simeq f_*\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f) \otimes \mathscr{L}^q)\big)\end{aligned}$$

and hence (b) is proven. Just as in [Kollár and Kovács 2010, Theorem 7.9], (a) follows from (b) by an argument similar to [Hartshorne 1977, Chapter III, Theorem 9.9].

Finally we prove (c). Since $f : X \to B$ is projective and B is affine, we may factor f as $X \xrightarrow{i} \mathbb{P}^n_B \xrightarrow{\pi} B$. It then suffices to show that

$$\varrho^{-i} : \big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{\mathbb{P}^n_B}}(\mathscr{I}_Z, \omega_\pi[n]))\big)_T \to \boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{\mathbb{P}^n_T}}(\mathscr{I}_{Z_T}, \omega_\pi[n]))$$

is an isomorphism. As in [Kollár and Kovács 2010, Theorem 7.9], we proceed by descending induction on i (the base case where $i \gg 0$ is obvious). We observe that $\mathscr{I}_Z$ is flat since so are $\mathscr{O}_X$ and $\mathscr{O}_Z$ and assume that $\varrho^{-(i+1)}$ is an isomorphism by induction. Since $\boldsymbol{h}^{-i}\big(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{\mathbb{P}^n_B}}(\mathscr{I}_Z, \omega_\pi[n])\big)$ is flat, by (a), we may apply [Altman and Kleiman 1980, Theorem 1.9], which completes the proof. $\square$

The following is the analog of [Kollár and Kovács 2010, Theorem 7.11] for pairs.

Theorem 5.3. *Let $f : X \to B$ be a flat projective morphism between schemes of finite type over $\mathbb{C}$. Assume that B is smooth and let $Z \subseteq X$ be a subscheme that is flat over B. Let $x \in X$ be a closed point and let $b = f(x)$. Then $\mathscr{I}_{Z_b} \subseteq \mathscr{O}_{X_b}$ is S_k at x if and only if*

$$\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))\big)_y = 0 \tag{5.3.1}$$

for $i < \min(k+\dim\overline{\{y\}}, \dim_x X)$ and for all $y \in X_b$ such that $x \in \overline{\{y\}}$. In particular, $\mathscr{I}_{Z_b}$ is S_k if and only if (5.3.1) holds for $i < \min(k+\dim\overline{\{y\}}, \dim_x X)$ and for all $y \in X_b$ (not restricted to closed points).

First we prove a lemma.

Lemma 5.4. *Let X be a scheme that admits a dualizing complex $\omega^{\bullet}_X$. Let $x \in X$ and let $\mathscr{F}$ be a coherent sheaf on X. Then $\mathscr{F}$ is S_k at $x \in X$ if and only if*

$$\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{F}, \omega^{\bullet}_X))\big)_y = 0$$

for $i < \min(k, \dim\mathscr{F}_y) + \dim\overline{\{y\}}$ and for all $y \in X$ such that $x \in \overline{\{y\}}$.

Proof. This is a consequence of local duality [Hartshorne 1966] and the cohomological criterion for depth; see for instance [Kovács 2011b, Proposition 3.2]. $\square$

Proof of Theorem 5.3. By the lemma, $\mathscr{I}_{Z_b} \subseteq \mathscr{O}_{X_b}$ is S_k at x if and only if

$$\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{X_b}}(\mathscr{I}_{Z_b}, \omega^{\bullet}_{X_b}))\big)_y = 0$$

for $i < \min(k, \dim(\mathscr{I}_{Z_b})_y) + \dim\overline{\{y\}} = \min(k+\dim\overline{\{y\}}, \dim_x X)$ and for all $y \in X_b$ such that $x \in \overline{\{y\}}$. By Theorem 5.2,

$$\begin{aligned}\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_{X_b}}(\mathscr{I}_{Z_b}, \omega^{\bullet}_{X_b}))\big)_y &\simeq \Big(\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))\big)_b\Big)_y \\ &\simeq \Big(\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))\big)_y\Big)_b.\end{aligned}$$

But notice that the right side is zero if and only if $\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z, \omega^{\bullet}_f))\big)_y$ is zero by Nakayama's lemma. This implies the desired statement. $\square$

Finally, we describe how the S_k condition behaves for pairs in families where the fibers are Du Bois.

Theorem 5.5. *Let $f:(X,Z)\to B$ be a flat projective family with $\mathscr{O}_Z$ (and hence $\mathscr{I}_Z$) flat over B as well. Assume that all the fiber pairs (X_b, Z_b) are Du Bois. Assume also that B is connected and the generic fibers $(\mathscr{I}_Z)_{\text{gen}}$ are S_k. Then all the fibers $(\mathscr{I}_Z)_b$ are S_k.*

Proof. By working with one component of B at a time, we may assume that B is irreducible and hence that X is equidimensional. If $(\mathscr{I}_Z)_b \simeq \mathscr{I}_{Z_b}$ (by flatness of $\mathscr{O}_Z$) is not S_k at some point $y\in X_b$, then by Theorem 5.3, $\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^{\bullet}_f))\neq 0$ near y for some $i<\min(k+\dim\overline{\{y\}},\dim X)$. Fix an irreducible component $W\subseteq \operatorname{supp}\big(\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^{\bullet}_f))\big)$ and observe that $\dim W_b$ is constant for $b\in B$ since $\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^{\bullet}_f))$ is flat by Theorem 5.2(a). However, in that case it follows that $\boldsymbol{h}^{-i}(\mathcal{R}\mathcal{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{I}_Z,\omega^{\bullet}_f))$ is nonzero near some point $\eta\in X_{\text{gen}}$ such that $\dim\overline{\{\eta\}}=\dim\overline{\{y\}}$, which contradicts the assumption that the generic fiber is S_k by Theorem 5.3. $\square$

Corollary 5.6. *Let $f:(X,Z)\to B$ be a flat projective family with $\mathscr{O}_Z$ (and hence $\mathscr{I}_Z$) flat over B as well. Assume that all the fiber pairs (X_b,Z_b) are Du Bois. Assume also that B is connected and the generic fibers $(\mathscr{I}_Z)_{\text{gen}}$ are Cohen–Macaulay. Then all the fibers $(\mathscr{I}_Z)_b$ are Cohen–Macaulay.* $\square$

At this point it is natural to ask the next question.

Question 5.7. Assume that (X,Z) is a pair and that $H\subseteq X$ is a Cartier divisor such that $(H, Z\cap H)$ is a Du Bois pair. If $\mathscr{I}_Z|_{X\setminus H}$ is Cohen–Macaulay, does it follow that $\mathscr{I}_Z$ is Cohen–Macaulay?

In the case that $Z=\varnothing$, the analogous result holds in characteristic $p>0$ for F-injective singularities by [Horiuchi et al. 2014, Appendix by K. Schwede and A. K. Singh].

6. Generalizing Kovács–Schwede–Smith to pairs

The goal of this section is to prove the analog of the main result of [Kovács et al. 2010] for pairs (X,Z).

Lemma 6.1. *Let X be a normal d-dimensional variety, $Z\subsetneq X$ a reduced closed subscheme and $\Sigma\subsetneq X$ a codimension ≥ 2 subset containing the singular locus of X. Let $\pi:\widetilde{X}\to X$ be a log resolution of $(X,\Sigma\cup Z)$ with $E=\pi^{-1}(\Sigma\cup Z)_{\text{red}}$. Then*

(a) $\underline{\Omega}^0_{X,\Sigma\cup Z}\simeq \mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}}(-E)$, *and*

(b) $\boldsymbol{h}^{-d}(\underline{\omega}^{\bullet}_{X,Z})\simeq \pi_*\omega_{\widetilde{X}}(E)$.

Proof. First we claim that both $\mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}}(-E)$ and $\pi_*\omega_{\widetilde{X}}(E)$ are independent of the choice of π. This was proved for $\mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}}(-E)$ on pages 67–68 in the proof of [Kovács and Schwede 2011b, Theorem 6.4] and for $\pi_*\omega_{\widetilde{X}}(E)$ in [Kovács et al.

2010, Lemma 3.12]. Therefore we are free to choose π and hence we may assume that it is an isomorphism outside of $\Sigma \cup Z$. We have the distinguished triangle

$$\begin{tikzcd} \underline{\Omega}^0_X \arrow[r] \arrow[d, "\simeq"'] & \mathcal{R}\pi_*\underline{\Omega}^0_{\widetilde{X}} \oplus \underline{\Omega}^0_{\Sigma\cup Z} \arrow[r] \arrow[d, "\simeq"'] & \mathcal{R}\pi_*\underline{\Omega}^0_E \arrow[r, "+1"] \arrow[d, "\simeq"'] & {} \\ \underline{\Omega}^0_X \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}} \oplus \underline{\Omega}^0_{\Sigma\cup Z} \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_E \arrow[r, "+1"] & {} \end{tikzcd}$$

The isomorphisms follow since $\widetilde{X}$ and E are Du Bois. In the next diagram the first two rows are distinguished triangles by definition; see (2.7.1). The third row is simply the pushforward of a natural short exact sequence from $\widetilde{X}$. The previous diagram and [Kollár and Kovács 2010, Lemma 2.1] (or simply the octahedral axiom) imply that α below is an isomorphism. The other two isomorphisms again follow since $\widetilde{X}$ and E are Du Bois. Note that the columns are *not* exact.

$$\begin{tikzcd} \underline{\Omega}^0_{X,\Sigma\cup Z} \arrow[r] \arrow[d, "\simeq"', "\alpha"] & \underline{\Omega}^0_X \arrow[r] \arrow[d] & \underline{\Omega}^0_{\Sigma\cup Z} \arrow[r, "+1"] \arrow[d] & {} \\ \mathcal{R}\pi_*\underline{\Omega}^0_{\widetilde{X},E} \arrow[r] \arrow[d, dotted] & \mathcal{R}\pi_*\underline{\Omega}^0_{\widetilde{X}} \arrow[r] \arrow[d, "\simeq"'] & \mathcal{R}\pi_*\underline{\Omega}^0_E \arrow[r, "+1"] \arrow[d, "\simeq"'] & {} \\ \mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}}(-E) \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_{\widetilde{X}} \arrow[r] & \mathcal{R}\pi_*\mathscr{O}_E \arrow[r, "+1"] & {} \end{tikzcd}$$

It follows that the dotted arrow, and hence its composition with α, are also isomorphisms. This proves (a).

In order to prove (b), consider the map $\underline{\Omega}^0_{X,\Sigma\cup Z} \to \underline{\Omega}^0_{X,Z}$ obtained in

$$\begin{tikzcd} \underline{\Omega}^0_{X,\Sigma\cup Z} \arrow[r] \arrow[d, dotted] & \underline{\Omega}^0_X \arrow[r] \arrow[d, equal] & \underline{\Omega}^0_{\Sigma\cup Z} \arrow[r, "+1"] \arrow[d] & {} \\ \underline{\Omega}^0_{X,Z} \arrow[r] & \underline{\Omega}^0_X \arrow[r] & \underline{\Omega}^0_Z \arrow[r, "+1"] & {} \end{tikzcd}$$

Now we have a distinguished triangle

$$\underline{\Omega}^0_{X,\Sigma\cup Z} \to \underline{\Omega}^0_{X,Z} \to C^\bullet \xrightarrow{+1} . \tag{6.1.1}$$

Claim 6.2. *With the above notation*, $0 = \boldsymbol{h}^{-d}(\boldsymbol{D}(C^\bullet)) = \boldsymbol{h}^{-d+1}(\boldsymbol{D}(C^\bullet))$.

Proof of claim. Consider the following diagram with distinguished triangles as rows and columns:

$$\begin{tikzcd}
\underline{\Omega}^0_{\Sigma\cup Z}[-1] \arrow[r] \arrow[d] & \underline{\Omega}^0_{Z}[-1] \arrow[r] \arrow[d] & \underline{\Omega}^0_{\Sigma\cup Z, Z} \arrow[r, "+1"] & {} \\
\underline{\Omega}^0_{X,\Sigma\cup Z} \arrow[r] \arrow[d] & \underline{\Omega}^0_{X,Z} \arrow[r] \arrow[d] & C^\bullet \arrow[r, "+1"] & {} \\
\underline{\Omega}^0_{X} \arrow[d, "+1"'] & \underline{\Omega}^0_{X} \arrow[d, "+1"'] & & \\
{} & {} & &
\end{tikzcd}$$

It follows from [Kovács 2013, Theorem B.1] that $C^\bullet \simeq \underline{\Omega}^0_{\Sigma\cup Z, Z}$. On the other hand, by Lemma 2.17(d), $\underline{\Omega}^0_{\Sigma\cup Z, Z} \simeq \underline{\Omega}^0_{\Sigma, \Sigma\cap Z}$ and hence $C^\bullet \simeq \underline{\Omega}^0_{\Sigma, \Sigma\cap Z}$.

Next recall that by Theorem 3.2 there exists a natural injective map

$$\boldsymbol{h}^{-j}\big(\boldsymbol{D}(\underline{\Omega}^0_{\Sigma,\Sigma\cap Z})\big) \hookrightarrow \boldsymbol{h}^{-j}\big(\mathcal{R}\mathcal{H}om^\bullet_{\mathcal{O}_\Sigma}(\mathcal{I}_{(\Sigma\cap Z)\subseteq\Sigma}, \omega^\bullet_\Sigma)\big). \tag{6.2.1}$$

Since $\dim \Sigma \le d-2$, the right hand side of (6.2.1) is zero for $j \ge d-1$, establishing the claim. □

Grothendieck duality and part (a) imply that $\boldsymbol{h}^{-d}(\underline{\omega}^\bullet_{X,\Sigma\cup Z}) \simeq \pi_*\omega_{\widetilde{X}}(E)$ and it follows from Claim 6.2 that $\boldsymbol{h}^{-d}(\underline{\omega}^\bullet_{X,\Sigma\cup Z}) \simeq \boldsymbol{h}^{-d}(\underline{\omega}^\bullet_{X,Z})$, which in turn implies part (b). □

Theorem 6.3. *Let X be a normal variety and $Z \subseteq X$ a divisor. Let $\pi : \widetilde{X} \to X$ be a log resolution of (X, Z) with $E = \pi^{-1}(Z)_{\mathrm{red}} \vee \operatorname{exc}(\pi)$. If $\mathcal{I}_Z$ is Cohen–Macaulay, then (X, Z) is Du Bois if and only if*

$$\pi_*\omega_{\widetilde{X}}(E) \simeq \omega_X(Z).$$

Proof. Since $\mathcal{I}_Z$ is Cohen–Macaulay, $\mathcal{R}\mathcal{H}om^\bullet_{\mathcal{O}_X}(\mathcal{I}_Z, \omega^\bullet_X) \simeq \mathcal{H}om_{\mathcal{O}_X}(\mathcal{I}_Z, \omega_X)[\dim X]$ by the local dual of the local cohomology criterion for Cohen–Macaulayness. Because the map

$$\underline{\omega}^\bullet_{X,Z} \to \mathcal{R}\mathcal{H}om^\bullet_{\mathcal{O}_X}(\mathcal{I}_Z, \omega^\bullet_X) \tag{6.3.1}$$

is injective on cohomology by Theorem 3.2, it follows that $\boldsymbol{h}^i(\underline{\omega}^\bullet_{X,Z}) = 0$ for $i \ne -\dim X$ and hence (X, Z) is Du Bois if and only if

$$\boldsymbol{h}^{-\dim X}(\underline{\omega}^\bullet_{X,Z}) \to \mathcal{H}om_{\mathcal{O}_X}(\mathcal{I}_Z, \omega_X) \simeq \omega_X(Z)$$

is an isomorphism. But $\boldsymbol{h}^{-\dim X}(\underline{\omega}^\bullet_{X,Z}) \simeq \pi_*\omega_{\widetilde{X}}(E)$ by Lemma 6.1, so the statement follows. □

7. An inversion of adjunction for rational and Du Bois pairs

In this final section of the paper, we will prove the following theorem.

Theorem 7.1. *Let $f : X \to B$ be a flat projective family with geometrically integral fibers over a smooth connected base B, $A \subseteq B$ a smooth closed subscheme containing no component of B and $H = f^{-1}(A) = X \times_B A$ (with the induced scheme-theoretic structure). Let D be a reduced codimension-1 subscheme of X which is flat over B. Assume that for every $s \in A$, (X_s, D_s) is Du Bois and that $(X \setminus H, D \setminus H)$ is a rational pair. Then (X, D) is a rational pair.*

Remark 7.2. In the introduction, A was assumed to be a closed point. This version is more general and more convenient for our proof.

Remark 7.3. The assumptions also imply the following auxiliary conditions:

(a) Since $X \to B$ has geometrically integral fibers and H is obtained by base change with a smooth subscheme, H is reduced.

(b) $\mathscr{I}_D$ is flat over B and no component of D contains a fiber of f. In particular D and H have no common components.

(c) As for any $s \in A$, $H_s = X_s$, it follows that $(H, D \cap H)$ is Du Bois by Corollary 4.4.

(d) $X \setminus H$ is normal by the definition of a rational pair.

Before embarking on proving the theorem, we will first prove several lemmas that show that our situation is simpler than it might first appear.

First we show that we may assume that A is a divisor in B.

Lemma 7.4. *In order to prove Theorem 7.1 it is sufficient to assume that A is a smooth Cartier divisor in B.*

Proof. The statement is local over the base so we may assume that B is affine. Additionally, since we only need to work in a neighborhood of a point $a \in A$, we may assume that (X, D) is Du Bois and all the fibers (X_b, D_b) for all $b \in B$ are Du Bois by Corollaries 4.4 and 4.6. Choose a general hypersurface G containing A and note that since A is smooth we may assume that G is smooth. Then the hypotheses of the theorem are satisfied for G replacing A as well since $X \setminus f^{-1}(G) \subseteq X \setminus f^{-1}(A)$ and since we already assumed that all the fibers (X_b, D_b) over all points $b \in B$ were Du Bois. $\square$

From this point forward, we will assume that B is a smooth affine scheme, A is a smooth hypersurface in B and $H = f^*A$.

Lemma 7.5. *Under the assumptions of Theorem 7.1, X is normal and thus D is also a divisor.*

Proof. Since H is reduced, every point $\eta \in H$ has depth at least $\min(1, \dim \mathscr{O}_{H,\eta})$. Because $f : X \to B$ is flat, the local defining equation of H is a regular element in $\mathscr{O}_X$, so any point $\eta \in X$ that lies in H has depth at least $\min(2, \dim \mathscr{O}_{X,\eta})$. Since $X \setminus H$ is normal it is S_2 and so X is S_2 everywhere. Finally observe that H is reduced, hence generically regular and $X \setminus H$ is R_1. As H is Cartier, this implies that X is also R_1 and therefore normal. □

Now observe that the fact that $(H, D \cap H)$ is Du Bois (see Remark 7.3(b)) says something about the structure of D on X.

Lemma 7.6. *With notation as in Theorem 7.1, no stratum of the snc locus of (X, D) can be contained inside H.*

Proof. Assume to the contrary that there exists a stratum Z of the snc locus of (X, D) contained in H. Let η be the generic point of Z. By assumption $\eta \in H$ and (X, D) is snc at η, so $\mathscr{O}_{X,\eta}$ is a regular ring. Let $n = \dim \mathscr{O}_{X,\eta}$. Replace X by Spec $\mathscr{O}_{X,\eta}$ and H and D by their pullbacks to this local scheme (in this step we lose projectivity, but we will not need that for now). Note that D is now Cartier and in fact snc. Furthermore $D + H$ has $n + 1$ irreducible components containing η, so $(X, D+H)$ cannot be Du Bois (or equivalently log canonical since X is Gorenstein). But as we observed, $(H, D \cap H)$ is a Du Bois pair. Then by Theorem 4.2 again we see that $(X, D + H)$ is Du Bois as well. This is a contradiction. □

Next we setup the notation for the proof of Theorem 7.1. Let Σ denote the non-snc locus of (X, D). Observe that as X is normal and D is a reduced divisor by Lemma 7.5, we have that $\operatorname{codim}_X(\Sigma) \geq 2$.

Additionally assume that $\pi : Y \to X$ is a log resolution of $(X, D \cup H \cup \Sigma)$ that simultaneously gives a thrifty resolution of (X, D). To see that such a π exists, first take a thrifty resolution (U, D_U) of (X, D) and then perform a log resolution of the scheme-theoretic preimages of H and Σ on U (while keeping the strict transform D_U snc). The result can be assumed to be a thrifty resolution of (X, D) since the preimages of Σ and H do not contain any strata of (U, D_U) by Lemma 7.6.

Set $\overline{H}$ and $\overline{D}$ to be the reduced total transforms of H and D respectively, set D_Y to be the strict transform of D and set E to be $(\pi^{-1}(\Sigma))_{\mathrm{red}}$.

Proof of Theorem 7.1. Clearly (X, D) is a Du Bois pair and all the fibers (X_b, D_b) are Du Bois by Corollaries 4.4 and 4.6 (possibly after shrinking the base B around A). By Corollary 5.6, we know that $\mathscr{O}_X(-D)$ is Cohen–Macaulay. Thus by the local dual version of the local cohomological criterion for Cohen–Macaulayness, $\mathscr{R}\mathscr{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{O}_X(-D), \omega^{\bullet}_X)$ has cohomology only in one term. In particular,

$$\begin{aligned}\mathscr{R}\mathscr{H}om^{\bullet}_{\mathscr{O}_X}(\mathscr{O}_X(-D), \omega^{\bullet}_X) &\simeq \mathscr{H}om_{\mathscr{O}_X}(\mathscr{O}_X(-D), \omega_X)[\dim X] \\ &\simeq \omega_X(D)[\dim X]. \qquad (7.6.1)\end{aligned}$$

Therefore by Proposition 2.7 it suffices to show that $\omega_X(D) \simeq \pi_*\omega_Y(D_Y)$.

Next observe that $(H, D|_H)$ is a Du Bois pair by Corollary 4.4 and hence by Lemma 2.18 we see that $(X, D \cup H) = (X, D + H)$ is a Du Bois pair.

Claim 7.7. *With notation as above,* $\pi_*\omega_Y(D_Y \vee \overline{H} \vee E) \simeq \pi_*\omega_Y(D_Y + \overline{H})$.

Note that $D_Y + \overline{H} = D_Y \vee \overline{H}$ since the divisors have no common components.

Proof of claim. The containment $\supseteq$ is obvious since D and H do not share a component (see Remark 7.3(a)), so choose $f \in \pi_*\omega_Y(D_Y \vee \overline{H} \vee E)$. We observe that

$$\mathrm{div}_Y(f) + K_Y + D_Y \vee \overline{H} \vee E = \mathrm{div}_Y(f) + K_Y + D_Y + \overline{H} \vee E \geq 0.$$

Working on $U = Y \setminus \overline{H} = \pi^{-1}(X \setminus H)$ we see that $\mathrm{div}_U(f) + K_U + D_Y|_U + E|_U \geq 0$. But since $(X \setminus H, D \setminus H)$ is a rational pair,

$$\pi_*\omega_U(D_Y|_U) = \pi_*\omega_U(D_Y|_U + E) = \omega_{X\setminus H}(D|_{X\setminus H}),$$

so $\mathrm{div}_U(f) + K_U + D_Y|_U + E|_U \geq 0$ is equivalent to $\mathrm{div}_U(f) + K_U + D_Y|_U \geq 0$. Because the components of E that lie over H are also components of $\overline{H}$, it follows that $\mathrm{div}_Y(f) + K_Y + D_Y + \overline{H} \geq 0$, proving the claim. □

By Lemma 6.1 we see that $\boldsymbol{h}^{-\dim X}(\underline{\omega}^{\bullet}_{X,D+H}) \simeq \pi_*\omega_Y(D_Y \vee \overline{H} \vee E)$, which agrees with $\pi_*\omega_Y(D_Y + \overline{H})$ by the claim. Since $(X, D + H)$ is a Du Bois pair, $\boldsymbol{h}^{-\dim X}(\underline{\omega}^{\bullet}_{X,D+H}) \simeq \omega_X(D + H)$ and so in conclusion we have that

$$\omega_X(D + H) \simeq \pi_*\omega_Y(D_Y + \overline{H}).$$

Twisting both sides by $-H$ and using the projection formula we see that

$$\omega_X(D) \simeq \pi_*\omega_Y(D_Y - (\pi^*H - \overline{H})) \subseteq \pi_*\omega_Y(D_Y),$$

since $\pi^*H - \overline{H}$ is effective. But $\pi_*\omega_Y(D_Y) \subseteq \omega_X(D)$ for any normal pair (X, D) and so $\omega_X(D) \simeq \pi_*\omega_Y(D_Y)$ as desired. □

Setting $D = 0$ we obtain the following.

Corollary 7.8. *Let $f : X \to B$ be a flat projective family over a smooth base B and $H = f^{-1}(0)$ a special fiber. Assume that H has Du Bois singularities and that $X \setminus H$ has rational singularities. Then X has rational singularities.*

There is a variant of our inversion of adjunction theorem that we would also like to prove (even in the $D = 0$ case).

Conjecture 7.9. *Assume (X, D) is a pair with D a reduced Weil divisor. Further assume that H is a Cartier divisor on X, not having any components in common with D, such that $(H, D \cap H)$ is Du Bois and such that $(X \setminus H, D \setminus H)$ is a rational pair. Then (X, D) is a rational pair.*

The only place where our proof above does not work in this situation is when we prove that $\mathscr{O}_X(-D)$ is Cohen–Macaulay. In particular, to accomplish this generalization, we would simply need a version of Corollary 5.6 that is not tied to a projective or proper family. What is missing is exactly a positive answer to Question 5.7.

Acknowledgements

The authors thank Zsolt Patakfalvi for several useful discussions and the referee for numerous helpful comments and corrections.

References

[Altman and Kleiman 1980] A. B. Altman and S. L. Kleiman, "Compactifying the Picard scheme", *Adv. Math.* **35**:1 (1980), 50–112. MR Zbl

[Du Bois 1981] P. Du Bois, "Complexe de de Rham filtré d'une variété singulière", *Bull. Soc. Math. France* **109**:1 (1981), 41–81. MR Zbl

[Du Bois and Jarraud 1974] P. Du Bois and P. Jarraud, "Une propriété de commutation au changement de base des images directes supérieures du faisceau structural", *C. R. Acad. Sci. Paris Sér. A* **279** (1974), 745–747. MR Zbl

[Elkik 1978] R. Elkik, "Singularités rationnelles et déformations", *Invent. Math.* **47**:2 (1978), 139–147. MR Zbl

[Erickson 2014] L. Erickson, "Deformation invariance of rational pairs", preprint, 2014. arXiv

[Esnault 1990] H. Esnault, "Hodge type of subvarieties of $\mathbb{P}^n$ of small degrees", *Math. Ann.* **288**:3 (1990), 549–551. MR Zbl

[Fedder and Watanabe 1989] R. Fedder and K.-I. Watanabe, "A characterization of F-regularity in terms of F-purity", pp. 227–245 in *Commutative algebra* (Berkeley, CA, 1987), edited by M. Hochster et al., Mathematical Sciences Research Institute Publications **15**, Springer, New York, NY, 1989. MR Zbl

[Graf and Kovács 2014] P. Graf and S. J. Kovács, "Potentially Du Bois spaces", *J. Singul.* **8** (2014), 117–134. MR Zbl

[Grauert and Riemenschneider 1970] H. Grauert and O. Riemenschneider, "Verschwindungssätze für analytische Kohomologiegruppen auf komplexen Räumen", *Invent. Math.* **11** (1970), 263–292. MR Zbl

[Greb et al. 2011] D. Greb, S. Kebekus, S. J. Kovács, and T. Peternell, "Differential forms on log canonical spaces", *Publ. Math. Inst. Hautes Études Sci.* **114** (2011), 87–169. MR Zbl

[Grothendieck 1963] A. Grothendieck, "Éléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, II", *Inst. Hautes Études Sci. Publ. Math.* **17** (1963), 5–91. MR Zbl

[Hacon 2014] C. D. Hacon, "On the log canonical inversion of adjunction", *Proc. Edinb. Math. Soc.* (2) **57**:1 (2014), 139–143. MR Zbl

[Hacon and Xu 2013] C. D. Hacon and C. Xu, "Existence of log canonical closures", *Invent. Math.* **192**:1 (2013), 161–195. MR Zbl

[Hartshorne 1966] R. Hartshorne, *Residues and duality*, Lecture Notes in Mathematics **20**, Springer, Berlin, 1966. MR Zbl

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

[Horiuchi et al. 2014] J. Horiuchi, L. E. Miller, and K. Shimomoto, "Deformation of F-injectivity and local cohomology", *Indiana Univ. Math. J.* **63**:4 (2014), 1139–1157. MR Zbl

[Karu 2000] K. Karu, "Minimal models and boundedness of stable varieties", *J. Algebraic Geom.* **9**:1 (2000), 93–109. MR Zbl

[Kawakita 2007] M. Kawakita, "Inversion of adjunction on log canonicity", *Invent. Math.* **167**:1 (2007), 129–133. MR Zbl

[Kollár 1997] J. Kollár, "Singularities of pairs", pp. 221–287 in *Algebraic geometry* (Santa Cruz, 1995), edited by J. Kollár et al., Proceedings of Symposia in Pure Mathematics **62**:1, American Mathematical Society, Providence, RI, 1997. MR Zbl

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics **200**, Cambridge University Press, 2013. MR Zbl

[Kollár and Kovács 2010] J. Kollár and S. J. Kovács, "Log canonical singularities are Du Bois", *J. Amer. Math. Soc.* **23**:3 (2010), 791–813. MR Zbl

[Kollár and Mori 1998] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics **134**, Cambridge University Press, 1998. MR Zbl

[Kollár and Shepherd-Barron 1988] J. Kollár and N. I. Shepherd-Barron, "Threefolds and deformations of surface singularities", *Invent. Math.* **91**:2 (1988), 299–338. MR Zbl

[Kovács 1999] S. J. Kovács, "Rational, log canonical, Du Bois singularities: on the conjectures of Kollár and Steenbrink", *Compos. Math.* **118**:2 (1999), 123–133. MR Zbl

[Kovács 2000] S. J. Kovács, "Rational, log canonical, Du Bois singularities, II: Kodaira vanishing and small deformations", *Compos. Math.* **121**:3 (2000), 297–304. MR Zbl

[Kovács 2011a] S. J. Kovács, "Du Bois pairs and vanishing theorems", *Kyoto J. Math.* **51**:1 (2011), 47–69. MR Zbl

[Kovács 2011b] S. J. Kovács, "Irrational centers", *Pure Appl. Math. Q.* **7**:4 (2011), 1495–1515. MR Zbl

[Kovács 2013] S. J. Kovács, "Singularities of stable varieties", pp. 159–203 in *Handbook of moduli*, vol. 2, edited by G. Farkas and I. Morrison, Advanced Lectures in Mathematics **25**, International Press, Somerville, MA, 2013. MR Zbl

[Kovács and Schwede 2011a] S. J. Kovács and K. Schwede, "Du Bois singularities deform", preprint, 2011. To appear in *Advanced Studies in Pure Mathematics*. arXiv

[Kovács and Schwede 2011b] S. J. Kovács and K. Schwede, "Hodge theory meets the minimal model program: a survey of log canonical and Du Bois singularities", pp. 51–94 in *Topology of stratified spaces*, edited by G. Friedman et al., Mathematical Sciences Research Institute Publications **58**, Cambridge University Press, 2011. MR Zbl

[Kovács et al. 2010] S. J. Kovács, K. Schwede, and K. E. Smith, "The canonical sheaf of Du Bois singularities", *Adv. Math.* **224**:4 (2010), 1618–1640. MR Zbl

[Lunts 2012] V. A. Lunts, "Categorical resolutions, poset schemes, and Du Bois singularities", *Int. Math. Res. Not.* **2012**:19 (2012), 4372–4420. MR Zbl

[Ma 2015] L. Ma, "F-injectivity and Buchsbaum singularities", *Math. Ann.* **362**:1-2 (2015), 25–42. MR Zbl

[Patakfalvi 2015] Z. Patakfalvi, "Semi-negativity of Hodge bundles associated to Du Bois families", *J. Pure Appl. Algebra* **219**:12 (2015), 5387–5393. MR Zbl

[Saito 2000] M. Saito, "Mixed Hodge complexes on algebraic varieties", *Math. Ann.* **316**:2 (2000), 283–331. MR Zbl

[Schumacher 2012] G. Schumacher, "Positivity of relative canonical bundles and applications", *Invent. Math.* **190**:1 (2012), 1–56. MR Zbl

[Schwede 2007] K. Schwede, "A simple characterization of Du Bois singularities", *Compos. Math.* **143**:4 (2007), 813–828. MR Zbl

[Schwede 2009] K. Schwede, "F-adjunction", *Algebra Number Theory* **3**:8 (2009), 907–950. MR Zbl

[Schwede and Takagi 2008] K. Schwede and S. Takagi, "Rational singularities associated to pairs", *Michigan Math. J.* **57** (2008), 625–658. MR Zbl

[Serre 1956] J.-P.. Serre, "Géométrie algébrique et géométrie analytique", *Ann. Inst. Fourier* (*Grenoble*) **6** (1956), 1–42. MR Zbl

Communicated by Ravi Vakil
Received 2015-01-23 Revised 2016-03-17 Accepted 2016-04-27

skovacs@uw.edu — *Department of Mathematics, University of Washington, Box 354350, Padelford C-138, Seattle, WA 98195, United States*

schwede@math.utah.edu — *Department of Mathematics, The University of Utah, 155 S 1400 E Room 233, Salt Lake City, UT 84112-0090, United States*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.1001

Hochschild cohomology commutes with adic completion

Liran Shaul

For a flat commutative $\Bbbk$-algebra A such that the enveloping algebra $A \otimes_{\Bbbk} A$ is noetherian, given a finitely generated bimodule M, we show that the adic completion of the Hochschild cohomology module $\mathrm{HH}^n(A/\Bbbk, M)$ is naturally isomorphic to $\mathrm{HH}^n(\hat{A}/\Bbbk, \hat{M})$. To show this, we make a detailed study of derived completion as a functor $\mathrm{D}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{Mod}\, \hat{A})$ over a nonnoetherian ring A, prove a flat base change result for weakly proregular ideals, and prove that Hochschild cohomology and analytic Hochschild cohomology of complete noetherian local rings are isomorphic, answering a question of Buchweitz and Flenner. Our results make it possible for the first time to compute the Hochschild cohomology of $\Bbbk[[t_1, \dots, t_n]]$ over any noetherian ring $\Bbbk$, and open the door for a theory of Hochschild cohomology over formal schemes.

All rings in this paper are assumed to be commutative and unital.

Introduction

Hochschild cohomology [1945] has been the prominent cohomology theory for associative algebras since its introduction. In commutative algebra and algebraic geometry, its importance was first demonstrated by the celebrated theorem of

The author acknowledges the support of the European Union for ERC Grant No. 257004-HHNcdMir.
MSC2010: primary 13D03; secondary 13J10, 14B15, 16E45, 13B35.
Keywords: Hochschild cohomology, adic completion.

Hochschild, Kostant and Rosenberg. See [Ionescu 2001] for a survey of the use of Hochschild cohomology in commutative algebra.

The aim of this paper is to initiate the study of Hochschild cohomology in the category of adic rings. Adic rings, the affine pieces of the theory of formal schemes, are by definition commutative noetherian rings A which are $\mathfrak{a}$-adically complete with respect to some ideal $\mathfrak{a} \subseteq A$. In the survey just cited, Ionescu states:

> In our survey no results about Hochschild cohomology of a topological algebra w[ere] mentioned. This is because this kind of results is missing completely.

As far as we know, little has changed regarding this statement since then. One of the main difficulties in developing such a theory can be already observed in the most simple example of an adic ring: Let $\Bbbk$ be a field of characteristic 0, and let $A = \Bbbk[[t]]$. The construction of Hochschild cohomology involves the enveloping algebra of A. But even in this simple case, the enveloping algebra $\Bbbk[[t]] \otimes_{\Bbbk} \Bbbk[[t]]$ is a nonnoetherian ring of infinite Krull dimension, so it is very difficult to do homological algebra over it. Passing to the completion of this enveloping algebra, one obtains the much more manageable completed tensor product

$$\Bbbk[[t]] \mathbin{\hat{\otimes}_{\Bbbk}} \Bbbk[[t]] \cong \Bbbk[[t_1, t_2]].$$

However, from a homological point of view, this step is highly nontrivial, as it involves the ring map from $\Bbbk[[t]] \otimes_{\Bbbk} \Bbbk[[t]]$ to its completion. Completions of nonnoetherian rings are in general poorly behaved (for instance, they need not be flat). In this paper we develop the homological tools needed to overcome this difficulty, and use them to study the Hochschild cohomology of such adic algebras.

Here is a more detailed description of the content of this paper. First, in Section 1 we review some preliminaries on Hochschild cohomology and about the derived torsion and derived completion functors. In particular, we recall the notion of a weakly proregular ideal in a commutative ring. Weak proregularity is the right condition in order for the derived torsion and derived completion functors to possess good behavior.

In Section 2, we prove that in most enveloping algebras of adic rings occurring in nature, the ideal of definition of the adic topology is weakly proregular. This result also has interesting implications in derived algebraic geometry of formal schemes. See Corollary 2.9.

Given a (not necessarily noetherian) ring A and a weakly proregular ideal $\mathfrak{a} \subseteq A$, we study in Section 3 the derived functors of the functors

$$\hat{\Gamma}_{\mathfrak{a}}(M) := \varinjlim \operatorname{Hom}_A(A/\mathfrak{a}^n, M) : \operatorname{Mod} A \to \operatorname{Mod} \hat{A}$$

and

$$\hat{\Lambda}_{\mathfrak{a}}(M) := \varprojlim A/\mathfrak{a}^n \otimes_A M : \operatorname{Mod} A \to \operatorname{Mod} \hat{A},$$

where $\hat{A}$ is the $\mathfrak{a}$-adic completion of A. To cope with the possible lack of flatness of the completion map $A \to \hat{A}$, we use DG-homological algebra techniques. From the results of this section, we deduce the following generalized Greenlees–May duality:

Corollary 3.9. *Let A be a commutative ring, let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal, and let $M, N \in \mathrm{D}(\mathrm{Mod}\, A)$. Then there are isomorphisms*

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_A(M, N)) \cong \mathrm{R}\operatorname{Hom}_A(\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(M), N) \cong \mathrm{R}\operatorname{Hom}_A(M, \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(N))$$

of functors

$$\mathrm{D}(\mathrm{Mod}\, A) \times \mathrm{D}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{Mod}\, \hat{A}).$$

Using the results of Sections 2 and 3, the main results of this paper are obtained in Section 4. First, in Theorem 4.1 we provide formulas which describe the effect of applying the derived completion functor to the Hochschild cohomology complex of a not necessarily adic algebra. The next major result reduces the problem of computing the Hochschild cohomology of an adic algebra to a problem over noetherian rings:

Corollary 4.3. *Let $\Bbbk$ be a commutative ring, and let A be a flat noetherian $\Bbbk$-algebra. Assume $\mathfrak{a} \subseteq A$ is an ideal, such that A is $\mathfrak{a}$-adically complete, and such that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$. Let $I := \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}$, and set $A \hat{\otimes}_{\Bbbk} A := \Lambda_I(A \otimes_{\Bbbk} A)$. Then for any $M \in \mathrm{Mod}\, A \otimes_{\Bbbk} A$ which is I-adically complete (for example, any $\mathfrak{a}$-adically complete A-module, or more particularly, any finitely generated A-module), there is a functorial isomorphism*

$$\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, M) \cong \mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(A, M)$$

in $\mathrm{D}(\mathrm{Mod}\, A)$, and the ring $A \hat{\otimes}_{\Bbbk} A$ is noetherian.

Using the results of [Buchweitz and Flenner 2006], as a corollary of this result, we are able to prove in Corollary 4.5 that Hochschild cohomology and analytic Hochschild cohomology of complete noetherian local algebras coincide, answering a question of Buchweitz and Flenner. Finally, Section 4 ends with a theorem which proves the result mentioned in the title of the paper. More precisely, we show:

Theorem 4.13. *Let $\Bbbk$ be a commutative ring, and let A be a flat noetherian $\Bbbk$-algebra such that $A \otimes_{\Bbbk} A$ is noetherian. Let $\mathfrak{a} \subseteq A$ be an ideal, and let M be a finitely generated $(A \otimes_{\Bbbk} A)$-module. Then for any $n \in \mathbb{N}$, there is a functorial isomorphism*

$$\Lambda_{\mathfrak{a}}\big(\operatorname{Ext}^n_{A \otimes_{\Bbbk} A}(A, M)\big) \cong \operatorname{Ext}^n_{\hat{A} \otimes_{\Bbbk} \hat{A}}(\hat{A}, \hat{M}).$$

If, moreover, either

(1) *$\Bbbk$ is a field, or*

(2) *A is projective over $\Bbbk$, $\mathfrak{a}$ is a maximal ideal, and M is a finitely generated A-module,*

then there is also a functorial isomorphism

$$\Lambda_{\mathfrak{a}}(\mathrm{HH}^n(A/\Bbbk, M)) \cong \mathrm{HH}^n(\hat{A}/\Bbbk, \hat{M}).$$

In the short and final Section 5, we briefly discuss analogous results for Hochschild homology.

Warning. Contrary to the convention in many papers in the field, unless stated otherwise, we do not assume that rings are noetherian.

1. Preliminaries on completion, torsion and Hochschild cohomology

Given a commutative ring A, we denote by $\operatorname{Mod} A$ the abelian category of A-modules, and by $\mathrm{D}(\operatorname{Mod} A)$ its (unbounded) derived category. If A is noetherian, we will denote by $\mathrm{D}_{\mathrm{f}}(\operatorname{Mod} A)$ the triangulated subcategory made of complexes with finitely generated cohomologies. We will freely use resolutions of unbounded complexes, following [Spaltenstein 1988].

Completion and torsion. References for the material in this section are [Alonso Tarrío et al. 1997; 1999; Greenlees and May 1992; Porta et al. 2014b; 2015; Schenzel 2003; Simon 1990; Yekutieli 2011]. See [Porta et al. 2014b, Remark 7.14] for a brief discussion on the history of this material. Let A be a commutative ring, and let $\mathfrak{a} \subseteq A$ be a finitely generated ideal. The $\mathfrak{a}$-torsion functor $\Gamma_{\mathfrak{a}}(-) : \operatorname{Mod} A \to \operatorname{Mod} A$ is defined by

$$\Gamma_{\mathfrak{a}}(M) := \varinjlim \operatorname{Hom}_A(A/\mathfrak{a}^n, M).$$

This functor is a left exact additive functor. We denote its (total) right derived functor by $\mathrm{R}\Gamma_{\mathfrak{a}} : \mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{Mod} A)$. It is computed using K-injective resolutions. See [Brodmann and Sharp 2013] for a detailed study of the $\mathfrak{a}$-torsion functor and its derived functor in the noetherian case. More important in this paper is the $\mathfrak{a}$-adic completion functor, defined by

$$\Lambda_{\mathfrak{a}}(-) : \operatorname{Mod} A \to \operatorname{Mod} A, \qquad \Lambda_{\mathfrak{a}}(M) := \varprojlim A/\mathfrak{a}^n \otimes_A M.$$

This functor is additive, but in general is neither left exact nor right exact (even when A is noetherian; see [Yekutieli 2011, Example 3.20]). It does however preserve surjections. We denote by $\mathrm{L}\Lambda_{\mathfrak{a}} : \mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{Mod} A)$ its left derived functor. By [Alonso Tarrío et al. 1997, Section 1], it can be computed using K-flat resolutions. Both of the functors $\Gamma_{\mathfrak{a}}(-)$, $\Lambda_{\mathfrak{a}}(-)$ are idempotent [Yekutieli 2011, Corollary 3.6].

For any ring A, the A-module $\Lambda_{\mathfrak{a}}(A)$ has the structure of a commutative A-algebra, called the completion of A. If A is noetherian then $\Lambda_{\mathfrak{a}}(A)$ is flat over A, but if A is not noetherian this does not always holds. For example, if A is any countable ring which is not coherent, then the completion map $A[x] \to A[\![x]\!]$ is not flat [Stacks 2005–, Tag 0AL8]. The ring $\Lambda_{\mathfrak{a}}(A)$ is noetherian if and only if the ring $A/\mathfrak{a}$ is noetherian [Stacks 2005–, Tag 05GH]. If A is noetherian and M is a finitely generated A-module, then there is an isomorphism of functors $\Lambda_{\mathfrak{a}}(M) \cong \Lambda_{\mathfrak{a}}(A) \otimes_A M$, so in particular in that case, $\Lambda_{\mathfrak{a}}(-)$ is exact on the category of finitely generated A-modules [Stacks 2005–, Tag 00MB]. We will sometimes denote by $\hat{A}$ the A-algebra $\Lambda_{\mathfrak{a}}(A)$ and by $\hat{M}$ the A-module $\Lambda_{\mathfrak{a}}(M)$. For any ring A, the A-modules $\Gamma_{\mathfrak{a}}(M)$ and $\Lambda_{\mathfrak{a}}(M)$ carry naturally the structure of $\hat{A}$-modules, and so one may view the $\mathfrak{a}$-torsion and $\mathfrak{a}$-completion functors as functors $\operatorname{Mod} A \to \operatorname{Mod} \hat{A}$. Section 3 is dedicated to a study of the functors obtained from this observation.

Given a ring A, and an element $\mathfrak{a} \in A$, the infinite dual Koszul complex associated to it is

$$\mathrm{K}^{\vee}_{\infty}(A; (a)) := \big(\cdots \to 0 \to A \to A[a^{-1}] \to 0 \to \cdots\big)$$

concentrated in degrees 0, 1. If $(a_1, \ldots, a_n)$ is a finite sequence of elements in A, then the infinite dual Koszul complex associated to it is

$$\mathrm{K}^{\vee}_{\infty}(A; (a_1, \ldots, a_n)) := \mathrm{K}^{\vee}_{\infty}(A; (a_1)) \otimes_A \cdots \otimes_A \mathrm{K}^{\vee}_{\infty}(A; (a_n)).$$

It is a bounded complex of flat A-modules. Given an ideal $\mathfrak{a} \subseteq A$, and a finite sequence $\boldsymbol{a}$ of elements of A that generate $\mathfrak{a}$, by [Porta et al. 2014b, Corollary 4.26], there is a morphism of functors

$$\mathrm{R}\Gamma_{\mathfrak{a}}(-) \to \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A -.$$

The sequence $\boldsymbol{a}$ is called weakly proregular if this morphism is an isomorphism of functors. This notion is actually independent of $\boldsymbol{a}$, and depends only on the ideal $\mathfrak{a}$ generated by it [Schenzel 2003, Lemma 3.3]. Hence, we say a finitely generated ideal $\mathfrak{a}$ is weakly proregular if some (or, equivalently, any) finite sequence that generates it is weakly proregular. In a noetherian ring, any ideal and any finite sequence are weakly proregular, but there are examples of finitely generated (even principal) ideals in nonnoetherian rings which are not weakly proregular.

Given a ring A and a finite sequence $\boldsymbol{a}$ of elements of A, the infinite dual Koszul complex has an explicit free resolution, called the telescope complex and denoted by $\operatorname{Tel}(A; \boldsymbol{a})$. This resolution is a bounded complex of countably generated free A-modules [Porta et al. 2014b, Lemma 5.7]. In particular, if the ideal $\mathfrak{a}$ generated by $\boldsymbol{a}$ is weakly proregular, then there is also an isomorphism of functors

$$\mathrm{R}\Gamma_{\mathfrak{a}}(-) \cong \operatorname{Tel}(A; \boldsymbol{a}) \otimes_A -.$$

Moreover, in this case, by [Porta et al. 2014b, Corollary 5.25], there is also an isomorphism of functors

$$\mathrm{L}\Lambda_{\mathfrak{a}}(-) \cong \operatorname{Hom}_A(\operatorname{Tel}(A; \boldsymbol{a}), -).$$

It follows that if A is a commutative ring, and $\mathfrak{a}$ is a weakly proregular ideal, then both of the functors $\mathrm{R}\Gamma_{\mathfrak{a}}$, $\mathrm{L}\Lambda_{\mathfrak{a}}$ have finite cohomological dimension, and there is a bifunctorial isomorphism, the Greenlees–May duality,

$$\mathrm{R}\operatorname{Hom}_A(\mathrm{R}\Gamma_{\mathfrak{a}}(M), N) \cong \mathrm{R}\operatorname{Hom}_A(M, \mathrm{L}\Lambda_{\mathfrak{a}}(N))$$

for any $M, N \in \mathrm{D}(\operatorname{Mod} A)$.

Both the infinite dual Koszul complex and the telescope complex enjoy the following base change property: if A is a ring, $\boldsymbol{a}$ is a finite sequence of elements in A, $A \to B$ is a ring map, and $\boldsymbol{b}$ is the image of $\boldsymbol{a}$ under this map, then there are isomorphisms

$$\mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A B \cong \mathrm{K}^{\vee}_{\infty}(B; \boldsymbol{b}), \quad \operatorname{Tel}(A; \boldsymbol{a}) \otimes_A B \cong \operatorname{Tel}(B; \boldsymbol{b})$$

of complexes of B-modules.

For any complex $M \in \mathrm{D}(\operatorname{Mod} A)$, there are canonical maps

$$\mathrm{R}\Gamma_{\mathfrak{a}}(M) \to M, \tag{1.1}$$

and

$$M \to \mathrm{L}\Lambda_{\mathfrak{a}}(M). \tag{1.2}$$

The complex M is called cohomologically $\mathfrak{a}$-torsion (resp. cohomologically $\mathfrak{a}$-adically complete) if the map (1.1) (resp. (1.2)) is an isomorphism. If $\mathfrak{a}$ is weakly proregular then the functors $\mathrm{R}\Gamma_{\mathfrak{a}}$ and $\mathrm{L}\Lambda_{\mathfrak{a}}$ are idempotent [Porta et al. 2014b, Corollary 4.30, Proposition 7.10] and it follows that in this case the collection of all cohomologically $\mathfrak{a}$-torsion (resp. cohomologically $\mathfrak{a}$-adically complete) complexes is a triangulated subcategory of $\mathrm{D}(\operatorname{Mod} A)$ which is equal to the essential image of the functor $\mathrm{R}\Gamma_{\mathfrak{a}}$ (resp. $\mathrm{L}\Lambda_{\mathfrak{a}}$). Moreover, by [Porta et al. 2014b, Theorem 7.11], in this case these categories are equivalent (the Matlis–Greenlees–May equivalence).

Hochschild cohomology. Let $\Bbbk$ be a commutative ring, and let A be a commutative $\Bbbk$-algebra. We let

$$A^{\otimes^n_{\Bbbk}} := \underbrace{A \otimes_{\Bbbk} \cdots \otimes_{\Bbbk} A}_{n},$$

and denote by $\mathcal{B}$ the bar resolution

$$\cdots \to A^{\otimes^n_{\Bbbk}} \to \cdots \to A^{\otimes^2_{\Bbbk}} \to A \to 0.$$

Given an A-bimodule M, the n-th Hochschild cohomology module of A over $\Bbbk$ with coefficients in M is given by

$$\mathrm{HH}^n(A/\Bbbk, M) := H^n \operatorname{Hom}_{A \otimes_{\Bbbk} A}(\mathcal{B}, M).$$

See [Cartan and Eilenberg 1956, Chapter IX], [Loday 1998, Chapter 1] and [Weibel 1994, Chapter 9] for more details on this classical construction. If A is projective (resp. flat) over $\Bbbk$, then $\mathcal{B}$ is a projective (resp. flat) resolution of A over the enveloping algebra $A \otimes_{\Bbbk} A$. Hence, in the projective case, the natural map

$$\mathrm{HH}^n(A/\Bbbk, M) \to \mathrm{Ext}^n_{A \otimes_{\Bbbk} A}(A, M)$$

is an isomorphism. When A is only flat over $\Bbbk$, but not necessarily projective, this map might fail in general to be an isomorphism. Nevertheless, the modules on the right-hand side are interesting on their own, and are sometimes referred to in the literature as the derived Hochschild (or Shukla) cohomology modules of A over $\Bbbk$. In this paper we will focus mostly on these modules[1] and, a bit more generally, on the complex $\mathrm{R}\,\mathrm{Hom}_{A \otimes_{\Bbbk} A}(A, M)$. Somewhat imprecisely, we will refer to

$$\mathrm{R}\,\mathrm{Hom}_{A \otimes_{\Bbbk} A}(A, M)$$

as the Hochschild complex of A with coefficients in M even when A is only flat over $\Bbbk$. We will however use the notation $\mathrm{HH}^n(A/\Bbbk, M)$ to denote only the classical Hochschild cohomology modules.

2. Weak proregularity and flat base change

Let $\Bbbk$ be a base commutative ring, and let A, B be two flat $\Bbbk$-algebras. Assume that A and B are equipped with adic topologies, generated by finitely generated ideals $\mathfrak{a} \subseteq A$ and $\mathfrak{b} \subseteq B$. In that case, the tensor product $A \otimes_{\Bbbk} B$ is also naturally equipped with an adic topology. It is generated by the finitely generated ideal $\mathfrak{a} \otimes_{\Bbbk} B + A \otimes_{\Bbbk} \mathfrak{b} \subseteq A \otimes_{\Bbbk} B$. The aim of this section is to discuss the question of when this ideal is weakly proregular. We allow A to be different from B, although we will only use the case $A = B$ in the rest of the paper.

Recall that a ring $\Bbbk$ is called absolutely flat (or Von Neumann regular) if every $\Bbbk$-module is flat. Over such rings, the above question is easy:

Proposition 2.1. *Let $\Bbbk$ be an absolutely flat ring. Let A, B be two $\Bbbk$-algebras, and let $\mathfrak{a} \subseteq A$ and $\mathfrak{b} \subseteq B$ be weakly proregular ideals. Then the ideal*

$$\mathfrak{a} \otimes_{\Bbbk} B + A \otimes_{\Bbbk} \mathfrak{b} \subseteq A \otimes_{\Bbbk} B$$

is weakly proregular.

Proof. In the case where $\Bbbk$ is a field, and A and B are noetherian and complete with respect to the adic topology, this is shown in [Porta et al. 2014b, Example 4.35], and the proof there remains true under the above assumptions. □

Remark 2.2. Assume A is a ring, $\mathfrak{a} \subseteq A$ is a weakly proregular ideal, B is a flat A-algebra, and $\mathfrak{b} = \mathfrak{a} \cdot B$. Then by [Alonso Tarrío et al. 1997, Example 3.0(B)], the ideal $\mathfrak{b}$ is also weakly proregular.

[1]See, however, Corollary 4.5 and Theorem 4.13 where even in the possible absence of projectivity we will discuss classical Hochschild cohomology.

Recall that if A is a noetherian ring, with $\mathfrak{a} \subseteq A$ an ideal, and if I is an injective A-module, then $\Gamma_{\mathfrak{a}}(I)$ is also an injective A-module (for example, by [Hartshorne 1977, Lemma 3.2]). We now state and prove a weaker form of this fact in the case when A is not necessarily noetherian, but $\Lambda_{\mathfrak{a}}(A)$ is.

If A is a ring, with $\mathfrak{a} \subseteq A$ a finitely generated ideal, and if M is an A-module, then M is called $\mathfrak{a}$-flasque if, for each $k > 0$, we have $\mathrm{H}^k_{\mathfrak{a}}(M) = 0$, where $\mathrm{H}^k_{\mathfrak{a}}(M) := \mathrm{H}^k(\mathrm{R}\Gamma_{\mathfrak{a}}(M))$. Any injective module is $\mathfrak{a}$-flasque. If M is $\mathfrak{a}$-flasque, then the canonical morphism $\Gamma_{\mathfrak{a}}(M) \to \mathrm{R}\Gamma_{\mathfrak{a}}(M)$ is an isomorphism. By [Brodmann and Sharp 2013, Theorem 3.4.10], the direct limit of $\mathfrak{a}$-flasque modules is $\mathfrak{a}$-flasque.

Lemma 2.3. *Let A be a ring, and let $\mathfrak{a}, \mathfrak{b} \subseteq A$ be two finitely generated ideals. Suppose that the ring $\hat{A} = \Lambda_{\mathfrak{a}}(A)$ is noetherian. Let $\hat{\mathfrak{b}} = \mathfrak{b}\hat{A}$. Then for any injective A-module I, the $\hat{A}$-module $\hat{\Gamma}_{\mathfrak{a}} I$ is $\hat{\mathfrak{b}}$-flasque.*

Proof. Let $A_j = A/\mathfrak{a}^{j+1}$. Since $\mathfrak{a}$ is finitely generated, there is an isomorphism $A_j \cong \hat{A}/(\mathfrak{a}\hat{A})^{j+1}$. Note that, by assumption, A_j is noetherian. Let $\hat{\mathfrak{b}}_j$ be the image of $\hat{\mathfrak{b}}$ in A_j. Let $I_j = \mathrm{Hom}_A(A_j, I)$. Then $\hat{\Gamma}_{\mathfrak{a}} I = \varinjlim I_j$, so it is enough to show that I_j is $\hat{\mathfrak{b}}$-flasque. Note also that I_j is an injective A_j-module. Let $k > 0$, let $\hat{\boldsymbol{b}}$ be a finite sequence generating $\hat{\mathfrak{b}}$, and let $\boldsymbol{b}_j$ be its image in A_j. Since $\hat{A}$ is noetherian, $\hat{\mathfrak{b}}$ is weakly proregular, so that

$$H^k_{\hat{\mathfrak{b}}}(I_j) \cong H^k\big(\mathrm{K}^\vee_\infty(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} I_j\big) \cong H^k\big(\mathrm{K}^\vee_\infty(A_j; \boldsymbol{b}_j) \otimes_{A_j} I_j\big) \cong H^k_{\hat{\mathfrak{b}}_j}(I_j),$$

where the last isomorphism follows from the fact that A_j is noetherian, so that $\hat{\mathfrak{b}}_j$ is weakly proregular. Since I_j is injective over A_j, it follows that $H^k_{\hat{\mathfrak{b}}_j}(I_j) = 0$ for all $k > 0$, which proves the claim. □

Proposition 2.4. *Let A be a commutative ring, and let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal such that $\Lambda_{\mathfrak{a}}(A)$ is noetherian. Let $\mathfrak{b} \subseteq A$ be an ideal containing $\mathfrak{a}$. Then $\mathfrak{b}$ is also weakly proregular.*

Proof. We keep the notation of Lemma 2.3. It is clear that $A/\mathfrak{b}$ is noetherian. Let $\boldsymbol{a}$ be a finite sequence generating $\mathfrak{a}$, and let $\boldsymbol{b}$ be a finite sequence generating $\mathfrak{b}$. Let I be an injective A-module. By [Schenzel 2003, Theorem 1.1], it is enough to show that

$$H^k(\mathrm{Tel}(A; \boldsymbol{b}) \otimes_A I) = 0$$

for all $k \neq 0$.

Since $\mathfrak{a} \subseteq \mathfrak{b}$, the ideal generated by the concatenated sequence $(\boldsymbol{a}, \boldsymbol{b})$ is equal to the ideal generated by $\boldsymbol{b}$, so there is a homotopy equivalence $\mathrm{Tel}(A; (\boldsymbol{a}, \boldsymbol{b})) \cong \mathrm{Tel}(A; \boldsymbol{b})$. Hence, there is an isomorphism

$$\mathrm{Tel}(A; \boldsymbol{b}) \otimes_A I \cong \mathrm{Tel}(A; (\boldsymbol{a}, \boldsymbol{b})) \otimes_A I \cong \mathrm{Tel}(A; \boldsymbol{b}) \otimes_A \mathrm{Tel}(A; \boldsymbol{a}) \otimes_A I$$

in $\mathrm{D}(\mathrm{Mod}\, A)$. Since $\boldsymbol{a}$ is a weakly proregular sequence, I is an injective A-module, and $\mathrm{Tel}(A; \boldsymbol{b})$ is a bounded complex of flat modules, the latter is isomorphic in $\mathrm{D}(\mathrm{Mod}\, A)$ to

$$\mathrm{Tel}(A; \boldsymbol{b}) \otimes_A \Gamma_{\mathfrak{a}} I.$$

Thus, it is enough to show that all the cohomologies (except the zeroth) of the complex of A-modules $\mathrm{Tel}(A; \boldsymbol{b}) \otimes_A \Gamma_{\mathfrak{a}} I$ vanish. Let

$$\mathrm{Rest}_{\hat{A}/A} : \mathrm{D}(\mathrm{Mod}\, \hat{A}) \to \mathrm{D}(\mathrm{Mod}\, A)$$

be the forgetful functor, and let $\hat{\boldsymbol{b}}$ be the image of the sequence $\boldsymbol{b}$ in $\hat{A}$. Consider the complex

$$\mathrm{Tel}(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} \hat{\Gamma}_{\mathfrak{a}} I \in \mathrm{D}(\mathrm{Mod}\, \hat{A}).$$

We claim that

$$\mathrm{Rest}_{\hat{A}/A}\big(\mathrm{Tel}(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} \hat{\Gamma}_{\mathfrak{a}} I\big) = \mathrm{Tel}(A; \boldsymbol{b}) \otimes_A \Gamma_{\mathfrak{a}} I. \tag{2.5}$$

Indeed, by the base change property of the telescope complex, we have an isomorphism

$$\mathrm{Tel}(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} \hat{\Gamma}_{\mathfrak{a}} I \cong \mathrm{Tel}(A; \boldsymbol{b}) \otimes_A \hat{\Gamma}_{\mathfrak{a}} I$$

of complexes in $\mathrm{D}(\mathrm{Mod}\, \hat{A})$. So using the fact that

$$\mathrm{Rest}_{\hat{A}/A}(\hat{\Gamma}_{\mathfrak{a}} I) = \Gamma_{\mathfrak{a}} I$$

we obtain (2.5). Since for a complex $M \in \mathrm{D}(\mathrm{Mod}\, \hat{A})$ we have

$$H^k(M) = 0 \quad \text{if and only if} \quad H^k\big(\mathrm{Rest}_{\hat{A}/A}(M)\big) = 0,$$

it is enough to show that $H^k\big(\mathrm{Tel}(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} \hat{\Gamma}_{\mathfrak{a}} I\big) = 0$ for all $k \neq 0$. By weak proregularity of the sequence $\hat{\boldsymbol{b}}$, there is an isomorphism

$$\mathrm{Tel}(\hat{A}; \hat{\boldsymbol{b}}) \otimes_{\hat{A}} \hat{\Gamma}_{\mathfrak{a}} I \cong \mathrm{R}\Gamma_{\hat{\mathfrak{b}}} \hat{\Gamma}_{\mathfrak{a}} I$$

in $\mathrm{D}(\mathrm{Mod}\, \hat{A})$. By Lemma 2.3, this is isomorphic in $\mathrm{D}(\mathrm{Mod}\, \hat{A})$ to $\Gamma_{\hat{\mathfrak{b}}} \hat{\Gamma}_{\mathfrak{a}} I$. Since this complex is clearly concentrated in degree zero, it follows that all of its cohomologies except the zeroth vanish, which proves the result. ☐

Here is the main result of this section:

Theorem 2.6. *Let $\Bbbk$ be a commutative ring, let A be a flat noetherian $\Bbbk$-algebra, and let $\mathfrak{a} \subseteq A$ be an ideal such that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$. Let B be a flat noetherian $\Bbbk$-algebra, and let $\mathfrak{b} \subseteq B$ be an ideal. Then the ideal*

$$I := \mathfrak{a} \otimes_{\Bbbk} B + A \otimes_{\Bbbk} \mathfrak{b} \subseteq A \otimes_{\Bbbk} B$$

is weakly proregular.

Proof. According to Remark 2.2, the ideal $I_1 := \mathfrak{a} \otimes_{\Bbbk} B \subseteq A \otimes_{\Bbbk} B$ is weakly proregular. Since B is flat over $\Bbbk$, we have $(A \otimes_{\Bbbk} B)/I_1 \cong A/\mathfrak{a} \otimes_{\Bbbk} B$, and as $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$, it follows that $(A \otimes_{\Bbbk} B)/I_1$ is noetherian. Hence, $\Lambda_{I_1}(A \otimes_{\Bbbk} B)$ is also noetherian. Since $I_1 \subseteq I$, the result follows from Proposition 2.4. □

Remark 2.7. The assumption that A is noetherian in the above result can be relaxed: it is enough to assume that $\mathfrak{a}$ is weakly proregular. It is an open problem to the author if the above result remains true without the assumption that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$ (as in Proposition 2.1).

Remark 2.8. Let $\Bbbk$ be a commutative ring, and let A, B be two commutative noetherian $\Bbbk$-algebras which are adically complete with respect to ideals $\mathfrak{a} \subseteq A$, $\mathfrak{b} \subseteq B$. In this situation, Grothendieck and Dieudonné [1960, Section 10.7] defined the fiber product of the two affine formal schemes $\operatorname{Spf} A$, $\operatorname{Spf} B$ over $\Bbbk$ to be the formal spectrum of the ring $\Lambda_{\mathfrak{a}\otimes_{\Bbbk}B+A\otimes_{\Bbbk}\mathfrak{b}}(A \otimes_{\Bbbk} B)$.

Now, we switch to the point of view of derived algebraic geometry, and assume that A and B are flat over $\Bbbk$. Forgetting the adic structure on A, B, the flatness assumption ensures that, in this situation, the usual fiber product $\operatorname{Spec}(A \otimes_{\Bbbk} B)$ of the schemes $\operatorname{Spec}(A)$ and $\operatorname{Spec}(B)$ coincides with their derived fiber product. Returning to the adic situation, Lurie defined [2011, Section 4.2] a notion of a derived completion of an (E_∞) ring, and showed [2011, Section 4.3] that if the ring is noetherian then its derived completion coincides with its ordinary completion. Recently, using our [Porta et al. 2014a, Theorem 4.2], it was shown in [Braunling et al. 2015, Proposition 5.4] that if R is any commutative ring, and if $I \subseteq R$ is a weakly proregular ideal, then the derived I-completion of R coincides with its ordinary I-adic completion. Hence, the results of this section imply the following:

Corollary 2.9. *Let $\Bbbk$ be a commutative ring, and let A, B be noetherian flat $\Bbbk$-algebras which are adically complete with respect to ideals $\mathfrak{a} \subseteq A$, $\mathfrak{b} \subseteq B$. Assume further that either $\Bbbk$ is an absolutely flat ring (e.g., a field) or that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$. Then the derived fiber product of the formal schemes* $\operatorname{Spf} A$, $\operatorname{Spf} B$ *over $\Bbbk$ is equal to the formal spectrum of* $\Lambda_{\mathfrak{a}\otimes_{\Bbbk}B+A\otimes_{\Bbbk}\mathfrak{b}}(A \otimes_{\Bbbk} B)$.

3. The functors $\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}$, $\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}$

Let A be a commutative ring, let $\mathfrak{a} \subseteq A$ be a finitely generated ideal, and let $\hat{A}$ be the $\mathfrak{a}$-adic completion of A. For any A-module M, the A-modules $\Gamma_{\mathfrak{a}}(M)$ and $\Lambda_{\mathfrak{a}}(M)$ carry naturally $\hat{A}$-module structures, and one obtains additive functors $\hat{\Gamma}_{\mathfrak{a}}, \hat{\Lambda}_{\mathfrak{a}} : \operatorname{Mod} A \to \operatorname{Mod} \hat{A}$, defined by the same formulas as $\Gamma_{\mathfrak{a}}$ and $\Lambda_{\mathfrak{a}}$. These functors have derived functors $\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}, \mathrm{L}\hat{\Lambda}_{\mathfrak{a}} : \mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{Mod} \hat{A})$, calculated using K-injective and K-flat resolutions, respectively. This section is dedicated to a study of these functors.

Keeping an eye towards the main goal of this text, we must avoid assuming that A is noetherian. Hence, we do not know if the completion map $A \to \hat{A}$ is flat. We overcome this issue by using DG-algebras, which will be assumed to be (graded-) commutative. We refer the reader to [Avramov 1998; Keller 1994; Mac Lane 1963; Yekutieli 2016] for information about DG-algebras and their derived categories. For a DG-algebra A, we denote by $\operatorname{DGMod} A$ the category of DG-modules over A, and by $\mathrm{D}(\operatorname{DGMod} A)$ the derived category over A.

We shall need the following well known result from DG-homological algebra:

Proposition 3.1. *Let $A \to B$ be a quasiisomorphism between two commutative DG-algebras, and let*

$$\operatorname{Rest}_{B/A} : \mathrm{D}(\operatorname{DGMod} B) \to \mathrm{D}(\operatorname{DGMod} A)$$

be the forgetful functor.

(1) *There is an isomorphism*

$$1_{\mathrm{D}(\operatorname{DGMod} B)} \cong B \otimes^{\mathrm{L}}_A \operatorname{Rest}_{B/A}(-)$$

of functors $\mathrm{D}(\operatorname{DGMod} B) \to \mathrm{D}(\operatorname{DGMod} B)$.

(2) *There is an isomorphism*

$$1_{\mathrm{D}(\operatorname{DGMod} B)} \cong \mathrm{R}\operatorname{Hom}_A(B, \operatorname{Rest}_{B/A}(-))$$

of functors $\mathrm{D}(\operatorname{DGMod} B) \to \mathrm{D}(\operatorname{DGMod} B)$.

Proof. Part (1) follows immediately from [Stacks 2005–, Tag 09S6], or [Yekutieli 2016, Proposition 2.5(1)], while part (2) follows immediately from [Shaul 2016, Lemma 2.2], or [Yekutieli 2016, Proposition 2.5(2)]. □

As far as we know, the next results are new even in the case where A is noetherian. In the noetherian case, one does not need DG-algebras in the proof of the next result.

Theorem 3.2. *Let A be a commutative ring, let $\mathfrak{a} \subseteq A$ be a finitely generated ideal, and let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$. Assume that $\mathfrak{a}$ is weakly proregular. Then there is an isomorphism of functors*

$$\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(-) \cong \hat{A} \otimes^{\mathrm{L}}_A (\mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A -).$$

Proof. Set $\hat{A} := \Lambda_{\mathfrak{a}}(A)$. Consider the completion map $A \to \hat{A}$. Since A is not necessarily noetherian, this map might fail to be flat, so let $A \xrightarrow{f} \tilde{A} \xrightarrow{g} \hat{A}$ be a K-flat DG-algebra resolution of $A \to \hat{A}$. That is, $f : A \to \tilde{A}$ is a K-flat DG-algebra map, $g : \tilde{A} \to \hat{A}$ is a quasiisomorphism of DG-algebras, and $g \circ f$ is equal to the completion map $A \to \hat{A}$. We denote by

$$\operatorname{Rest}_{\hat{A}/\tilde{A}} : \mathrm{D}(\operatorname{Mod} \hat{A}) \to \mathrm{D}(\operatorname{DGMod} \tilde{A}) \quad \text{and} \quad \operatorname{Rest}_{\tilde{A}/A} : \mathrm{D}(\operatorname{DGMod} \tilde{A}) \to \mathrm{D}(\operatorname{Mod} A)$$

the corresponding forgetful functors. Set

$$\mathrm{R}\tilde{\Gamma}_{\mathfrak{a}}(-) := \mathrm{Rest}_{\hat{A}/\tilde{A}} \circ \mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(-) : \mathrm{D}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{DGMod}\, \tilde{A}).$$

Let $M \in \mathrm{D}(\mathrm{Mod}\, A)$. Let $P \to M$ be a K-flat resolution of M, and let $M \to I$ be a K-injective resolution of M. The map $f : A \to \tilde{A}$ induces a map $1_P \otimes_A f : P \to P \otimes_A \tilde{A}$. Let $P \otimes_A \tilde{A} \to J$ be a K-injective resolution of $P \otimes_A \tilde{A}$ over $\tilde{A}$. Because f is flat, J is also a K-injective resolution of $P \otimes_A \tilde{A}$ over A. There is a unique map $\phi : I \to J$ in $\mathrm{K}(\mathrm{Mod}\, A)$, which makes the diagram

$$\begin{tikzcd} P \arrow[r, "1_P \otimes_A f"] \arrow[d] & P \otimes_A \tilde{A} \arrow[d] \\ I \arrow[r, "\phi"] & J \end{tikzcd} \tag{3.3}$$

commutative, and it induces a map $\Gamma_{\mathfrak{a}}(\phi) : \Gamma_{\mathfrak{a}}(I) \to \Gamma_{\mathfrak{a}}(J)$. Our goal is to show that $\Gamma_{\mathfrak{a}}(\phi)$ is a quasiisomorphism. The morphism of functors

$$\alpha(-) : \Gamma_{\mathfrak{a}}(-) \to \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A -$$

that was constructed in [Porta et al. 2014b, Equation (4.19)], and the map ϕ induce the commutative diagram

$$\begin{tikzcd} \Gamma_{\mathfrak{a}}(I) \arrow[r, "\Gamma_{\mathfrak{a}}(\phi)"] \arrow[d, "\alpha_I"] & \Gamma_{\mathfrak{a}}(J) \arrow[d, "\alpha_J"] \\ \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A I \arrow[r, "1 \otimes_A \phi"] & \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A J \end{tikzcd} \tag{3.4}$$

Because $\boldsymbol{a}$ is weakly proregular, the two vertical maps are quasiisomorphisms. We claim that the bottom horizontal map is also a quasiisomorphism. To see this, consider the following commutative diagram in $\mathrm{K}(\mathrm{Mod}\, A)$:

$$\begin{tikzcd} \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A I \arrow[r, "1"] & \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A J \\ \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A P \arrow[u, "2"] \arrow[r, "3"] & \mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A P \otimes_A \tilde{A} \arrow[u, "4"'] \\ \mathrm{Tel}(A; \boldsymbol{a}) \otimes_A P \arrow[u, "6"] \arrow[r, "5"] \arrow[d, "7"] & \mathrm{Tel}(A; \boldsymbol{a}) \otimes_A P \otimes_A \tilde{A} \arrow[u, "8"'] \arrow[d, "9"] \\ \mathrm{Tel}(A; \boldsymbol{a}) \otimes_A P \otimes_A \mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), A) \arrow[r, "10"] & \mathrm{Tel}(A; \boldsymbol{a}) \otimes_A P \otimes_A \hat{A} \end{tikzcd}$$

The top square in this diagram is induced from the square (3.3), the middle square is induced from the quasiisomorphism

$$\operatorname{Tel}(A;\boldsymbol{a}) \to \mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}),$$

and the bottom square is induced from the commutative diagram [Porta et al. 2014b, Equation (5.26)]. By [Porta et al. 2014b, Corollary 5.23], $\operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a}), A) \to \hat{A}$ is a quasiisomorphism. Since $\operatorname{Tel}(A;\boldsymbol{a})$ and P are both K-flat, it follows that (10) is a quasiisomorphism. K-flatness of P also implies that (9) is a quasiisomorphism. The map (7), which is induced by the map $A \to \operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a}), A)$ is a quasiisomorphism by [Alonso Tarrío et al. 1997, Corollary after Theorem (0.3)*] (or the proof of [Porta et al. 2014b, Lemma 7.6]). Hence, the map (5) is also a quasiisomorphism. It is clear that (6) and (8) are quasiisomorphisms, so that (3) is also a quasiisomorphism. As (2) and (4) are also quasiisomorphisms, we deduce that (1) is a quasiisomorphism. Returning to the commutative diagram (3.4), we deduce that the map

$$\Gamma_{\mathfrak{a}}(\phi) : \Gamma_{\mathfrak{a}}(I) \to \Gamma_{\mathfrak{a}}(J)$$

is a quasiisomorphism.

There are functorial isomorphisms in $\mathrm{D}(\operatorname{DGMod} \tilde{A})$:

$$\mathrm{R}\tilde{\Gamma}_{\mathfrak{a}}(M) = \operatorname{Rest}_{\hat{A}/\tilde{A}}(\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(M)) \cong \operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(I)).$$

Since the map $\Gamma_{\mathfrak{a}}(\phi) : \Gamma_{\mathfrak{a}}(I) \to \Gamma_{\mathfrak{a}}(J)$ is a quasiisomorphism, it follows that

$$\operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(\phi)) : \operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(I)) \to \operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(J))$$

is also a quasiisomorphism. The DG $\tilde{A}$-module $\operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(J))$ is a sub-DG-module of J, and the inclusion map induces a map

$$\operatorname{Rest}_{\hat{A}/\tilde{A}}(\hat{\Gamma}_{\mathfrak{a}}(J)) \to \mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}) \otimes_A J. \tag{3.5}$$

Applying the forgetful functor $\operatorname{Rest}_{\tilde{A}/A}$ to the map in (3.5) yields the quasiisomorphism

$$\Gamma_{\mathfrak{a}}(J) \cong \mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}) \otimes_A J,$$

so that the map in (3.5) is also a quasiisomorphism. Hence,

$$\mathrm{R}\tilde{\Gamma}_{\mathfrak{a}}(M) \cong \mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}) \otimes_A J \cong \mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}) \otimes_A M \otimes_A \tilde{A}.$$

By Proposition 3.1, there is an isomorphism of functors

$$1_{\mathrm{D}(\operatorname{DGMod} \hat{A})} \cong \hat{A} \otimes^{\mathrm{L}}_{\tilde{A}} \operatorname{Rest}_{\hat{A}/\tilde{A}}(-).$$

Hence,

$$\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(-) \cong \hat{A} \otimes^{\mathrm{L}}_{\tilde{A}} \mathrm{R}\tilde{\Gamma}_{\mathfrak{a}}(-),$$

which implies that

$$\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(M) \cong \hat{A} \otimes^{\mathrm{L}}_{A} (\mathrm{K}^{\vee}_{\infty}(A; \boldsymbol{a}) \otimes_A M). \qquad \square$$

Dually, we have the next result for the $\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}$ functor. Note however that, in this case, even if A is noetherian, we have to use DG-algebra resolutions, because $\hat{A}$ is almost never projective over A (see, for example, [Buchweitz and Flenner 2006, Theorem 2.1]).

Theorem 3.6. *Let A be a commutative ring, let $\mathfrak{a} \subseteq A$ be a finitely generated ideal, and let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$. Assume that $\mathfrak{a}$ is weakly proregular. Then there is an isomorphism of functors*

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(-) \cong \mathrm{R}\operatorname{Hom}_A(\hat{A} \otimes_A \operatorname{Tel}(A; \boldsymbol{a}), -).$$

Proof. We use notation as in the proof of Theorem 3.2. Let $A \xrightarrow{f} \tilde{A} \xrightarrow{g} \hat{A}$ be a K-projective DG-algebra resolution of $A \to \hat{A}$, and set

$$\mathrm{L}\tilde{\Lambda}_{\mathfrak{a}}(-) := \operatorname{Rest}_{\hat{A}/\tilde{A}} \circ \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(-) : \mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{DGMod} \tilde{A}).$$

Let $M \in \mathrm{D}(\operatorname{Mod} A)$. Let $P \to M$ be a K-projective resolution, and let $M \to I$ be a K-injective resolution. The map $f : A \to \tilde{A}$ induces a map $\operatorname{Hom}_A(f, 1) : \operatorname{Hom}_A(\tilde{A}, I) \to I$. Let $Q \to \operatorname{Hom}_A(\tilde{A}, I)$ be a K-projective resolution of $\operatorname{Hom}_A(\tilde{A}, I)$ over $\tilde{A}$. There is a unique map $\phi : Q \to P$ in $\mathrm{K}(\operatorname{Mod} A)$ making the diagram

$$\begin{tikzcd} \operatorname{Hom}_A(\tilde{A}, I) \arrow[r, "\operatorname{Hom}_A(f,1)"] & I \\ Q \arrow[u] \arrow[r, "\phi"] & P \arrow[u] \end{tikzcd} \tag{3.7}$$

commutative. The morphism of functors

$$\beta(-) : \operatorname{Hom}_A(\operatorname{Tel}(A; \boldsymbol{a}), -) \to \Lambda_{\mathfrak{a}}(-)$$

that was constructed in [Porta et al. 2014b, Definition 5.16] and the map $\phi : Q \to P$ induce a commutative diagram

$$\begin{tikzcd} \Lambda_{\mathfrak{a}}(Q) \arrow[r, "\Lambda_{\mathfrak{a}}(\phi)"] & \Lambda_{\mathfrak{a}}(P) \\ \operatorname{Hom}_A(\operatorname{Tel}(A; \boldsymbol{a}), Q) \arrow[u, "\beta_Q"] \arrow[r, "\operatorname{Hom}_A(1,\phi)"] & \operatorname{Hom}_A(\operatorname{Tel}(A; \boldsymbol{a}), P) \arrow[u, "\beta_P"'] \end{tikzcd} \tag{3.8}$$

and, because of weak proregularity of $\boldsymbol{a}$, the two vertical maps in this diagram are quasiisomorphisms. We will show that the bottom horizontal map is also a quasi-isomorphism, which will imply that the top horizontal map is a quasiisomorphism.

To see this, consider the commutative diagram

$$\begin{tikzcd}
\mathrm{Hom}_A(T, Q) \arrow[rr, "1"] \arrow[d, "2"] & & \mathrm{Hom}_A(T, P) \arrow[d, "3"] \\
\mathrm{Hom}_A(T, \mathrm{Hom}_A(\tilde{A}, I)) \arrow[rr, "4"] & & \mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), I) \\
\mathrm{Hom}_A(T, \mathrm{Hom}_A(\hat{A}, I)) \arrow[rr, "6"] \arrow[u, "5"] & & \mathrm{Hom}_A(T, \mathrm{Hom}_A(\mathrm{Hom}_A(T, A), I)) \arrow[u, "7"']
\end{tikzcd}$$

in $\mathrm{K}(\mathrm{Mod}\, A)$, where we have set $T := \mathrm{Tel}(A; \boldsymbol{a})$. The top square of this diagram is induced from the square (3.7), while the bottom square is induced from the commutative diagram of [Porta et al. 2014b, Equation (5.26)]. Weak proregularity of $\boldsymbol{a}$ implies that the map (6) is a quasiisomorphism. The fact that I is K-injective and that $\mathrm{Tel}(A; \boldsymbol{a})$ is K-projective implies that (5) is a quasiisomorphism. The hom-tensor adjunction and [Alonso Tarrío et al. 1997, Corollary after Theorem (0.3)*] (or the proof of [Porta et al. 2014b, Lemma 7.6]) shows that (7) is also a quasiisomorphism. Hence, (4) is a quasiisomorphism. As (2) and (3) are clearly quasiisomorphisms, we deduce that (1) is a quasiisomorphism. Returning to the commutative diagram (3.8), we deduce that the map

$$\Lambda_{\mathfrak{a}}(\phi) : \Lambda_{\mathfrak{a}}(Q) \to \Lambda_{\mathfrak{a}}(P)$$

is a quasiisomorphism.

There are functorial isomorphisms

$$\mathrm{L}\tilde{\Lambda}_{\mathfrak{a}}(M) = \mathrm{Rest}_{\hat{A}/\tilde{A}}(\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(M)) \cong \mathrm{Rest}_{\hat{A}/\tilde{A}}(\hat{\Lambda}_{\mathfrak{a}}(P))$$

in $\mathrm{D}(\mathrm{DGMod}\, \tilde{A})$, and since the map $\Lambda_{\mathfrak{a}}(\phi) : \Lambda_{\mathfrak{a}}(Q) \to \Lambda_{\mathfrak{a}}(P)$ is a quasiisomorphism, it follows that the map

$$\mathrm{Rest}_{\hat{A}/\tilde{A}}(\hat{\Lambda}_{\mathfrak{a}}(\phi)) : \mathrm{Rest}_{\hat{A}/\tilde{A}}(\hat{\Lambda}_{\mathfrak{a}}(Q)) \to \mathrm{Rest}_{\hat{A}/\tilde{A}}(\hat{\Lambda}_{\mathfrak{a}}(P))$$

is also a quasiisomorphism. By [Porta et al. 2014b, Corollary 5.23], there is an A-linear quasiisomorphism

$$\beta_Q : \mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), Q) \to \Lambda_{\mathfrak{a}}(Q),$$

and it is easy to verify that the same construction gives rise to an $\tilde{A}$-linear quasi-isomorphism

$$\mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), Q) \to \mathrm{Rest}_{\hat{A}/\tilde{A}}(\hat{\Lambda}_{\mathfrak{a}}(Q)).$$

Hence, there are isomorphisms of functors

$$\begin{aligned}\mathrm{L}\tilde{\Lambda}_{\mathfrak{a}}(M) &\cong \operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a}), Q)\\ &\cong \operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a}), \operatorname{Hom}_A(\tilde{A}, I))\\ &\cong \mathrm{R}\operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a})\otimes_A \tilde{A}, M).\end{aligned}$$

By Proposition 3.1, there is an isomorphism of functors

$$1_{\mathrm{D}(\mathrm{DGMod}\,\hat{A})} \cong \mathrm{R}\operatorname{Hom}_{\tilde{A}}(\hat{A}, \operatorname{Rest}_{\hat{A}/\tilde{A}}(-)).$$

Hence,

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(-) \cong \mathrm{R}\operatorname{Hom}_{\tilde{A}}(\hat{A}, \mathrm{L}\tilde{\Lambda}_{\mathfrak{a}}(-)),$$

which implies that

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(M) \cong \mathrm{R}\operatorname{Hom}_A(\operatorname{Tel}(A;\boldsymbol{a})\otimes_A \hat{A}, M).$$ □

Corollary 3.9. *Let A be a commutative ring, let $\mathfrak{a}\subseteq A$ be a weakly proregular ideal, and let $M, N \in \mathrm{D}(\mathrm{Mod}\,A)$. Then there are isomorphisms*

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_A(M,N)) \cong \mathrm{R}\operatorname{Hom}_A(\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(M), N) \cong \mathrm{R}\operatorname{Hom}_A(M, \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(N))$$

of functors

$$\mathrm{D}(\mathrm{Mod}\,A)\times \mathrm{D}(\mathrm{Mod}\,A) \to \mathrm{D}(\mathrm{Mod}\,\hat{A}).$$

Proof. Let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$. The hom-tensor adjunction and the quasiisomorphism $\mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a}) \cong \operatorname{Tel}(A;\boldsymbol{a})$ show that there are bifunctorial isomorphisms

$$\mathrm{R}\operatorname{Hom}_A\big(\operatorname{Tel}(A;\boldsymbol{a})\otimes_A \hat{A}, \mathrm{R}\operatorname{Hom}_A(M,N)\big) \cong \mathrm{R}\operatorname{Hom}_A\big(\mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a})\otimes_A \hat{A}\otimes^{\mathrm{L}}_A M, N\big)$$

and

$$\mathrm{R}\operatorname{Hom}_A\big(\mathrm{K}^{\vee}_{\infty}(A;\boldsymbol{a})\otimes_A \hat{A}\otimes^{\mathrm{L}}_A M, N\big) \cong \mathrm{R}\operatorname{Hom}_A\big(M, \mathrm{R}\operatorname{Hom}_A\big(\operatorname{Tel}(A;\boldsymbol{a})\otimes_A \hat{A}, N\big)\big),$$

so the result follows from Theorems 3.2 and 3.6. □

Remark 3.10. If $A = \hat{A}$, then the above corollary collapses to the Greenlees–May duality (see [Alonso Tarrío et al. 1997, Theorem (0.3)], or [Porta et al. 2014b, Theorem 7.12]).

The next two corollaries will be applied in the next section to study relations between the derived completion functor and Hochschild cohomology.

Corollary 3.11. *Let A be a commutative ring, and let $\mathfrak{a}\subseteq A$ be a weakly proregular ideal. Given a ring map $\hat{A}\to B$, and a complex $M \in \mathrm{D}(\mathrm{Mod}\,B)$, there is an isomorphism*

$$\mathrm{R}\operatorname{Hom}_A(M, \mathrm{L}\Lambda_{\mathfrak{a}}(-)) \cong \mathrm{R}\operatorname{Hom}_{\hat{A}}(M, \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(-))$$

of functors

$$\mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{Mod} B).$$

Proof. Let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$. Let $N \in \mathrm{D}(\operatorname{Mod} A)$. By Theorem 3.6, there is an isomorphism of functors

$$\mathrm{R}\operatorname{Hom}_{\hat{A}}(M, \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(N)) \cong \mathrm{R}\operatorname{Hom}_{\hat{A}}\big(M, \mathrm{R}\operatorname{Hom}_A\big(\operatorname{Tel}(A; \boldsymbol{a}) \otimes_A \hat{A}, N\big)\big).$$

Applying the hom-tensor adjunction twice, we get an isomorphism

$$\mathrm{R}\operatorname{Hom}_{\hat{A}}\big(M, \mathrm{R}\operatorname{Hom}_A\big(\operatorname{Tel}(A; \boldsymbol{a}) \otimes_A \hat{A}, N\big)\big) \cong \mathrm{R}\operatorname{Hom}_A\big(M, \operatorname{Hom}_A(\operatorname{Tel}(A; \boldsymbol{a}), N)\big),$$

which proves the claim. $\square$

Corollary 3.12. *Let A be a commutative ring, and let B be a commutative A-algebra. Let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal, let $\mathfrak{b} = \mathfrak{a} \cdot B$, and assume that $\mathfrak{b}$ is also weakly proregular. Then there is an isomorphism*

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{b}} \mathrm{R}\operatorname{Hom}_A(B, -) \cong \mathrm{R}\operatorname{Hom}_A(\hat{B}, \mathrm{L}\Lambda_{\mathfrak{a}}(-))$$

of functors

$$\mathrm{D}(\operatorname{Mod} A) \to \mathrm{D}(\operatorname{Mod} \hat{B}),$$

where $\hat{B} := \Lambda_{\mathfrak{b}}(B)$.

Proof. Let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$, and let $\boldsymbol{b}$ be its image in B. By Theorem 3.6, given $M \in \mathrm{D}(\operatorname{Mod} A)$, there is a functorial isomorphism

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{b}} \mathrm{R}\operatorname{Hom}_A(B, M) \cong \mathrm{R}\operatorname{Hom}_B\big(\hat{B} \otimes_B \operatorname{Tel}(B; \boldsymbol{b}), \mathrm{R}\operatorname{Hom}_A(B, M)\big),$$

so, by the derived hom-tensor adjunction,

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{b}} \mathrm{R}\operatorname{Hom}_A(B, M) \cong \mathrm{R}\operatorname{Hom}_A\big(\hat{B} \otimes_B \operatorname{Tel}(B; \boldsymbol{b}), M\big).$$

By the base change property of the telescope complex, we have

$$\operatorname{Tel}(B; \boldsymbol{b}) \cong B \otimes_A \operatorname{Tel}(A; \boldsymbol{a}),$$

so that

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{b}} \mathrm{R}\operatorname{Hom}_A(B, M) \cong \mathrm{R}\operatorname{Hom}_A\big(\hat{B} \otimes_A \operatorname{Tel}(A; \boldsymbol{a}), M\big).$$

Hence, using adjunction again, and the fact that $\mathfrak{a}$ is weakly proregular, we obtain the result. $\square$

Dually to these two corollaries, we have the following results which will apply to the study of Hochschild homology. We omit the very similar proofs.

Corollary 3.13. *Let A be a commutative ring, and let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal. Given a ring map $\hat{A} \to B$, and a complex $M \in \mathrm{D}(\mathrm{Mod}\, B)$, there is an isomorphism*

$$M \otimes^{\mathrm{L}}_{A} \mathrm{R}\Gamma_{\mathfrak{a}}(-) \cong M \otimes^{\mathrm{L}}_{\hat{A}} \mathrm{R}\hat{\Gamma}_{\mathfrak{a}}(-)$$

of functors

$$\mathrm{D}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{Mod}\, B).$$

Corollary 3.14. *Let A be a commutative ring, and let B be a commutative A-algebra. Let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal, let $\mathfrak{b} = \mathfrak{a} \cdot B$, and assume that $\mathfrak{b}$ is also weakly proregular. Then there is an isomorphism*

$$\mathrm{R}\hat{\Gamma}_{\mathfrak{b}}(B \otimes^{\mathrm{L}}_{A} -) \cong \hat{B} \otimes^{\mathrm{L}}_{A} \mathrm{R}\Gamma_{\mathfrak{a}}(-))$$

of functors

$$\mathrm{D}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{Mod}\, \hat{B}),$$

where $\hat{B} := \Lambda_{\mathfrak{b}}(B)$.

4. Hochschild cohomology and derived completion

We now turn to the main theme of this paper: relations between adic completion (and its derived functor) and Hochschild cohomology. The results of this section rely heavily on the tools developed in the previous sections.

Our first result describes the effect of applying the derived completion functor to the Hochschild cohomology complex in a rather general situation. We will later specialize further to obtain more explicit results.

Theorem 4.1. *Let $\Bbbk$ be a commutative ring, let A be a flat noetherian $\Bbbk$-algebra, and let $\mathfrak{a} \subseteq A$ be an ideal. Assume further that at least one of the following holds:*

(1) *The ring $\Bbbk$ is an absolutely flat ring (e.g., a field).*
(2) *The ring $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$.*
(3) *The ideal $I := \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a} \subseteq A \otimes_{\Bbbk} A$ is weakly proregular.*

Set $\hat{A} := \Lambda_{\mathfrak{a}}(A)$ and $A \hat{\otimes}_{\Bbbk} A := \Lambda_I(A \otimes_{\Bbbk} A)$. Then there are isomorphisms

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}} \mathrm{R}\,\mathrm{Hom}_{A\otimes_{\Bbbk}A}(A, -) \cong \mathrm{R}\,\mathrm{Hom}_{A\otimes_{\Bbbk}A}(\hat{A}, \mathrm{L}\Lambda_I(-)) \cong \mathrm{R}\,\mathrm{Hom}_{A\hat{\otimes}_{\Bbbk}A}(\hat{A}, \mathrm{L}\hat{\Lambda}_I(-))$$

of functors

$$\mathrm{D}(\mathrm{Mod}\, A \otimes_{\Bbbk} A) \to \mathrm{D}(\mathrm{Mod}\, \hat{A}).$$

Proof. If $\Bbbk$ is absolutely flat, then I is weakly proregular by Proposition 2.1, while if $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$, then I is weakly proregular by Theorem 2.6. Since the image of I in A is equal to $\mathfrak{a}$, and since A being noetherian implies that

$\mathfrak{a}$ is weakly proregular, given $M \in \mathrm{D}(\operatorname{Mod} A \otimes_{\Bbbk} A)$, by Corollary 3.12, there is a functorial isomorphism

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M) \cong \mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(\hat{A}, \mathrm{L}\Lambda_I(M)).$$

Note that, considered as an $(A \otimes_{\Bbbk} A)$-module, the I-adic completion of A is equal to $\hat{A}$. Hence, we may apply Corollary 3.11, and deduce that there is a functorial isomorphism

$$\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(\hat{A}, \mathrm{L}\Lambda_I(M)) \cong \mathrm{R}\operatorname{Hom}_{A\hat{\otimes}_{\Bbbk}A}(\hat{A}, \mathrm{L}\hat{\Lambda}_I(M))$$

in $\mathrm{D}(\operatorname{Mod} \hat{A})$. This proves the result. □

The next lemma might seem trivial at first glance. However, the possible lack of flatness of the completion map makes it a little more difficult.

Lemma 4.2. *Let A be a commutative ring, and let $\mathfrak{a} \subseteq A$ be a weakly proregular ideal. Assume that $\hat{A} := \Lambda_{\mathfrak{a}}(A)$ is noetherian, and let $M \in \operatorname{Mod} \hat{A}$ be an $\mathfrak{a}$-adically complete $\hat{A}$-module. Let*

$$\operatorname{Rest}_{\hat{A}/A} : \mathrm{D}(\operatorname{Mod} \hat{A}) \to \mathrm{D}(\operatorname{Mod} A)$$

be the forgetful functor. Then there is a functorial isomorphism

$$\mathrm{L}\Lambda_{\mathfrak{a}}(\operatorname{Rest}_{\hat{A}/A}(M)) \cong \operatorname{Rest}_{\hat{A}/A}(M)$$

in $\mathrm{D}(\operatorname{Mod} A)$, *and a functorial isomorphism*

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\operatorname{Rest}_{\hat{A}/A}(M)) \cong M$$

in $\mathrm{D}(\operatorname{Mod} \hat{A})$.

Proof. Let $\boldsymbol{a}$ be a finite sequence that generates $\mathfrak{a}$, and let $\hat{\boldsymbol{a}}$ be its image in $\hat{A}$. Since $\hat{A}$ is noetherian and M is $\mathfrak{a}$-adically complete, it follows from [Porta et al. 2015, Theorem 1.21] that M is cohomologically $\mathfrak{a}\hat{A}$-adically complete, so there is an isomorphism

$$M \cong \mathrm{R}\operatorname{Hom}_{\hat{A}}(\operatorname{Tel}(\hat{A}; \hat{\boldsymbol{a}}), M).$$

The base change property of the telescope complex implies that $\operatorname{Tel}(A; \boldsymbol{a}) \otimes_A \hat{A} \cong \operatorname{Tel}(\hat{A}; \hat{\boldsymbol{a}})$. Using this fact and the hom-tensor adjunction, we deduce that

$$\operatorname{Rest}_{\hat{A}/A}(M) \cong \mathrm{R}\operatorname{Hom}_A\bigl(\operatorname{Tel}(A; \boldsymbol{a}), \operatorname{Rest}_{\hat{A}/A}(M)\bigr),$$

so that

$$\mathrm{L}\Lambda_{\mathfrak{a}}\bigl(\operatorname{Rest}_{\hat{A}/A}(M)\bigr) \cong \operatorname{Rest}_{\hat{A}/A}(M).$$

Note that

$$\operatorname{Rest}_{\hat{A}/A}\bigl(\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\bigl(\operatorname{Rest}_{\hat{A}/A}(M)\bigr)\bigr) \cong \mathrm{L}\Lambda_{\mathfrak{a}}\bigl(\operatorname{Rest}_{\hat{A}/A}(M)\bigr) \cong \operatorname{Rest}_{\hat{A}/A}(M).$$

Hence, the complex of $\hat{A}$-modules

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\big(\mathrm{Rest}_{\hat{A}/A}(M)\big)$$

is concentrated in degree 0. Letting $P \to \mathrm{Rest}_{\hat{A}/A}(M)$ be a projective resolution, we deduce that the map

$$\hat{\Lambda}_{\mathfrak{a}}(P) \to \hat{\Lambda}_{\mathfrak{a}}\big(\mathrm{Rest}_{\hat{A}/A}(M)\big) \cong M$$

is an $\hat{A}$-linear quasiisomorphism, so the result follows from the fact that

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\big(\mathrm{Rest}_{\hat{A}/A}(M)\big) \cong \hat{\Lambda}_{\mathfrak{a}}(P). \qquad \square$$

Using this lemma, and as a first corollary of Theorem 4.1, we obtain the next result which reduces the problem of computing the Hochschild cohomology of an adically complete ring to a problem over noetherian rings.

Corollary 4.3. *Let $\Bbbk$ be a commutative ring, and let A be a flat noetherian $\Bbbk$-algebra. Assume $\mathfrak{a} \subseteq A$ is an ideal such that A is $\mathfrak{a}$-adically complete, and such that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$. Let $I = \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}$, and set $A \hat{\otimes}_{\Bbbk} A := \Lambda_I(A \otimes_{\Bbbk} A)$. Then for any $M \in \mathrm{Mod}\, A \otimes_{\Bbbk} A$ which is I-adically complete (for example, any $\mathfrak{a}$-adically complete A-module, or, more particularly, any finitely generated A-module), there is a functorial isomorphism*

$$\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, M) \cong \mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(A, M)$$

in $\mathrm{D}(\mathrm{Mod}\, A)$, and the ring $A \hat{\otimes}_{\Bbbk} A$ is noetherian.

Proof. That $A \hat{\otimes}_{\Bbbk} A$ is noetherian follows from the fact that $(A \otimes_{\Bbbk} A)/I \cong A/\mathfrak{a} \otimes_{\Bbbk} A/\mathfrak{a}$ is noetherian, and I is finitely generated. Since $A = \hat{A}$, according to Theorem 4.1, there is a functorial isomorphism

$$\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, \mathrm{L}\Lambda_I(M)) \cong \mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(A, \mathrm{L}\hat{\Lambda}_I(M)).$$

By Lemma 4.2, we have

$$\mathrm{L}\Lambda_I(M) \cong M \quad \text{in } \mathrm{D}(\mathrm{Mod}\, A \otimes_{\Bbbk} A)$$

and

$$\mathrm{L}\hat{\Lambda}_I(M) \cong M \quad \text{in } \mathrm{D}(\mathrm{Mod}\, A \hat{\otimes}_{\Bbbk} A).$$

Using these two facts, we obtain the functorial isomorphism

$$\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, M) \cong \mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(A, M). \qquad \square$$

Remark 4.4. In [Buchweitz and Flenner 2006, Section 3, page 113], the authors defined the analytic Hochschild cohomology and discussed its relation to ordinary Hochschild cohomology. The setup there is as follows: $(A, \mathfrak{m}_A)$ and $(B, \mathfrak{m}_B)$ are two complete noetherian local rings, and there is a flat local map $A \to B$ such that the induced map $A/\mathfrak{m}_A \to B/\mathfrak{m}_B$ is an isomorphism. In this situation, the authors defined the analytic bar resolution $\hat{\mathcal{B}}$ by replacing $B^{\otimes_A^n}$ with its completion with respect to the maximal ideal $\ker(B^{\otimes_A^n} \to B/\mathfrak{m}_B)$ in the ordinary bar resolution. Using this complex, the authors defined the analytic Hochschild cohomology by

$$\widehat{\mathrm{HH}}^n(B/A, M) := H^n \operatorname{Hom}_{B \hat{\otimes}_A B}(\hat{\mathcal{B}}, M),$$

observed that there is a canonical map

$$\widehat{\mathrm{HH}}^n(B/A, M) \to \mathrm{HH}^n(B/A, M),$$

and asked if these modules are isomorphic. Using Corollary 4.3 and the results of [Buchweitz and Flenner 2006], we may now obtain a positive answer to this question:

Corollary 4.5. *Let $(A, \mathfrak{m}_A) \to (B, \mathfrak{m}_B)$ be a flat local homomorphism of complete noetherian local rings such that the induced map of residue fields $A/\mathfrak{m}_A \to B/\mathfrak{m}_B$ is an isomorphism. Then for any $\mathfrak{m}_B$-adically complete B-module M (in particular, for any finitely generated B-module M), there is a natural B-module isomorphism*

$$\widehat{\mathrm{HH}}^n(B/A, M) \cong \mathrm{HH}^n(B/A, M).$$

Proof. Note that M is $\ker(B \otimes_A B \to B/\mathfrak{m}_B)$-adically complete, so by [Buchweitz and Flenner 2006, Proposition 3.1], the natural map

$$\mathrm{HH}^n(B/A, M) \to \operatorname{Ext}^n_{B \otimes_A B}(B, M)$$

is an isomorphism. Similarly, by [Buchweitz and Flenner 2006, Proposition 3.2], the natural map

$$\widehat{\mathrm{HH}}^n(B/A, M) \to \operatorname{Ext}^n_{B \hat{\otimes}_A B}(B, M)$$

is an isomorphism. Finally, by Corollary 4.3, there is a natural isomorphism

$$\operatorname{Ext}^n_{B \otimes_A B}(B, M) \cong \operatorname{Ext}^n_{B \hat{\otimes}_A B}(B, M),$$

so the result follows. $\square$

Our next goal is to give the complex appearing in Theorem 4.1 a description in terms of the Hochschild complex of the ring $\hat{A}$. First we need a lemma.

Lemma 4.6. *Let $\Bbbk$ be a commutative ring, let A be a flat noetherian $\Bbbk$-algebra, and let $\mathfrak{a} \subseteq A$ be an ideal. Let*

$$I := \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a} \subseteq A \otimes_{\Bbbk} A,$$

let $\hat{\mathfrak{a}} := \mathfrak{a} \cdot \Lambda_{\mathfrak{a}}(A)$, and let

$$J = \hat{\mathfrak{a}} \otimes_{\Bbbk} \hat{A} + \hat{A} \otimes_{\Bbbk} \hat{\mathfrak{a}} = I \cdot (\hat{A} \otimes_{\Bbbk} \hat{A}) \subseteq \hat{A} \otimes_{\Bbbk} \hat{A}.$$

If $\tau_A : A \to \hat{A}$ is the completion map, then the ring map

$$\Lambda_I(\tau_A \otimes_k \tau_A) : \Lambda_I(A \otimes_{\Bbbk} A) \to \Lambda_J(\hat{A} \otimes_{\Bbbk} \hat{A})$$

is an isomorphism.

Proof. Notice that for each n we have

$$I^n = \sum_{i=0}^{n} \mathfrak{a}^i \otimes \mathfrak{a}^{n-i},$$

where we have set $\mathfrak{a}^0 = A$. Hence,

$$\mathfrak{a}^{2n} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}^{2n} \subseteq I^{2n} \subseteq \mathfrak{a}^n \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}^n.$$

It follows that the two sequences of ideals $\{I^n\}$ and $\{\mathfrak{a}^n \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}^n\}$ are cofinal in each other, so that

$$\varprojlim (A \otimes_{\Bbbk} A)/I^n \cong \varprojlim (A \otimes_{\Bbbk} A)/(\mathfrak{a}^n \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}^n) \cong \varprojlim A/\mathfrak{a}^n \otimes_{\Bbbk} A/\mathfrak{a}^n.$$

Since A is noetherian, $\hat{A}$ is flat over A, so it is also flat over $\Bbbk$. Hence, in the exact same manner,

$$\varprojlim (\hat{A} \otimes_{\Bbbk} \hat{A})/J^n \cong \varprojlim \hat{A}/(\hat{\mathfrak{a}})^n \otimes_{\Bbbk} A/(\hat{\mathfrak{a}})^n.$$

Hence, the result follows from the fact that $\hat{A}/(\hat{\mathfrak{a}})^n \cong A/\mathfrak{a}^n$, and from the observation that the maps $\Lambda_I(\tau_A \otimes_k \tau_A)$ and

$$\varprojlim ((\tau_A \otimes_k \tau_A)/(\mathfrak{a}^n \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a}^n))$$

are equal. $\square$

Remark 4.7. In the notation of the above lemma, note that composing the completion map

$$\hat{A} \otimes_{\Bbbk} \hat{A} \to \Lambda_J(\hat{A} \otimes_{\Bbbk} \hat{A})$$

with the isomorphism

$$\Lambda_J(\hat{A} \otimes_{\Bbbk} \hat{A}) \to \Lambda_I(A \otimes_{\Bbbk} A)$$

we obtain a ring map

$$\hat{A} \otimes_{\Bbbk} \hat{A} \to \Lambda_I(A \otimes_{\Bbbk} A).$$

Using this map, given $M \in \operatorname{Mod}(A \otimes_{\Bbbk} A)$, we will be able to regard

$$\hat{\Lambda}_I(M) \in \operatorname{Mod}(\Lambda_I(A \otimes_{\Bbbk} A))$$

as an $(\hat{A} \otimes_{\Bbbk} \hat{A})$-module which is J-adically complete.

We may now refine Theorem 4.1, and present the derived completion of the Hochschild cohomology complex as a Hochschild cohomology complex over the completion:

Theorem 4.8. *Let $\Bbbk$ be a commutative ring, and let A be a flat noetherian $\Bbbk$-algebra such that $A \otimes_{\Bbbk} A$ is noetherian. Let $\mathfrak{a} \subseteq A$ be an ideal, and let M be a finitely generated $(A \otimes_{\Bbbk} A)$-module. Set*

$$I := \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a} \subseteq A \otimes_{\Bbbk} A.$$

Then there is a functorial isomorphism

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, M)) \cong \mathrm{R}\operatorname{Hom}_{\hat{A} \otimes_{\Bbbk} \hat{A}}(\hat{A}, \hat{M})$$

in $\mathrm{D}(\operatorname{Mod} \hat{A})$, *where* $\hat{A} := \hat{\Lambda}_{\mathfrak{a}}(A)$ *and* $\hat{M} := \hat{\Lambda}_I(M)$.

Proof. Set $A \mathbin{\hat{\otimes}_{\Bbbk}} A := \Lambda_I(A \otimes_{\Bbbk} A)$. By assumption, $A \otimes_{\Bbbk} A$ is noetherian, so I is weakly proregular. According to Theorem 4.1, there is a functorial isomorphism

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_{A \otimes_{\Bbbk} A}(A, M)) \cong \mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(\hat{A}, \mathrm{L}\hat{\Lambda}_I(M)). \tag{4.9}$$

Because $A \otimes_{\Bbbk} A$ is noetherian and M is a finitely generated module, it follows that

$$\mathrm{L}\hat{\Lambda}_I(M) \cong \hat{\Lambda}_I(M) = \hat{M}$$

in $\mathrm{D}(\operatorname{Mod} A \mathbin{\hat{\otimes}_{\Bbbk}} A)$. Hence, by Lemma 4.6, there is a functorial isomorphism

$$\mathrm{R}\operatorname{Hom}_{A \hat{\otimes}_{\Bbbk} A}(\hat{A}, \hat{M}) \cong \mathrm{R}\operatorname{Hom}_{\hat{A} \hat{\otimes}_{\Bbbk} \hat{A}}(\hat{A}, \hat{M}) \tag{4.10}$$

in $\mathrm{D}(\operatorname{Mod} \hat{A})$. Let $\hat{\mathfrak{a}} := \mathfrak{a} \cdot \hat{A} \subseteq \hat{A}$. Since A is noetherian, the map $A \to \hat{A}$ is flat, so the map $A \otimes_{\Bbbk} A \to \hat{A} \otimes_{\Bbbk} \hat{A}$ is also flat. Hence, by Remark 2.2, the ideal $J := I \cdot (\hat{A} \otimes_{\Bbbk} \hat{A}) = \hat{\mathfrak{a}} \otimes_{\Bbbk} \hat{A} + \hat{A} \otimes_{\Bbbk} \hat{\mathfrak{a}}$ is weakly proregular.

Because $A \otimes_{\Bbbk} A$ is noetherian, its completion $\Lambda_I(A \otimes_{\Bbbk} A)$ is also noetherian, so, again by Lemma 4.6, the ring $\Lambda_J(\hat{A} \otimes_{\Bbbk} \hat{A})$ is noetherian. Hence, by Lemma 4.2, $\mathrm{L}\Lambda_J(\hat{M}) \cong \mathrm{L}\hat{\Lambda}_J(\hat{M}) \cong \hat{M}$. Applying Theorem 4.1 to the ring $\hat{A}$, we have a functorial isomorphism

$$\mathrm{R}\operatorname{Hom}_{\hat{A} \otimes_{\Bbbk} \hat{A}}(\hat{A}, \mathrm{L}\Lambda_J(M)) \cong \mathrm{R}\operatorname{Hom}_{\hat{A} \hat{\otimes}_{\Bbbk} \hat{A}}(\hat{A}, \mathrm{L}\hat{\Lambda}_J(M)).$$

Composing this isomorphism with the isomorphisms of (4.9) and (4.10), we obtain the result. ☐

Our final goal in this section is to show that when taking cohomology in the above theorem, derived completion may be replaced with ordinary completion. The next simple lemma is needed for the proof of Proposition 4.12.

Lemma 4.11. *Let A be a noetherian ring, let P be a bounded complex of free A-modules, and let M be a finitely generated A-module. Then the canonical map*

$$\mathrm{Hom}_A(P, A) \otimes_A M \to \mathrm{Hom}_A(P, M)$$

is an isomorphism of complexes.

Proof. It is enough to show this in the case where P is a single free A-module. In that case, note that $\mathrm{Hom}_A(P, A)$ is a direct product of copies of A, and since A is noetherian this is a flat A-module. Thus, both of the functors $\mathrm{Hom}_A(P, A) \otimes_A -$ and $\mathrm{Hom}_A(P, -)$ are exact, and if M is a finitely generated free A-module, then the canonical map $\mathrm{Hom}_A(P, A) \otimes_A M \to \mathrm{Hom}_A(P, M)$ is obviously an isomorphism. Hence, the result of the lemma follows from the standard finite presentation trick. □

In the case where M is bounded above, the next proposition is [Frankild 2003, Proposition 2.7]. We, however, wish to apply this result to the Hochschild complex which is bounded below, so we give a proof that works for complexes without any boundedness condition.

Proposition 4.12. *Let A be a noetherian ring, and let $\mathfrak{a} \subseteq A$ be an ideal. Then there is an isomorphism*

$$\mathrm{L}\Lambda_{\mathfrak{a}}(M) \cong \hat{A} \otimes_A M$$

of functors

$$\mathrm{D_f}(\mathrm{Mod}\, A) \to \mathrm{D}(\mathrm{Mod}\, A).$$

Proof. There is a sequence of morphisms of functors

$$\hat{A} \otimes_A M \cong \mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), A) \otimes_A M \to \mathrm{Hom}_A(\mathrm{Tel}(A; \boldsymbol{a}), M) \cong \mathrm{L}\Lambda_{\mathfrak{a}}(M).$$

Because A is assumed to be noetherian, $\hat{A}$ is flat over A. Hence, both $\mathrm{L}\Lambda_{\mathfrak{a}}(-)$ and $- \otimes_A \hat{A}$ are functors of finite cohomological dimension. Thus, by [Hartshorne 1966, Proposition I.7.1], it is enough to show that the above morphism is an isomorphism in the case where M is a finitely generated A-module, and this follows from Lemma 4.11. □

Here is the result promised in the title of this paper:

Theorem 4.13. *Let $\Bbbk$ be a commutative ring, and let A be a flat noetherian $\Bbbk$-algebra such that $A \otimes_{\Bbbk} A$ is noetherian. Let $\mathfrak{a} \subseteq A$ be an ideal, and let M be a finitely generated $(A \otimes_{\Bbbk} A)$-module. Then for any $n \in \mathbb{N}$, there is a functorial isomorphism*

$$\Lambda_{\mathfrak{a}}\big(\mathrm{Ext}^n_{A\otimes_{\Bbbk} A}(A, M)\big) \cong \mathrm{Ext}^n_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{M}).$$

If, moreover, either

(1) $\Bbbk$ *is a field, or*

(2) *A is projective over $\Bbbk$, $\mathfrak{a}$ is a maximal ideal, and M is a finitely generated A-module,*

then there is also a functorial isomorphism

$$\Lambda_{\mathfrak{a}}\big(\mathrm{HH}^n(A/\Bbbk, M)\big) \cong \mathrm{HH}^n(\hat{A}/\Bbbk, \hat{M}).$$

Proof. The assumptions of the theorem ensure that

$$\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M) \in \mathrm{D}_{\mathrm{f}}(\operatorname{Mod} A),$$

so since A is noetherian, we have a functorial isomorphism

$$\Lambda_{\mathfrak{a}}\big(\operatorname{Ext}^n_{A\otimes_{\Bbbk}A}(A, M)\big) \cong \hat{A} \otimes_A H^n(\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M)).$$

Flatness of $\hat{A}$ over A implies (for example, by [Porta et al. 2014b, Corollary 2.12]) that there is a natural isomorphism

$$\hat{A} \otimes_A H^n(\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M)) \cong H^n(\hat{A} \otimes_A \mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M)).$$

Hence, by Proposition 4.12, it is enough to compute the n-th cohomology of the complex

$$\mathrm{L}\Lambda_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M)).$$

Letting

$$\operatorname{Rest}_{\hat{A}/A} : \mathrm{D}(\operatorname{Mod}\hat{A}) \to \mathrm{D}(\operatorname{Mod} A)$$

be the forgetful functor, the above complex is equal to

$$\operatorname{Rest}_{\hat{A}/A} \circ \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}(\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M)).$$

By Theorem 4.8, there is a functorial isomorphism

$$\mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\mathrm{R}\operatorname{Hom}_{A\otimes_{\Bbbk}A}(A, M) \cong \mathrm{R}\operatorname{Hom}_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{M})$$

in $\mathrm{D}(\operatorname{Mod}\hat{A})$, so applying the forgetful functor we obtain an A-linear natural isomorphism

$$\Lambda_{\mathfrak{a}}\big(\operatorname{Ext}^n_{A\otimes_{\Bbbk}A}(A, M)\big) \cong \operatorname{Ext}^n_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{M}).$$

(Actually, any A-linear map between $\hat{A}$-modules is automatically $\hat{A}$-linear, so this isomorphism is even an isomorphism of $\hat{A}$-modules.) This establishes the first claim of the theorem. If $\Bbbk$ is a field then the second claim obviously follows from the first one. Assume now that A is projective over $\Bbbk$, that $\mathfrak{a}$ is a maximal ideal, and that M is a finitely generated A-module. Let ϕ be the composition of the diagonal map $\hat{A} \otimes_{\Bbbk} \hat{A} \to \hat{A}$ with the map $\hat{A} \to \hat{A}/\mathfrak{a}\hat{A}$. Then $m = \ker(\phi) \subseteq \hat{A} \otimes_{\Bbbk} \hat{A}$ is a maximal ideal, and the image of m in $\hat{A}$ is equal to $\mathfrak{a}\hat{A}$. Hence $\hat{M}$ is m-adically complete, so by [Buchweitz and Flenner 2006, Proposition 3.1] the canonical map

$$\mathrm{HH}^n(\hat{A}/\Bbbk, \hat{M}) \to \mathrm{Ext}^n_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{M})$$

is an isomorphism. This proves the second claim. □

The above result allows us to compute the Hochschild cohomology of power series rings, which is new as far as we know.

Example 4.14. Let $\Bbbk$ be a noetherian ring. Let $A = \Bbbk[x_1, \dots, x_n]$, $\mathfrak{a} = (x_1, \dots, x_n)$, and $M = A$. Note that A is projective over $\Bbbk$, and that $\Lambda_{\mathfrak{a}}(A) = \Bbbk[\![x_1, \dots, x_n]\!]$. Hence, by the above theorem, we have

$$\mathrm{Ext}^i_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{A}) \cong \Lambda_{\mathfrak{a}}\big(\mathrm{HH}^i(\Bbbk[x_1, \dots, x_n]/\Bbbk, \Bbbk[x_1, \dots, x_n])\big).$$

By the Hochschild–Kostant–Rosenberg theorem, the right-hand side is equal to

$$\Lambda_{\mathfrak{a}} \wedge^i (\Bbbk[x_1, \dots, x_n]^n),$$

so we have an isomorphism

$$\mathrm{Ext}^i_{\hat{A}\otimes_{\Bbbk}\hat{A}}(\hat{A}, \hat{A}) \cong \wedge^i(\Bbbk[\![x_1, \dots, x_n]\!]^n).$$

If, moreover, $\Bbbk$ is a field, we obtain that

$$\mathrm{HH}^i(\Bbbk[\![x_1, \dots, x_n]\!]/\Bbbk, \Bbbk[\![x_1, \dots, x_n]\!]) \cong \wedge^i(\Bbbk[\![x_1, \dots, x_n]\!]^n).$$

We remark that one can also use Corollary 4.3 to make this calculation.

Example 4.15. Let A be a noetherian ring, and let $\mathfrak{a} \subseteq A$ be an ideal. Then, by Theorem 4.8, we have

$$\mathrm{R}\operatorname{Hom}_{\hat{A}\otimes_A\hat{A}}(\hat{A}, \hat{A}) \cong \mathrm{L}\hat{\Lambda}_{\mathfrak{a}}\mathrm{R}\operatorname{Hom}_A(A, A) = \hat{A}.$$

Assume now that $\mathfrak{a}$ is a maximal ideal. Then, by Theorem 4.13,

$$\mathrm{HH}^n(\hat{A}/A, \hat{A}) = 0$$

for all $n \neq 0$, and $\mathrm{HH}^0(\hat{A}/A, \hat{A}) = \hat{A}$. Specializing to the case where $A = \mathbb{Z}$, and $\mathfrak{a} = (p)$ for some prime number p, we have computed the absolute Hochschild cohomology of the ring of p-adic integers $\mathbb{Z}_p$.

Remark 4.16. The above examples are both particular cases of an adic Hochschild–Kostant–Rosenberg theorem which applies for any formally smooth adic algebra A such that $A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$. This can be shown using Corollary 4.3 and by studying the local structure of the completed diagonal $\ker(A \mathbin{\hat{\otimes}_{\Bbbk}} A \to A)$. A full proof of this will appear elsewhere.

5. Hochschild homology and derived torsion

In this short and final section we discuss relations between Hochschild homology and the derived torsion functor.

Theorem 5.1. *Let $\Bbbk$ be a commutative ring, let A be a flat noetherian $\Bbbk$-algebra, and let $\mathfrak{a} \subseteq A$ be an ideal. Assume further that at least one of the following holds:*

(1) *The ring $\Bbbk$ is an absolutely flat ring (e.g., a field).*

(2) *$A/\mathfrak{a}$ is essentially of finite type over $\Bbbk$.*

(3) *The ideal $I := \mathfrak{a} \otimes_{\Bbbk} A + A \otimes_{\Bbbk} \mathfrak{a} \subseteq A \otimes_{\Bbbk} A$ is weakly proregular.*

Set $\hat{A} := \Lambda_{\mathfrak{a}}(A)$ and $A \hat{\otimes}_{\Bbbk} A := \Lambda_I(A \otimes_{\Bbbk} A)$. Then there are isomorphisms

$$\mathrm{R}\hat{\Gamma}_{\mathfrak{a}}\big(A \otimes^{\mathrm{L}}_{A \otimes_{\Bbbk} A} -\big) \cong \hat{A} \otimes^{\mathrm{L}}_{A \otimes_{\Bbbk} A} \mathrm{R}\Gamma_I(-) \cong \hat{A} \otimes^{\mathrm{L}}_{A \hat{\otimes}_{\Bbbk} A} \mathrm{R}\hat{\Gamma}_I(-)$$

of functors

$$\mathrm{D}(\mathrm{Mod}\, A \otimes_{\Bbbk} A) \to \mathrm{D}(\mathrm{Mod}\, \hat{A}).$$

Proof. As in the proof of Theorem 4.1, the first two conditions imply the third one, so we may assume I is weakly proregular. The first isomorphism then follows from Corollary 3.14, while the second isomorphism follows from Corollary 3.13. □

Remark 5.2. In view of Theorem 4.13, it is natural to ask if Hochschild homology also commutes with adic completion. Here, the answer is false, even in simple situations. Indeed, let $\Bbbk$ be a field of characteristic 0, let $A = \Bbbk[x]$, and let $\mathfrak{a} = (x)$. Then

$$\mathrm{HH}_1(A/\Bbbk, A) \cong A,$$

so that $\Lambda_{\mathfrak{a}}(\mathrm{HH}_1(A/\Bbbk, A)) \cong \Bbbk[[x]]$. On the other hand, $\mathrm{HH}_1(\hat{A}/\Bbbk, \hat{A}) \cong \Omega^1_{\Bbbk[[x]]/\Bbbk}$ is an infinitely generated $\Bbbk[[x]]$-module.

Remark 5.3. As an alternative to the badly behaved Hochschild homology of commutative adic algebras, Hübl [1989] developed a theory of adic Hochschild homology by studying the cohomologies of the functor

$$A \otimes^{\mathrm{L}}_{A \hat{\otimes}_{\Bbbk} A} -.$$

As our Corollary 4.3 shows, in the case of Hochschild cohomology, usual Hochschild cohomology coincides with adic Hochschild cohomology, but by the previous remark we see that for Hochschild homology this is not the case.

Acknowledgments

The author would like to thank Amnon Yekutieli for some helpful suggestions. The author is grateful for the anonymous referees for their helpful comments.

References

[Alonso Tarrío et al. 1997] L. Alonso Tarrío, A. Jeremías López, and J. Lipman, "Local homology and cohomology on schemes", *Ann. Sci. École Norm. Sup.* (4) **30**:1 (1997), 1–39. MR 1422312 Zbl 0894.14002

[Alonso Tarrío et al. 1999] L. Alonso Tarrío, A. Jeremías López, and J. Lipman, *Duality and flat base change on formal schemes*, Contemporary Mathematics **244**, American Mathematical Society, Providence, RI, 1999. Correction at *Proc. Amer. Math. Soc.* **131**:2 (2003), 351–357. MR 1716706 Zbl 0927.00024

[Avramov 1998] L. L. Avramov, "Infinite free resolutions", pp. 1–118 in *Six lectures on commutative algebra* (Bellaterra, 1996), edited by J. Elias et al., Progress in Mathematics **166**, Birkhäuser, Basel, 1998. MR 1648664 Zbl 0934.13008

[Braunling et al. 2015] O. Braunling, M. Groechenig, and J. Wolfson, "A generalized Contou-Carrère symbol and its reciprocity laws in higher dimensions", preprint, 2015. arXiv 1410.3451v2

[Brodmann and Sharp 2013] M. P. Brodmann and R. Y. Sharp, *Local cohomology: an algebraic introduction with geometric applications*, 2nd ed., Cambridge Studies in Advanced Mathematics **136**, Cambridge University Press, 2013. MR 3014449 Zbl 1263.13014

[Buchweitz and Flenner 2006] R.-O. Buchweitz and H. Flenner, "Power series rings and projectivity", *Manuscripta Math.* **119**:1 (2006), 107–114. MR 2194381 Zbl 1084.13005

[Cartan and Eilenberg 1956] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton University Press, 1956. MR 0077480 Zbl 0075.24305

[Frankild 2003] A. Frankild, "Vanishing of local homology", *Math. Z.* **244**:3 (2003), 615–630. MR 1992028 Zbl 1020.13003

[Greenlees and May 1992] J. P. C. Greenlees and J. P. May, "Derived functors of I-adic completion and local homology", *J. Algebra* **149**:2 (1992), 438–453. MR 1172439 Zbl 0774.18007

[Grothendieck and Dieudonne 1960] A. Grothendieck and J. Dieudonne, "Éléments de géométrie algébrique, I: Le langage des schémas", *Inst. Hautes Études Sci. Publ. Math.* **4** (1960), 5–214. MR 0217083 Zbl 0118.36206

[Hartshorne 1966] R. Hartshorne, *Residues and duality*, Lecture Notes in Mathematics **20**, Springer, Berlin, 1966. MR 0222093 Zbl 0212.26101

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

[Hochschild 1945] G. Hochschild, "On the cohomology groups of an associative algebra", *Ann. of Math.* (2) **46** (1945), 58–67. MR 0011076 Zbl 0063.02029

[Hübl 1989] R. Hübl, *Traces of differential forms and Hochschild homology*, Lecture Notes in Mathematics **1368**, Springer, Berlin, 1989. MR 995670 Zbl 0675.13019

[Ionescu 2001] C. Ionescu, "Hochschild (co)homology in commutative algebra: a survey", *An. Ştiinţ. Univ. Ovidius Constanţa Ser. Mat.* **9**:2 (2001), 87–96. MR 2075931 Zbl 1090.13502

[Keller 1994] B. Keller, "Deriving DG categories", *Ann. Sci. École Norm. Sup.* (4) **27**:1 (1994), 63–102. MR 1258406 Zbl 0799.18007

[Loday 1998] J.-L. Loday, *Cyclic homology*, 2nd ed., Grundlehren der Mathematischen Wissenschaften **301**, Springer, Berlin, 1998. MR 1600246 Zbl 0885.18007

[Lurie 2011] J. Lurie, "Derived algebraic geometry, XII: Proper morphisms, completions, and the Grothendieck existence theorem", preprint, 2011, http://www.math.harvard.edu/~lurie/papers/DAG-XII.pdf.

[Mac Lane 1963] S. Mac Lane, *Homology*, Grundlehren der Mathematischen Wissenschaften **114**, Springer, Berlin, 1963. Reprinted in Classics in Mathematics, 1995. MR 0156879 Zbl 0133.26502

[Porta et al. 2014a] M. Porta, L. Shaul, and A. Yekutieli, "Completion by derived double centralizer", *Algebr. Represent. Theory* **17**:2 (2014), 481–494. MR 3181733 Zbl 06306474

[Porta et al. 2014b] M. Porta, L. Shaul, and A. Yekutieli, "On the homology of completion and torsion", *Algebr. Represent. Theory* **17**:1 (2014), 31–67. MR 3160712 Zbl 06306453

[Porta et al. 2015] M. Porta, L. Shaul, and A. Yekutieli, "Cohomologically cofinite complexes", *Comm. Algebra* **43**:2 (2015), 597–615. MR 3274024 Zbl 06422500

[Schenzel 2003] P. Schenzel, "Proregular sequences, local cohomology, and completion", *Math. Scand.* **92**:2 (2003), 161–180. MR 1973941 Zbl 1023.13011

[Shaul 2016] L. Shaul, "Tensor product of dualizing complexes over a field", preprint, 2016. To appear in *J. Commut. Algebra*. arXiv 1412.3759v2

[Simon 1990] A.-M. Simon, "Some homological properties of complete modules", *Math. Proc. Cambridge Philos. Soc.* **108**:2 (1990), 231–246. MR 1074711 Zbl 0719.13007

[Spaltenstein 1988] N. Spaltenstein, "Resolutions of unbounded complexes", *Compositio Math.* **65**:2 (1988), 121–154. MR 932640 Zbl 0636.18006

[Stacks 2005–] "The Stacks Project", electronic reference, 2005–, http://stacks.math.columbia.edu.

[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

[Yekutieli 2011] A. Yekutieli, "On flatness and completion for infinitely generated modules over Noetherian rings", *Comm. Algebra* **39**:11 (2011), 4221–4245. MR 2855123 Zbl 1263.13028

[Yekutieli 2016] A. Yekutieli, "The squaring operation for commutative DG rings", *J. Algebra* **449** (2016), 50–107. MR 3448166

Communicated by Hubert Flenner
Received 2015-05-23 Revised 2016-03-04 Accepted 2016-05-16

Liran.Shaul@uantwerpen.be *Departement Wiskunde-Informatica, Universiteit Antwerpen, Middelheim Campus, Middelheimlaan 1, 2020 Antwerp, Belgium*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.1031

Bifurcations, intersections, and heights

Laura DeMarco

We prove the equivalence of dynamical stability, preperiodicity, and canonical height 0, for algebraic families of rational maps $f_t : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$, parameterized by t in a quasiprojective complex variety. We use this to prove one implication in the if-and-only-if statement of a certain conjecture on unlikely intersections in the moduli space of rational maps (see "Special curves and postcritically finite polynomials", *Forum Math. Pi* **1** (2013), e3). We present the conjecture here in a more general form.

1. Introduction

Let $f : V \times \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$ be an *algebraic family* of rational maps of degree $d \geq 2$. That is, V is an irreducible quasiprojective complex variety, and f is a morphism such that $f_t := f(t, \cdot) : \mathbb{P}^1 \to \mathbb{P}^1$ has degree d for all $t \in V$. Fix a morphism $a : V \to \mathbb{P}^1$, which we view as a *marked point* on $\mathbb{P}^1$. When V is a curve, we will alternatively view f as a rational function defined over the function field $k = \mathbb{C}(V)$, with $a \in \mathbb{P}^1(k)$. In this article, we study the relation between dynamical stability of the pair (f, a), preperiodicity of the point a, and the canonical height of a (defined over the field k). In the final section, we present the general form of a conjecture on density of "special points" in this setting of dynamics on $\mathbb{P}^1$ (which includes as a special case some known statements about points on elliptic curves) — see Conjecture 6.1; compare [Baker and DeMarco 2013, Conjecture 1.10]. We finish the article with the proof of one part of the conjecture, as an application of this study of stability in algebraic families.

Stability. The pair (f, a) is said to be *stable* if the sequence of iterates

$$\{t \mapsto f_t^n(a(t))\}_{n \geq 1}$$

forms a normal family on V. (Recall that a family of holomorphic maps is normal if it is precompact in the topology of uniform convergence on compact subsets; i.e., any sequence contains a locally uniformly convergent subsequence.) The pair (f, a) is *preperiodic* if there exist integers $m > n \geq 0$ such that $f_t^m(a(t)) = f_t^n(a(t))$ for

MSC2010: primary 37P30; secondary 37F45, 11G05.
Keywords: dynamics of rational maps, canonical height, stability.

all $t \in V$. The pair (f, a) is *isotrivial* if there exists a branched cover $W \to V$ and an algebraic family of Möbius transformations $M : W \times \mathbb{P}^1 \to \mathbb{P}^1$ such that $M_t \circ f_t \circ M_t^{-1} : \mathbb{P}^1 \to \mathbb{P}^1$ and $M_t(a(t)) \in \mathbb{P}^1$ are independent of t.

It is immediate from the definitions that either preperiodicity or isotriviality will imply stability. In this article, we prove the converse:

Theorem 1.1. *Let f be an algebraic family of rational maps of degree $d \geq 2$, and let a be a marked point. Suppose (f, a) is stable. Then either (f, a) is isotrivial or it is preperiodic.*

This is a generalization of [Dujardin and Favre 2008, Theorem 2.5], which itself extends [McMullen 1987, Theorem 2.2], treating the case where a is a critical point of f. In the study of complex dynamics, it is well known that a holomorphic family $f : X \times \mathbb{P}^1 \to \mathbb{P}^1$, for X any complex manifold, is dynamically stable if and only if the pair (f, c) is stable for all critical points c of f [Lyubich 1983; Mañé et al. 1983; McMullen 1994, Chapter 4]. For nonisotrivial algebraic families $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$, McMullen [1987, Lemma 2.1] proved that dynamical stability on all of V implies that all critical points are preperiodic. Combining this with Thurston's rigidity theorem, he concluded that a nonisotrivial stable family must be a family of flexible Lattès maps (i.e., covered by an endomorphism of a nonisotrivial family of elliptic curves).

Canonical height. One step in the proof of Theorem 1.1 provides an elementary geometric proof of Baker's theorem [2009, Theorem 1.6] on the finiteness of rational points with small height, for the canonical height $\hat{h}_f$ associated to the function field $\mathbb{C}(V)$, when the variety V has dimension 1. In fact, we obtain his statement under the weaker hypothesis that f is not isotrivial over $k = \mathbb{C}(V)$ (rather than assuming f is nonisotrivial over any extension of k); see [Baker 2009, Remark 1.7(i)]. The map f is *isotrivial over* $k = \mathbb{C}(V)$ if there exists an algebraic family of Möbius transformations $M : V \times \mathbb{P}^1 \to \mathbb{P}^1$ such that $M_t \circ f_t \circ M_t^{-1}$ is independent of t.

Theorem 1.2. *Suppose f is a rational function defined over the function field $k = \mathbb{C}(V)$, of degree ≥ 2, and assume that V has dimension 1. Let $\hat{h}_f$ be the canonical height of f. If f is not isotrivial over k, then there exists a $b > 0$ such that the set $\{a \in \mathbb{P}^1(k) : \hat{h}_f(a) < b\}$ is finite.*

Remark 1.3. Theorem 1.2 contains as a special case the corresponding result about rational points on an elliptic curve E over k, equipped with the Néron–Tate height, generally attributed to Lang and Néron [1959]; see [Silverman 1994, Theorem III.5.4] for a proof. It was a step in proving the Mordell–Weil theorem for function fields. (To treat this case, we project E to $\mathbb{P}^1$ and let f be the rational function induced by multiplication-by-2 on E.) In this setting, one can say more: the set of points in $\mathbb{P}^1(k)$ with height $< b$ that lift to rational points in $E(k)$ is finite for *all* $b > 0$; see [Baker 2009, Theorem B.9], where this is deduced from the conclusion of Theorem 1.2.

The canonical height $\hat{h}_f$ was introduced in [Call and Silverman 1993], and it satisfies $\hat{h}_f(f(a)) = d\hat{h}_f(a)$ when f has degree d. So Theorem 1.2 implies that rational points of height 0 are preperiodic unless f is isotrivial over k. This was proved for polynomials f in [Benedetto 2005]. When k is a number field, this was observed in [Call and Silverman 1993], and the conclusion of Theorem 1.2 holds for all $b > 0$. In the function field setting, the conclusion of Theorem 1.2 *cannot* hold for all $b > 0$, since, for example, the union of all constant points $a \in \mathbb{P}^1(\mathbb{C})$ will form an infinite set of bounded canonical height for any f. On page 1040, we provide explicit examples of functions f for which we compute the sharp bound b and the total number of rational preperiodic points. Also, note that examples do exist of rational functions $f \in k(z)$ that are isotrivial but *not* isotrivial over $k = \mathbb{C}(V)$. A necessary condition is a nontrivial automorphism group of f; see Example 2.2.

Combining Theorem 1.1 with Theorem 1.2, we have:

Theorem 1.4. *Suppose $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ is a nonisotrivial algebraic family of rational maps, where V has dimension* 1. *Let $\hat{h}_f : \mathbb{P}^1(\bar{k}) \to \mathbb{R}$ be a canonical height of f, defined over the function field $k = \mathbb{C}(V)$. For each $a \in \mathbb{P}^1(\bar{k})$, the following are equivalent*:

(1) *The pair (f, a) is stable.*

(2) $\hat{h}_f(a) = 0$.

(3) *(f, a) is preperiodic.*

Moreover, the set $\{a \in \mathbb{P}^1(k) : (f, a)$ is stable$\}$ is finite.

Application to intersection theory. Combining Theorem 1.1 with Montel's theorem on normal families, we obtain another argument for the "easy" implication of the Masser–Zannier theorems [2010; 2012] on anomalous torsion for elliptic curves.

Proposition 1.5. *Suppose that E is any nonisotrivial elliptic curve defined over a function field $k = \mathbb{C}(V)$, where V has dimension* 1, *and let P be a point of $E(k)$. Then the set of $t \in V$ for which the point P_t is torsion on E_t is infinite.*

The harder part of the Masser–Zannier theorems is the following statement: if two points P and Q in $E(k)$ are independent on E and neither is torsion, meaning that they do not satisfy a relation of the form $mP + nQ = 0$ with integers m and n not both zero, then the set of $t \in V$ for which P_t and Q_t are both torsion on E_t is finite. (In [DeMarco et al. 2016], we gave a dynamical proof of this harder implication for the Legendre family E_t.)

The theorems of [Masser and Zannier 2012], followed by a series of analogous results in the dynamical setting (e.g., [Baker and DeMarco 2011; Ghioca et al.

2013; 2015; DeMarco et al. 2015]), led to the development of a general conjecture about rational maps and marked points — addressing a question first posed by Umberto Zannier, but also encompassing a case of intrinsic dynamical interest, where the marked points are the critical points of the map. A precise statement of this conjecture appears as Conjecture 6.1 in Section 6. In [Baker and DeMarco 2013], we formulated the conjecture in the setting of marked critical points; an error in one of our definitions is corrected here (Remark 6.3). In this article, we use Theorem 1.1 to give a proof of one implication of the more general statement of Conjecture 6.1. This implication reduces to proving the following statement.

Every algebraic family $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ induces a (regular) projection $V \to \mathrm{M}_d$ from the parameter space to the moduli space M_d of conformal conjugacy classes of maps. We say the family f has *dimension N in moduli* if the image of V under this projection has dimension N in M_d. Since V is irreducible, f has dimension 0 in moduli if and only if f is isotrivial.

Theorem 1.6. *Let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of rational maps of degree $d \geq 2$, of dimension $N > 0$ in moduli. Let $a_1, \dots, a_k$, with $k \leq N$, be any marked points. Then the set*

$$S(a_1, \dots, a_k) = \bigcap_{i=1}^{k} \{t \in V : a_i(t) \textit{ is preperiodic for } f_t\}$$

is Zariski-dense in V.

Remark 1.7. Conjecture 6.1 asserts that the set $S(a_1, \dots, a_k)$ with $k > N$ will be Zariski-dense if and only if at most N of the points $a_1, \dots, a_k$ are dynamically "independent" on V.

Idea of the proof of Theorem 1.1. The proofs of [McMullen 1987, Theorem 2.2] and [Dujardin and Favre 2008, Theorem 2.5] use crucially that the point is critical. The first ingredient of our proof is similar to their proofs, building upon the fact that there are only finitely many nonconstant morphisms from a quasiprojective algebraic curve V to $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. This leads to the proof of Theorem 1.2. As a special case of Theorem 1.2, we have:

Proposition 1.8. *Let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of rational maps of degree $d \geq 2$, and assume that V has complex dimension 1. Fix a marked point $a : V \to \mathbb{P}^1$, and define $g_n : V \to \mathbb{P}^1$ by $g_n(t) = f_t^n(a(t))$. If f is not isotrivial, and if the degrees of $\{g_n\}$ are bounded, then there exist integers $m > n \geq 0$ such that*

$$f_t^m(a(t)) = f_t^n(a(t)) \quad \textit{for all } t \in V.$$

Remark 1.9. If f is isotrivial, then the degrees of $g_n(t) = f_t^n(a(t))$ are bounded if and only if the pair (f, a) is isotrivial. (See Proposition 2.3.)

For Theorem 1.1, it suffices to treat the case where V has complex dimension 1 and so is a finitely punctured Riemann surface. Then each $g_n(t) = f_t^n(a(t))$ extends to a holomorphic map on the compactification of V. In light of Proposition 1.8, it remains to prove that stability on V implies the degrees of $\{g_n\}$ are bounded.

The second ingredient of the proof is a study of normality and escape rates near the punctures of V. The arguments given in Section 3 are inspired by the methods of [DeMarco et al. 2015] and [DeMarco et al. 2016].

From a geometric point of view, the idea to show that the degrees of $\{g_n\}$ are bounded is as follows. Let X denote (the normalization of) a compactification of V. Under iteration, one would typically expect that $\deg g_n \approx d \deg g_{n-1}$, where $d = \deg f$. Viewing g_n as a curve in $X \times \mathbb{P}^1$ (by identifying the function with its graph), the expected degree growth fails when the graph of g_n passes through the indeterminacy points of $(t, z) \mapsto (t, f_t(z))$ and thus its image contains "vertical components" over the punctures of V. Lemma 3.3 shows that the multiplicity of the vertical component in the image of any curve is uniformly bounded by some integer q. On the other hand, normality on V implies that the graphs of g_n are converging over compact subsets of V. Therefore, if $\deg g_n \to \infty$, the graph of g_n must be fluctuating wildly near the punctures of V when n is large (Lemma 4.1). But this fluctuation is controlled by Proposition 3.1.

2. Isotriviality and Theorem 1.2

Throughout this section, we assume that V is an irreducible quasiprojective complex variety of dimension 1; i.e., V is obtained from a compact Riemann surface by removing finitely many points. Let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of rational maps of degree $d \geq 2$. We prove Theorem 1.2. We also prove Proposition 2.3, to provide a characterization of isotriviality which is used in the proof of Theorem 1.1; it is not needed for the proof of Theorem 1.2.

Isotrivial maps. By definition, f is isotrivial if there exists a family $\{M_t\}$ of Möbius transformations, regular over a branched cover $p : W \to V$, such that $M_t \circ f_{p(t)} \circ M_t^{-1}$ is constant in t. For a marked point $a : V \to \mathbb{P}^1$, the pair (f, a) is isotrivial if, in addition, the function $M_t(a(p(t)))$ is constant on W. (Note that this is well defined, even if the family M_t is not uniquely determined.) The map f (or the pair (f, a)) is isotrivial over $k = \mathbb{C}(V)$ if M can be chosen to be an algebraic family that is regular on V.

Lemma 2.1. *Let V have dimension* 1. *If (f, a) is isotrivial, and if $\{a, f(a), f^2(a)\}$ is a set of three distinct functions on V, then (f, a) is isotrivial over k.*

Proof. Suppose (f, a) is isotrivial. Let $M : W \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of Möbius transformations, with branched cover $p : W \to V$, such that $R =$

$M_w \circ f_{p(w)} \circ M_w^{-1} : \mathbb{P}^1 \to \mathbb{P}^1$ and $b = M_w(a(p(w)))$ are independent of $w \in W$. By removing finitely many points from V, we may assume that $p : W \to V$ is a covering map.

Fix a basepoint $t_0 \in V$ and choose a point $w_0 \in p^{-1}(t_0)$. The choices of basepoint determine a representation

$$\rho_f : \pi_1(V, t_0) \to \mathrm{Aut}(f_{t_0}) \subset \mathrm{PSL}_2(\mathbb{C})$$

that is trivial if and only if (f, a) is isotrivial over k. Indeed, choose any $\gamma \in \pi_1(V, t_0)$ and let $\eta : [0, 1] \to W$ be a lift of γ with $\eta(0) = w_0$. Write w_t for $\eta(t)$. Then the equality $f_{p(w_0)} = f_{p(w_1)} = f_{t_0}$ and isotriviality imply that $M_{w_0} f_{t_0} M_{w_0}^{-1} = M_{w_1} f_{t_0} M_{w_1}^{-1}$, so $\rho_f(\gamma) := M_{w_1}^{-1} M_{w_0}$ is an automorphism of f_{t_0}. The triviality of ρ_f is equivalent to the statement that $M_{w_0} = M_{w_1}$ for all such paths, so that M descends to a regular map $M : V \times \mathbb{P}^1 \to \mathbb{P}^1$.

Now choose the basepoint t_0 so that $\{a(t_0), f_{t_0}(a(t_0)), f_{t_0}^2(a(t_0))\}$ are three distinct points in $\mathbb{P}^1$. Fix any $\gamma \in \pi_1(V, t_0)$. For each $n \geq 0$, we have

$$\begin{aligned} \rho_f(\gamma)(f_{t_0}^n(a(t_0))) &= f_{t_0}^n(\rho_f(\gamma)(a(t_0))) = f_{t_0}^n M_{w_1}^{-1} M_{w_0}(a(p(w_0))) \\ &= f_{t_0}^n M_{w_1}^{-1}(b) = f_{t_0}^n(a(p(w_1)) = f_{t_0}^n(a(t_0)), \end{aligned}$$

so the full orbit of $a(t_0)$ under f_{t_0} lies in the fixed set of $\rho_f(\gamma)$. By the assumption on a we deduce that $\rho_f(\gamma)$ is the identity. Therefore, (f, a) is isotrivial over k. $\square$

Example 2.2. Consider the rational function $f_1(z) = z + 1/z$ or the cubic polynomial $P_1(z) = z^3 - 3z$. Both of these functions have $z \mapsto -z$ as an automorphism. Conjugating by $M_t(z) = tz$ and setting $s = t^2$, we see that the families

$$f_s(z) = z + \frac{1}{sz} \quad \text{and} \quad P_s(z) = sz^3 - 3z$$

are isotrivial over a degree-2 extension of $k = \mathbb{C}(s)$. On the other hand, neither f nor P is isotrivial over k. This can be seen by computing the critical points ($= \pm\sqrt{1/s}$ in both examples), and observing that the critical points are interchanged by a nontrivial loop in $V = \mathbb{C} \setminus \{0\}$.

Proposition 2.3. *Suppose V has dimension* 1 *and f is isotrivial. Let $a : V \to \mathbb{P}^1$ be any marked point. The following are equivalent*:

(1) *The pair (f, a) is isotrivial.*

(2) *The pair (f, a) is stable.*

(3) *The degrees of $g_n(t) = f_t^n(a(t))$ are bounded.*

Proof. Since f is isotrivial, there exist a finite branched cover $p : W \to V$, an algebraic family of Möbius transformations $M : W \times \mathbb{P}^1 \to \mathbb{P}^1$, and a map $R : \mathbb{P}^1 \to \mathbb{P}^1$ such that $M_t \circ f_{p(t)} \circ M_t^{-1} = R$ for all $t \in W$. Set $s = p(t)$.

If the pair (f, a) is isotrivial, then $b = M_t(a(s))$ is also independent of t, so the degrees of $g_n(t) = f_s^n(a(s)) = M_t^{-1}(R^n(b))$ are clearly bounded. Thus (1) implies (3). In addition, note that any sequence in the set $\{R^n(b)\}_{n\geq 1} \subset \mathbb{P}^1$ has a convergent subsequence. This implies the normality of $\{M_t^{-1}(R^n(b))\}_{n\geq 1}$ on W which in turn implies the normality of $\{f_t^n(a(t))\}_n$ on V. Thus (1) implies (2).

Now suppose that (f, a) is not isotrivial, so that $b(t) = M_t(a(s))$ is nonconstant on W. We observe first that the degrees of $\{g_n\}$ must be unbounded, showing that (3) implies (1). Indeed, since b is nonconstant, it extends to a surjective map of finite degree from a compactification $\overline{W}$ to $\mathbb{P}^1$. Choose any point $z_0 \in \mathbb{P}^1$ which is nonexceptional for R, that is, such that the set of preimages $R^{-n}(z_0)$ is growing in cardinality as $n \to \infty$. For any $D > 0$, choose n so that the size of the set $R^{-n}(z_0)$ is larger than D. Then there is a set $P \subset \overline{W}$ of cardinality $|P| \geq D$ such that $b(t)$ is in $R^{-n}(z_0)$ for all $t \in P$. Then $R^n(b(t)) = z_0$ for all $t \in P$. Taking D as large as desired, this shows that the degrees of $\{t \mapsto R^n(b(t))\}_n$ are unbounded. This in turn implies that the degrees of $g_n(t) = M_t^{-1}(R^n(b(t)))$ are unbounded.

Continuing to assume that (f, a) is not isotrivial, we also see that $b(t) = M_t(a(s))$ has only finitely many critical points in W and the image $b(W)$ omits at most finitely many points in $\mathbb{P}^1$. The map R has infinitely many repelling cycles in its Julia set, so there must exist a $t_0 \in W$ such that $b'(t_0) \neq 0$ and $b(t_0)$ is a repelling periodic point of R for all $t \in P$. It follows that the sequence of derivatives $\frac{\partial}{\partial t} R^n(b(t))|_{t=t_0}$ is unbounded; so the sequence $\{R^n(b(t)) = M_t(f_s^n(a(s)))\}_n$ cannot be a normal family on all of W. Therefore, $\{g_n\}_n$ also fails to be normal, and so we have proved that (2) implies (1). □

Finiteness of nonconstant maps. As in the proofs of Theorem 2.2 (and specifically Proposition 4.3) in [McMullen 1987] and of Theorem 2.5 in [Dujardin and Favre 2008], we will need the following statement for our proof of Theorem 1.2.

Lemma 2.4. *Let Λ be any quasiprojective, complex algebraic curve. There are only finitely many nonconstant holomorphic maps from Λ to the triply punctured sphere $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. The bound depends only on the Euler characteristic $\chi(\Lambda)$.*

Proof. Any holomorphic map $h : \Lambda \to \mathbb{P}^1 \setminus \{0, 1, \infty\}$ extends to a meromorphic function on a (smooth) compactification X of Λ. From Riemann–Hurwitz, the degree of h is bounded by the Euler characteristic $-\chi(\Lambda)$. Any meromorphic function on X is determined by its zeros, poles, and ones; indeed, the ratio of two functions h_1 and h_2 with the same zeros and poles must be constant on X, and if $h_1(x) = 1 = h_2(x)$ for some x, then $h_1 \equiv h_2$. Thus, there are only finitely many combinatorial possibilities for h. □

Proof of Theorem 1.2. Let $d = \deg f \geq 2$. The canonical height $\hat{h}_f(a)$ computes the growth rate of the degrees of $t \mapsto f_t^n(a(t))$ as $n \to \infty$. Precisely, each $a \in \mathbb{P}^1(\bar{k})$

determines a meromorphic function $W \to \mathbb{P}^1$ on a branched cover $p_a : W \to V$, where the topological degree of p_a coincides with the algebraic degree of the field extension $k(a)$ over k. The canonical height is computed as

$$\hat{h}_f(a) = \frac{1}{\deg p_a} \lim_{n\to\infty} \frac{1}{d^n} \deg g_n$$

for the maps $g_n : W \to \mathbb{P}^1$ defined by $g_n(s) = f^n_{p_a(s)}(a(s))$. This height function $\hat{h}_f$ is characterized by two conditions [Call and Silverman 1993, Theorem 1.1]:

(1) The difference $|\hat{h}_f(a) - \deg a/\deg p_a|$ is uniformly bounded on $\mathbb{P}^1(\bar{k})$.

(2) $\hat{h}_f(f(a)) = d\hat{h}_f(a)$ for all $a \in \mathbb{P}^1(\bar{k})$.

In particular, the degrees of $\{g_n\}$ are growing to infinity if and only if the canonical height of a is positive.

Suppose there is a sequence of rational points $a_m \in \mathbb{P}^1(k),\ m \geq 1$, such that

$$1 > \hat{h}_f(a_m) \to 0$$

as $m \to \infty$. For each point $a \in \mathbb{P}^1(k)$, the *length* of the orbit of a is defined to be the cardinality of the set $\{f^n(a) : n \geq 0\}$ in $\mathbb{P}^1(k)$. Suppose further that only finitely many of the a_m have infinite orbit, and that the finite orbit lengths are uniformly bounded. In this case, all but finitely many of the a_m satisfy a finite number of equations of the form $f^n(a) = f^\ell(a)$ with $n \neq \ell$; thus the set $\{a_m\}$ will be finite. If this holds for any such sequence, then the theorem is proved.

We can assume, therefore, that the orbit lengths of the a_m are tending to infinity with m or are equal to infinity for all m.

From property (1) of the height function, there exists a degree D such that $\deg a \geq D$ with $a \in \mathbb{P}^1(k)$ implies $\hat{h}_f(a) \geq 1$. For each m, choose an integer $N_m \geq 0$ so that the orbit of a_m has length greater than N_m and so that

$$\deg f^i(a_m) \leq D$$

for all $i \leq N_m$. Property (2) of the height function and the condition $\hat{h}_f(a_m) \to 0$ imply that we may take $N_m \to \infty$ as $m \to \infty$. We will deduce that f must be isotrivial over k.

Suppose now that f_t has at least three distinct fixed points for general $t \in V$. Remove the finitely many parameters in V where these three fixed points have collisions. Then there exists a branched cover $p : W \to V$ and an algebraic family of Möbius transformations $M : W \times \mathbb{P}^1 \to \mathbb{P}^1$ such that

$$R_t = M_t \circ f_{p(t)} \circ M_t^{-1}$$

has its fixed points at 0, 1, ∞ for all $t \in W$.

Set $b_m(t) = M_t(a_m(p(t)))$ for each m, so that $R^i(b_m) = M(f^i(a_m))$ for all iterates. The uniform bound of D on the degrees of $\bigcup_m \{f^i(a_m) : i \leq N_m\}$ implies that there is a uniform bound of D' on the degrees of $\bigcup_m \{R^i(b_m) : i \leq N_m\}$. For each point b_m, we define

$$S_{0,m}(n) = \{t \in W : R_t^n(b_m(t)) = 0\},$$
$$S_{1,m}(n) = \{t \in W : R_t^n(b_m(t)) = 1\},$$
$$S_{\infty,m}(n) = \{t \in W : R_t^n(b_m(t)) = \infty\}.$$

Since $\{0, 1, \infty\}$ are fixed points of R_t for all t, we have

$$S_{0,m}(n) \subset S_{0,m}(n+1)$$

for all m and all n; similarly for $S_{1,m}(n)$ and $S_{\infty,m}(n)$. Let

$$S_m = S_{0,m}(N_m) \cup S_{1,m}(N_m) \cup S_{\infty,m}(N_m).$$

For each m and each $n \leq N_m$, the iterate $R^n(b_m)$ determines a holomorphic map

$$W \setminus S_m \longrightarrow \mathbb{P}^1 \setminus \{0, 1, \infty\}.$$

By construction, the degree of $R^{N_m}(b_m)$ is bounded by D', so we have $|S_m| \leq 3D'$ for all m. Therefore, there is a uniform bound B (independent of m) on the number of nonconstant maps from $W \setminus S_m$ to the triply punctured sphere $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ (Lemma 2.4). In other words, for each m, at most B of the first N_m iterates of b_m are nonconstant. Therefore, since $N_m \to \infty$, there exists an m_0 such that b_{m_0} has at least $2d + 2$ consecutive iterates that are constant and distinct.

Lemma 2.5. *Suppose A and B are rational functions of degree d such that we have $A(x_i) = B(x_i)$ for a sequence of $2d + 1$ distinct points $x_1, \ldots, x_{2d+1}$. Then $A = B$.*

Proof. By postcomposing A and B with a Möbius transformation, we may assume that the values $A(x_i)$ and $B(x_i)$ are finite for each i. Consider the difference $F = A - B$. Then F is a rational function of degree $\leq 2d$. But F vanishes in at least $2d + 1$ distinct points, so $F \equiv 0$. □

Set $x_i = R^i(b_{m_0})$ for the consecutive indices i for which x_i is constant in t. Then $R_t(x_i) = x_{i+1}$ for all t along a set of $2d + 1$ distinct points $x_i \in \mathbb{P}^1(\mathbb{C})$. Applying Lemma 2.5, we conclude that the rational function R_t of degree d is independent of t. In other words, f is isotrivial.

It remains to show that f is in fact isotrivial over k, but this follows from Lemma 2.1. Indeed, for the point b_{m_0} in the preceding paragraph, we had $x_i = R^i(b_{m_0})$ independent of t. In other words, the pair $(f, f^i(a_{m_0}))$ is isotrivial. In addition, the orbit length of $f^i(a_{m_0})$ is at least $2d + 1 \geq 3$. Lemma 2.1 states that the pair $(f, f^i(a_{m_0}))$ must be isotrivial over k, so f itself is isotrivial over k.

Finally, suppose that f_t has only 1 or 2 fixed points, generally in V. For a general parameter t_0, choose a forward-invariant set of fixed points and preimages, consisting of at least three distinct points. Pass to a branched cover $W \to V$ on which these points can be marked holomorphically, excluding the finitely many points where collisions occur. For each of these three points, we define the sets $S_{i,m}(n)$ as above. If the i-th point is mapped to the j-th point by f_t, then $S_i(n) \subset S_j(n+1)$ for all n. The rest of the proof goes through exactly the same. This completes the proof of Theorem 1.2.

Two examples: computing canonical height and the number of rational preperiodic points. Consider $Q_t(z) = z^2 + t$, the family of quadratic polynomials. This family defines a nonisotrivial rational function Q over the function field $k = \mathbb{C}(t)$. There is a unique point in $\mathbb{P}^1(k)$ with finite orbit for Q, namely the point $a = \infty$. Indeed, writing $a(t) = a_1(t)/a_2(t)$ for $a \in \mathbb{P}^1(k) \setminus \{\infty\}$, we can compute explicitly that

$$\deg(Q(a)) = \begin{cases} 2\deg a & \text{if } \deg a_1 > \deg a_2, \\ 2\deg a + 1 & \text{if } \deg a_1 \le \deg a_2. \end{cases}$$

In both cases, the image $Q(a)$ will satisfy the hypothesis of the first case. Inductively then, we have $\deg Q^n(a) = 2^{n-1} \deg Q(a)$. Consequently, the largest possible b in the statement of Theorem 1.2 is $b = \frac{1}{2}$, since the set $\left\{a \in \mathbb{P}^1(k) : \hat{h}_Q(a) = \frac{1}{2}\right\}$ is precisely the constant points $a \in \mathbb{C}$, while

$$\left|\left\{a \in \mathbb{P}^1(k) : \hat{h}_Q(a) < \tfrac{1}{2}\right\}\right| = 1.$$

As a second example, consider the family of flexible Lattès maps,

$$L_t(z) = \frac{(z^2 - t)^2}{4z(z-1)(z-t)},$$

defining a nonisotrivial L over the field $k = \mathbb{C}(t)$. This family is the quotient of the endomorphism $P \mapsto P + P$ on the Legendre family of elliptic curves $E_t = \{y^2 = x(x-1)(x-t)\}$. (See also the beginning of Section 5, where this is discussed further.) For this example, Proposition 1.4 of [DeMarco et al. 2016] shows there are exactly 4 rational preperiodic points, namely $\{0, 1, t, \infty\}$. We also explicitly computed the height of any starting point $a \in \mathbb{P}^1(k)$ in Proposition 3.1 of the same work. The constant points $a \in \mathbb{C} \setminus \{0, 1\}$ form an infinite set of points of canonical height $\frac{1}{2}$, while

$$\left|\left\{a \in \mathbb{P}^1(k) : \hat{h}_L(a) < \tfrac{1}{2}\right\}\right| = 4.$$

Again, $b = \frac{1}{2}$ is the largest possible constant in the statement of Theorem 1.2.

3. Escape rate at a degenerate parameter

In this section, we construct a "good" escape-rate function associated to a pair of holomorphic maps $f : \mathbb{D}^* \times \mathbb{P}^1 \to \mathbb{P}^1$ and $a : \mathbb{D}^* \to \mathbb{P}^1$, on the punctured unit disk $\mathbb{D}^* = \{t \in \mathbb{C} : 0 < |t| < 1\}$. The construction follows that of [DeMarco et al. 2015; 2016]. I am indebted to Hexi Ye for his assistance in the proof of Proposition 3.1.

The setting. Throughout this section, we work in homogeneous coordinates on $\mathbb{P}^1$. We assume we are given a family of homogeneous polynomial maps

$$F_t : \mathbb{C}^2 \to \mathbb{C}^2,$$

of degree $d \geq 2$, parameterized by $t \in \mathbb{D} = \{t \in \mathbb{C} : |t| < 1\}$, such that the coefficients of F_t are holomorphic in t. Each F_t is given by a pair of homogeneous polynomials (P_t, Q_t), and we define $\mathrm{Res}(F_t)$ to be the homogeneous resultant of the polynomials P_t and Q_t. Recall that the resultant is a polynomial function of the coefficients of F_t, vanishing if and only if P_t and Q_t share a root in $\mathbb{P}^1$. See [Silverman 2007, §2.4] for more information. We assume further that $\mathrm{Res}(F_t) = 0$ if and only if $t = 0$, and also that at least one coefficient of F_0 is nonzero.

We use the norm

$$\|(z_1, z_2)\| = \max\{|z_1|, |z_2|\} \quad \text{on } \mathbb{C}^2.$$

The escape-rate function. Let $A : \mathbb{D} \to \mathbb{C}^2 \setminus \{(0, 0)\}$ be any holomorphic map. Write A_t for $A(t)$.

For each $n \geq 0$, the iterate $F_t^n(A_t)$ is a pair of holomorphic functions in t; we define

$$a_n = \mathrm{ord}_{t=0} \, F_t^n(A_t)$$

to be the minimum of the order of vanishing of the two coordinate functions at $t = 0$; so $a_0 = 0$ and a_n is a nonnegative integer for all $n \geq 1$. Set

$$F_n(t) = t^{-a_n} F_t^n(A_t)$$

so that F_n is a holomorphic map from $\mathbb{D}$ to $\mathbb{C}^2 \setminus \{(0, 0)\}$ for each n. Our main goal in this section is to prove the following statement.

Proposition 3.1. *The functions*

$$G_n(t) = \frac{1}{d^n} \log \|F_n(t)\|$$

converge locally uniformly on the punctured disk $\mathbb{D}^*$ to a continuous function G satisfying

$$G(t) = o(\log |t|) \quad \text{as } t \to 0.$$

Remark 3.2. In [DeMarco et al. 2015; 2016], we used explicit expressions for F_t to deduce that the function G was *continuous* at $t = 0$ for our examples. It remains an interesting open question to determine necessary and sufficient conditions for the functions G_n to converge uniformly to a continuous function G on a neighborhood of $t = 0$.

Order of vanishing. Let F_t and $\operatorname{Res}(F_t)$ be defined as in page 1041.

Lemma 3.3. *Let* $q = \operatorname{ord}_{t=0} \operatorname{Res}(F_t)$. *There are constants* $0 < \alpha < 1 < \beta$ *and* $\delta > 0$ *such that*

$$\alpha |t|^q \leq \frac{\| F_t(z_1, z_2) \|}{\| (z_1, z_2) \|^d} \leq \beta,$$

for all $(z_1, z_2) \in \mathbb{C}^2 \setminus \{(0, 0)\}$ *and* $0 < |t| < \delta$.

Proof. The statement of Lemma 3.3 is essentially the content of [Baker and Rumely 2010, Lemma 10.1], letting k be the field of Laurent series in t, equipped with the nonarchimedean valuation measuring the order of vanishing at $t = 0$. But to obtain our estimate with the Euclidean norm, we work directly with their proof.

By the homogeneity of F_t, it suffices to prove the estimate assuming $\|(z_1, z_2)\| = 1$ with either $z_1 = 1$ or $z_2 = 1$. The upper bound is immediate from the presentation of F, with bounded coefficients on compact subsets of $\mathbb{D}$.

Write $F = (F_1(x, y), F_2(x, y))$. The resultant $\operatorname{Res}(F)$ is a nonzero element of the valuation ring $\mathcal{O}_k = \{z \in k : \operatorname{ord}_{t=0} z \geq 0\}$. From basic properties of the resultant (e.g., [Silverman 2007, Proposition 2.13]), there exist polynomials $g_1, g_2, h_1, h_2 \in \mathcal{O}_k[x, y]$ such that

$$g_1(x, y) F_1(x, y) + g_2(x, y) F_2(x, y) = \operatorname{Res}(F) x^{2d-1} \tag{3-1}$$

and

$$h_1(x, y) F_1(x, y) + h_2(x, y) F_2(x, y) = \operatorname{Res}(F) y^{2d-1}. \tag{3-2}$$

Setting $x = 1$, equation (3-1) shows that

$$\min\{\operatorname{ord} F_1(1, z_2), \operatorname{ord} F_2(1, z_2)\} \leq q$$

for any choice of $z_2 \in \mathcal{O}_k$. In fact, taking

$$M = 2 \max\Bigl\{\sup\bigl\{|g_1(1, y)| : |t| \leq \tfrac{1}{2}, |y| \leq 1\bigr\}, \sup\bigl\{|g_2(1, y)| : |t| \leq \tfrac{1}{2}, |y| \leq 1\bigr\}\Bigr\}$$

we may find $\alpha_1 > 0$ and $0 < \delta_1 < \frac{1}{2}$ such that

$$\min\{\inf\{|F_1(1, z_2)| : |z_2| \leq 1\}, \inf\{|F_2(1, y)| : |z_2| \leq 1\}\} \geq |\operatorname{Res}(F)|/M \geq \alpha_1 |t|^q$$

for all $|t| < \delta_1$.

Similarly, setting $y = 1$ in equation (3-2), we may define the analogous α_2 and δ_2 to estimate $F_1(z_1, 1)$ and $F_2(z_1, 1)$ for any $|z_1| \leq 1$; the conclusion follows by setting $\alpha = \min\{\alpha_1, \alpha_2\}$ and $\delta = \min\{\delta_1, \delta_2\}$. □

Proof of Proposition 3.1. It is a standard convergence argument in complex dynamics that the functions $d^{-n} \log \|F_t^n(A_t)\|$ converge locally uniformly in the region where $\mathrm{Res}(F_t) \neq 0$, exactly as in [Hubbard and Papadopol 1994], [Fornæss and Sibony 1994], or [Branner and Hubbard 1988]. To see that the functions G_n converge locally uniformly on $\mathbb{D}^*$, we must look at the growth of the orders $\{a_n\}$ as $n \to \infty$. From the definition of a_n, we have $a_0 = 0$ and

$$a_{n+1} = da_n + \mathrm{ord}_{t=0} F_t(F_n(t)) \tag{3-3}$$

for all n. Hence by Lemma 3.3, noting that $z = F_n(t)$ has norm bounded away from both 0 and ∞ as $t \to 0$, we find that

$$0 \leq k_{n+1} := a_{n+1} - d \cdot a_n \leq q.$$

Consequently, the sequence $a_n/d^n = \sum_{i=1}^{n} k_i/d^i$ has a finite limit. In particular, we may conclude that the sequence

$$G_n(t) = \frac{1}{d^n} \log \|F_n(t)\| = \frac{1}{d^n} \log \|F_t^n(A_t)\| - \frac{a_n}{d^n} \log |t|$$

converges locally uniformly (to $G(t)$) in the punctured unit disk.

To show that $G(t) = o(\log |t|)$ it suffices to show that, for any $\varepsilon > 0$, there is a constant C and a $\delta > 0$ such that

$$|G(t)| \leq \varepsilon |\log |t|| + C$$

for all t in the disk of radius δ.

Fix a positive integer N, and define

$$b_n := \mathrm{ord}_{t=0} F_t^{n-N}(F_N(t))$$

for $n \geq N$, so that $b_N = 0$ and

$$0 \leq \ell_{n+1} := b_{n+1} - d \cdot b_n \leq q$$

by Lemma 3.3. In particular, we have

$$\frac{b_n}{d^n} = \sum_{i=N+1}^{n} \frac{\ell_i}{d^i} \leq \sum_{i=N+1}^{\infty} \frac{q}{d^i}$$

for all $n > N$. By increasing N if necessary, we can assume that

$$\sum_{i=N+1}^{\infty} \frac{q}{d^i} < \varepsilon.$$

Therefore (recalling the constants $0 < \alpha < 1$ and $\delta > 0$ from Lemma 3.3),

$$\begin{aligned}
&\frac{1}{d^n}\log\|F_n(t)\| + \frac{b_n}{d^n}\log|t| \\
&\quad = \frac{1}{d^n}\log\|F_t^{n-N}(F_N(t))\| \\
&\quad = \sum_{i=1}^{n-N}\left(\frac{1}{d^{i+N}}\log\|F_t^i(F_N(t))\| - \frac{1}{d^{i+N-1}}\log\|F_t^{i-1}(F_N(t))\|\right) + \frac{1}{d^N}\log\|F_N(t))\| \\
&\quad = \sum_{i=1}^{n-N}\frac{1}{d^{i+N}}\log\frac{\|F_t^i(F_N(t))\|}{\|F_t^{i-1}(F_N(t))\|^d} + \frac{1}{d^N}\log\|F_N(t))\| \\
&\quad \geq \sum_{i=1}^{n-N}\frac{1}{d^{i+N}}(\log|t|^q + \log\alpha) + \frac{1}{d^N}\log\|F_N(t))\| \\
&\quad \geq \varepsilon\log|t| + \frac{1}{d^N}\log\|F_N(t))\| + \sum_{i=N}^{\infty}\frac{1}{d^i}\log\alpha.
\end{aligned}$$

Let $C = \sup\left|\log\|F_N(t))\|/d^N\right|$ for t in the disk of radius $\frac{1}{2}$. Then

$$\frac{1}{d^n}\log\|F_n(t)\| \geq \varepsilon\log|t| - C + \sum_{i=N}^{\infty}\frac{1}{d^i}\log\alpha - \frac{b_n}{d^n}\log|t| \geq \varepsilon\log|t| - C + \sum_{i=N}^{\infty}\frac{1}{d^i}\log\alpha,$$

for all $|t| < \delta$ and all $n \geq N$.

For the reverse estimate, we have

$$\begin{aligned}
\frac{1}{d^n}\log\|F_n(t)\| + \frac{b_n}{d^n}\log|t| &= \sum_{i=1}^{n-N}\frac{1}{d^{i+N}}\log\frac{\|F_t^i(F_N(t))\|}{\|F_t^{i-1}(F_N(t))\|^d} + \frac{1}{d^N}\log\|F_N(t))\| \\
&\leq \sum_{i=N}^{\infty}\frac{1}{d^i}\log\beta + \frac{1}{d^N}\log\|F_N(t))\|,
\end{aligned}$$

where $\beta > 1$ is the constant from Lemma 3.3. With the same C as above, we conclude that

$$\frac{1}{d^n}\log\|F_n(t)\| \leq -\frac{b_n}{d^n}\log|t| + \sum_{i=N}^{\infty}\frac{1}{d^i}\log\beta + C \leq -\varepsilon\log|t| + \sum_{i=N}^{\infty}\frac{1}{d^i}\log\beta + C$$

for all $|t| < \delta$ and all $n \geq N$. Passing to the limit as $n \to \infty$, we conclude that $G(t) = o(\log|t|)$ for t near 0. This concludes the proof of Proposition 3.1.

Stability. We now gather some consequences of Proposition 3.1 that will be used in the proof of Theorem 1.1. Let F_t be given as on page 1041, so it induces a family of rational maps f of degree d, parameterized by the punctured disk $\mathbb{D}^*$. Let $a : \mathbb{D} \to \mathbb{P}^1$ be a holomorphic map with a holomorphic lift $A : \mathbb{D} \to \mathbb{C}^2 \setminus \{(0, 0)\}$.

We define the functions F_n and G_n as on page 1041, with

$$G(t) = \lim_{n\to\infty} \frac{1}{d^n} \log \|F_n(t)\|.$$

Recall that the pair (f, a) is stable on $\mathbb{D}^*$ if the sequence of holomorphic functions $\{g_n(t) := f_t^n(a(t))\}$ forms a normal family on $\mathbb{D}^*$.

Corollary 3.4. *Suppose the pair (f, a) is stable on the punctured disk $\mathbb{D}^*$. Then there exists a choice of holomorphic lift $A : \mathbb{D} \to \mathbb{C}^2 \setminus \{(0, 0)\}$ of a such that $G \equiv 0$.*

Proof. Stability of (f, a) on $\mathbb{D}^*$ implies that G is harmonic where $t \neq 0$ for any choice of holomorphic lift A of a [DeMarco 2003, Theorem 9.1]. Indeed, take a subsequence of $\{g_n\}$ that converges uniformly on a small neighborhood U in $\mathbb{D}^*$ to a holomorphic map h into $\mathbb{P}^1$. Shrinking U if necessary, we may select the norm on $\mathbb{C}^2$ so that $\log \|s(\cdot)\|$ is harmonic on a region containing the image $h(U)$ in $\mathbb{P}^1$, where s is any holomorphic section of $\mathbb{C}^2 \setminus \{(0, 0)\} \to \mathbb{P}^1$. Then the corresponding subsequence of the harmonic functions G_n are converging uniformly to a harmonic limit.

Fix a choice of $A : \mathbb{D} \to \mathbb{C}^2 \setminus \{(0, 0)\}$, and construct the escape-rate function G. The bound on G from Proposition 3.1 implies that G extends to a harmonic function on the entire disk. Indeed, by a standard argument in complex analysis, we fix a small disk of radius r and let h be the unique harmonic function on this disk with $h = G$ on the boundary circle. For each $\varepsilon > 0$, consider

$$u_\varepsilon(t) = G(t) - h(t) + \varepsilon \log |t|$$

for t in the punctured disk. The function u_ε extends to an upper-semicontinuous function, setting $u_\varepsilon(0) = -\infty$, and so u_ε is subharmonic on the disk because it satisfies the sub-mean-value property. Thus, $u_\varepsilon \le \varepsilon \log r$ on the disk by the maximum principle. Letting $\varepsilon \to 0$, we deduce that $G \le h$ on the punctured disk. Applying the same reasoning to

$$v_\varepsilon(t) = h(t) - G(t) + \varepsilon \log |t|$$

we obtain the reverse inequality, that $h \le G$, and therefore, $h = G$.

The harmonic function G can now be expressed locally as $\operatorname{Re} \eta$ for a holomorphic function η on $\mathbb{D}$. Now replace A_t with $\tilde{A}_t = e^{-\eta(t)} A_t$. Then

$$F_t^n(\tilde{A}_t) = e^{-d^n \eta(t)} F_t^n(A_t),$$

so the order of vanishing at $t = 0$ is unchanged. We obtain a new escape-rate function

$$\begin{aligned}\tilde{G}(t) &= \lim_{n\to\infty} \frac{1}{d^n} \log \|t^{-a_n} F_t^n(\tilde{A}_t)\| = \lim_{n\to\infty} \frac{1}{d^n} \log \|t^{-a_n} e^{-d^n \eta(t)} F_t^n(A_t)\| \\ &= \lim_{n\to\infty} \frac{1}{d^n} \log \|e^{-d^n \eta(t)} F_n(t)\| = G(t) + \log |e^{-\eta}| = G(t) - G(t) \equiv 0,\end{aligned}$$

completing the proof of the corollary. □

Lemma 3.5. *Suppose that $G \equiv 0$. The functions $\{F_n(t) = t^{-a_n} F_t^n(A_t)\}$ are uniformly bounded in $\mathbb{C}^2 \setminus \{(0,0)\}$ on compact subsets of $\mathbb{D}^*$.*

Proof. Recall that the "usual" escape rate of F_t is defined by

$$G_{F_t}(z) = \lim_{n\to\infty} \frac{1}{d^n} \log \|F_t^n(z)\|$$

for z in $\mathbb{C}^2$. The local uniform convergence of the limit (on $\mathbb{D}^* \times \mathbb{C}^2 \setminus \{(0,0)\}$) implies that $G_{F_t}(z)$ is continuous as a function of (t,z); it is proper in $z \in \mathbb{C}^2 \setminus \{(0,0)\}$, since the function satisfies

$$G_{F_t}(\alpha z) = G_{F_t}(z) + \log|\alpha|$$

for all $\alpha \in \mathbb{C}^*$ and all (t,z). So our desired result follows if we can show that, for each compact subset K of $\mathbb{D}^*$, there exist constants $-\infty < c < C < \infty$ such that

$$c \le G_{F_t}(F_n(t)) \le C$$

for all n and all $t \in K$.

Indeed, note that $G(t) = 0$ implies that $G_{F_t}(A_t) = \eta \log|t|$ for

$$\eta = \lim \frac{a_n}{d^n} = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{k_i}{d^i}$$

as in the proof of Proposition 3.1, with $0 \le k_i \le q$ for all i. Therefore,

$$\begin{aligned} G_{F_t}(F_n(t)) &= d^n G_{F_t}(A_t) - a_n \log|t| = (d^n\eta - a_n)\log|t| \\ &= \left(\sum_{i=n+1}^{\infty} \frac{k_i}{d^{i-n}}\right) \log|t| \ge \left(\sum_{i=0}^{\infty} \frac{q}{d^i}\right) \log|t|. \end{aligned}$$

On the other hand, the sequence a_n/d^n increases to η, so $(d^n\eta - a_n)\log|t| \le 0$ for all n and all $t \in \mathbb{D}^*$; therefore, $G_{F_t}(F_n(t)) \le 0$ for all t and all n. □

4. Proof of Theorem 1.1

Let f be an algebraic family of rational maps of degree $d \ge 2$, parameterized by the irreducible quasiprojective complex variety V. Let $a : V \to \mathbb{P}^1$ be a marked point. In this section, we prove Theorem 1.1.

It suffices to prove the theorem when V is one-dimensional. For, if t_0 is any parameter in V at which $a(t_0)$ is not preperiodic, taking any one-dimensional slice through t_0 on which (f,a) is not isotrivial, we conclude that (f,a) will not be stable on this slice. Therefore, the family of iterates cannot be normal on all of V.

If f is isotrivial, then the result follows immediately from Proposition 2.3. If f is not isotrivial and if the degrees of $g_n(t) := f_t^n(a(t))$ are bounded, then the conclusion follows immediately from Proposition 1.8.

For the rest of the proof, assume that the degrees of $\{g_n\}$ are unbounded. Suppose also that (f, a) is stable and f is not isotrivial. We will derive a contradiction.

Let C denote the normalization of a compactification of V, so that we may view f as a family defined over the punctured Riemann surface $C \setminus \{x_1, \ldots, x_n\}$. The stability of (f, a) implies that $\{g_n\}$ is normal on V. As such, there exists a subsequence $\{g_{n_k}\}$ with unbounded degree that converges locally uniformly on V to a holomorphic function $h : V \to \mathbb{P}^1$. Note that h might have finite degree or it may have essential singularities at the punctures x_i of V. In either case, we find:

Lemma 4.1. *There exists a puncture of V such that, for any neighborhood U of this puncture, and for any point $b \in \mathbb{P}^1$ (with at most one exception), the cardinality of $g_{n_k}^{-1}(b) \cap U$ (counted with multiplicities) is unbounded as $n_k \to \infty$.*

Proof. We apply the argument principle. Fix any $b \in \mathbb{P}^1$ such that $h \not\equiv b$. Choose coordinates on $\mathbb{P}^1$ such that $b = 0$ and $h \not\equiv \infty$. Choose a small loop γ_j around each puncture x_j of V on which h has no zeros or poles.

Consider the integral

$$N(\gamma_j) = \frac{1}{2\pi i} \int_{\gamma_j} \frac{h'}{h} \in \mathbb{Z}$$

computing the winding number of the loop $h \circ \gamma_j$ around the origin. By uniform convergence of $g_{n_k} \to h$ on γ_j, and since g_{n_k} is meromorphic, the number $N(\gamma_j)$ is equal to the difference between the number of zeros and number of poles of g_{n_k} inside the circle for all n_k sufficiently large. But since $\deg g_{n_k} \to \infty$ and the functions converge uniformly to h outside these small loops, the actual count of zeros and poles must be growing to infinity inside one of these circles. □

Fix a puncture x_i of V satisfying the condition of Lemma 4.1. Choose local coordinate t on C on a small disk D around the puncture x_i of V. Choose coordinates on $\mathbb{P}^1$ such that the conclusion of Lemma 4.1 holds for $b = 0$ and $b = \infty$. Let B be a small annulus in the disk D of the form

$$B = \{r_0 < |t| < r_1\}$$

with $0 < 2r_0 < r_1 - r_0 < 1$. Passing to a further subsequence if necessary, let

$$\kappa(n_k) = \min\left\{|g_{n_k}^{-1}(0) \cap D_{r_0}|, |g_{n_k}^{-1}(\infty) \cap D_{r_0}|\right\}$$

so that $\kappa(n_k) \to \infty$ with n_k.

As in Section 3, choose a homogeneous polynomial lift F_t of f_t to $\mathbb{C}^2$, normalized so that the coefficients of F_t are holomorphic in t and not all 0 at $t = 0$. By Proposition 3.1 and Corollary 3.4, we may choose a holomorphic lift A of a with values in $\mathbb{C}^2 \setminus \{(0, 0)\}$ such that the escape-rate function G satisfies $G(t) \equiv 0$. From Lemma 3.5, we deduce that the sequence $\{F_n(t)\}$ is uniformly bounded — away from $(0,0)$ and ∞ — on the closed annulus $\bar{B} = \{r_0 \le |t| \le r_1\}$. In other words, there exist constants $0 < c \le C < \infty$ such that

$$c \le \|F_n(t)\| \le C$$

for all n and all $t \in B$.

Write

$$F_{n_k}(t) = (P_{n_k}(t) R_{n_k}(t),\ Q_{n_k}(t) S_{n_k}(t))$$

for holomorphic P_{n_k}, R_{n_k}, Q_{n_k}, S_{n_k} where P_{n_k}, $Q_{n_k} \approx t^{\kappa(n_k)}$ on the disk D. More precisely, there exist factors

$$P_{n_k}(t) = \prod_{i=1}^{\kappa(n_k)} (t - t_i) \quad \text{and} \quad Q_{n_k}(t) = \prod_{i=1}^{\kappa(n_k)} (t - s_i)$$

for two disjoint sets of roots $\{t_i\}$ and $\{s_j\}$ contained in the small disk D_{r_0}. Note that

$$|P_{n_k}(t)|, |Q_{n_k}(t)| \le (2r_0)^{\kappa(n_k)}$$

for $|t| = r_0$, and

$$|P_{n_k}(t)|, |Q_{n_k}(t)| \ge (r_1 - r_0)^{\kappa(n_k)}$$

for $|t| = r_1$. The uniform bounds on $F_n(t)$ imply that

$$|R_{n_k}(t)|, |S_{n_k}(t)| \le \frac{C}{(r_1 - r_0)^{\kappa(n_k)}}$$

on the circle $|t| = r_1$ and

$$\max\{|R_{n_k}(t)|, |S_{n_k}(t)|\} \ge \frac{c}{(2r_0)^{\kappa(n_k)}}$$

for each t on the circle $|t| = r_0$ and all n_k. But $\kappa(n_k) \to \infty$ with n_k and $2r_0 < r_1 - r_0$, so for large n_k these estimates will violate the maximum principle applied to the holomorphic function P'_{n_k} or Q'_{n_k}.

The contradiction obtained shows that if (f, a) is stable on V with f not isotrivial, then the degrees of $\{g_n\}$ must be bounded, returning us to the setting treated by Proposition 1.8. This completes the proof of Theorem 1.1.

5. Density of intersections

In this section, we prove Proposition 1.5 and Theorem 1.6.

Elliptic curves. We begin by explaining the connection between Proposition 1.5 and the theme of this article. Let E_t be a family of smooth elliptic curves, parameterized by a quasiprojective algebraic curve V. The equivalence relation $x \sim -x$ on E_t induces a projection to $\mathbb{P}^1$. Via this projection, the multiplication-by-2 map on E_t descends to a rational function f_t on $\mathbb{P}^1$ of degree 4, called a *Lattès map*. (A formula for the resulting f_t is shown for the Legendre family E_t at the bottom of page 1040, defined there as L_t.) The family f_t is nonisotrivial if and only if the family E_t is nonisotrivial. A point $P_t \in E_t$ projects to a preperiodic point for f_t if and only if P_t is torsion.

In Theorem 1.2, we also refer to the canonical height function: by its definition, the Néron–Tate height on an elliptic curve is equal to $\frac{1}{2}$ times the canonical height for the associated multiplication-by-2 Lattès map. Therefore, height 0 on the elliptic curve coincides with height 0 for the rational function.

Proof of Proposition 1.5. Let E be a nonisotrivial elliptic curve defined over a function field $k = \mathbb{C}(X)$ for an irreducible complex algebraic curve X. We view E as a family E_t of smooth elliptic curves, for all but finitely many $t \in X$; alternatively, we view E as a complex surface, equipped with an elliptic fibration $E \to X$. Fix $P \in E(k)$. Then P determines a section $P : X \to E$. Composing this section P with the degree-two quotient from each E_t, $t \in X$, to $\mathbb{P}^1$, we obtain a marked point $a_P : X \to \mathbb{P}^1$ and a nonisotrivial algebraic family of Lattès maps $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ on a Zariski-open subset $V \subset X$.

From Theorem 1.1, we know that the pair (f, a_P) is stable if and only if the pair (f, a_P) is preperiodic. So either $a_P(t)$ is preperiodic for f_t for all $t \in V$ (in which case P is torsion on E/k), or the pair (f, a_P) is not stable.

In this way, the proposition is a consequence of the following statement, which is a direct application of Montel's theorem on normal families; see, e.g., [Milnor 2006] for background on Montel's theorem.

Given a pair (f, a), the *stable set* $\Omega(f, a) \subset V$ is the largest open set on which $\{t \mapsto f_t^n(a(t))\}$ forms a normal family. The *bifurcation locus* $B(f, a) \subset V$ is the complement of $\Omega(f, a)$ in V.

Proposition 5.1. *Suppose $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ is an algebraic family of rational maps of degree $d \geq 2$. Let $a : V \to \mathbb{P}^1$ be a marked point, and suppose that the bifurcation locus $B(f, a)$ is nonempty. Then for each open $U \subset V$ intersecting $B(f, a)$, there are infinitely many $t \in U$ where $a(t)$ is preperiodic to a repelling cycle of f_t.*

Proof. Fix an open set U having nonempty intersection with $B(f, a)$. Choose a point t_0 in $B(f, a) \cap U$. Choose three distinct repelling periodic points $z_1(t_0)$, $z_2(t_0)$, $z_3(t_0)$ of f_{t_0} that are not in the forward orbit of $a(t_0)$. Shrinking U if necessary, the implicit function theorem implies that these periodic points can be holomorphically parameterized by $t \in U$. By Montel's theorem, the failure of normality of $\{t \mapsto f_t^n(a(t))\}$ on U implies that there exists a parameter $t_1 \in U$ and an integer $n_1 > 0$ such that $f_{t_1}^{n_1}(a(t_1))$ is an element of the set $\{z_1(t_1), z_2(t_1), z_3(t_1)\}$. In particular, $a(t_1)$ is preperiodic for f_{t_1}. Shrinking the neighborhood U, we may find infinitely many such parameters. $\square$

Remark 5.2. In [DeMarco et al. 2016], we studied the *distribution* of the parameters $t \in X$ for which a marked point $P_t \in E_t$ is torsion, with E_t the Legendre family of elliptic curves. The set of such parameters is dense in the parameter space (in the usual analytic topology).

Proof of Theorem 1.6. By hypothesis, the family $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ has dimension N in moduli; this means that the image of the induced projection $V \to \mathrm{M}_d$ to the moduli space of rational maps has dimension N.

To prove Zariski density, we need to show that, for any algebraic subvariety $Y \subset V$ (possibly reducible), the complement $\Lambda = V \setminus Y$ contains a parameter t at which all points $a_1(t), \ldots, a_k(t)$ are preperiodic. Note that Λ is itself an irreducible, quasiprojective complex algebraic variety, so that $f : \Lambda \times \mathbb{P}^1 \to \mathbb{P}^1$ is again an algebraic family of rational maps, of dimension N in moduli.

Consider the marked point a_1. Since f projects to an N-dimensional family in the moduli space, with $N > 0$, it follows that (f, a_1) is not isotrivial on Λ. By Theorem 1.1, the pair (f, a_1) is either preperiodic or it fails to be stable on Λ. If (f, a_1) is preperiodic, we set $\Lambda_1 = \Lambda$. If (f, a_1) is unstable, then Proposition 5.1 shows that there exists a parameter $t_1 \in \Lambda$ where $a_1(t_1)$ is preperiodic to a repelling cycle of f_{t_1}. If a_1 satisfies the equation $f^{n_1}(a_1) = f^{m_1}(a_1)$ at the parameter t_1, we define $\Lambda_1 \subset \Lambda$ to be an irreducible component of the subvariety defined by the equation $f_t^{n_1}(a_1(t)) = f_t^{m_1}(a_1(t))$ that contains t_1. Then Λ_1 is a nonempty quasiprojective variety, of codimension 1 in Λ. Furthermore, since the cycle persists under perturbation, the condition defining Λ_1 will also cut out a codimension-1 subvariety in the moduli space. In other words, Λ_1 must project to a family of dimension $N - 1$ in the moduli space. By construction, (f, a_1) is preperiodic on Λ_1.

We continue inductively. Fix $1 \leq i < k$. Suppose Λ_i is a quasiprojective subvariety of dimension $\geq N - i$ in moduli on which $(f, a_1), \ldots, (f, a_i)$ are preperiodic. Since $N - i > N - k \geq 0$, the pair (f, a_{i+1}) is not isotrivial on Λ_i. As above, we combine Theorem 1.1 with Proposition 5.1 to find a parameter $t_{i+1} \in \Lambda_i$ where a_{i+1} is preperiodic. We define $\Lambda_{i+1} \subset \Lambda_i$ so that (f, a_{i+1}) is preperiodic on Λ_{i+1}, and the family $f : \Lambda_{i+1} \times \mathbb{P}^1 \to \mathbb{P}^1$ has dimension at least $N - i - 1$ in moduli. In conclusion, all of the points $(f, a_1), \ldots, (f, a_k)$ are preperiodic on Λ_k, and Λ_k has dimension at least $N - k \geq 0$ in moduli. In particular, Λ_k is nonempty, and the theorem is proved.

6. A conjecture on intersections and dynamical relations

We conclude this article with a revised statement of the conjecture from [Baker and DeMarco 2013] on "unlikely intersections" and density of "special points" — and we provide the proof of one implication, as an application of Theorem 1.1. Specifically, we look at algebraic families $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ of dimension $N > 0$

in moduli. We prove that if an $(N+1)$-tuple of marked points is dynamically related, then the set of parameters $t \in V$ where they are simultaneously preperiodic is Zariski-dense in V. We conclude the article with an explanation of how this implies one implication of [Baker and DeMarco 2013, Conjecture 1.10].

Density of special points. In [Baker and DeMarco 2013, Conjecture 1.10], we formulated a conjecture about the arrangement of postcritically finite maps ("special points") in the moduli space of rational maps of degree $d \geq 2$. It was presented as a dynamical analog of the André–Oort conjecture in arithmetic geometry, with the aim of characterizing the "special subvarieties" of the moduli space, meaning the algebraic families $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ with a Zariski-dense subset of postcritically finite maps. Roughly speaking, the special subvarieties should be those that are defined by (a general notion of) critical orbit relations. We proved special cases of the conjecture, for certain families of polynomial maps, and we sketched the proof of one implication in the general case.

If we formulate the conjecture to handle arbitrary marked points, not only critical points, then the statement encompasses recent results about elliptic curves, as in the work of Masser and Zannier (and therefore has overlap with the Pink and Zilber conjectures); see [Masser and Zannier 2012] and the references therein. Evidence towards the more general result is given by [Baker and DeMarco 2013, Theorem 1.3] and the results of [Ghioca et al. 2013; 2015]. Conjecture 6.1 presented here is, therefore, more than just an analogy with statements in arithmetic geometry.

Let V be an irreducible, quasiprojective complex algebraic variety, and let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of rational maps of degree $d \geq 2$. For a collection of n marked points $a_1, \ldots, a_n : V \to \mathbb{P}^1$, we define

$$S(a_1, \ldots, a_n) = \bigcap_{i=1}^{n} \{t \in V : a_i(t) \text{ is preperiodic for } f_t\}.$$

We say the marked points $a_1, \ldots, a_n$ are *coincident* along V if there exists a marked point a_i and a Zariski-open subset $V' \subset V$ such that

$$S(a_1, \ldots, a_n) \cap V' = S(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \cap V'.$$

In other words, if $\{a_1(t), \ldots, a_{i-1}(t), a_{i+1}(t), \ldots, a_n(t)\}$ are all preperiodic for f_t at a parameter $t \in V'$, then the remaining point $a_i(t)$ must also be preperiodic for f_t. For example, if a pair (f, a) is preperiodic on V, then any collection of points $\{a_1, \ldots, a_n\}$ containing a will be coincident.

A stronger notion than coincidence is that of the dynamical relation, requiring an f-invariant algebraic relation between the points $\{a_1, \ldots, a_n\}$. A formal definition is given below.

Conjecture 6.1. *Let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be an algebraic family of rational maps of degree $d \geq 2$, of dimension $N > 0$ in moduli. Let $a_0, \dots, a_N$ be any collection of $N+1$ marked points. The following are equivalent*:

(1) *The set $S(a_0, \dots, a_N)$ is Zariski-dense in V.*

(2) *The points $a_0, \dots, a_N$ are coincident along V.*

(3) *The points $a_0, \dots, a_N$ are dynamically related along V.*

Theorem 6.2. *We have* (3) $\implies$ (2) *and* (2) $\implies$ (1) *in Conjecture 6.1.*

The implication (3) $\implies$ (2) will be a formal consequence of the definitions, while (2) $\implies$ (1) is presented below as an application of Theorem 1.1. The remaining challenge is to show that (1) implies (3). We expect that (1) $\implies$ (2) should be a consequence of "arithmetic equidistribution" as in the proofs of [Baker and DeMarco 2011; 2013; Ghioca et al. 2013; 2015; DeMarco et al. 2016] when V is a curve.

Dynamical relations. The basic example of a dynamical relation between two marked points $a, b : V \to \mathbb{P}^1$ is an orbit relation: the existence of integers n, m such that

$$f_t^n(a(t)) = f_t^m(b(t))$$

for all $t \in V$. To allow for complicated symmetries, we will say that N marked points $a_1, \dots, a_N$ are *dynamically related* along V if there exists a (possibly reducible) algebraic subvariety

$$X \subset (\mathbb{P}^1)^N$$

defined over the function field $k = \mathbb{C}(V)$, such that three conditions are satisfied:

(R1) $(a_1, \dots, a_N) \in X$.

(R2) (invariance) $F(X) \subset X$, where $F = (f, f, \dots, f) : (\mathbb{P}^1)^N \to (\mathbb{P}^1)^N$.

(R3) (nondegeneracy) There exists an $i \in \{1, \dots, N\}$ and a Zariski-open subset $V' \subset V$ such that the projection from the specialization X_t to the i-th coordinate hyperplane in $(\mathbb{P}^1_{\mathbb{C}})^N$ is a finite map for all $t \in V'$.

Remark 6.3. In [Baker and DeMarco 2013], after stating Conjecture 1.10, we offhandedly remarked that one implication of the conjecture "follows easily from an argument mimicking the proof of Proposition 2.6 and the following observation". The proof of Theorem 1.6 in this article is the argument we had in mind, mimicking [Baker and DeMarco 2013, Proposition 2.6], but the stated "observation" was not formulated correctly. The inclusion of condition (R3) and the argument in the proof of Theorem 6.2 on the next page are an attempt to correct that error.

To illustrate the dynamical relation, observe that any N points $a_1, \dots, a_N$ are dynamically related if one of the pairs (f, a_i) is preperiodic. Indeed, if a_i satisfies

$f_t^n(a_i(t)) = f_t^m(a_i(t))$ for all $t \in V$, for some pair of integers $n \neq m \geq 0$, then we could take X to be the hypersurface

$$\{x \in (\mathbb{P}^1)^N : f^n(x_i) = f^m(x_i)\}$$

defined over the field $k = \mathbb{C}(V)$. As a nontrivial example, we look at the relation arising in the Masser–Zannier theorems [2012]. Let $f_{[k]}$ be the Lattès map induced from multiplication by $k \in \mathbb{N}$ on a nonisotrivial elliptic curve E over $k = \mathbb{C}(V)$. Let a_p, a_q be the projections to $\mathbb{P}^1$ of two points p and q in $E(k)$. The linear relation $n \cdot p = m \cdot q$ between points p and q on E, for integers n and m, translates into a dynamical relation in $(\mathbb{P}^1)^2$ defined by

$$f_{[n]}(x_1) = f_{[m]}(x_2).$$

This relation satisfies condition (R2) for $F = (f_{[k]}, f_{[k]})$ because all Lattès maps descended from the same elliptic curve must commute.

Since the writing of [Baker and DeMarco 2013], we have learned about the results in [Medvedev 2007] which significantly simplify the form of possible dynamical relations. In particular, Medvedev has shown that the varieties X satisfying condition (R2) should depend nontrivially on only two input variables. In other words, the rational function $f : \mathbb{P}^1 \to \mathbb{P}^1$ will be *disintegrated* in the sense of [Medvedev and Scanlon 2014, Definition 2.20]; see the first theorem in the introduction of the same paper, treating the case where f is a polynomial. An affirmative answer to the following question would provide a further refinement — and simplification — to the notion of dynamical relation, extending the results of [Medvedev and Scanlon 2014] beyond the polynomial setting. (The work of Medvedev and Scanlon relied on Ritt's decomposition theory [1922] for polynomials; the analogous decomposition theory for rational functions is not completely understood.)

Question 6.4. Assume that f is not isotrivial, and suppose that points $a_1, \dots, a_N$ are dynamically related. Does there always exist a pair of indices i, j (allowing possibly $i = j$) such that the point $(a_1, \dots, a_N)$ satisfies a relation of the form

$$A(x_i) = B(x_j), \tag{6-1}$$

where $A, B \in k(z)$ are nonconstant rational functions that commute with an iterate of f?

Hypersurfaces in $(\mathbb{P}^1)^N$ defined by relations of the form (6-1) satisfy condition (R2) in the definition of the dynamical relation because f commutes with A and B; they satisfy condition (R3) taking either coordinate i or j.

Proof of Theorem 6.2. We begin by proving (3) $\Longrightarrow$ (2); namely, that a dynamical relation among the points $a_0, \dots, a_N$ implies that the points are coincident. This follows from the definition of dynamical relation, and it does not depend on the number of points.

Lemma 6.5. *Let $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ be any algebraic family of rational maps. Suppose marked points $a_0, a_1, \ldots, a_n$ are dynamically related along V. Then the points $a_0, \ldots, a_n$ are coincident along V.*

Proof. Let X denote the $(f, \ldots, f)$-invariant subvariety in $(\mathbb{P}^1)^{n+1}$ for the point $(a_1, \ldots, a_n)$, given in the definition of the dynamical relation. Suppose the points are labeled so that x_0 is the coordinate satisfying condition (R3). Then the projection from X_t to $(\mathbb{P}^1_{\mathbb{C}})^n$, forgetting the 0-th coordinate, is finite, for all t in the Zariski-open subset $V' \subset V$.

Now let t_0 be any parameter in V' at which $a_1(t_0), \ldots, a_n(t_0)$ are preperiodic for f_{t_0}. The point $a_0(t_0)$ must lie in the fiber of $X_{t_0} \to (\mathbb{P}^1)^N$ over $(a_1(t_0), \ldots, a_n(t_0))$. Invariance of X implies the invariance of X_{t_0}, so that $f^m_{t_0}(a_0(t_0))$ lies in the fiber over $(f^m(a_1(t_0)), \ldots, f^m(a_N(t_0)))$ for all $m \geq 1$. The preperiodicity of the points guarantees that there are only finitely many points in the base in the orbit of $(a_1(t_0), \ldots, a_n(t_0))$, so the orbit of $a_0(t_0)$ must be contained in a finite set. In other words, $a_0(t_0)$ is preperiodic. This completes the proof. □

Now assume (2), that the given points $a_0, \ldots, a_N$ are coincident. Assume the points are labeled so that a_0 is the dependent point, in the sense that

$$S(a_0, \ldots, a_N) \cap V' = S(a_1, \ldots, a_N) \cap V'$$

for some Zariski-open subset $V' \subset V$. Since V has dimension N in moduli, Theorem 1.6 tells us that the set $S(a_1, \ldots, a_N)$ is Zariski-dense in V. Therefore, so is $S(a_0, \ldots, a_N)$, and the implication (2) $\Longrightarrow$ (1) is proved.

***Proof of one implication of* [Baker and DeMarco 2013, Conjecture 1.10].** Suppose that $f : V \times \mathbb{P}^1 \to \mathbb{P}^1$ is an algebraic family of rational maps of degree $d \geq 2$ and dimension $N > 0$ in moduli. We assume that all $2d - 2$ critical points of f are marked. Conjecture 1.10 of [Baker and DeMarco 2013] states: *the map f_t is postcritically finite for a Zariski-dense set of $t \in V$ if and only if there are at most N dynamically independent critical points.*

Let $c_1, \ldots, c_{2d-2}$ denote the marked critical points. Assume that f has at most N dynamically independent critical points; in other words, given any $n > N$ marked critical points $c_{i_1}, \ldots, c_{i_n}$, there is a dynamical relation among them.

Note that $S(c_1, \ldots, c_{2d-2})$ is precisely the set of parameters t for which f_t is postcritically finite. Applying Lemma 6.5 repeatedly, and reordering the points as needed, there exists a Zariski-open subset $V' \subset V$ such that

$$S(c_1, \ldots, c_{2d-2}) \cap V' = S(c_1, \ldots, c_{2d-3}) \cap V' = \cdots = S(c_1, \ldots, c_N) \cap V'.$$

From Theorem 1.6, we know that $S(c_1, \ldots, c_N)$ is Zariski-dense in V. This proves that the postcritically finite maps form a Zariski-dense subset of V.

Acknowledgements

I would like to thank Matt Baker, Xander Faber, Dragos Ghioca, Curt McMullen, Joe Silverman, Xiaoguang Wang, and Hexi Ye for helpful comments and discussions. Hexi Ye provided ideas for the details in Section 3. The research was supported by the National Science Foundation, DMS-1517080.

References

[Baker 2009] M. Baker, "A finiteness theorem for canonical heights attached to rational maps over function fields", *J. Reine Angew. Math.* **626** (2009), 205–233. MR 2492995 Zbl 1187.37133

[Baker and DeMarco 2011] M. Baker and L. DeMarco, "Preperiodic points and unlikely intersections", *Duke Math. J.* **159**:1 (2011), 1–29. MR 2817647 Zbl 1242.37062

[Baker and DeMarco 2013] M. Baker and L. DeMarco, "Special curves and postcritically finite polynomials", *Forum Math. Pi* **1** (2013), e3, 35. MR 3141413 Zbl 1320.37022

[Baker and Rumely 2010] M. Baker and R. Rumely, *Potential theory and dynamics on the Berkovich projective line*, Mathematical Surveys and Monographs **159**, American Mathematical Society, Providence, RI, 2010. MR 2599526 Zbl 1196.14002

[Benedetto 2005] R. L. Benedetto, "Heights and preperiodic points of polynomials over function fields", *Int. Math. Res. Not.* **2005**:62 (2005), 3855–3866. MR 2202175 Zbl 1114.14018

[Branner and Hubbard 1988] B. Branner and J. H. Hubbard, "The iteration of cubic polynomials, I: The global topology of parameter space", *Acta Math.* **160**:3-4 (1988), 143–206. MR 945011 Zbl 0668.30008

[Call and Silverman 1993] G. S. Call and J. H. Silverman, "Canonical heights on varieties with morphisms", *Compositio Math.* **89**:2 (1993), 163–205. MR 1255693 Zbl 0826.14015

[DeMarco 2003] L. DeMarco, "Dynamics of rational maps: Lyapunov exponents, bifurcations, and capacity", *Math. Ann.* **326**:1 (2003), 43–73. MR 1981611 Zbl 1032.37029

[DeMarco et al. 2015] L. DeMarco, X. Wang, and H. Ye, "Bifurcation measures and quadratic rational maps", *Proc. Lond. Math. Soc.* (3) **111**:1 (2015), 149–180. MR 3404779 Zbl 06464910

[DeMarco et al. 2016] L. DeMarco, X. Wang, and H. Ye., "Torsion points and the Lattès family", preprint, 2016, available at https://www.press.jhu.edu/journals/american_journal_of_mathematics/future_publications/pre_print_content/AJM-demarco-wang-ye-FINAL.pdf. To appear in *Amer. J. Math.*

[Dujardin and Favre 2008] R. Dujardin and C. Favre, "Distribution of rational maps with a preperiodic critical point", *Amer. J. Math.* **130**:4 (2008), 979–1032. MR 2427006 Zbl 1246.37071

[Fornæss and Sibony 1994] J. E. Fornæss and N. Sibony, "Complex dynamics in higher dimensions", pp. 131–186 in *Complex potential theory* (Montreal, PQ, 1993), edited by P. M. Gauthier and G. Sabidussi, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **439**, Kluwer Acad. Publ., Dordrecht, 1994. MR 1332961 Zbl 0811.32019

[Ghioca et al. 2013] D. Ghioca, L.-C. Hsia, and T. J. Tucker, "Preperiodic points for families of polynomials", *Algebra Number Theory* **7**:3 (2013), 701–732. MR 3095224 Zbl 1323.37056

[Ghioca et al. 2015] D. Ghioca, L.-C. Hsia, and T. J. Tucker, "Preperiodic points for families of rational maps", *Proc. Lond. Math. Soc.* (3) **110**:2 (2015), 395–427. MR 3335283 Zbl 1317.37112

[Hubbard and Papadopol 1994] J. H. Hubbard and P. Papadopol, "Superattractive fixed points in $\mathbf{C}^n$", *Indiana Univ. Math. J.* **43**:1 (1994), 321–365. MR 1275463 Zbl 0858.32023

[Lang and Néron 1959] S. Lang and A. Néron, "Rational points of abelian varieties over function fields", *Amer. J. Math.* **81** (1959), 95–118. MR 0102520 Zbl 0099.16103

[Lyubich 1983] M. Y. Lyubich, "Some typical properties of the dynamics of rational mappings", *Uspekhi Mat. Nauk* **38**:5(233) (1983), 197–198. In Russian; translated in *Russ. Math. Surv.* **38**:5 (1983), 154–155. MR 718838

[Mañé et al. 1983] R. Mañé, P. Sad, and D. Sullivan, "On the dynamics of rational maps", *Ann. Sci. École Norm. Sup.* (4) **16**:2 (1983), 193–217. MR 732343 Zbl 0524.58025

[Masser and Zannier 2010] D. Masser and U. Zannier, "Torsion anomalous points and families of elliptic curves", *Amer. J. Math.* **132**:6 (2010), 1677–1691. MR 2766181 Zbl 1225.11078

[Masser and Zannier 2012] D. Masser and U. Zannier, "Torsion points on families of squares of elliptic curves", *Math. Ann.* **352**:2 (2012), 453–484. MR 2874963 Zbl 1306.11047

[McMullen 1987] C. McMullen, "Families of rational maps and iterative root-finding algorithms", *Ann. of Math.* (2) **125**:3 (1987), 467–493. MR 890160 Zbl 0634.30028

[McMullen 1994] C. T. McMullen, *Complex dynamics and renormalization*, Annals of Mathematics Studies **135**, Princeton University Press, 1994. MR 1312365 Zbl 0822.30002

[Medvedev 2007] A. Medvedev, *Minimal sets in ACFA*, Ph.D. thesis, University of California, Berkeley, 2007, available at http://search.proquest.com/docview/304900205.

[Medvedev and Scanlon 2014] A. Medvedev and T. Scanlon, "Invariant varieties for polynomial dynamical systems", *Ann. of Math.* (2) **179**:1 (2014), 81–177. MR 3126567 Zbl 06260586

[Milnor 2006] J. Milnor, *Dynamics in one complex variable*, 2nd ed., Annals of Mathematics Studies **160**, Princeton University Press, 2006. MR 2193309 Zbl 1085.30002

[Ritt 1922] J. F. Ritt, "Prime and composite polynomials", *Trans. Amer. Math. Soc.* **23**:1 (1922), 51–66. MR 1501189 Zbl 48.0079.01

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, New York, 1994. MR 1312368 Zbl 0911.14015

[Silverman 2007] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics **241**, Springer, New York, 2007. MR 2316407 Zbl 1130.37001

Communicated by Joseph H. Silverman
Received 2015-06-16 Revised 2016-02-10 Accepted 2016-03-10

demarco@math.northwestern.edu *Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208-2730, United States*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.1057

Frobenius and valuation rings

Rankeya Datta and Karen E. Smith

The behavior of the Frobenius map is investigated for valuation rings of prime characteristic. We show that valuation rings are always F-pure. We introduce a generalization of the notion of strong F-regularity, which we call F-pure regularity, and show that a valuation ring is F-pure regular if and only if it is Noetherian. For valuations on function fields, we show that the Frobenius map is finite if and only if the valuation is Abhyankar; in this case the valuation ring is Frobenius split. For Noetherian valuation rings in function fields, we show that the valuation ring is Frobenius split if and only if Frobenius is finite, or equivalently, if and only if the valuation ring is excellent.

1. Introduction

Classes of singularities defined using Frobenius — F-purity, Frobenius splitting, and the various variants of F-regularity — have played a central role in commutative algebra and algebraic geometry over the past forty years. The goal of this paper is a systematic study of these F-singularities in the novel, but increasingly important *non-Noetherian* setting of *valuation rings*.

Datta was partially supported by the Juha Heinonen Memorial Graduate Fellowship at the University of Michigan. The second author was partially supported by NSF grants DMS-1001764 and DMS-1501625.

MSC2010: primary 13A35; secondary 13F30, 14B05.

Keywords: valuation rings, Abhyankar valuations, characteristic p commutative algebra, F-pure, F-regular, Frobenius split.

Let R be a commutative ring of prime characteristic p. The Frobenius map is the ring homomorphism $R \xrightarrow{F} R$ sending each element to its p-th power. While simple enough, the Frobenius map reveals deep structural properties of a Noetherian ring of prime characteristic, and is a powerful tool for proving theorems for rings containing an arbitrary field (or varieties, say, over $\mathbb{C}$) by standard reduction to characteristic p techniques. Theories such as Frobenius splitting [Mehta and Ramanathan 1985] and tight closure [Hochster and Huneke 1990] are well-developed in the Noetherian setting, often under the additional assumption that the Frobenius map is finite. Since classically most motivating problems were inspired by algebraic geometry and representation theory, these assumptions seemed natural and not very restrictive. Now, however, good reasons are emerging to study F-singularities in certain non-Noetherian settings as well.

One such setting is cluster algebras [Fomin and Zelevinsky 2002]. An upper cluster algebra over $\mathbb{F}_p$ need not be Noetherian, but recently it was shown that it is always Frobenius split, and indeed, admits a "cluster canonical" Frobenius splitting [Benito et al. 2015]. Likewise valuation rings are enjoying a resurgence of popularity despite rarely being Noetherian, with renewed interest in non-Archimedean geometry [Conrad 2008], the development of tropical geometry [Gubler et al. 2016], and the valuative tree [Favre and Jonsson 2004], to name just a few examples, as well as fresh uses in higher dimensional birational geometry (e.g., [Cutkosky 2004; Fernández de Bobadilla and Pereira 2012; Boucksom 2014]).

For a Noetherian ring R, the Frobenius map is flat if and only if R is regular, by a famous theorem of Kunz [1969]. As we observe in Theorem 3.1, the Frobenius map is always flat for a valuation ring. So in some sense, a valuation ring of characteristic p might be interpreted as a "non-Noetherian regular ring."

On the other hand, some valuation rings are decidedly more like the local rings of smooth points on varieties than others. For example, for a variety X (over, say, an algebraically closed field of characteristic p), the Frobenius map is always finite. For valuation rings of the function field of X, however, we show that the Frobenius is finite if and only the valuation is Abhyankar; see Theorem 5.1. In particular, for discrete valuations, finiteness of Frobenius is equivalent to the valuation being *divisorial* — that is, given by the order of vanishing along a prime divisor on some birational model. Abhyankar valuations might be considered the geometrically most interesting ones (see [Ein et al. 2003]), so it is fitting that their valuation rings behave the most like the rings of smooth points on a variety. Indeed, recently, the local uniformization problem for Abhyankar valuations was settled in positive characteristic [Knaf and Kuhlmann 2005].

One can weaken the demand that Frobenius is flat and instead require only that the Frobenius map is *pure* (see Section 2.5). Hochster and Roberts observed that this condition, which they dubbed *F-purity*, is often sufficient for controlling singularities of a Noetherian local ring, an observation at the heart of their famous theorem on the Cohen–Macaulayness of invariant rings [Hochster and Roberts 1976; 1974]. We show in Corollary 3.3 that any valuation ring of characteristic p is F-pure. Purity of a map is equivalent to its splitting under suitable finiteness hypotheses, but at least for valuation rings (which rarely satisfy said hypotheses), the purity of Frobenius seems to be better behaved and more straightforward than its splitting. Example 4.5.1 shows that not all valuation rings are Frobenius split, even in the Noetherian case.

Frobenius splitting has well known deep local and global consequences for algebraic varieties. In the local case, Frobenius splitting has been said to be a "characteristic p analog" of log canonical singularities for complex varieties, whereas related properties correspond to other singularities in the minimal model program [Hara and Watanabe 2002; Schwede 2009b; Smith 1997; Takagi 2008]. For projective varieties, Frobenius splitting is related to positivity of the anticanonical bundle; see [Brion and Kumar 2005; Mehta and Ramanathan 1985; Smith 2000; Schwede and Smith 2010]. Although valuation rings are always F-pure, the question of their Frobenius splitting is subtle. Abhyankar valuations in function fields are Frobenius split (Theorem 5.1), but a discrete valuation ring is Frobenius split if and only if it is *excellent* in the sense of Grothendieck (Corollary 4.2.2). Along the way, we prove a simple characterization of the finiteness of Frobenius for a Noetherian domain in terms of excellence, which gives a large class of Noetherian domains in which Frobenius splitting implies excellence; see Section 2.6 for details.

Closely related to F-purity and Frobenius splitting are the various variants of F-regularity. Strong F-regularity was introduced by Hochster and Huneke [1989] as a proxy for weak F-regularity — the property that all ideals are tightly closed — because it is easily shown to pass to localizations. Whether or not a weakly F-regular ring remains so after localization is a long standing open question in tight closure theory, as is the equivalence of weak F-regularity and strong F-regularity. Strong F-regularity has found many applications beyond tight closure, and is closely related to Ramanathan's notion of "Frobenius split along a divisor" [Ramanathan 1991; Smith 2000]. A smattering of applications might include [Aberbach and Leuschke 2003; Benito et al. 2015; Blickle 2008; Brion and Kumar 2005; Gongyo et al. 2015; Hacon and Xu 2015; Patakfalvi 2014; Schwede and Tucker 2012; Schwede 2009a; Schwede and Smith 2010; Smith and Van den Bergh 1997; Smith and Zhang 2015; Smith 2000].

Traditionally, strong F-regularity has been defined only for Noetherian rings in which Frobenius is finite. To clarify the situation for valuation rings, we introduce a new definition which we call *F-pure regularity* (see Definition 6.1.1) requiring *purity* rather than *splitting* of certain maps. We show that F-pure regularity is better suited for arbitrary rings, but equivalent to strong F-regularity under the standard finiteness hypotheses; it also agrees with another generalization of strong F-regularity proposed by Hochster [2007] (using tight closure) in the local Noetherian case. Likewise, we show that F-pure regularity is a natural and straightforward generalization of strong F-regularity, satisfying many expected properties — for example, regular rings are F-pure regular. Returning to valuation rings, in Theorem 6.5.1 we characterize F-pure regular valuation rings as precisely those that are Noetherian.

Finally, in Section 6.6, we compare our generalization of strong F-regularity with the obvious competing generalization, in which the standard definition in terms of splitting certain maps is naively extended without assuming any finiteness conditions. To avoid confusion,[1] we call this *split F-regularity*. We characterize split F-regular valuation rings (at least in a certain large class of fields) as precisely those that are Frobenius split, or equivalently *excellent;* see Corollary 6.6.3. But we also point out that there are regular local rings that fail to be split F-regular, so perhaps split F-regularity is not a reasonable notion of "singularity."

2. Preliminaries

Throughout this paper, all rings are assumed to be commutative, and of prime characteristic p unless explicitly stated otherwise. By a local ring, we mean a ring with a unique maximal ideal, *not necessarily Noetherian.*

2.1. *Valuation rings.* We recall some basic facts and definitions about valuation rings (of arbitrary characteristic), while fixing notation. See [Bourbaki 1989, Chapter VI] or [Matsumura 1989, Chapter 4] for proofs and details.

The symbol Γ denotes an ordered abelian group. Recall that such an abelian group is torsion free. The rational rank of Γ, denoted $\operatorname{rat.rank}\Gamma$, is the dimension of the $\mathbb{Q}$-vector space $\mathbb{Q} \otimes_{\mathbb{Z}} \Gamma$.

Let K be a field. A *valuation on K* is a homomorphism

$$v : K^{\times} \to \Gamma$$

from the group of units of K, satisfying

$$v(x+y) \geq \min\{v(x), v(y)\}$$

for all $x, y \in K^{\times}$. We say that v is *defined over a subfield k of K*, or that v is a *valuation on K/k*, if v takes the value 0 on elements of k.

There is no loss of generality in assuming that v is surjective, in which case we say Γ (or Γ_v) is the *value group* of v. Two valuations v_1 and v_2 on K are said to

[1] An earlier version of this paper used the terms *pure F-regularity* and *split F-regularity* for the two generalizations of classical strong F-regularity, depending upon whether maps were required to be pure or split. The names were changed at Karl Schwede's urging to avoid confusion with terminology for pairs in [Takagi 2004].

be *equivalent* if there is an order preserving isomorphism of their value groups identifying $v_1(x)$ and $v_2(x)$ for all $x \in K^\times$. Throughout this paper, we identify equivalent valuations.

The *valuation ring of* v is the subring $R_v \subseteq K$ consisting of all elements $x \in K^\times$ such that $v(x) \geq 0$ (together with the zero element of K). Two valuations on a field K are equivalent if and only if they determine the same valuation ring. Hence a valuation ring of K is essentially the same thing as an equivalence class of valuations on K.

The valuation ring of v is local, with *maximal ideal* m_v consisting of elements of strictly positive values (and zero). The *residue field* R_v/m_v is denoted $\kappa(v)$. If v is a valuation over k, then both R_v and $\kappa(v)$ are k-algebras.

A valuation ring V of K can be characterized directly, without reference to a valuation, as a subring with the property that for every $x \in K$, either $x \in V$ or $x^{-1} \in V$. The valuation ring V uniquely determines a valuation v on K (up to equivalence), whose valuation ring in turn recovers V. Indeed, it is easy to see that the set of ideals of a valuation ring is totally ordered by inclusion, so the set of principal ideals Γ^+ forms a monoid under multiplication, ordered by $(f) \leq (g)$ whenever f *divides* g. Thus, Γ can be taken to be the ordered abelian group generated by the principal ideals, and the valuation $v : K^\times \to \Gamma$ is induced by the monoid map sending each nonzero $x \in V$ to the ideal generated by x. Clearly, the valuation ring of v is V. See [Matsumura 1989, Chapter 4].

2.2. *Extension of valuations.* Consider an extension of fields $K \subseteq L$. By definition, a valuation w on L is an *extension* of a valuation v on K if the restriction of w to the subfield K is v. Equivalently, w extends v if R_w *dominates* R_v, meaning that $R_v = R_w \cap K$ with $m_w \cap R_v = m_v$. In this case, there is an induced map of residue fields

$$\kappa(v) \hookrightarrow \kappa(w).$$

The *residue degree of* w *over* v, denoted by $f(w/v)$, is the degree of the residue field extension $\kappa(v) \hookrightarrow \kappa(w)$.

If w extends v, there is a natural injection of ordered groups $\Gamma_v \hookrightarrow \Gamma_w$, since Γ_v is the image of w restricted to the subset K. The *ramification index of* w *over* v, denoted by $e(w/v)$, is the index of Γ_v in Γ_w.

If $K \hookrightarrow L$ is a finite extension, then both the ramification index $e(w/v)$ and the residue degree $f(w/v)$ are *finite*. Indeed, if $K \subseteq L$ is a degree n extension, then

$$e(w/v)f(w/v) \leq n. \tag{2.2.0.1}$$

More precisely:

Proposition 2.2.1 [Bourbaki 1989, VI.8]. *Let $K \subseteq L$ be an extension of fields of finite degree n. For a valuation v on K, consider the set $\mathcal{S}$ of all extensions* (*up to* *equivalence*) w *of* v *to* L. *Then*

$$\sum_{w_i \in \mathcal{S}} e(w_i/v) f(w_i/v) \leq n.$$

In particular, the set $\mathcal{S}$ *is finite. Furthermore, equality holds if and only if the integral closure of* R_v *in* L *is a finitely generated* R_v*-module.*

2.3. ***Abhyankar valuations.*** Fix a field K finitely generated over a fixed ground field k, and let v be a valuation on K/k. By definition, the *transcendence degree* of v is the transcendence degree of the field extension

$$k \hookrightarrow \kappa(v).$$

The main result about the transcendence degree of valuations is due to Abhyankar [1956]. See also [Bourbaki 1989, VI.10.3, Corollary 1].

Theorem 2.3.1 (Abhyankar's inequality). *Let* K *be a finitely generated field extension of* k, *and let* v *be a valuation on* K/k. *Then*

$$\text{trans.}\deg v + \text{rat.}\operatorname{rank} \Gamma_v \leq \text{trans.}\deg K/k. \tag{2.3.1.1}$$

Moreover if equality holds, then Γ_v *is a finitely generated abelian group, and* $\kappa(v)$ *is a finitely generated extension of* k.

We say v is an *Abhyankar valuation* if equality holds in Abhyankar's inequality (2.3.1.1). Note that an Abyhankar valuation has a finitely generated value group, and its residue field is finitely generated over the ground field k.

Example 2.3.2. Let K/k be the function field of a normal algebraic variety X of dimension n over a ground field k. For a prime divisor Y of X, consider the local ring $\mathcal{O}_{X,Y}$ of rational functions on X regular at Y. The ring $\mathcal{O}_{X,Y}$ is a discrete valuation ring, corresponding to a valuation v (the order of vanishing along Y) on K/k; this valuation is of rational rank one and transcendence degree $n-1$ over k, hence Abhyankar. Such a valuation is called a *divisorial valuation*. Conversely, every rational rank one Abhyankar valuation is divisorial: for such a v, there exists some normal model X of K/k and a divisor Y such that v is the order of vanishing along Y [Zariski and Samuel 1960, VI, §14, Theorem 31].

Proposition 2.3.3. *Let* $K \subseteq L$ *be a finite extension of finitely generated field extensions of* k, *and suppose that* w *is valuation on* L/k *extending a valuation* v *on* K/k. *Then* w *is Abhyankar if and only if* v *is Abhyankar.*

Proof. Since L/K is finite, L and K have the same transcendence degree over k. On the other hand, the extension $\kappa(v) \subseteq \kappa(w)$ is also finite by (2.2.0.1), and so

$\kappa(v)$ and $\kappa(w)$ also have the same transcendence degree over k. Again by (2.2.0.1), since Γ_w/Γ_v is a finite abelian group, $\mathbb{Q}\otimes_{\mathbb{Z}}\Gamma_w/\Gamma_v=0$. By exactness of

$$0\to\mathbb{Q}\otimes_{\mathbb{Z}}\Gamma_v\to\mathbb{Q}\otimes_{\mathbb{Z}}\Gamma_w\to\mathbb{Q}\otimes_{\mathbb{Z}}\Gamma_w/\Gamma_z\to 0$$

we conclude that Γ_w and Γ_v have the same rational rank. The result is now clear from the definition of an Abhyankar valuation. $\square$

2.4. ***Frobenius.*** Let R be a ring of prime characteristic p. The Frobenius map $R\xrightarrow{F}R$ is defined by $F(x)=x^p$. We can denote the target copy of R by F_*R and view it as an R-module via restriction of scalars by F; thus F_*R is both a ring (indeed, it is precisely R) and an R-module in which the action of $r\in R$ on $x\in F_*R$ produces r^px. With this notation, the Frobenius map $F:R\to F_*R$ and its iterates $F^e:R\to F^e_*R$ are ring maps, as well as R-module maps. See [Smith and Zhang 2015, p. 294] for a further discussion of this notation.

We note that F^e_* gives us an exact covariant functor from the category of R-modules to itself. This is nothing but the usual restriction of scalars functor associated to the ring homomorphism $F^e:R\to R$.

For an ideal $I\subset R$, the notation $I^{[p^e]}$ denotes the ideal generated by the p^e-th powers of the elements of I. Equivalently, $I^{[p^e]}$ is the expansion of I under the Frobenius map, that is, $I^{[p^e]}=I\,F^e_*R$ as subsets of R.

The image of F^e is the subring $R^{p^e}\subset R$ of p^e-th powers. If R is reduced (which is equivalent to the injectivity of Frobenius), statements about the R-module F^e_*R are equivalent to statements about the R^{p^e}-module R.

Definition 2.4.1. A ring R of characteristic p is *F-finite* if $F:R\to F_*R$ is a finite map of rings, or equivalently, if R is a finitely generated R^p-module. Note that $F:R\to F_*R$ is a finite map if and only if $F^e:R\to F^e_*R$ is a finite map for all $e>0$.

F-finite rings are ubiquitous. For example, every perfect field is F-finite, and a finitely generated algebra over an F-finite ring is F-finite. Furthermore, F-finiteness is preserved under homomorphic images, localization and completion. This means that nearly every ring classically arising in algebraic geometry is F-finite. However, valuation rings even of F-finite fields are often not F-finite.

2.5. ***F-purity and Frobenius splitting.*** We first review purity and splitting for maps of modules over an arbitrary commutative ring A, not necessarily Noetherian or of prime characteristic. A map of A-modules $M\xrightarrow{\varphi}N$ is *pure* if for any A-module Q, the induced map

$$M\otimes_A Q\to N\otimes_A Q$$

is injective. The map $M \xrightarrow{\varphi} N$ is *split* if φ has a left inverse in the category of A-modules. Clearly, a split map is always pure. Although it is not obvious, the converse holds under a weak hypothesis:

Lemma 2.5.1 [Hochster and Roberts 1976, Corollary 5.2]. *Let $M \xrightarrow{\varphi} N$ be a pure map of A-modules where A is a commutative ring. Then φ is split if the cokernel $N/\varphi(M)$ is finitely presented.*

Definition 2.5.2. Let R be an arbitrary commutative ring of prime characteristic p.

(a) The ring R is *Frobenius split* if the map $F : R \to F_*R$ splits as a map of R-modules, that is, there exists an R-module map $F_*R \to R$ such that the composition

$$R \xrightarrow{F} F_*R \to R$$

is the identity map.

(b) The ring R is *F-pure* if $F : R \to F_*R$ is a pure map of R-modules.

A Frobenius split ring is always F-pure. The converse is also true under modest hypothesis:

Corollary 2.5.3. *A Noetherian F-finite ring of characteristic p is Frobenius split if and only if it is F-pure.*

Proof. The F-finiteness hypothesis implies that F_*R is a finitely generated R-module. So a quotient of F_*R is also finitely generated. Since a finitely generated module over a Noetherian ring is finitely presented, the result follows from Lemma 2.5.1. □

2.6. *F-finiteness and excellence.* Although we are mainly concerned with non-Noetherian rings in this paper, it is worth pointing out the following curiosity for readers familiar with Grothendieck's concept of an *excellent ring,* a particular kind of Noetherian ring expected to be the most general setting for many algebro-geometric statements [EGA IV$_2$ 1965, définition 7.8.2].

Proposition 2.6.1. *A Noetherian domain is F-finite if and only if it is excellent and its fraction field is F-finite.*

Proof. If R is F-finite with fraction field K, then also $R \otimes_{R^p} K^p \cong K$ is finite over K^p, so the fraction field of R is F-finite. Furthermore, Kunz [1976, Theorem 2.5] showed that F-finite Noetherian rings are excellent.

We need to show that an excellent Noetherian domain with F-finite fraction field is F-finite. We make use of the following well-known property[2] of an excellent domain A: the integral closure of A in any finite extension of its fraction field is *finite* as an A-module [EGA IV$_2$ 1965, IV, 7.8.3(vi)]. The ring R^p is excellent

[2]sometimes called the *Japanese* or *N2* property.

because it is isomorphic to R, and its fraction field is K^p. Since $K^p \hookrightarrow K$ is finite, the integral closure S of R^p in K is a finite R^p-module. But clearly $R \subset S$, so R is also a finitely generated R^p module, since submodules of a Noetherian module over a Noetherian ring are Noetherian. That is, R is F-finite. □

Using this observation, we can clarify the relationship between F-purity and Frobenius splitting in an important class of rings.

Corollary 2.6.2. *For an excellent Noetherian domain whose fraction field is F-finite, Frobenius splitting is equivalent to F-purity.*

Proof. Our hypothesis implies F-finiteness, so splitting and purity are equivalent by Lemma 2.5.1. □

3. Flatness and purity of Frobenius in valuation rings

Kunz [1969, Theorem 2.1] showed that for a Noetherian ring of characteristic p, the Frobenius map is flat if and only if the ring is regular. In this section, we show how standard results on valuations yield the following result.

Theorem 3.1. *Let V be a valuation ring of characteristic p. Then the Frobenius map $F : V \to F_*V$ is faithfully flat.*

This suggests that we can imagine a valuation ring to be "regular" in some sense. Of course, a Noetherian valuation ring is either a field or a one dimensional regular local ring, but because valuation rings are rarely Noetherian, Theorem 3.1 is not a consequence of Kunz's theorem.

Theorem 3.1 follows from the following general result, whose proof we include for the sake of completeness.

Lemma 3.2 [Bourbaki 1989, VI.3.6, Lemma 1]. *A finitely generated, torsion-free module over a valuation ring is free. In particular, a torsion free module over a valuation ring is flat.*

Proof. Let $M \neq 0$ be a finitely generated, torsion-free V-module. Choose a minimal set of generators $\{m_1, \dots, m_n\}$. If there is a nontrivial relation among these generators, then there exists $v_1, \dots, v_n \in V$ (not all zero) such that $v_1 m_1 + \cdots + v_n m_n = 0$. Re-ordering if necessary, we may assume that v_1 is minimal among (nonzero) coefficients, that is, $(v_i) \subset (v_1)$ for all $i \in \{1, \dots, n\}$. Then for each $i > 1$, there exists $a_i \in V$ such that $v_i = a_i v_1$. This implies that

$$v_1(m_1 + a_2 m_2 + \cdots + a_n m_n) = 0.$$

Since $v_1 \neq 0$ and M is torsion free, we get

$$m_1 + a_2 m_2 + \cdots + a_n m_n = 0.$$

Then $m_1 = -(a_2 m_2 + \cdots + a_n m_n)$. So M can be generated by the smaller set $\{m_2, \ldots, m_n\}$ which contradicts the minimality of n. Hence $\{m_1, \ldots, m_n\}$ must be a free generating set.

The second statement follows by considering a torsion-free module as a directed union of its finitely generated submodules, since a directed union of flat modules is flat [Bourbaki 1989, I.2.7 Proposition 9] □

Proof of Theorem 3.1. Observe that $F_* V$ is a torsion free V-module. So by Lemma 3.2, the module $F_* V$ is flat, which means the Frobenius map is flat. To see that Frobenius is faithfully flat, we need only check that $m F_* V \neq F_* V$ for m the maximal ideal of V [Bourbaki 1989, I.3.5 Proposition 9(e)]. But this is clear: the element $1 \in F_* V$ is not in $m F_* V$, since $1 \in V$ is not in the ideal $m^{[p]}$. □

Corollary 3.3. *Every valuation ring of characteristic p is F-pure.*

Proof. Fix a valuation ring V of characteristic p. We have already seen that the Frobenius map $V \to F_* V$ is faithfully flat (Theorem 3.1). But any faithfully flat map of rings $A \to B$ is pure as a map of A-modules [Bourbaki 1989, I.3.5 Proposition 9(c)]. □

4. F-finite valuation rings

In this section, we investigate F-finiteness in valuation rings. We first prove Theorem 4.1.1 characterizing F-finite valuation rings as those V for which $F_* V$ is a *free* V-module. We then prove a numerical characterization of F-finiteness in terms of ramification index and residue degree for extensions of valuations under Frobenius in Theorem 4.3.1. This characterization is useful for constructing interesting examples, and later for showing that F-finite valuations are Abhyankar.

4.1. ***Finiteness and freeness of Frobenius.*** For any domain R of characteristic p, we have already observed (see the proof of Proposition 2.6.1) that a necessary condition for F-finiteness is the F-finiteness of its fraction field. For this reason, we investigate F-finiteness of valuation rings only in F-finite ambient fields.

Theorem 4.1.1. *Let K be an F-finite field. A valuation ring V of K is F-finite if and only if $F_* V$ is a free V-module.*

Proof. First assume $F_* V$ is free over V. Since $K \otimes_R F_* V \cong F_* K$ as K-vector spaces, the rank of $F_* V$ over V must be the same as the rank of $F_* K$ over K, namely the degree $[F_* K : K] = [K : K^p]$. Since K is F-finite, this degree is finite, and so $F_* V$ is a free V-module of finite rank. In particular, V is F-finite.

Conversely, suppose that V is F-finite. Then $F_* V$ is a finitely generated, torsion-free V-module. So it is free by Lemma 3.2. □

Corollary 4.1.2. *An F-finite valuation ring is Frobenius split.*

Proof. One of the rank one free summands of F_*V is the copy of V under F, so this copy of V splits off F_*V. Alternatively, since $V \to F_*V$ is pure, we can use Lemma 2.5.1: the cokernel of $V \to F_*V$ is finitely presented because it is finitely generated (being a quotient of the finitely generated V-module F_*V) and the module of relations is finitely generated (by $1 \in F_*V$). □

Remark 4.1.3. The same argument shows that any module finite extension $V \hookrightarrow S$ splits — in other words, every valuation ring is a *splinter* in the sense of [Ma 1988]; see also [Ma 1988, Lemma 1.2].

4.2. ***Frobenius splitting in the Noetherian case.*** We can say more for Noetherian valuation rings. First we make a general observation about F-finiteness in Noetherian rings.

Theorem 4.2.1. *For a Noetherian domain whose fraction field is F-finite, Frobenius splitting implies F-finiteness (and hence excellence).*

Before embarking on the proof, we point out a consequence for valuation rings.

Corollary 4.2.2. *For a discrete valuation ring V whose fraction field is F-finite, the following are equivalent:*

(i) *V is Frobenius split;*

(ii) *V is F-finite;*

(iii) *V is excellent.*

Proof. A DVR is Noetherian, so equivalence of (i) and (ii) follows from combining Theorem 4.2.1 and Corollary 4.1.2. The equivalence with excellence follows from Proposition 2.6.1. □

Remark 4.2.3. We have proved that all valuation rings are F-pure. However, not all valuation rings, even discrete ones on $\mathbb{F}_p(x, y)$, are Frobenius split, as Example 4.5.1 below shows.

The proof of Theorem 4.2.1 relies on the following lemma.

Lemma 4.2.4. *A Noetherian domain with F-finite fraction field is F-finite if and only if there exists $\phi \in \operatorname{Hom}_{R^p}(R, R^p)$ such that $\phi(1) \neq 0$.*

Proof. Assuming such ϕ exists, we first observe that the canonical map $R \to R^{\vee\vee}$ is injective, where $R^{\vee\vee} := \operatorname{Hom}_{R^p}(\operatorname{Hom}_{R^p}(R, R^p), R^p)$ is the double dual. Indeed, let $x \in R$ be a nonzero element. It suffices to show that there exists $f \in R^\vee := \operatorname{Hom}_{R^p}(R, R^p)$ such that $f(x) \neq 0$. Let $f = \phi \circ x^{p-1}$, where x^{p-1} is the R^p-linear map $R \to R$ given by multiplication by x^{p-1}. Then $f(x) = \phi(x^p) = x^p\phi(1) \neq 0$. This shows that the double dual map is injective.

Now, to show that R is a finitely generated R^p-module, it suffices to show that the larger module $R^{\vee\vee}$ is finitely generated. For this it suffices to show that $R^\vee$ is a finitely generated R^p-module, since the dual of a finitely generated module is finitely generated.

We now show that $R^\vee$ is finitely generated. Let M be a maximal free R^p-submodule of R. Note that M has finite rank (equal to $[K : K^p]$, where K is the fraction field of R) and that R/M is a torsion R^p-module. Since the dual of a torsion module is zero, dualizing the exact sequence $0 \to M \to R \to R/M \to 0$ induces an injection

$$R^\vee := \mathrm{Hom}_{R^p}(R, R^p) \hookrightarrow \mathrm{Hom}_{R^p}(M, R) = M^\vee.$$

Since M is a finitely generated R^p-module, also $M^\vee$, and hence its submodule $R^\vee$ is finitely generated (R is Noetherian). This completes the proof that R is F-finite.

For the converse, fix any K^p-linear splitting $\psi : K \to K^p$. Restricting to R produces an R^p-linear map to K^p. Since R is finitely generated over R^p, we can multiply by some nonzero element c of R^p to produce a nonzero map $\phi : R \to R^p$ such that $\phi(1) = c \neq 0$, completing the proof. □

Proof of Theorem 4.2.1. Let R be a domain with F-finite fraction field. A Frobenius splitting is a map $\phi \in \mathrm{Hom}_{R^p}(R, R^p)$ such that $\phi(1) = 1$. Theorem 4.2.1 then follows immediately from Lemma 4.2.4. □

4.3. ***A numerical criterion for F-finiteness.*** Consider the extension

$$K^p \subseteq K$$

where K any field of characteristic p. For any valuation v on K, let v^p denote the restriction to K^p. We next characterize F-finite valuations in terms of the ramification index and residue degree of v over v^p.

Theorem 4.3.1. *A valuation ring V of an F-finite field K of prime characteristic p is F-finite if and only if*

$$e(v/v^p)f(v/v^p) = [K : K^p],$$

where v is the corresponding valuation on K and v^p is its restriction to K^p.

Proof. First note that v is the only valuation of K extending v^p. Indeed, v is uniquely determined by its values on elements of K^p, since $v(x^p) = pv(x)$ and the value group of v is torsion-free. Furthermore, the valuation ring of v^p is easily checked to be V^p.

Observe that V is the integral closure of V^p in K. Indeed, since V is a valuation ring, it is integrally closed in K, but it is also obviously integral over V^p. We now apply Proposition 2.2.1. Since there is only one valuation extending v^p, the inequality

$$e(v/v^p)f(v/v^p) \leq [K : K^p]$$

will be an equality if and only if the integral closure of V^p in K, namely V, is finite over V^p. □

The following simple consequence has useful applications to the construction of interesting examples of F-finite and non-F-finite valuations.

Corollary 4.3.2. *Let V be a valuation ring of an F-finite field K of characteristic p. If $e(v/v^p) = [K : K^p]$ or $f(v/v^p) = [K : K^p]$, then V is F-finite.*

Remark 4.3.3. Theorem 4.3.1 and its corollary are easy to apply, because the ramification index and residue degree for the extension $V^p \hookrightarrow V$ can be computed in practice. Indeed, since Γ_{v^p} is clearly the subgroup $p\Gamma_v$ of Γ_v, we see that

$$e(v/v^p) = [\Gamma_v : p\Gamma_v]. \tag{4.3.3.1}$$

Also, the local map $V^p \hookrightarrow V$ induces the residue field extension $\kappa(v^p) \hookrightarrow \kappa(v)$, which identifies the field $\kappa(v^p)$ with the subfield $(\kappa(v))^p$. This means that

$$f(v/v^p) = [\kappa(v) : \kappa(v)^p]. \tag{4.3.3.2}$$

4.4. *Examples of Frobenius split valuations.* We can use our characterization of F-finite valuations to easily give examples of valuations on $\mathbb{F}_p(x, y)$ that are nondiscrete but Frobenius split.

Example 4.4.1. Consider the rational function field $K = k(x, y)$ over a perfect field k of characteristic p. For an irrational number $\alpha \in \mathbb{R}$, let Γ be the ordered additive subgroup of $\mathbb{R}$ generated by 1 and α. Consider the unique valuation $v : K^\times \to \Gamma$ determined by

$$v(x^i y^j) = i + j\alpha,$$

and let V be the corresponding valuation ring. Since $\Gamma \cong \mathbb{Z} \oplus \mathbb{Z}$ via the map which sends $a + b\alpha \mapsto (a, b)$, we see that the value group of v^p is $p\Gamma \cong p(\mathbb{Z} \oplus \mathbb{Z})$. Hence

$$e(v/v^p) = [\Gamma : p\Gamma] = p^2 = [K : K^p].$$

So V is F-finite by Corollary 4.3.2. Thus V is also Frobenius split by Corollary 4.1.2.

Example 4.4.2. Consider the lex valuation on the rational function field $K = k(x_1, x_2, \ldots, x_n)$ over a perfect field k of characteristic p. This is the valuation $v : K^\times \to \mathbb{Z}^n$ on K/k defined by sending a monomial $x_1^{a_1} \cdots x_n^{a_n}$ to $(a_1, \ldots, a_n) \in \mathbb{Z}^{\oplus n}$, where $\Gamma = \mathbb{Z}^{\oplus n}$ is ordered lexicographically. Let V be the corresponding valuation ring. The value group of V^p is $p\Gamma$, so $e(v/v^p) = [\Gamma : p\Gamma] = p^n = [K : K^p]$. As in the previous example, Corollary 4.3.2 implies that V is F-finite, and so again F-split.

4.5. ***Example of a non-Frobenius split valuation.*** Our next example shows that discrete valuation rings are not always F-finite, even in the rational function field $\mathbb{F}_p(x, y)$. This is adapted from [Zariski and Samuel 1960, Example, p. 62], where it is credited to F. K. Schmidt.

Example 4.5.1. Let $\mathbb{F}_p((t))$ be the fraction field of the discrete valuation ring $\mathbb{F}_p[[t]]$ of power series in one variable. Since the field of rational functions $\mathbb{F}_p(t)$ is countable, the uncountable field $\mathbb{F}_p((t))$ cannot be algebraic over $\mathbb{F}_p(t)$. So we can find some power series

$$f(t) = \sum_{n=1}^{\infty} a_n t^n$$

in $\mathbb{F}_p[[t]]$ transcendental over $\mathbb{F}_p(t)$.

Since t and $f(t)$ are algebraically independent, there is an *injective* ring map

$$\mathbb{F}_p[x, y] \hookrightarrow \mathbb{F}_p[[t]] \quad \text{such that} \quad x \mapsto t \text{ and } y \mapsto f(t)$$

which induces an extension of fields

$$\mathbb{F}_p(x, y) \hookrightarrow \mathbb{F}_p((t)).$$

Restricting the t-adic valuation on $\mathbb{F}_p((t))$ to the subfield $\mathbb{F}_p(x, y)$ produces a discrete valuation v of $\mathbb{F}_p(x, y)$. Let V denote its valuation ring.

We claim that V is not F-finite, a statement we can verify with Theorem 4.3.1. Note that $L = \mathbb{F}_p(x, y)$ is F-finite, with $[L : L^p] = p^2$. Since the value group Γ_v is $\mathbb{Z}$, we see that

$$e(v/v^p) = [\Gamma_v : p\Gamma_v] = p.$$

On the other hand, to compute the residue degree $f(v/v^p)$, we must understand the field extension $\kappa(v)^p \hookrightarrow \kappa(v)$. Observe that for an element $u \in \mathbb{F}_p(x, y)$ to be in V, its image in $\mathbb{F}_p((t))$ must be a power series of the form

$$\sum_{n=0}^{\infty} b_n t^n,$$

where $b_n \in \mathbb{F}_p$. Clearly

$$v(u - b_0) > 0,$$

which means that the class of $u = \sum_{n=0}^{\infty} b_n t^n$ in $\kappa(v)$ is equal to the class of b_0 in $\kappa(v)$. This implies that $\kappa(v) \cong \mathbb{F}_p$, so that $[\kappa(v) : \kappa(v)^p] = 1$. That is, $f(v/v^p) = 1$.

Finally, we then have that

$$e(v/v^p) f(v/v^p) = p \neq p^2 = [\mathbb{F}_p(x, y) : (\mathbb{F}_p(x, y))^p].$$

So V *cannot* be F-finite by Theorem 4.3.1. Thus this Noetherian ring is neither Frobenius split nor excellent by Corollary 4.2.2.

4.6. ***Finite extensions.*** Frobenius properties of valuations are largely preserved under finite extension. First note that if $K \hookrightarrow L$ is a finite extension of F-finite fields, then $[L : L^p] = [K : K^p]$; this follows immediately from the commutative diagram of fields

$$\begin{tikzcd} L & K \arrow[l, hook'] \\ L^p \arrow[u, hook] & K^p \arrow[l, hook'] \arrow[u, hook] \end{tikzcd}$$

To wit, $[L : K^p] = [L : K][K : K^p] = [L : L^p][L^p : K^p]$ and $[L : K] = [L^p : K^p]$, so that $[L : L^p] = [K : K^p]$. Moreover, we have:

Proposition 4.6.1. *Let $K \hookrightarrow L$ be a finite extension of F-finite fields of characteristic p. Let v be a valuation on K and w an extension of v to L. Then*:

(i) *The ramification indices $e(v/v^p)$ and $e(w/w^p)$ are equal.*

(ii) *The residue degrees $f(v/v^p)$ and $f(w/w^p)$ are equal.*

(iii) *The valuation ring for v is F-finite if and only if the valuation ring for w is F-finite.*

Proof. By (2.2.0.1), we have

$$[\Gamma_w : \Gamma_v][\kappa(w) : \kappa(v)] \le [L : K],$$

so both $[\Gamma_w : \Gamma_v]$ and $[\kappa(w) : \kappa(v)]$ are finite. Of course, we also know that the ramification indices $e(w/w^p) = [\Gamma_w : p\Gamma_w]$ and $e(v/v^p) = [\Gamma_v : p\Gamma_v]$ are finite, as are the residue degrees $f(w/w^p) = [\kappa(w) : \kappa(w)^p]$ and $f(v/v^p) = [\kappa(v) : \kappa(v)^p]$.

(i) In light of (4.3.3.1), we need to show that $[\Gamma_w : p\Gamma_w] = [\Gamma_v : p\Gamma_v]$. Since Γ_w is torsion-free, multiplication by p induces an isomorphism $\Gamma_w \cong p\Gamma_w$, under which the subgroup Γ_v corresponds to $p\Gamma_v$. Thus $[p\Gamma_w : p\Gamma_v] = [\Gamma_w : \Gamma_v]$. Using the commutative diagram of finite index abelian subgroups

$$\begin{tikzcd} \Gamma_w & \Gamma_v \arrow[l, hook'] \\ p\Gamma_w \arrow[u, hook] & p\Gamma_v \arrow[l, hook'] \arrow[u, hook] \end{tikzcd}$$

we see that $[\Gamma_w : p\Gamma_w][p\Gamma_w : p\Gamma_v] = [\Gamma_w : \Gamma_v][\Gamma_v : p\Gamma_v]$. Whence $[\Gamma_w : p\Gamma_w] = [\Gamma_v : p\Gamma_v]$.

(ii) In light of (4.3.3.2), we need to show that $[\kappa(w):\kappa(w)^p]=[\kappa(v):\kappa(v)^p]$. We have $[\kappa(w)^p:\kappa(v)^p]=[\kappa(w):\kappa(v)]$, so the result follows from computing the extension degrees in the commutative diagram of finite field extensions:

$$\begin{tikzcd} \kappa(w) & \kappa(v) \arrow[l, hook'] \\ \kappa(w)^p \arrow[u, hook] & \kappa(v)^p \arrow[l, hook'] \arrow[u, hook] \end{tikzcd}$$

(iii) By (i) and (ii) we get $e(w/w^p)=e(v/v^p)$ and $f(w/w^p)=f(v/v^p)$. Therefore

$$e(w/w^p)f(w/w^p)=e(v/v^p)f(v/v^p).$$

Since also $[L:L^p]=[K:K^p]$, we see using Theorem 4.3.1 that w is F-finite if and only if v is F-finite. □

5. F-finiteness in function fields

An important class of fields are *function fields over a ground field* k. By definition, a field K is a function field over k if it is a finitely generated field extension of k. These are the fields that arise as function fields of varieties over a (typically algebraically closed) ground field k. What more can be said about valuation rings in this important class of fields?

We saw in Example 4.5.1 that not every valuation of an F-finite function field is F-finite. However, the following theorem gives a nice characterization of those that are.

Theorem 5.1. *Let K be a finitely generated field extension of an F-finite ground field k. The following are equivalent for a valuation v on K/k:*

(i) *The valuation v is Abhyankar.*

(ii) *The valuation ring R_v is F-finite.*

(iii) *The valuation ring R_v is a free R_v^p-module.*

Furthermore, when these equivalent conditions hold, it is also true that R_v is Frobenius split.

Since Abhyankar valuations have finitely generated value groups and residue fields, the following corollary holds.

Corollary 5.2. *An F-finite valuation of a function field over an F-finite field k has a finitely generated value group and its residue field is a finitely generated field extension of k.*

For example, valuations whose value groups are $\mathbb{Q}$ can never be F-finite.

Remark 5.3. In light of Proposition 2.6.1, we could add a fourth item to the list of equivalent conditions in Theorem 5.1 in the *Noetherian* case: the valuation R_v is excellent. The theorem says that the only discrete valuation rings (of function fields) that are F-finite are the divisorial valuation rings or equivalently, the excellent DVRs.

To prove Theorem 5.1, first recall that the equivalence of (ii) and (iii) was already established in Theorem 4.1.1. The point is to connect these conditions with the Abyhankar property. Our strategy is to use Theorem 4.3.1, which tells us that a valuation v on K is F-finite if and only if

$$e(v/v^p)f(v/v^p) = [K : K^p].$$

We do this by proving two propositions, one comparing the rational rank of v to the ramification index $e(v/v^p)$, and the other comparing the transcendence degree of v to the residue degree $f(v/v^p)$.

Proposition 5.4. *Let v be a valuation of rational rank s on an F-finite field K. Then*

$$e(v/v^p) \leq p^s,$$

with equality when the value group Γ_v is finitely generated.

Proof. To see that equality holds when Γ_v is finitely generated, note that in this case, $\Gamma_v \cong \mathbb{Z}^{\oplus s}$. So $\Gamma_v/p\Gamma_v \cong (\mathbb{Z}/p\mathbb{Z})^{\oplus s}$, which has cardinality p^s. That is, $e(v/v^p) = p^s$.

It remains to consider the case where Γ may not be finitely generated. Nonetheless, since $e(v/v^p)$ is finite (see (2.2.0.1)), we do know that $[\Gamma_v : p\Gamma_v] = e(v/v^p)$ is finite. So the proof of Proposition 5.4 comes down to the following simple lemma about abelian groups.

Lemma 5.5. *Let Γ be a torsion free abelian group of rational rank s. Then*

$$[\Gamma : p\Gamma] \leq p^s.$$

It suffices to show that $\Gamma/p\Gamma$ is a vector space of dimension $\leq s$ over $\mathbb{Z}/p\mathbb{Z}$. So let $t_1, \ldots, t_n$ be elements of Γ whose classes modulo $p\Gamma$ are linearly independent over $\mathbb{Z}/p\mathbb{Z}$. Then we claim that the t_i are $\mathbb{Z}$-independent elements of Γ. Assume to the contrary that there is some nontrivial relation $a_1t_1 + \cdots + a_nt_n = 0$, for some integers a_i. Since Γ is torsion-free, we can assume without loss of generality, that at least one a_j is not divisible by p. But now modulo $p\Gamma$, this relation produces a nontrivial relation on classes of the t_i in $\Gamma/p\Gamma$, contrary to the fact that these are linearly independent. This shows that any $\mathbb{Z}/p\mathbb{Z}$-linearly independent subset of $\Gamma/p\Gamma$ must have cardinality at most s. Thus the lemma, and hence Proposition 5.4, is proved. $\square$

Proposition 5.6. *Let K be a finitely generated field extension of an F-finite ground field k. Let v be a valuation of transcendence degree t on K over k. Then*

$$f(v/v^p) \leq p^t[k : k^p],$$

with equality when $\kappa(v)$ is finitely generated over k.

Proof. The second statement follows immediately from the following well-known fact, whose proof is an easy computation.

Lemma 5.7. *A finitely generated field L of characteristic p and transcendence degree n over k satisfies $[L : L^p] = [k : k^p]p^n$.*

It remains to consider the case where $\kappa(v)$ may not be finitely generated. Because K/k is a function field, Abhyankar's inequality (2.3.1.1) guarantees that the transcendence degree of $\kappa(v)$ over k is finite. Let $x_1, \dots, x_t$ be a transcendence basis. There is a factorization

$$k \hookrightarrow k(x_1, \dots, x_t) \hookrightarrow \kappa(v),$$

where the second inclusion is algebraic. The proposition follows immediately from the next lemma.

Lemma 5.8. *If $L' \subseteq L$ is an algebraic extension of F-finite fields, then $[L : L^p] \leq [L' : L'^p]$.*

To prove this lemma, recall that Proposition 4.6.1 ensures that $[L : L^p] = [L' : L'^p]$ when $L' \subseteq L$ is *finite.* So suppose L is algebraic but not necessarily finite over L'. Fix a basis $\{\alpha_1, \dots, \alpha_n\}$ for L over L^p, and consider the intermediate field

$$L' \hookrightarrow L'(\alpha_1, \dots, \alpha_n) \hookrightarrow L.$$

Since each α_i is algebraic over L', it follows that $\tilde{L} := L'(\alpha_1, \dots, \alpha_n)$ is finite over L', so again $[\tilde{L} : \tilde{L}^p] = [L' : L'^p]$ by Proposition 4.6.1. Now observe that $\tilde{L}^p \subset L^p$, and so the L^p-linearly independent set $\{\alpha_1, \dots, \alpha_n\}$ is also linearly independent over $\tilde{L}^p$. This means that $[L : L^p] \leq [\tilde{L} : \tilde{L}^p]$ and hence $[L : L^p] \leq [L' : L'^p]$. This proves Lemma 5.8.

Finally, Proposition 5.6 is proved by applying Lemma 5.8 to the inclusion

$$L' = k(x_1, \dots, x_t) \hookrightarrow L = \kappa(v).$$

(Note that $\kappa(v)$ is F-finite, because $[\kappa(v) : (\kappa(v))^p] = f(v/v^p) \leq [K : K^p]$ from the general inequality (2.2.0.1).) So we get

$$f(v/v^p) = [\kappa(v) : (\kappa(v))^p] \leq [k(x_1, \dots, x_t) : (k(x_1, \dots, x_t))^p] = p^t[k : k^p]. \quad \square$$

Proof of Theorem 5.1. It only remains to prove the equivalence of (i) and (ii). First assume v is Abhyankar. Then its value group Γ_v is finitely generated and its residue field $\kappa(v)$ is finitely generated over k. According to Proposition 5.4, we have $e(v/v^p) = p^s$, where s is the rational rank of v. According to Proposition 5.6, we have $f(v/v^p) = p^t[k : k^p]$, where t is the transcendence degree of v. By definition of Abhyankar, $s + t = n$, where n is the transcendence degree of K/k. But then

$$e(v/v^p)f(v/v^p) = (p^s)(p^t)[k : k^p] = p^n[k : k^p] = [K : K^p].$$

By Theorem 4.3.1, we can conclude that v is F-finite.

Conversely, we want to prove that a valuation v with F-finite valuation ring R_v is Abhyankar. Let s denote the rational rank and t denote the transcendence degree of v. From Theorem 4.3.1, the F-finiteness of v gives

$$e(v/v^p)f(v/v^p) = [K : K^p] = p^n[k : k^p].$$

Using the bounds $p^s \geq e(v/v^p)$ and $p^t[k : k^p] \geq f(v/v^p)$ provided by Propositions 5.4 and 5.6, respectively, we substitute to get

$$p^s p^t[k : k^p] \geq e(v/v^p)f(v/v^p) = p^n[k : k^p].$$

It follows that $s + t \geq n$. Then $s + t = n$ by (2.3.1.1), and v is Abhyankar. □

6. F-regularity

An important class of F-pure rings are the strongly F-regular rings. Originally, strongly F-regular rings were defined only in the Noetherian F-finite case. By definition, a Noetherian F-finite reduced ring R of prime characteristic p is strongly F-regular if for every non-zerodivisor c, there exists e such that the map

$$R \to F^e_* R, \quad 1 \mapsto c$$

splits in the category of R-modules [Hochster and Huneke 1989]. In this section, we show that by replacing the word "splits" with the words "is pure" in the above definition, we obtain a well-behaved notion of F-regularity in a broader setting. Hochster and Huneke [1994, Remark 5.3] themselves suggested, but never pursued, this possibility.

Strong F-regularity first arose as a technical tool in the theory of *tight closure*: Hochster and Huneke [1994] made use of it in their deep proof of the existence of test elements. Indeed, the original motivation for (and the name of) strong F-regularity was born from a desire to better understand *weak F-regularity*, the property of a Noetherian ring characterized by all ideals being tightly closed. In many contexts, strong and weak F-regularity are known to be equivalent (see, e.g., [Lyubeznik and Smith 1999] for the graded case, [Hochster and Huneke 1989] for the Gorenstein case), but it is becoming clear that at least for many applications, strong F-regularity is the more useful and flexible notion. Applications beyond tight closure include commutative algebra more generally [Aberbach and Leuschke 2003; Blickle 2008; Schwede and Tucker 2012; Schwede 2009a; Smith and Zhang 2015], algebraic geometry [Gongyo et al. 2015; Hacon and Xu 2015; Patakfalvi 2014; Schwede and Smith 2010; Smith 2000], representation theory [Brion and Kumar 2005; Mehta and Ramanathan 1985; Ramanathan 1991; Smith and Van den Bergh 1997] and combinatorics [Benito et al. 2015].

6.1. ***Basic properties of F-pure regularity.*** We propose the following definition, intended to be a generalization of strong F-regularity to arbitrary commutative rings of characteristic p, not necessarily F-finite or Noetherian.

Definition 6.1.1. Let c be an element in a ring R of prime characteristic p. Then R is said to be *F-pure along c* if there exists $e > 0$ such that the R-linear map

$$\lambda_c^e : R \to F_*^e R, \quad 1 \mapsto c$$

is a pure map of R-modules. We say R is *F-pure regular* if it is F-pure along every non-zerodivisor.

A ring R is F-pure if and only if it is F-pure along the element 1. Thus F-pure regularity is a substantial strengthening of F-purity, requiring F-purity along *all* non-zerodivisors instead of just along the unit.

Remark 6.1.2. (i) If R is Noetherian and F-finite, then the map $\lambda_c^e : R \to F_*^e R$ is pure if and only if it splits (by Lemma 2.5.1). So F-pure regularity for a Noetherian F-finite ring is the same as strong F-regularity.

(ii) If c is a zerodivisor, then the map λ_c^e is never injective for any $e \geq 1$. In particular, a ring is never F-pure along a zerodivisor.

(iii) The terminology "F-pure along c" is chosen to honor Ramanathan's [1991] closely related notion of "Frobenius splitting along a divisor". See [Smith 2000].

The following proposition gathers up some basic properties of F-pure regularity for arbitrary commutative rings.

Proposition 6.1.3. *Let R be a commutative ring of characteristic p, not necessarily Noetherian or F-finite.*

(a) *If R is F-pure along some element, then R is F-pure. More generally, if R is F-pure along a product cd, then R is F-pure along the factors c and d.*

(b) *If R is F-pure along some element, then R is reduced.*

(c) *If R is an F-pure regular ring with finitely many minimal primes, and $S \subset R$ is a multiplicative set, then $S^{-1}R$ is F-pure regular. In particular, F-pure regularity is preserved under localization in Noetherian rings, as well as in domains.*

(d) *Let $\varphi : R \to T$ be a pure ring map which maps non-zerodivisors of R to non-zerodivisors of T. If T is F-pure regular, then R is F-pure regular. In particular, if $\varphi : R \to T$ is faithfully flat and T is F-pure regular, then R is F-pure regular.*

(e) *Let $R_1, \ldots, R_n$ be rings of characteristic p. If $R_1 \times \cdots \times R_n$ is F-pure regular, then each R_i is F-pure regular.*

The proof of Proposition 6.1.3 consists mostly of applying general facts about purity to the special case of the maps λ_c^e. For the convenience of the reader, we gather these basic facts together in one lemma.

Lemma 6.1.4. *Let A be an arbitrary commutative ring A, not necessarily Noetherian nor of characteristic p.*

(a) *If $M \to N$ and $N \to Q$ are pure maps of A-modules, then the composition $M \to N \to Q$ is also pure.*

(b) *If a composition $M \to N \to Q$ of A-modules is pure, then $M \to N$ is pure.*

(c) *If B is an A-algebra and $M \to N$ is pure map of A-modules, then $B \otimes_A M \to B \otimes_A N$ is a pure map of B-modules.*

(d) *Let B be an A-algebra. If $M \to N$ is a pure map of B-modules, then it is also pure as a map of A-modules.*

(e) *An A-module map $M \to N$ is pure if and only if for all prime ideals $\mathcal{P} \subset A$, $M_{\mathcal{P}} \to N_{\mathcal{P}}$ is pure.*

(f) *A faithfully flat map of rings is pure.*

(g) *If $(\Lambda, \leq)$ is a directed set with a least element λ_0, and $\{N_\lambda\}_{\lambda \in \Lambda}$ is a direct limit system of A-modules indexed by Λ and $M \to N_{\lambda_0}$ is an A-linear map, then $M \to \varinjlim_\lambda N_\lambda$ is pure if and only if $M \to N_\lambda$ is pure for all λ.*

(h) *A map of modules $A \to N$ over a Noetherian local ring (A, m) is pure if and only if $E \otimes_A A \to E \otimes_A N$ is injective, where E is the injective hull of the residue field of R.*

Proof. Properties (a)–(d) follow easily from the definition of purity and elementary properties of tensor products. As an example, let us prove (d). If P is an A-module, we want to show that $P \otimes_A M \to P \otimes_A N$ is injective. The map of B-modules

$$(P \otimes_A B) \otimes_B M \to (P \otimes_A B) \otimes_B N$$

is injective by purity of $M \to N$ as a map of B-modules. Using the natural A-module isomorphisms $(P \otimes_A B) \otimes_B M \cong P \otimes_A M$ and $(P \otimes_A B) \otimes_B N \cong P \otimes_A N$, we conclude that $P \otimes_A M \to P \otimes_A N$ is injective in the category of A-modules.

Property (e) follows from (c) by tensoring with A_p and the fact that injectivity of a map of modules is a local property. Property (f) follows from [Bourbaki 1989, I.3.5, Proposition 9(c)]. Properties (g) and (h) are proved in [Hochster and Huneke 1995, Lemma 2.1]. $\square$

Proof of Proposition 6.1.3. (a) Multiplication by d is an R-linear map, so by restriction of scalars also

$$F^e_* R \xrightarrow{\times d} F^e_* R$$

is R-linear. Precomposing with λ^e_c we have

$$R \xrightarrow{\lambda^e_c} F^e_* R \xrightarrow{\times d} F^e_* R, \quad 1 \mapsto cd,$$

which is λ^e_{cd}. Our hypothesis that R is F-pure along cd means that there is some e for which this composition is pure. So by Lemma 6.1.4(b), it follows also that λ^e_c is pure. That is, R is F-pure along c (and since R is commutative, along d). The second statement follows since F-purity along the product $c \times 1$ implies R is F-pure along 1. So some iterate of Frobenius is a pure map, and so F-purity follows from Lemma 6.1.4(b).

(b) By (a) we see that R is F-pure. In particular, the Frobenius map is pure and hence injective, so R is reduced.

(c) Note that by (b), R is reduced. Let $\alpha \in S^{-1}R$ be a non-zerodivisor. Because R has finitely many minimal primes, a standard prime avoidance argument shows that there exists a non-zerodivisor $c \in R$ and $s \in S$ such that $\alpha = c/s$ (a minor modification of [Hochster 2007, Proposition, p. 57]). By hypothesis, R is F-pure along c. Hence there exists $e > 0$ such that the map $\lambda^e_c : R \to F^e_* R$ is pure. Then the map

$$\lambda^e_{c/1} : S^{-1}R \longrightarrow F^e_*(S^{-1}R), \quad 1 \mapsto c/1$$

is pure by 6.1.4(e) and the fact that $S^{-1}(F^e_* R) \cong F^e_*(S^{-1}R)$ as $S^{-1}R$-modules (the isomorphism $S^{-1}(F^e_* R) \cong F^e_*(S^{-1}R)$ is given by $r/s \mapsto r/s^{p^e}$). Now the $S^{-1}R$-linear map

$$\ell_{1/s} : S^{-1}R \to S^{-1}R, \quad 1 \mapsto 1/s$$

is an isomorphism. Applying F^e_*, we see that

$$F^e_*(\ell_{1/s}) : F^e_*(S^{-1}R) \to F^e_*(S^{-1}R), \quad 1 \mapsto 1/s$$

is also an isomorphism of $S^{-1}R$-modules. In particular, $F^e_*(\ell_{1/s})$ is a pure map of $S^{-1}R$-modules. So purity of

$$F^e_*(\ell_{1/s}) \circ \lambda^e_{c/1}$$

follows by 6.1.4(a). But $F^e_*(\ell_{1/s}) \circ \lambda^e_{c/1}$ is precisely the map

$$\lambda^e_{c/s} : S^{-1}R \to F^e_*(S^{-1}R), \quad 1 \mapsto c/s.$$

(d) Let $c \in R$ be a non-zerodivisor. Then $\varphi(c)$ is a non-zerodivisor in T by hypothesis. Pick $e > 0$ such that the map $\lambda^e_{\varphi(c)} : T \to F^e_*T$ is a pure map of T-modules. By 6.1.4(f) and 6.1.4(a),

$$R \xrightarrow{\varphi} T \xrightarrow{\lambda^e_{\varphi(c)}} F^e_*T$$

is a pure map of R-modules. We have commutative diagram of R-linear maps

$$\begin{tikzcd} R \arrow[r, "\varphi"] \arrow[d, "\lambda^e_c"'] & T \arrow[d, "\lambda^e_{\varphi(c)}"] \\ F^e_*R \arrow[r, "F^e_*(\varphi)"] & F^e_*T \end{tikzcd}$$

The purity of λ^e_c follows by 6.1.4(b). Note that if φ is faithfully flat, then it is pure by 6.1.4(f) and maps non-zerodivisors to non-zerodivisors.

(e) Let $R := R_1 \times \cdots \times R_n$. Consider the multiplicative set

$$S := R_1 \times \cdots \times R_{i-1} \times \{1\} \times R_{i+1} \times \cdots \times R_n.$$

Since $S^{-1}R \cong R_i$, it suffices to show that $S^{-1}R$ is F-pure regular. So let $\alpha \in S^{-1}R$ be a non-zerodivisor. Note that we can select $u \in R$ and $s \in S$ such that u is a non-zerodivisor and $\alpha = u/s$. So we can now repeat the proof of (c) verbatim to see that $S^{-1}R$ must be pure along α. $\square$

Remark 6.1.5. It is worth observing that in Definition 6.1.1, if the map λ^e_c is a pure map, then λ^f_c is also a pure map for all $f \geq e$. Indeed, to see this note that it suffices to show that λ^{e+1}_c is pure. We know R is F-pure by 6.1.3(a). So Frobenius

$$F : R \to F_*R$$

is a pure map of R-modules. By hypothesis,

$$\lambda^e_c : R \to F^e_*R$$

is pure. Hence 6.1.4(d) tell us that

$$F_*(\lambda^e_c) : F_*R \to F_*(F^e_*R)$$

is a pure map of R-modules. Hence the composition

$$R \xrightarrow{F} F_*R \xrightarrow{F_*(\lambda_c^e)} F_*(F_*^eR), \quad 1 \mapsto c$$

is a pure map of R-modules by 6.1.4(a). But $F_*(F_*^eR)$ as an R-module is precisely $F_*^{e+1}R$. So

$$\lambda_c^{e+1} : R \to F_*^{e+1}R.$$

is pure.

Example 6.1.6. The polynomial ring over $\mathbb{F}_p$ in infinitely many variables (localized at the obvious maximal ideal) is an example of a F-pure regular ring which is not Noetherian.

6.2. *Relationship of F-pure regularity to other singularities.* We show that our generalization of strong F-regularity continues to enjoy many important properties of the more restricted version.

Theorem 6.2.1 [Hochster and Huneke 1989, Theorem 3.1(c)]. *A regular local ring, not necessarily F-finite, is F-pure regular.*

Proof. Let (R, m) be a regular local ring. By Krull's intersection theorem we know that

$$\bigcap_{e>0} m^{[p^e]} = 0.$$

Since R is a domain, the non-zerodivisors are precisely the nonzero elements of R. So let $c \in R$ be a nonzero element. Choose e such that $c \notin m^{[p^e]}$. We show that the map

$$\lambda_c^e : R \to F_*^eR, \quad 1 \mapsto c$$

is pure.

By Lemma 6.1.4, it suffices to check that for the injective hull E of the residue field of R, the induced map

$$\lambda_c^e \otimes \mathrm{id}_E : R \otimes_R E \to F_*^eR \otimes_R E$$

is injective, and for this, in turn, we need only check that the socle generator is not in the kernel.

Recall that E is the direct limit of the injective maps

$$R/(x_1, \dots, x_n) \xrightarrow{x} R/(x_1^2, \dots, x_n^2) \xrightarrow{x} R/(x_1^3, \dots, x_n^3) \xrightarrow{x} \cdots$$

where $x_1, \dots, x_n$ is a minimal set of generators for m, and the maps are given by multiplication by $x = \Pi_{i=1}^d x_i$. So the module $F_*^eR \otimes_R E$ is the direct limit of the maps

$$R/(x_1^{p^e},\dots,x_n^{p^e}) \xrightarrow{x^{p^e}} R/(x_1^{2p^e},\dots,x_n^{2p^e}) \xrightarrow{x^{p^e}} R/(x_1^{3p^e},\dots,x_n^{3p^e}) \xrightarrow{x^{p^e}} \cdots$$

which remains injective by the faithful flatness of F_*^eR. The induced map $\lambda_c^e \otimes \mathrm{id}_E : E \to F_*^eR \otimes E$ sends the socle (namely the image of 1 in R/m) to the class of c in $R/m^{[p^e]}$, so it is nonzero provided $c \notin m^{[p^e]}$. Thus for every nonzero c in a regular local (Noetherian) ring, we have found an e, such that the map λ_c^e is pure. So regular local rings are F-pure regular. □

Proposition 6.2.2. *An F-pure regular ring is normal, that is, it is integrally closed in its total quotient ring.*

Proof. Take a fraction r/s in the total quotient ring integral over R. Then clearing denominators in an equation of integral dependence, we have $r \in \overline{(s)}$, the integral closure of the ideal (s). This implies that there exists an h such that $(r,s)^{n+h} = (s)^n(r,s)^h$ for all n [Matsumura 1989, p. 64]. Setting $c = s^h$, this implies $cr^n \in (s)^n$ for all large n. In particular, taking $n = p^e$, we see that the class of r modulo (s) is in the kernel of the map induced by tensoring the map

$$R \to F_*^eR, \quad 1 \mapsto c \tag{6.2.2.1}$$

with the quotient module $R/(s)$. By purity of the map (6.2.2.1), it follows that $r \in (s)$. We conclude that r/s is in R and that R is normal. □

6.3. ***Connections with tight closure.*** In his lecture notes on tight closure, Hochster [2007] suggests another way to generalize strong F-regularity to non-F-finite (but Noetherian) rings using tight closure. We show here that his generalized strong F-regularity is the same as F-pure regularity for local Noetherian rings.

Although Hochster and Huneke introduced tight closure only in Noetherian rings, we can make the same definition in general for an arbitrary ring of prime characteristic p. Let $N \hookrightarrow M$ be an inclusion of R-modules. The *tight closure* of N in M is an R-module N_M^* containing N. By definition, an element $x \in M$ is in N_M^* if there exists $c \in R$, not in any minimal prime, such that for all sufficiently large e, the element $c \otimes x \in F_*^eR \otimes_R M$ belongs to the image of the module $F_*^eR \otimes_R N$ under the natural map $F_*^eR \otimes_R N \to F_*^eR \otimes_R M$ induced by tensoring the inclusion $N \hookrightarrow M$ with the R-module F_R^e. We say that N is tightly closed in M if $N_M^* = N$.

Definition 6.3.1. Let R be a Noetherian ring of prime characteristic p. We say that R is *strongly F-regular* in the sense of Hochster if, for any inclusion of R modules $N \hookrightarrow M$, $N_M^* = N$.

The next result compares F-pure regularity with strong F-regularity in the sense of Hochster.

Proposition 6.3.2. *Let R be an arbitrary commutative ring of prime characteristic. If R is F-pure regular, then N is tightly closed in M for any pair of R modules with $N \subset M$. The converse also holds if R is Noetherian and local.*

Proof. Suppose $x \in N^*_M$. Equivalently the class $\bar{x}$ of x in M/N is in $0^*_{M/N}$. So there exists c not in any minimal prime such that $c \otimes \bar{x} = 0$ in $F^e_* R \otimes_R M/N$ for all large e. But this means that the map

$$R \to F^e_* R, \quad 1 \mapsto c$$

is not pure for any e, since the naturally induced map

$$R \otimes M/N \to F^e_* R \otimes M/N$$

has $1 \otimes \bar{x}$ in its kernel.

For the converse, let $c \in R$ be not in any minimal prime. We need to show that there exists some e such that the map $R \to F^e_* R$ sending 1 to c is pure. Let E be the injective hull of the residue field of R. According to Lemma 6.1.4(i), it suffices to show that there exists an e such that after tensoring E, the induced map

$$R \otimes E \to F^e_* R \otimes E$$

is injective. But if not, then a generator η for the socle of E is in the kernel for every e, that is, for all e, $c \otimes \eta = 0$ in $F^e_* R \otimes E$. In this case, $\eta \in 0^*_E$, contrary to our hypothesis that all modules are tightly closed. □

Remark 6.3.3. We do not know whether Proposition 6.3.2 holds in the nonlocal case. Indeed, we do not know if F-pure regularity is a local property: if R_m is F-pure regular for all maximal ideals m of R, does it follow that R is F-pure regular? If this were the case, then our argument above extends to arbitrary Noetherian rings.

Remark 6.3.4. A Noetherian ring of characteristic p is *weakly F-regular* if N is tightly closed in M for any pair of *Noetherian* R modules with $N \subset M$. Clearly F-pure regular implies weakly F-regular. The converse is a long standing open question in the F-finite Noetherian case. For valuation rings, however, our arguments show that weak and pure F-regularity are equivalent (and both are equivalent to the valuation ring being Noetherian); See Corollary 6.5.4.

6.4. ***Elements along which F-purity fails.*** We now observe an analog of the splitting prime of Aberbach and Enescu [2005]; See also [Tucker 2012, Lemma 4.7].

Proposition 6.4.1. *Let R be a ring of characteristic p. The set*

$$\mathcal{I} := \{c \in R \ : \ R \textit{ is not F-pure along } c\}.$$

is closed under multiplication by R, and $R \setminus \mathcal{I}$ is multiplicatively closed. Thus, if $\mathcal{I}$ is closed under addition, then $\mathcal{I}$ is a prime ideal (or the whole ring R).

Proof. We first note that $\mathcal{I}$ is closed under multiplication by elements of R. Indeed, suppose that $c \in \mathcal{I}$ and $r \in R$. Then if $rc \notin \mathcal{I}$, we have that R is F-pure along rc, but this implies R is F-pure along c by Proposition 6.1.3(a), contrary to $c \in \mathcal{I}$.

We next show that the complement $R \setminus \mathcal{I}$ is a multiplicatively closed set (if nonempty). To wit, take $c, d \notin \mathcal{I}$. Because R is F-pure along both c and d, we have that there exist e and f such that the maps

$$R \xrightarrow{\lambda_c^e} F_*^e R, \quad 1 \mapsto c, \quad \text{and} \quad R \xrightarrow{\lambda_d^f} F_*^f R, \quad 1 \mapsto d$$

are both pure. Since purity is preserved by restriction of scalars (Lemma 6.1.4(d)), we also have that

$$F_*^e R \xrightarrow{F_*^e(\lambda_d^f)} F_*^e F_*^f R = F_*^{e+f} R$$

is pure. Hence the composition

$$R \xrightarrow{\lambda_c^e} F_*^e R \xrightarrow{\lambda_d^f} F_*^e F_*^f R, \quad 1 \mapsto c^{p^e} d$$

is pure as well (Lemma 6.1.4(a)). This means that $c^{p^e} d$ is not in $\mathcal{I}$, and since $\mathcal{I}$ is closed under multiplication, neither is cd. Note also that if $R \setminus \mathcal{I}$ is nonempty, then $1 \in R \setminus \mathcal{I}$ by Proposition 6.1.3(a). Thus $R \setminus \mathcal{I}$ is a multiplicative set.

Finally, if $\mathcal{I}$ is closed under addition (and $\mathcal{I} \neq R$), we conclude that $\mathcal{I}$ is a prime ideal since it is an ideal whose complement is a multiplicative set. □

Remark 6.4.2. If R is a Noetherian local domain, then the set $\mathcal{I}$ of Proposition 6.4.1 can be checked to be closed under addition (see, for example, [Tucker 2012, Lemma 4.7] for the F-finite case). Likewise, for valuation rings, the set $\mathcal{I}$ is also an ideal: we construct it explicitly in the next section. However, for an arbitrary ring, $\mathcal{I}$ can fail to be an ideal. For example, under suitable hypothesis, the set $\mathcal{I}$ is also the union of the centers of F-purity in the sense of Schwede, hence in this case, $\mathcal{I}$ is a finite union of ideals but not necessarily an ideal in the nonlocal case; see [Schwede 2010].

6.5. *F-pure regularity and valuation rings.* In this subsection we characterize valuation rings that are F-pure regular, as summarized by the following main result.

Theorem 6.5.1. *A valuation ring is F-pure regular if and only if it is Noetherian. Equivalently, a valuation ring is F-pure regular if and only if it is a field or a DVR.*

A key ingredient in the proof is the following theorem about the set of elements along which V fails to be F-pure (see Definition 6.1.1).

Theorem 6.5.2. *The set of elements c along which a valuation ring (V, m) fails to be F-pure is the prime ideal*

$$\mathcal{Q} := \bigcap_{e>0} m^{[p^e]}.$$

Proof. First, take any $c \in \mathcal{Q}$. We need to show that V is not F-pure along c, that is, that the map

$$\lambda_c^e : V \to F_*^e V, \quad 1 \mapsto c$$

is not pure for any e. Because $c \in m^{[p^e]}$, we see that tensoring with $\kappa := V/m$ produces the zero map. So λ_c^e is not pure for any e, which means V is not F-pure along c.

For the other inclusion, let $c \notin m^{[p^e]}$ for some $e > 0$. We claim that $\lambda_c^e : V \to F_*^e V$ is pure. Apply Lemma 6.1.4(g) to the set Σ of finitely generated submodules of $F_*^e V$ which contain c. Note that Σ is a directed set under inclusion with a least element, namely the V-submodule of $F_*^e V$ generated by c, and $F_*^e V$ is the direct limit of the elements of Σ. It suffices to show that if $T \in \Sigma$, then

$$\lambda_T : V \to T, \quad 1 \mapsto c$$

is pure. But T is free since it is a finitely generated, torsion-free module over a valuation ring (Lemma 3.2). Since $c \notin m^{[p^e]}$, by the V module structure on T, we get $c \notin mT$. By Nakayama's lemma, we know c is part of a free basis for T. So λ_T splits, and is pure in particular.

Now that we know that the set of elements along which R is not F-pure is an ideal, it follows that it is a prime ideal from Proposition 6.4.1. □

Corollary 6.5.3. *For a valuation ring (V, m) of characteristic p, define*

$$\mathcal{Q} := \bigcap_{e>0} m^{[p^e]}.$$

Then the quotient $V/\mathcal{Q}$ is an F-pure regular valuation ring. Furthermore, V is F-pure regular if and only if $\mathcal{Q}$ is zero.

Proof. The second statement follows immediately from Theorem 6.5.2. For the first, observe that $V/\mathcal{Q}$ is a domain since $\mathcal{Q}$ is prime. So ideals of $V/\mathcal{Q}$ inherit the total ordering under inclusion from V, and $V/\mathcal{Q}$ is a valuation ring whose maximal ideal $\overline{m}$ satisfies $\bigcap_{e>0} \overline{m}^{[p^e]} = 0$. So $V/\mathcal{Q}$ is F-pure regular. □

Corollary 6.5.4. *For a valuation ring, F-pure regularity is equivalent to all ideals (equivalently, the maximal ideal) being tightly closed.*

Proof. Proposition 6.3.2 ensures that F-pure regularity implies all ideals are tightly closed. For the converse, note that if there is some nonzero c in $\bigcap_{e>0} m^{[p^e]}$, then $1 \in m^*$. So for any proper ideal m, the condition that $m^* = m$ implies that $\bigcap_{e>0} m^{[p^e]} = 0$. In particular, if the maximal ideal of a valuation ring V is tightly closed, then Corollary 6.5.3 implies that V is F-pure regular. □

Proof of Theorem 6.5.1. First observe that if V is a field or DVR, then it is F-pure regular. Indeed, every map of modules over a field is pure (since all vector space maps split). And a DVR is a one dimensional regular local ring, so it is F-pure regular by Theorem 6.2.1.

Conversely, we show that if (V, m) is F-pure regular, its dimension is at most one. Suppose (V, m) admits a nonzero prime ideal $P \neq m$. Choose $x \in m \setminus P$, and a nonzero element $c \in P$. The element c *cannot* divide x^n in V, since in that case we would have $x^n \subset (c) \subset P$, but P is a prime ideal not containing x. It then follows from the definition of a valuation ring that x^n divides c for all n. This means in particular that $c \in (x)^{[p^e]} \subset m^{[p^e]}$ for all e. So $c \in Q$. According to Theorem 6.5.2, R is not F-pure regular.

It remains to show that an F-pure regular valuation ring V of dimension one is discrete. Recall that the value group Γ of V is (order isomorphic to) an additive subgroup of $\mathbb{R}$ [Matsumura 1989, Theorem 10.7].

We claim that Γ has a least positive element. To see this, let η be the greatest lower bound of all positive elements in Γ. First observe that η is strictly positive. Indeed, for fixed $c \in m$, the sequence $v(c)/p^e$ consists of positive real numbers approaching zero as e gets large. If Γ contains elements of arbitrarily small positive values, then we could find $x \in V$ such that

$$0 < v(x) < \frac{v(c)}{p^e}.$$

But then $0 < v(x^{p^e}) < v(c)$, which says that $c \in (x)^{[p^e]} \subset m^{[p^e]}$ for all e. This contradicts our assumption that V is F-pure along c (again, using Theorem 6.5.2).

Now that we know the greatest lower bound η of Γ is positive, it remains to show that $\eta \in \Gamma$. Choose ϵ such that $0 < \epsilon < \eta$. If $\eta \notin \Gamma$, we know $\eta < v(y)$ for all $y \in m$. Since η is the greatest lower bound, we can find y such that

$$\eta < v(y) < \eta + \epsilon,$$

as well as x such that

$$\eta < v(x) < v(y) < \eta + \epsilon.$$

Then

$$0 < v(y/x) < \epsilon < \eta,$$

contradicting the fact that η is a lower bound for Γ. We conclude that $\eta \in \Gamma$, and that Γ has a least positive element.

It is now easy to see, using the Archimedean axiom for real numbers, that the ordered subgroup Γ of $\mathbb{R}$ is generated by its least positive element η. In particular, Γ is order isomorphic to $\mathbb{Z}$. We conclude that V is a DVR. $\square$

Remark 6.5.5. For a valuation ring (V, m) of dimension $n \geq 1$, our results show that in general $Q = \bigcap_{e \in \mathbb{N}} m^{[p^e]}$ is a prime ideal of height at least $n - 1$. It is easy to see that the situation where V/Q is a DVR arises if and only if m is principal, which in turn is equivalent to the value group Γ having a least positive element. For example, this is the case for the lex valuation in Example 4.4.2. It is not hard to check that $\mathcal{Q}$ is a uniformly F-compatible ideal in the sense of Schwede [2010] (see also [Smith and Zhang 2015, §3A] for further discussion of uniformly F-compatible ideals), generalizing of course to the non-Noetherian and non-F-finite setting. A general investigation of uniformly F-compatible ideals appears to be fruitful, and is being undertaken by the first author.

6.6. ***Split F-regularity.*** Of course, there is another obvious way[3] to adapt Hochster and Huneke's definition of strongly F-regular to arbitrary rings of prime characteristic p:

Definition 6.6.1. A ring R is *split F-regular* if for all nonzero divisors c, there exists e such that the map $R \to F_*^e R$ sending 1 to c splits as a map of R-modules.

Since split maps are pure, a split F-regular ring is F-pure regular. Split F-regular rings are also clearly Frobenius split. On the other hand, Example 4.5.1 shows that a discrete valuation ring need not be Frobenius split, so split F-regularity is strictly stronger than F-pure regularity. In particular, not every regular local ring is split F-regular, so split F-regularity should not really be considered a class of "singularities" even for Noetherian rings.

Remark 6.6.2. In Noetherian rings, split F-regularity is very close to F-pure regularity. For example, if R is an F-pure regular Noetherian domain whose fraction field is F-finite, then the only obstruction to split F-regularity is the splitting of Frobenius. This is a consequence of Lemma 4.2.4, which tells us R is F-finite if it is Frobenius split, and Lemma 2.5.1, which tells us F-split and F-pure are the same in F-finite Noetherian rings.

Corollary 6.6.3. *For a discrete valuation ring V whose fraction field is F-finite, the following are equivalent*:

(i) *V is split F-regular.*

(ii) *V is Frobenius split.*

(iii) *V is F-finite.*

(iv) *V is free over V^p.*

(v) *V is excellent.*

Moreover, if K is a function field over an F-finite ground field k, and V is a valuation of K/k, then (i)–(v) *are equivalent to V being a divisorial valuation ring.*

[3] This generalization is used for cluster algebras in [Benito et al. 2015] for example.

Proof. All this has been proved already. Recall that a DVR is a regular local ring, so it is always F-pure regular and hence split F-regular if it is F-finite. Also, the final statement follows from Theorem 5.1 because an Abyhankar valuation of rational rank one is necessarily divisorial, and a divisorial valuation of a functional field over an F-finite field is necessarily F-finite. □

To summarize: a valuation ring is F-pure regular if and only if it is Noetherian, and split F-regular (under the additional assumption that its fraction field is F-finite) if and only if it is excellent.

7. Concluding remarks

We have argued that for valuation rings, F-purity and F-pure regularity (a version of strong F-regularity defined using pure maps instead of split maps) are natural and robust properties. We have also seen that the conditions of Frobenius splitting and split F-regularity are more subtle, and that even regular rings can fail to satisfy these.

For Noetherian valuation rings in F-finite fields, we have seen that the Frobenius splitting property is equivalent to F-finiteness and also to excellence, but we do not know what happens in the non-Noetherian case: does there exist an example of a (necessarily non-Noetherian) Frobenius split valuation ring of an F-finite field that is *not* F-finite? By Corollary 5.2, a possible strategy could be to construct a Frobenius split valuation ring in a function field whose value group is infinitely generated. For example, can one construct an F-split valuation in $\mathbb{F}_p(x, y)$ with value group $\mathbb{Q}$? On the other hand, perhaps Frobenius splitting is equivalent to F-finiteness (just as in the Noetherian case). One might then ask whether a generalized version of Theorem 4.1.1 holds for arbitrary fields: is a valuation ring Frobenius split if and only if Frobenius is free?

We propose that F-pure regularity is a more natural generalization of strong F-regularity to the non-F-finite case than a suggested generalization of strong F-regularity using tight closure due to Hochster. We have seen that F-pure regularity implies Hochster's notion, and that they are equivalent for local Noetherian rings. However, we do not know whether F-pure regularity is a local notion: if R_m is F-pure regular for all maximal ideals, does it follow that R is F-pure regular? We expect this to be true, but the standard arguments are insufficient to prove it. (In the Noetherian F-finite case, this is well known; see [Hochster and Huneke 1994, Theorem 5.5(a)]. Furthermore, the answer is affirmative for excellent rings with F-finite total quotient rings, by Proposition 2.6.1.) If true, then F-pure regularity would be equivalent to all modules being tightly closed in the Noetherian case. More generally, might F-pure regularity be equivalent to the property that all modules are tightly closed even in the non-Noetherian case? Or even that all ideals are tightly closed? An affirmative answer to this last question would imply that strong and weak F-regularity are equivalent.

Acknowledgements

The authors gratefully acknowledge the conversation with Karl Schwede and Zsolt Patakfalvi that inspired our investigation of Frobenius in valuation rings, and especially Karl, with whom we later had many fruitful discussions. The authors also acknowledge helpful discussions with Mel Hochster, Daniel Hernandez, Mattias Jonsson, Eric Canton, Linquan Ma, and Juan Felipe Pérez, some of which occurred at the American Mathematical Society's 2015 Mathematics Research Community in commutative algebra, attended by the first author. In addition, Karl and Linquan made detailed comments on an earlier version, which greatly improved the paper. In particular, Linquan suggested Corollary 6.5.4. We also thank the referee for helpful suggestions.

References

[Aberbach and Enescu 2005] I. M. Aberbach and F. Enescu, "The structure of F-pure rings", *Math. Z.* **250**:4 (2005), 791–806. MR 2180375 Zbl 1102.13001

[Aberbach and Leuschke 2003] I. M. Aberbach and G. J. Leuschke, "The F-signature and strong F-regularity", *Math. Res. Lett.* **10**:1 (2003), 51–56. MR 1960123 Zbl 1070.13006

[Abhyankar 1956] S. Abhyankar, "On the valuations centered in a local domain", *Amer. J. Math.* **78** (1956), 321–348. MR 0082477 Zbl 0074.26301

[Benito et al. 2015] A. Benito, G. Muller, J. Rajchgot, and K. E. Smith, "Singularities of locally acyclic cluster algebras", *Algebra Number Theory* **9**:4 (2015), 913–936. MR 3352824

[Blickle 2008] M. Blickle, "Minimal γ-sheaves", *Algebra Number Theory* **2**:3 (2008), 347–368. MR 2407119 Zbl 1183.13005

[Fernández de Bobadilla and Pereira 2012] J. Fernández de Bobadilla and M. P. Pereira, "The Nash problem for surfaces", *Ann. of Math.* (2) **176**:3 (2012), 2003–2029. MR 2979864 Zbl 1264.14049

[Boucksom 2014] S. Boucksom, "Corps d'Okounkov: d'après Okounkov, Lazarsfeld–Mustaţă et Kaveh–Khovanskii", exposé no. 1059, vii, 1–41 in *Séminaire Bourbaki,* 2012/2013, Astérisque **361**, Société Mathématique de France, Paris, 2014. MR 3289276

[Bourbaki 1989] N. Bourbaki, *Commutative algebra: Chapters 1–7*, Springer, 1989. MR 979760 Zbl 0666.13001

[Brion and Kumar 2005] M. Brion and S. Kumar, *Frobenius splitting methods in geometry and representation theory*, Progress in Mathematics **231**, Birkhäuser, Boston, 2005. MR 2107324 Zbl 1072.14066

[Conrad 2008] B. Conrad, "Several approaches to non-Archimedean geometry", pp. 9–63 in *p-adic geometry* (10th Arizona Winter School, Tucson, AZ (2007)), edited by D. Savitt and D. S. Thakur, Univ. Lecture Ser. **45**, Amer. Math. Soc., Providence, RI, 2008. MR 2482345 Zbl 1153.14020

[Cutkosky 2004] S. D. Cutkosky, *Resolution of singularities*, Graduate Studies in Mathematics **63**, American Mathematical Society, Providence, RI, 2004. MR 2058431 Zbl 1076.14005

[EGA IV$_2$ 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR 0199181 Zbl 0135.39701

[Ein et al. 2003] L. Ein, R. Lazarsfeld, and K. E. Smith, "Uniform approximation of Abhyankar valuation ideals in smooth function fields", *Amer. J. Math.* **125**:2 (2003), 409–440. MR 1963690 Zbl 1033.14030

[Favre and Jonsson 2004] C. Favre and M. Jonsson, *The valuative tree*, Lecture Notes in Mathematics **1853**, Springer, Berlin, 2004. MR 2097722 Zbl 1064.14024

[Fomin and Zelevinsky 2002] S. Fomin and A. Zelevinsky, "Cluster algebras, I: Foundations", *J. Amer. Math. Soc.* **15**:2 (2002), 497–529. MR 1887642 Zbl 1021.16017

[Gongyo et al. 2015] Y. Gongyo, Z. Li, Z. Patakfalvi, K. Schwede, H. Tanaka, and R. Zong, "On rational connectedness of globally F-regular threefolds", *Adv. Math.* **280** (2015), 47–78. MR 3350212 Zbl 1329.14087

[Gubler et al. 2016] W. Gubler, J. Rabinoff, and A. Werner, "Skeletons and tropicalizations", *Adv. Math.* **294** (2016), 150–215. MR 3479562

[Hacon and Xu 2015] C. D. Hacon and C. Xu, "On the three dimensional minimal model program in positive characteristic", *J. Amer. Math. Soc.* **28**:3 (2015), 711–744. MR 3327534 Zbl 1326.14032

[Hara and Watanabe 2002] N. Hara and K.-I. Watanabe, "F-regular and F-pure rings vs. log terminal and log canonical singularities", *J. Algebraic Geom.* **11**:2 (2002), 363–392. MR 1874118 Zbl 1013.13004

[Hochster 2007] M. Hochster, "Foundations of tight closure theory", lecture notes, University of Michigan, 2007, Available at http://www.math.lsa.umich.edu/~hochster/711F07/fndtc.pdf.

[Hochster and Huneke 1989] M. Hochster and C. Huneke, "Tight closure and strong F-regularity", pp. 119–133 in *Colloque en l'honneur de Pierre Samuel* (Orsay, 1987), Mém. Soc. Math. Fr. **38**, 1989. MR 1044348 Zbl 0699.13003

[Hochster and Huneke 1990] M. Hochster and C. Huneke, "Tight closure, invariant theory, and the Briançon–Skoda theorem", *J. Amer. Math. Soc.* **3**:1 (1990), 31–116. MR 1017784

[Hochster and Huneke 1994] M. Hochster and C. Huneke, "F-regularity, test elements, and smooth base change", *Trans. Amer. Math. Soc.* **346**:1 (1994), 1–62. MR 1273534 Zbl 0844.13002

[Hochster and Huneke 1995] M. Hochster and C. Huneke, "Applications of the existence of big Cohen–Macaulay algebras", *Adv. Math.* **113**:1 (1995), 45–117. MR 1332808 Zbl 0834.13013

[Hochster and Roberts 1974] M. Hochster and J. L. Roberts, "Rings of invariants of reductive groups acting on regular rings are Cohen–Macaulay", *Advances in Math.* **13** (1974), 115–175. MR 0347810 Zbl 0289.14010

[Hochster and Roberts 1976] M. Hochster and J. L. Roberts, "The purity of the Frobenius and local cohomology", *Advances in Math.* **21**:2 (1976), 117–172. MR 0417172 Zbl 0348.13007

[Knaf and Kuhlmann 2005] H. Knaf and F.-V. Kuhlmann, "Abhyankar places admit local uniformization in any characteristic", *Ann. Sci. École Norm. Sup.* (4) **38**:6 (2005), 833–846. MR 2216832 Zbl 1159.13301

[Kunz 1969] E. Kunz, "Characterizations of regular local rings for characteristic p", *Amer. J. Math.* **91** (1969), 772–784. MR 0252389 Zbl 0188.33702

[Kunz 1976] E. Kunz, "On Noetherian rings of characteristic p", *Amer. J. Math.* **98**:4 (1976), 999–1013. MR 0432625 Zbl 0341.13009

[Lyubeznik and Smith 1999] G. Lyubeznik and K. E. Smith, "Strong and weak F-regularity are equivalent for graded rings", *Amer. J. Math.* **121**:6 (1999), 1279–1290. MR 1719806 Zbl 0970.13003

[Ma 1988] F. Ma, "Splitting in integral extensions, Cohen–Macaulay modules and algebras", *J. Algebra* **116**:1 (1988), 176–195. MR 944154 Zbl 0656.13006

[Matsumura 1989] H. Matsumura, *Commutative ring theory*, 2-nd ed., Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, 1989. MR 1011461 Zbl 0666.13002

[Mehta and Ramanathan 1985] V. B. Mehta and A. Ramanathan, "Frobenius splitting and cohomology vanishing for Schubert varieties", *Ann. of Math.* (2) **122**:1 (1985), 27–40. MR 799251 Zbl 0601.14043

[Patakfalvi 2014] Z. Patakfalvi, "Semi-positivity in positive characteristics", *Ann. Sci. Éc. Norm. Supér.* (4) **47**:5 (2014), 991–1025. MR 3294622 Zbl 1326.14015

[Ramanathan 1991] A. Ramanathan, "Frobenius splitting and Schubert varieties", pp. 497–508 in *Proceedings of the Hyderabad Conference on Algebraic Groups* (Hyderabad, 1989), edited by S. Ramanan et al., Manoj Prakashan, Madras, 1991. MR 1131324 Zbl 0785.14029

[Schwede 2009a] K. Schwede, "F-adjunction", *Algebra Number Theory* **3**:8 (2009), 907–950. MR 2587408 Zbl 1209.13013

[Schwede 2009b] K. Schwede, "F-injective singularities are Du Bois", *Amer. J. Math.* **131**:2 (2009), 445–473. MR 2503989 Zbl 1164.14001

[Schwede 2010] K. Schwede, "Centers of F-purity", *Math. Z.* **265**:3 (2010), 687–714. MR 2644316 Zbl 1213.13014

[Schwede and Smith 2010] K. Schwede and K. E. Smith, "Globally F-regular and log Fano varieties", *Adv. Math.* **224**:3 (2010), 863–894. MR 2628797 Zbl 1193.13004

[Schwede and Tucker 2012] K. Schwede and K. Tucker, "A survey of test ideals", pp. 39–99 in *Progress in commutative algebra, 2*, edited by C. Francisco et al., Walter de Gruyter, Berlin, 2012. MR 2932591 Zbl 1254.13007

[Smith 1997] K. E. Smith, "F-rational rings have rational singularities", *Amer. J. Math.* **119**:1 (1997), 159–180. MR 1428062 Zbl 0910.13004

[Smith 2000] K. E. Smith, "Globally F-regular varieties: applications to vanishing theorems for quotients of Fano varieties", *Michigan Math. J.* **48** (2000), 553–572. MR 1786505 Zbl 0994.14012

[Smith and Van den Bergh 1997] K. E. Smith and M. Van den Bergh, "Simplicity of rings of differential operators in prime characteristic", *Proc. London Math. Soc.* (3) **75**:1 (1997), 32–62. MR 1444312 Zbl 0948.16019

[Smith and Zhang 2015] K. E. Smith and W. Zhang, "Frobenius Splitting in Commutative Algebra", pp. 291–345 in *Commutative algebra and noncommutative algebraic geometry*, vol. 1: Expository articles, edited by D. Eisenbud et al., MSRI book series **67**, 2015.

[Takagi 2004] S. Takagi, "F-singularities of pairs and inversion of adjunction of arbitrary codimension", *Invent. Math.* **157**:1 (2004), 123–146. MR 2135186 Zbl 1121.13008

[Takagi 2008] S. Takagi, "A characteristic p analogue of plt singularities and adjoint ideals", *Math. Z.* **259**:2 (2008), 321–341. MR 2390084 Zbl 1143.13007

[Tucker 2012] K. Tucker, "F-signature exists", *Invent. Math.* **190**:3 (2012), 743–765. MR 2995185 Zbl 1271.13012

[Zariski and Samuel 1960] O. Zariski and P. Samuel, *Commutative algebra*, vol. II, van Nostrand, Princeton, 1960. MR 0120249 Zbl 0121.27801

Communicated by János Kollár
Received 2015-07-30 Accepted 2016-01-18

rankeya@umich.edu *Mathematics Department, University of Michigan, 2074 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, United States*

kesmith@umich.edu *Department of Mathematics, University of Michigan, 2074 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, United States*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.1091

Hoffmann's conjecture for totally singular forms of prime degree

Stephen Scully

One of the most significant discrete invariants of a quadratic form ϕ over a field k is its (*full*) *splitting pattern*, a finite sequence of integers which describes the possible isotropy behavior of ϕ under scalar extension to *arbitrary* overfields of k. A similarly important but more accessible variant of this notion is that of the *Knebusch splitting pattern* of ϕ, which captures the isotropy behavior of ϕ as one passes over a certain *prescribed tower* of k-overfields. We determine all possible values of this latter invariant in the case where ϕ is *totally singular*. This includes an extension of Karpenko's theorem (formerly Hoffmann's conjecture) on the possible values of the first Witt index to the totally singular case. Contrary to the existing approaches to this problem (in the nonsingular case), our results are achieved by means of a new structural result on the higher anisotropic kernels of totally singular quadratic forms. Moreover, the methods used here readily generalize to give analogous results for arbitrary Fermat-type forms of degree p over fields of characteristic $p > 0$.

1. Introduction

Let k be a field, let ϕ be a nonzero quadratic form on a (nonzero) k-vector space V of finite dimension and let $X_\phi \subseteq \mathbb{P}(V)$ denote the projective k-scheme defined by the vanishing of ϕ. Given a k-linear subspace W of V, we write $\phi|_W$ for the form obtained by restricting ϕ to W. In the case where $\phi|_W$ is the zero form, W is said to be *totally isotropic* (with respect to ϕ). The largest integer among the dimensions of all totally isotropic subspaces of V is called the *isotropy index* of ϕ, and is denoted by $\mathfrak{i}_0(\phi)$. In the special case where ϕ is *nonsingular* (i.e., where X_ϕ is smooth), $\mathfrak{i}_0(\phi)$ is more commonly known as the *Witt index* of ϕ, and is bounded from above by the integer part of $\frac{1}{2}\dim\phi$ (where $\dim\phi$ denotes the dimension of the k-vector space V). In the opposite extreme where ϕ is *totally singular* (i.e., where X_ϕ has no smooth points at all), $\mathfrak{i}_0(\phi)$ may take any value between 0 and $\dim\phi - 1$.

Assume now that ϕ is *anisotropic* (i.e., that $\mathfrak{i}_0(\phi) = 0$). A simple, yet fundamentally important invariant of ϕ is its (*full*) *splitting pattern*, which may be defined

MSC2010: primary 11E04; secondary 14E05, 15A03.
Keywords: quadratic forms, quasilinear p-forms, splitting patterns, canonical dimension.

as the increasing sequence of nonzero isotropy indices attained by ϕ under scalar extension to *every* overfield of k.[1] Although this sequence appears to be somewhat intractable in general, its first entry (assuming there is one) may be computed more explicitly as the isotropy index of the extension ϕ_L of ϕ to the function field $L = k(X_\phi)$ of the (integral) quadric X_ϕ. This (almost tautological) observation is the basic motivation underlying the following construction, originally due to Knebusch [1976]: let $k_0 = k$, $\phi_0 = \phi$, and inductively define $k_r = k_{r-1}(X_{\phi_{r-1}})$, $\phi_r = (\phi_{k_r})_{\mathrm{an}}$,[2] with the understanding that this (finite) process stops when we reach the first nonnegative integer $h(\phi)$ such that $\dim \phi_{h(\phi)} \leq 1$. The integer $h(\phi)$ and the tower of fields $k = k_0 \subset k_1 \subset \cdots \subset k_{h(\phi)}$ are known as the *height* and *Knebusch splitting tower* of ϕ, respectively. For $1 \leq r \leq h(\phi)$, the anisotropic form ϕ_r is called the *r-th higher anisotropic kernel* of ϕ. The *r-th higher isotropy index* of ϕ, denoted $\mathfrak{i}_r(\phi)$, is defined as the difference $\mathfrak{i}_0(\phi_{k_r}) - \mathfrak{i}_0(\phi_{k_{r-1}})$. The sequence $\mathfrak{i}(\phi) = (\mathfrak{i}_1(\phi), \ldots, \mathfrak{i}_{h(\phi)}(\phi))$ is called the *Knebusch splitting pattern* of ϕ.[3] Note that we have $\mathfrak{i}_r(\phi) = \mathfrak{i}_1(\phi_{r-1})$ for every $r \geq 2$ by the inductive nature of the construction.

If ϕ is nonsingular, then its full and Knebusch splitting patterns are easily seen to determine one another (see [Elman et al. 2008, Proposition 25.1]). This need not be the case for (totally) singular forms (see Example 2.47 below), but Knebusch's construction still offers a meaningful and practical way to preclassify quadratic forms according to some notion of "algebraic complexity". By virtue of its definition, the Knebusch splitting pattern thus embodies a fundamental link between intrinsic algebraic properties of quadratic forms and the geometry of the algebraic varieties which are naturally associated to them. In recent years, the advent of effective new tools with which to study algebraic cycles on projective homogeneous varieties has, in this way, led to dramatic progress on many long-standing problems within the algebraic theory of quadratic forms. The impact of these developments has been felt most deeply in characteristic $\neq 2$, where (1) anisotropic forms of dimension ≥ 2 are necessarily nonsingular, and (2) the geometric methods are better developed, even if we restrict our considerations to nonsingular forms only; see [Elman et al. 2008] for a thorough exposition of much of the recent work which has been done in this area.

One of the central problems in the investigation of splitting properties of quadratic forms over general fields is the following:

[1] This terminology is not standard; see, e.g., [Hoffmann and Laghribi 2004], where refined splitting pattern invariants are considered. Our definition does, however, agree (in content if not presentation) with those found in the literature when ϕ is nonsingular or totally singular (the only relevant cases here).

[2] For any form ψ over a field K, ψ_{an} denotes the *anisotropic kernel* of ψ, an anisotropic K-form uniquely determined up to isomorphism by the (refined) Witt decomposition of ψ; see [Hoffmann and Laghribi 2004, §2].

[3] The term *standard splitting pattern* is also used in the literature. Footnote 1 again applies here.

Question 1.1. Let ϕ be an anisotropic quadratic form of dimension ≥ 2 over a field. What are the possible values of the sequence $\mathfrak{i}(\phi)$?

Since $\mathfrak{i}_r(\phi) = \mathfrak{i}_1(\phi_{r-1})$ for all $2 \leq r \leq h(\phi)$, the natural first approximation to this problem is to determine the possible values of the invariant $\mathfrak{i}_1$ among all forms of a given dimension. To this end, we have the following general conjecture:

Conjecture 1.2 (Hoffmann[4]). Let ϕ be an anisotropic quadratic form of dimension ≥ 2 over a field. Then $\mathfrak{i}_1(\phi) - 1$ is the remainder modulo 2^s of $\dim\phi - 1$ for some $s < \log_2(\dim\phi)$.[5]

In characteristic $\neq 2$, the first major result in the direction of Conjecture 1.2 was established by Hoffmann himself [1995, Corollary 1], who showed that if $\dim\phi = 2^n + m$ for nonnegative integers n and $1 \leq m \leq 2^n$, then $\mathfrak{i}_1(\phi) \leq m$. A few years later, Izhboldin [2004, Corollary 5.12] proved that one cannot have $\mathfrak{i}_1(\phi) = m - 1$ here unless $m = 2$. Izhboldin's paper combined an elaboration of the algebraic methods conceived in [Hoffmann 1995] with emerging work of Vishik [1998; 1999], who developed a systematic approach to the study of the splitting pattern using the (integral) motive of the given quadric (see [Vishik 2004], where motivic methods were used to verify Hoffmann's conjecture in all dimensions ≤ 22 in this setting). Going further, Vishik later formulated a very general conjecture concerning the complete motivic decomposition of a smooth anisotropic quadric (the *excellent connections* conjecture) which subsumed the nonsingular case of Conjecture 1.2 in a conceptual way. Not long after this, Karpenko [2003] used a similar approach to prove the characteristic $\neq 2$ case of the conjecture in its entirety, the key new ingredient being the use of (reduced power) Steenrod operations on modulo-2 Chow groups. More recently, Vishik [2011, Theorem 1.3] proved the excellent connections conjecture in characteristic $\neq 2$, thus yielding another proof of Hoffmann's conjecture in this setting. Vishik's work also makes essential use of Steenrod squares in Chow theory.

In characteristic 2, the general picture is more complicated. In the *nonsingular* case, Karpenko and Vishik's approaches to Conjecture 1.2 are still valid, and progress is only hindered here by the fact that the total mod-2 Steenrod operation is not yet available when 2 is not invertible in the base field. To this end, weak forms of the first three Steenrod squares have been constructed by Haution [2013, Theorem 6.2; 2015, Theorem 5.8], and these suffice to prove nontrivial partial

[4]This conjecture was originally stated by Hoffmann under the additional hypothesis $\mathrm{char}(k) \neq 2$, but we expect that the assertion is also valid in characteristic 2.

[5]After passing to a purely transcendental extension of k if necessary, all values of $\mathfrak{i}_1(\phi)$ which are not excluded by the conjecture can be realized by making an appropriate choice of ϕ. In all cases, ϕ may, in fact, be chosen to be either nonsingular or (if $\mathrm{char}(k) = 2$) totally singular; see [Vishik 2004, §7.4] and Proposition 6.4 below, respectively.

results towards Conjecture 1.2. Haution's results are further supplemented by earlier work of Hoffmann and Laghribi [2006, Lemma 4.1], who extended Hoffmann's upper bound on $\mathfrak{i}_1(\phi)$ to the characteristic-2 setting, irrespective of whether ϕ is nonsingular or not. For *singular* forms, however, the situation is different, and almost nothing is known in the direction of Conjecture 1.2 beyond Hoffmann and Laghribi's bound. In fact, the only real exception to this general state of affairs lies in the extreme case where ϕ is *totally singular*. Here, it was recently shown in [Scully 2016, Theorem 9.4] that if $\dim\phi = 2^n + m$ for nonnegative integers n and $1 \le m \le 2^n$, and if $\mathfrak{i}_1(\phi) \ne m$ (that is, if Hoffmann and Laghribi's bound is *not* met), then $\mathfrak{i}_1(\phi) \le m/2$. In the present article, we will settle this case completely by proving:

Theorem 1.3. *Conjecture 1.2 is true in the case where ϕ is totally singular.*

Contrary to the existing approaches to the nonsingular case of Conjecture 1.2, our proof of Theorem 1.3 does not involve the study of Chow correspondences on the quadric X_ϕ. Indeed, although we also make use of the computation of the canonical dimension[6] of X_ϕ (see [Karpenko and Merkurjev 2003; Totaro 2008]), it is exploited here in a rather more direct and algebraic way. This point of view begins with the following observation:

Proposition 1.4 (see Proposition 4.3 below). *Let ϕ be an anisotropic totally singular quadratic form of dimension ≥ 2 over a field k of characteristic 2 and let $\psi \subset \phi$ be a subform of codimension $\mathfrak{i}_1(\phi)$. Suppose furthermore that $h(\psi) < h(\phi)$.[7] Then there exist a quasi-Pfister quadratic form π,[8] a subform $\sigma \subset \pi$, an element $\lambda \in k^*$ and a form τ over k such that $\psi \simeq \pi \otimes \tau$ and $\phi \simeq \psi \perp \lambda\sigma$.*

In the situation of Proposition 1.4, Hoffmann's conjecture is immediately verified. Indeed, (since $\sigma \subset \pi$) the integer $\dim\pi$ is a power of 2 strictly greater than $\mathfrak{i}_1(\phi) - 1 = \dim\sigma - 1$, and (since ψ is divisible by π) we have $\dim\phi - 1 = \dim\psi + \mathfrak{i}_1(\phi) - 1 \equiv \mathfrak{i}_1(\phi) - 1 \pmod{\dim\pi}$. It is not always possible, however, to decompose the form ϕ in the manner intimated by the proposition. In fact, Vishik (see [Totaro 2009, Lemma 7.1]) has given examples of 16-dimensional anisotropic quadratic forms in characteristic $\ne 2$ which have first higher isotropy index equal to 2, but which do not decompose in this way, and the same examples carry over into the totally singular setting (see Lemma 4.4 below). Thus, the picture is, in general, more complicated than that suggested by Proposition 1.4. Perhaps surprisingly, however, the main result of this paper shows that the next best thing happens:

[6]Here, the *canonical dimension* of an algebraic variety X over a field k should be understood as the minimal dimension $\mathrm{cdim}(X)$ of the image of a rational self-map $X \dashrightarrow X$.

[7]A weaker condition will suffice; see the statement of Proposition 4.3.

[8]That is, π is the diagonal part of a bilinear Pfister form over k; see Section 2C below.

Theorem 1.5. *Let ϕ be an anisotropic totally singular quadratic form of dimension ≥ 2 over a field of characteristic 2 and let s be the smallest nonnegative integer such that $2^s \geq \mathfrak{i}_1(\phi)$. Then ϕ_1 is divisible by an s-fold quasi-Pfister form.*

This is a new kind of statement of which no analogue is known in the nonsingular theory (even in characteristic $\neq 2$). As remarked above, a key ingredient needed for its proof is Totaro's computation [2008, Theorem 5.1] of the canonical dimension of a totally singular quadric. In [Scully 2013], this computation was extended to the wider class of Fermat-type hypersurfaces of degree p over fields of characteristic $p > 0$, and this enables us to also prove a direct analogue of Theorem 1.5 for totally singular forms of any prime degree $p > 2$ (known here as *quasilinear p-forms*); see Theorem 5.1 below. Subsequently, we also get an analogue of Theorem 1.3 in higher degrees. As a corollary, this yields a complete solution to the problem of determining the possible values of the canonical dimension of a degree-p Fermat-type hypersurface in characteristic $p > 0$ (Theorem 6.6); it is worth noting here that no such result is known for Fermat-type hypersurfaces of prime degree $p > 2$ over fields of characteristic *not* p. Returning to the case where $p = 2$, let us explain more precisely how Theorem 1.5 implies the totally singular case of Hoffmann's conjecture:

Proof of Theorem 1.3. Since ϕ is totally singular, we have $\dim \phi_1 = \dim \phi - \mathfrak{i}_1(\phi)$ (see Remarks 2.36(2) below). Thus, if s is as in Theorem 1.5, then

$$\dim \phi - 1 = \dim \phi_1 + \mathfrak{i}_1(\phi) - 1 \equiv \mathfrak{i}_1(\phi) - 1 \pmod{2^s}.$$

Since $\mathfrak{i}_1(\phi) - 1 < 2^s$, the result follows. □

Of course, our main result goes somewhat deeper than this. In fact, Theorem 1.5 (resp. Theorem 5.1 below) yields a complete answer to Question 1.1 in the totally singular case (resp. its analogue for arbitrary quasilinear p-forms). In other words, all restrictions on the possible values of the Knebusch splitting pattern of ϕ are explained here by the presence of certain divisibilities among its higher anisotropic kernels — precise statements are given in Section 6A below (see Theorem 6.1, Proposition 6.4). With this result in hand, it then becomes natural to try to understand the general discrepancy which exists between the Knebusch and *full* splitting patterns in the totally singular case. An immediate challenge here concerns the determination of all nontrivial restrictions which the former invariant imposes on the latter. In Section 7 below, we initiate this process by conjecturing that the gaps in the full splitting pattern established in characteristic $\neq 2$ by Vishik [2011, Proposition 2.6] (or see Theorem 7.1 below) as a consequence of his proof of the excellent connections conjecture are also present in the totally singular theory. A particular case of this conjecture was already proved in [Scully 2016, Theorem 9.2], and we provide some further evidence for its general veracity here.[9]

Finally, the main results of this paper should have valuable implications for the study of symmetric bilinear forms over fields of characteristic 2. Indeed, in characteristic 2, the diagonal parts of symmetric bilinear forms are nothing else but the totally singular quadratic forms discussed above. This will be investigated in a later text.

The remainder of this text is organized as follows. In Sections 2 and 3, we recall the basic theory of quasilinear p-forms and introduce the key notions and results which will be needed in the main part of the text. As a warm-up for the proof of our main result, we prove in Section 4 (a stronger version of) Proposition 1.4 and consider some situations in which it may be applied. The proof of Theorem 1.5 (and its generalization to higher degrees) is then given in Section 5, and, in Section 6, we apply this result to determine all possible Knebusch splitting patterns of quasilinear p-forms and settle another conjecture of Hoffmann concerning quasilinear p-forms with "maximal splitting". Lastly, in Section 7, we consider the aforementioned problem of establishing totally singular analogues of the results obtained in [Vishik 2011].

Notation and Terminology. Unless stated otherwise, p will denote an arbitrary prime integer and F will denote an arbitrary field of characteristic p. If L is a field of characteristic p and $a_1, \dots, a_n$ are elements of L, then $L_{a_1,\dots,a_n}$ will denote the field $L(\sqrt[p]{a_1}, \dots, \sqrt[p]{a_n})$. Finally, if k is a field and $T = (T_1, \dots, T_m)$ is a tuple of algebraically independent variables over k, then we will write $k[T]$ for the polynomial ring $k[T_1, \dots, T_n]$ and $k(T)$ for its fraction field.

2. Quasilinear p-forms and quasilinear p-hypersurfaces

The basic material presented in this section was originally developed in the series of papers [Hoffmann and Laghribi 2004; Laghribi 2004a; 2004b; 2006; Hoffmann 2004]. Additional elementary results which will be needed in the sequel are also included here. For any details which are omitted from our exposition of the basic theory, we refer the reader to [Hoffmann 2004].

2A. ***Basic notions.*** Let $\phi : V \to F$ be an F-valued form on a finite-dimensional F-vector space V. We say that ϕ is a *quasilinear p-form* (on V) if ϕ is homogeneous of degree p and the equation $\phi(v+w) = \phi(v)+\phi(w)$ holds for all $(v, w) \in V \times V$. By a *quasilinear p-form over F* (or sometimes simply a *form over F* or *F-form*), we will mean a quasilinear p-form on some finite-dimensional F-vector space. By a *quasilinear p-hypersurface over F* we will mean a projective hypersurface defined by the vanishing of a nonzero quasilinear p-form on some F-vector space

[9] To the author's knowledge, there is no conjectural description of the possible values of the full splitting pattern, even in the nonsingular case (in any characteristic).

of dimension ≥ 2. In the special case where $p = 2$, we will speak of *quasilinear quadratic forms* rather than quasilinear 2-forms.

Remark 2.1. Let $T = (T_1, \ldots, T_n)$ be a tuple of $n \geq 2$ algebraically independent variables over F, and let $f \in F[T] \setminus \{0\}$. Then the projective hypersurface $X_f = \{f = 0\} \subset \mathbb{P}^{n-1}$ is nowhere smooth if and only if $f \in F[T_1^p, \ldots, T_n^p]$. In particular, if $p = 2$, then a nonzero quadratic form of dimension ≥ 2 over F is totally singular (in the sense of Section 1) if and only if it is quasilinear.

Let ϕ be a quasilinear p-form over F. The underlying F-vector space of ϕ will be denoted by V_ϕ. Its dimension will be called the *dimension* of ϕ and will be denoted by $\dim \phi$. If $\dim \phi \geq 2$ and ϕ is nonzero, then the quasilinear p-hypersurface $\{\phi = 0\} \subset \mathbb{P}(V_\phi)$ (which is nowhere smooth by Remark 2.1) will be denoted by X_ϕ. The set $\{\phi(v) \mid v \in V_\phi\}$ of elements of F represented by ϕ will be denoted by $D(\phi)$. Given a field extension L of F, we will write ϕ_L for the unique quasilinear p-form on the L-vector space $V_\phi \otimes_F L$ such that $\phi_L(v \otimes 1) = \phi(v)$ for all $v \in V_\phi$. If R is a subring of L containing F, then $D(\phi_R)$ will denote the subset $\{\phi(w) \mid w \in V_\phi \otimes_F R\}$ of $D(\phi_L)$ (which lies in R). Given $a \in F$, we will write $a\phi$ for the form $v \mapsto a\phi(v)$ on the vector space V_ϕ.

Let ψ be another quasilinear p-form over F. If there exists an injective (resp. bijective) F-linear map $f : V_\psi \to V_\phi$ such that $\phi(f(v)) = \psi(v)$ for all $v \in V_\psi$, then we will say that ψ is a *subform* of (resp. is *isomorphic* to) ϕ and write $\psi \subset \phi$ (resp. $\psi \simeq \phi$). If $\psi \simeq a\phi$ for some $a \in F^*$, then we will say that ψ and ϕ are *similar*. The *sum* $\psi \oplus \phi$ (resp. *product* $\psi \otimes \phi$) is defined as the unique quasilinear p-form on $V_\psi \oplus V_\phi$ (resp. $V_\psi \otimes_F V_\phi$) such that $(\psi \oplus \phi)((v, w)) = \psi(v) + \phi(w)$ (resp. $(\psi \otimes \phi)(v \otimes w) = \psi(v)\phi(w)$) for all $(v, w) \in V_\psi \times V_\phi$. Given a positive integer n, we will let $n \cdot \phi$ denote the sum of n copies of ϕ (note that we have $n \cdot \phi \neq n\phi$ for $n > 1$). If there exists a form τ over F such that $\phi \simeq \psi \otimes \tau$, then we will say that ϕ is *divisible* by ψ.

Given elements $a_1, \ldots, a_n \in F$, we will write $\langle a_1, \ldots, a_n \rangle$ for the quasilinear p-form $(\lambda_1, \ldots, \lambda_n) \mapsto \sum_{i=1}^n a_i \lambda_i^p$ on the F-vector space $F^{\oplus n}$. By definition, every quasilinear p-form of dimension n over F is isomorphic to $\langle a_1, \ldots, a_n \rangle$ for some $a_i \in F$.

A vector $v \in V_\phi$ is said to be *isotropic* if $\phi(v) = 0$. We will say that ϕ is *isotropic* if V_ϕ contains a nonzero isotropic vector, and *anisotropic* otherwise. By the additivity of ϕ, the subset V_ϕ^0 of all isotropic vectors in V_ϕ is, in fact, an F-linear subspace of V_ϕ. Its dimension will be called the *isotropy index* of ϕ, and will be denoted by $\mathrm{i}_0(\phi)$ (note that in the case where $p = 2$ and ϕ is nonzero, this agrees with the definition given in Section 1). The additivity of ϕ also implies that $D(\phi)$ is a finite-dimensional F^p-linear subspace of F (where F is equipped with its natural F^p-vector space structure). Conversely, if U is a nonzero finite-dimensional

F^p-linear subspace of F, and if $a_1, \dots, a_n$ is a basis of U, then $\sigma = \langle a_1, \dots, a_n \rangle$ is a quasilinear p-form over F satisfying $D(\sigma) = U$. In fact, it is easy to see that, up to isomorphism, σ is the unique *anisotropic* quasilinear p-form with this property:

Lemma 2.2 (see [Hoffmann 2004, Proposition 2.12]). *Let U be a finite-dimensional F^p-linear subspace of F. Then, up to isomorphism, there exists a unique anisotropic quasilinear p-form ϕ over F such that $D(\phi) = U$.*

In particular, Pfister's quadratic subform theorem [Elman et al. 2008, Theorem 17.12] takes the following simplified form in this setting:

Proposition 2.3 (see [Hoffmann 2004, Proposition 2.6]). *Let ψ and ϕ be anisotropic quasilinear p-forms over F. Then $\psi \subset \phi$ if and only if $D(\psi) \subseteq D(\phi)$. In particular, $\psi \simeq \phi$ if and only if $D(\psi) = D(\phi)$.*

In view of these observations, we can define (up to isomorphism) the *anisotropic kernel* of ϕ as the unique anisotropic quasilinear p-form ϕ_{an} over F such that $D(\phi_{\mathrm{an}}) = D(\phi)$. If we view $D(\phi)$ as an F-vector space via the Frobenius $F \mapsto F^p$, then $\phi : V_\phi \to D(\phi)$ is a surjective F-linear map with kernel V_ϕ^0. We thus obtain:

Proposition 2.4 (see [Hoffmann 2004, Lemma 2.10]). *Let ϕ be a quasilinear p-form over F. Then ϕ is anisotropic if and only if $\phi \simeq \phi_{\mathrm{an}}$. If ϕ is isotropic, then $\phi \simeq \phi_{\mathrm{an}} \oplus \mathfrak{i}_0(\phi) \cdot \langle 0 \rangle$. In particular, $\dim \phi_{\mathrm{an}} = \dim \phi - \mathfrak{i}_0(\phi)$.*

In summary, we see that ϕ is determined up to isomorphism by the set $D(\phi)$ and the integer $\mathfrak{i}_0(\phi)$. If $\dim \phi_{\mathrm{an}} \leq 1$, then we will say that ϕ is *split*. Given another form ψ over F, we will write $\psi \sim \phi$ whenever $\psi_{\mathrm{an}} \simeq \phi_{\mathrm{an}}$. If F is perfect (that is, if $F = F^p$), then every form over F is split and the theory is vacuous. As such, we are essentially only interested in the case where F is imperfect. Unless indicated otherwise, we will assume henceforth that all quasilinear p-forms are *nonzero* (i.e., of anisotropic dimension ≥ 1).

Remark 2.5. Let ϕ be a quasilinear p-form over F. Choose a basis $v_1, \dots, v_n$ of V_ϕ and let $a_i = \phi(v_i)$ for all $1 \leq i \leq n$, so that $\phi \simeq \langle a_1, \dots, a_n \rangle$. Suppose that $v \in V_\phi$ is an isotropic vector, and write $v = \sum_{i=1}^n \lambda_i v_i$. If $\lambda_j \neq 0$ for $1 \leq j \leq n$, then the F^p-vector space $D(\phi)$ is spanned by the elements $a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n$. In particular, we have $\phi \sim \langle a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n \rangle$.

2B. ***Function fields of quasilinear p-hypersurfaces and their products.*** Let ϕ be a quasilinear p-form of dimension ≥ 2 over F. If ϕ is not split, then the quasilinear p-hypersurface X_ϕ is an integral scheme (see [Hoffmann 2004, Lemma 7.1]), as is its affine cone $\{\phi = 0\} \subset \mathbb{A}(V_\phi)$. In this case, we will write $F(\phi)$ for the function field of the former and $F[\phi]$ for that of the latter. If L is a field extension of F, then we will simply write $L(\phi)$ instead of $L(\phi_L)$ whenever it is defined. In general, given a finite collection $\phi_1, \dots, \phi_n$ of quasilinear p-forms of dimension ≥ 2 over F, we will write $F(\phi_1 \times \cdots \times \phi_n)$ for the function field of the scheme $X_{\phi_1} \times \cdots \times X_{\phi_n}$, provided that it is integral. This notation will be further simplified where possible; for example, if $\phi_1 = \cdots = \phi_n = \phi$, then we will simply write $F(\phi^{\times n})$ instead of $F(\phi_1 \times \cdots \times \phi_n)$.

Remarks 2.6. Let ϕ be a quasilinear p-form over F. Assume that ϕ is not split. We make the following basic observations:

(1) Let $a_0, \dots, a_n \in F$ be such that $\phi \simeq \langle a_0, \dots, a_n \rangle$ and $a_0, a_1 \neq 0$. Then we have F-isomorphisms

$$F(\phi) \simeq F(S)\Big(\sqrt[p]{a_1^{-1}\big(a_0 + a_2 S_2^p + \cdots + a_n S_n^p\big)}\Big)$$

and

$$F[\phi] \simeq \mathrm{Frac}\big(F[T]/\big(a_0 T_0^p + \cdots + a_n T_n^p\big)\big) \simeq F(U)\Big(\sqrt[p]{a_0^{-1}\big(a_1 U_1^p + \cdots + a_n U_n^p\big)}\Big),$$

where $S = (S_2, \dots, S_n)$, $T = (T_0, \dots, T_n)$ and $U = (U_1, \dots, U_n)$ are tuples of algebraically independent variables over F.

(2) $F[\phi]$ is F-isomorphic to a degree-1 purely transcendental extension of $F(\phi)$.

(3) $F(\phi)$ is F-isomorphic to a degree-$\mathrm{i}_0(\phi)$ purely transcendental extension of $F(\phi_{\mathrm{an}})$; see Proposition 2.4.

(4) The form $\phi_{F(\phi)}$ is evidently isotropic. Furthermore, if $a_1, \dots, a_n \in F$ are such that $\phi \simeq \langle a_1, \dots, a_n \rangle$, then consideration of the generic point in $X_\phi(F(\phi))$ shows that $\phi_{F(\phi)} \sim \langle a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n \rangle$ for every $1 \le i \le n$; see Remark 2.5.

2C. *Quasi-Pfister p-forms.* Let ϕ be a quasilinear p-form over F and let n be a positive integer. We say that ϕ is an *n-fold quasi-Pfister p-form* if there exist $a_1, \dots, a_n \in F$ such that $\phi \simeq \langle\langle a_1, \dots, a_n \rangle\rangle := \bigotimes_{i=1}^n \langle 1, a_i, a_i^2, \dots, a_i^{p-1} \rangle$. For convenience, we also say that ϕ is a *0-fold quasi-Pfister p-form* if $\phi \simeq \langle 1 \rangle$. Note, in particular, that if ϕ is an n-fold quasi-Pfister p-form for some $n \ge 0$, then we have $\dim \phi = p^n$. The basic observation concerning quasi-Pfister p-forms is found in the following proposition (which follows easily from Lemma 2.2):

Proposition 2.7 (see [Hoffmann 2004, Proposition 4.6]). *Let ϕ be a quasilinear p-form over F. Then ϕ_{an} is a quasi-Pfister p-form if and only if $D(\phi)$ is a subfield of F.*

Since the set of elements represented by an arbitrary quasi-Pfister p-form is, by definition, a subfield of the base field, we obtain:

Corollary 2.8 (see [Hoffmann 2004, Proposition 4.6]). *Let ϕ be a quasi-Pfister p-form over F. Then ϕ_{an} is a quasi-Pfister p-form. In particular, if ϕ is isotropic, then* $\dim \phi_{\mathrm{an}} = \frac{1}{p^k} \dim \phi$ *for some* $k \ge 1$.

Remark 2.9. More explicitly, let $\phi = \langle\langle a_1, \dots, a_n \rangle\rangle$ for some $n \geq 1$ and $a_i \in F$. Then $D(\phi) = F^p(a_1, \dots, a_n)$. Let m be such that $[F^p(a_1, \dots, a_n) : F^p] = p^m$. If $m = 0$ (i.e., $D(\phi) = F^p$), then $\phi_{\mathrm{an}} \simeq \langle 1 \rangle$. If $m \geq 1$, then $\phi_{\mathrm{an}} \simeq \langle\langle b_1, \dots, b_m \rangle\rangle$ for any m elements $b_1, \dots, b_m \in F$ such that $F^p(b_1, \dots, b_m) = F^p(a_1, \dots, a_n)$.

Quasi-Pfister p-forms have a central role to play in the general theory of quasilinear p-forms. As shown by Hoffmann [2004], these forms are distinguished here by the very same properties which distinguish the classical Pfister forms among nonsingular quadratic forms. For this reason, it will be useful to define the *divisibility index* of a given form ϕ, denoted $\mathfrak{d}_0(\phi)$, as the largest nonnegative integer s such that ϕ_{an} is divisible by an s-fold quasi-Pfister p-form. Clearly we have $\mathfrak{d}_0(\phi) \leq \log_p(\dim \phi_{\mathrm{an}})$, with equality holding if and only if ϕ_{an} is similar to a quasi-Pfister p-form. An alternative description of this invariant will be given in Corollary 2.19 below.

2D. *The norm form.* Let ϕ be a quasilinear p-form over F. The *norm field* of ϕ, denoted $N(\phi)$, is defined (see [Hoffmann 2004, Definition 4.1]) as the smallest subfield of F which contains all ratios of nonzero elements of $D(\phi)$. Note, in particular, that we have $N(a\phi) = N(\phi) = N(\phi_{\mathrm{an}})$ for all $a \in F^*$. In spite of its simple nature, this invariant has an important role to play in the whole theory. A more explicit description of the norm field may be given as follows:

Remark 2.10. If $a_1, \dots, a_n \in F$ are such that $\phi \simeq \langle a_1, \dots, a_n \rangle$ and $a_1 \neq 0$, then we have $N(\phi) = F^p\left(\frac{a_2}{a_1}, \dots, \frac{a_n}{a_1}\right)$.

In particular, we see that $N(\phi)$ is a nonzero finite-dimensional F^p-linear subspace of F. By Lemma 2.2, it follows that, up to isomorphism, there exists a unique anisotropic quasilinear p-form ϕ_{nor} over F such that $D(\phi_{\mathrm{nor}}) = N(\phi)$. The form ϕ_{nor} is called the *norm form* of ϕ (see [Hoffmann 2004, Definition 4.9]). By Proposition 2.7, ϕ_{nor} is a quasi-Pfister p-form. Its dimension (which is necessarily equal to a power of p) is called the *norm degree* of ϕ, and is denoted by $\mathrm{ndeg}(\phi)$ (see [Hoffmann 2004, Definition 4.1]). The following lemma characterizes the norm form as the smallest anisotropic quasi-Pfister p-form which contains ϕ_{an} as a subform up to multiplication by a scalar (again, this is a simple consequence of Proposition 2.3):

Lemma 2.11 (see [Scully 2016, Lemma 2.10]). *Let ϕ be a quasilinear p-form (resp. a quasilinear p-form such that $1 \in D(\phi)$) and π an anisotropic quasi-Pfister p-form over F. Then ϕ_{an} is similar to a subform of π (resp. $\phi_{\mathrm{an}} \subset \pi$) if and only if $\phi_{\mathrm{nor}} \subset \pi$. In particular, ϕ_{an} is similar to a subform of ϕ_{nor} (resp. $\phi_{\mathrm{an}} \subset \phi_{\mathrm{nor}}$).*

Example 2.12. Let ϕ be a quasilinear p-form over F. The following are equivalent:

(1) ϕ_{an} is similar (resp. isomorphic) to a quasi-Pfister p-form.

(2) $\phi_{\mathrm{nor}} \simeq a\phi_{\mathrm{an}}$ for some $a \in F^*$ (resp. $\phi_{\mathrm{nor}} \simeq \phi_{\mathrm{an}}$).

(3) $N(\phi) = aD(\phi)$ for some $a \in F^*$ (resp. $N(\phi) = D(\phi)$).

2E. *Similarity factors.* Let ϕ be a quasilinear p-form over F. By a *similarity factor* of ϕ, we mean an element $a \in F^*$ such that $a\phi \simeq \phi$. The set of all similarity factors of ϕ will be denoted by $G(\phi)^*$, and we will write $G(\phi)$ for the set $G(\phi)^* \cup \{0\}$. Note that $G(a\phi) = G(\phi) = G(\phi_{\mathrm{an}})$ for all $a \in F^*$, the second equality being an obvious consequence of Proposition 2.4. Thus, in view of Proposition 2.3, we have:

Lemma 2.13 (see [Hoffmann 2004, Lemma 6.3]). *Let ϕ be a quasilinear p-form over F and let $a \in F^*$. Then $a \in G(\phi)^*$ if and only if $aD(\phi) \subseteq D(\phi)$.*

Example 2.14. Let ϕ be a quasi-Pfister p-form over F. Then, since $D(\phi)$ is a subfield of F, we have $G(\phi) = D(\phi)$.

More generally, Lemma 2.13 immediately implies the following:

Corollary 2.15 (see [Hoffmann 2004, Proposition 6.4]). *Let ϕ be a quasilinear p-form over F. Then $G(\phi)$ is a subfield of $N(\phi)$ containing F^p.*

In particular, $G(\phi)$ is a nonzero finite-dimensional F^p-linear subspace of F. By Lemma 2.2, it follows that, up to isomorphism, there exists a unique anisotropic quasilinear p-form ϕ_{sim} over F such that $D(\phi_{\mathrm{sim}}) = G(\phi)$. The form ϕ_{sim} is called the *similarity form* of ϕ (see [Hoffmann 2004, Definition 6.5]). By Proposition 2.7, ϕ_{sim} is a quasi-Pfister p-form. Taken together, Examples 2.12 and 2.14 yield:

Example 2.16. Let ϕ be a quasilinear p-form over F. The following are equivalent:

(1) ϕ_{an} is similar (resp. isomorphic) to a quasi-Pfister p-form.

(2) $\phi_{\mathrm{sim}} \simeq \phi_{\mathrm{nor}} \simeq a\phi_{\mathrm{an}}$ for some $a \in F^*$ (resp. $\phi_{\mathrm{sim}} \simeq \phi_{\mathrm{nor}} \simeq \phi_{\mathrm{an}}$).

(3) $G(\phi) = N(\phi) = aD(\phi)$ for some $a \in F^*$ (resp. $G(\phi) = N(\phi) = D(\phi)$).

The basic observation concerning similarity factors is the following:

Proposition 2.17. *Let ϕ and ψ be quasilinear p-forms over F. Then $G(\psi) \subseteq G(\phi)$ if and only if ϕ_{an} is divisible by ψ_{sim}.*

Proof. We may assume that $1 \in D(\phi)$. Suppose $G(\psi) \subseteq G(\phi)$. By Corollary 2.15, $G(\psi)$ and $G(\phi)$ are subfields of F. By Lemma 2.13, $D(\phi)$ is naturally a (finite-dimensional) vector space over $G(\phi)$, and hence over $G(\psi)$. If $a_1, \dots, a_m$ is a basis of $D(\phi)$ over $G(\psi)$, then (since $D(\psi_{\mathrm{sim}}) = G(\psi)$), Lemma 2.2 implies that $\phi_{\mathrm{an}} \simeq \psi_{\mathrm{sim}} \otimes \langle a_1, \dots, a_m\rangle$. Conversely, if ϕ_{an} is divisible by ψ_{sim}, then it is clear that $G(\psi) \subseteq G(\phi)$, since $G(\psi) = D(\psi_{\mathrm{sim}}) = G(\psi_{\mathrm{sim}})$ by Example 2.16. □

We thus obtain the following characterization of the similarity form:

Corollary 2.18 (see [Hoffmann 2004, Proposition 6.4]). *Let ϕ be a quasilinear p-form and π an anisotropic quasi-Pfister p-form over F. Then ϕ_{an} is divisible by π if and only if ϕ_{sim} is divisible by π. In particular, ϕ_{an} is divisible by ϕ_{sim}.*

This enables us to reinterpret the divisibility index $\mathfrak{d}_0(\phi)$ (see Section 2C) as follows:

Corollary 2.19. *Let ϕ be a quasilinear p-form over F. Then we have $\mathfrak{d}_0(\phi) = \log_p(\dim \phi_{\mathrm{sim}}) = \log_p([G(\phi) : F^p])$.*

We also get the following:

Corollary 2.20. *Let ϕ and ψ be quasilinear p-forms over F. Then ϕ_{an} is divisible by ψ_{nor} if and only if $N(\psi) \subseteq G(\phi)$. If, additionally, $1 \in D(\psi)$, then the latter condition may be replaced by $D(\psi) \subseteq G(\phi)$.*

Proof. For the second statement, we simply recall that $N(\phi)$ is the smallest subfield of F containing all ratios of nonzero elements of $D(\phi)$ and that $G(\phi)$ is a subfield of F (Corollary 2.15). For the first, we can replace ψ by its norm form to arrive at the case where $N(\psi) = G(\psi)$ and $\psi_{\mathrm{nor}} \simeq \psi_{\mathrm{sim}}$ (see Example 2.16). The result is therefore a particular case of Proposition 2.17. □

2F. ***A criterion for a quasilinear p-form to be quasi-Pfister.*** Let ϕ be a quasilinear p-form over F such that $1 \in D(\phi)$, let L be a field extension of F and let $\alpha \in D(\phi_L) \setminus \{0\}$. Consider the set $S_\alpha = \{a \in F \mid \alpha a \in D(\phi_L)\}$. Since $D(\phi_L)$ is an L^p-linear subspace of L, we have the following observation:

$$\sum \lambda_i^p a_i \in S_\alpha \quad \text{for all } \lambda_i \in F \text{ and all } a_i \in S_\alpha. \tag{2-1}$$

Lemma 2.21. *In the above situation, let $P \in F^p[T]$ be a polynomial of degree $< p$ in a single variable T such that $P(b) \in S_\alpha$ for all $b \in D(\phi)$. Then $b^n \in S_\alpha$ for all $b \in D(\phi)$ and all $n \leq \deg(P)$.*

Proof. We proceed by induction on $d = \deg(P)$. Since $\alpha \in D(\phi_L)$, the case where $d = 0$ is trivial. Suppose now that $d > 0$, and let $\lambda \in F$ be such that $P(T+\lambda^p) = P(T) + Q(T)$ for some $Q \in F^p[T]$ of degree $d-1$. Since $F^p \subseteq D(\phi)$ by hypothesis, our assumption and (2-1) imply that $Q(b) = P(b+\lambda^p) - P(b) \in S_\alpha$ for all $b \in D(\phi)$. By the induction hypothesis, it follows that $b^n \in S_\alpha$ for all $b \in D(\phi)$ and all $n < d$. Finally, since $P(b) = \sum_{i=0}^{d} \lambda_i^p b^i$ for some $\lambda_i \in F$ with $\lambda_d \neq 0$, (2-1) implies that, for any $b \in D(\phi)$, we also have $b^d \in S_\alpha$. This proves the lemma. □

Suppose now that there exists a polynomial $P \in F^p[T]$ as in the statement of Lemma 2.21 with $\deg(P) \geq 2$ (in particular, we necessarily have $p > 2$). A first application of the lemma shows that we have $D(\phi) \subseteq S_\alpha$. Since $D(\phi_L)$ is spanned by $D(\phi)$ as an L^p-vector space, this implies that $\alpha D(\phi_L) \subseteq D(\phi_L)$, and hence (for dimension reasons) that $\alpha D(\phi_L) = D(\phi_L)$. Another application of Lemma 2.21

then shows that $b^n \in D(\phi_L)$ for all $b \in D(\phi)$ and all $n \leq \deg(P)$. In particular, since $\deg(P) \geq 2$, we have $2bc = (b+c)^2 - b^2 - c^2 \in D(\phi_L)$ for all $b, c \in D(\phi)$. Since $p > 2$, and since $D(\phi_L)$ is spanned by $D(\phi)$ as an L^p-vector space, this implies that $D(\phi_L)$ is closed under multiplication, i.e., that $D(\phi_L) = N(\phi_L)$ (see Remark 2.10). By Example 2.12, this means that $(\phi_L)_{\mathrm{an}} \simeq (\phi_L)_{\mathrm{nor}}$. We have thus proved:

Lemma 2.22. *Assume that $p > 2$. Let ϕ be a quasilinear p-form over F such that $1 \in D(\phi)$ and let L be a field extension of F. Suppose that there exists a polynomial $P \in F^p[T]$ in a single variable T, and an element $\alpha \in D(\phi_L) \setminus \{0\}$ such that $2 \leq \deg(P) < p$ and $\alpha P(b) \in D(\phi_L)$ for all $b \in D(\phi)$. Then $(\phi_L)_{\mathrm{an}}$ is a quasi-Pfister p-form.*

2G. *The Cassels–Pfister representation theorem.* Let ϕ be a quadratic form over a field k and let $f \in k[T]$ be a polynomial in a single variable T which is represented by the form $\phi_{k(T)}$. One of the foundational results of the classical algebraic theory of quadratic forms is the Cassels–Pfister representation theorem, which asserts that, in this case, ϕ already represents f over the polynomial ring $k[T]$ (see [Elman et al. 2008, Theorem 17.3]). In the present setting, the original argument of Cassels may be readily adapted to prove the analogous statement for quasilinear p-forms. However, as pointed out by Hoffmann [2004], the additivity property of these forms enables one to prove a stronger multivariable statement taking the following form:

Theorem 2.23 (see [Hoffmann 2004, Corollary 3.4]). *Let ϕ be a quasilinear p-form over F, let $T = (T_1, \ldots, T_m)$ be a tuple of algebraically independent variables over F and let $f \in F[T]$. Then $f \in D(\phi_{F(T)})$ if and only if $f \in D(\phi_{F[T]})$, if and only if $f \in D(\phi)[T_1^p, \ldots, T_m^p]$.*

Now, in the situation of Theorem 2.23, the $F(T)^p$-vector space $D(\phi_{F(T)})$ is (evidently) spanned by elements of $D(\phi)$. Thus, in view of Lemma 2.13, we immediately obtain the following result concerning rational similarity factors:

Corollary 2.24 (see [Hoffmann 2004, Proposition 6.7]). *Let ϕ be a quasilinear p-form over F, let $T = (T_1, \ldots, T_m)$ be a tuple of algebraically independent variables over F and let $f \in F[T]$. Then $f \in G(\phi_{F(T)})$ if and only if $f \in G(\phi)[T_1^p, \ldots, T_m^p]$.*

2H. *Isotropy of quasilinear p-forms under scalar extension.* We now collect some basic facts regarding the isotropy of quasilinear p-forms under scalar extension.

Let K and L be extensions of a field k. Recall that a *k-place* $K \dashrightarrow L$ is a pair (R, f) consisting of a valuation subring $k \subseteq R \subseteq K$ and a local k-algebra homomorphism $f : R \to L$. For example, given an inclusion $i : K \hookrightarrow L$, the pair (K, i) defines a k-place $K \dashrightarrow L$. If there exist k-places $K \dashrightarrow L$ and $L \dashrightarrow K$, then we say that K and L are *equivalent over k*, and write $K \sim_k L$. For instance, this is easily seen to be the case whenever L (resp. K) is a purely transcendental extension of K (resp. L) (see [Elman et al. 2008, §103] for further details). We have here the following basic lemma, which is a consequence of the completeness of X_ϕ (see [EGA II 1961, (7.3.8)]):

Lemma 2.25 (see [Scully 2016, Lemma 3.4]). *Let ϕ be a quasilinear p-form over F and let K and L be field extensions of F such that there exists an F-place $K \dashrightarrow L$. Then $\mathfrak{i}_0(\phi_L) \geq \mathfrak{i}_0(\phi_K)$. In particular, if $K \sim_F L$, then $\mathfrak{i}_0(\phi_K) = \mathfrak{i}_0(\phi_L)$.*

Note, in particular, that passage to rational extensions of the base field does not affect the isotropy index of a quasilinear p-form. By MacLane's theorem [Lang 2002, Proposition VIII.4], the same is, in fact, true of arbitrary *separable* extensions:[10]

Lemma 2.26 (see [Hoffmann 2004, Proposition 5.3]). *Let ϕ be an anisotropic quasilinear p-form over F and let L be a field extension of F. If L is separable over F, then ϕ_L is anisotropic and* $\mathrm{ndeg}(\phi_L) = \mathrm{ndeg}(\phi)$.

Thus, in order to study the isotropy behavior of quasilinear p-forms under scalar extension, we are effectively reduced to considering the case of purely inseparable algebraic extensions. In degree p, we have the following basic observations, all of which can be easily verified using the results which have been discussed thus far (recall here that, given $a_1, \ldots, a_n \in F$, we denote by $F_{a_1,\ldots,a_n}$ the field $F(\sqrt[p]{a_1}, \ldots, \sqrt[p]{a_n})$):

Lemma 2.27 (see [Hoffmann 2004, §5; Scully 2016, Lemma 3.8]). *Let ϕ be a quasilinear p-form over F and let $a \in F \setminus F^p$. Then*:

(1) $D(\phi_{F_a}) = D(\langle\langle a \rangle\rangle \otimes \phi) = \sum_{i=0}^{p-1} a^i D(\phi)$.

(2) $\mathfrak{i}_0(\phi_{F_a}) = \frac{1}{p}\mathfrak{i}_0(\langle\langle a \rangle\rangle \otimes \phi)$.

(3) $\mathrm{ndeg}(\phi_{F_a}) = \begin{cases} \frac{1}{p}\,\mathrm{ndeg}(\phi) & \text{if } a \in N(\phi), \\ \mathrm{ndeg}(\phi) & \text{if } a \notin N(\phi). \end{cases}$

(4) *If ϕ is anisotropic and $a \notin N(\phi)$, then ϕ_{F_a} is anisotropic.*

(5) $\dim\,(\phi_{F_a})_{\mathrm{an}} \geq \frac{1}{p} \dim \phi_{\mathrm{an}}$.

(6) *Equality holds in* (5) *if and only if ϕ_{an} is divisible by $\langle\langle a \rangle\rangle$, if and only if $a \in G(\phi)$.*

Remark 2.28. The second equivalence in (6) holds by Corollary 2.20.

As an application of the first part of the lemma, we have:

Corollary 2.29. *Let ϕ be a quasilinear p-form over F and let $a \in F \setminus F^p$. Then $G(\phi_{F_a}) = G(\langle\langle a \rangle\rangle \otimes \phi)$. In particular, $\mathfrak{d}_0(\langle\langle a \rangle\rangle \otimes \phi) = \mathfrak{d}_0(\phi_{F_a}) + 1$.*

[10] Recall that an extension of fields $k \subseteq L$ is called *separable* if, for any algebraic closure $\bar{k}$ of k, the ring $L \otimes_k \bar{k}$ has no nontrivial nilpotent elements.

Proof. More specifically, the first statement is an immediate consequence of Lemmas 2.27(1) and 2.13. The second then follows from Corollary 2.19. □

Suppose now that $\pi = \langle\langle a_1, \dots, a_n\rangle\rangle$ is an *anisotropic* quasi-Pfister p-form over F. By Remark 2.9, we have $[F^p(a_1, \dots, a_n) : F^p] = p^n$, which means that $a_i \notin F_{a_1,\dots,a_{i-1}}$ for every $1 \le i \le n$. Repeated applications of Lemma 2.27(2) and Corollary 2.29 therefore yield the following proposition:

Proposition 2.30. *Let ϕ be a quasilinear p-form over F, let π be as above and let $\psi = \pi \otimes \phi$. Then $\mathfrak{i}_0(\psi) = p^n \mathfrak{i}_0(\phi_{F_{a_1,\dots,a_n}})$ and $\mathfrak{d}_0(\psi) = \mathfrak{d}_0(\phi_{F_{a_1,\dots,a_n}}) + n$.*

Now, in view of Remarks 2.6(1), one may combine the above results in order to study the isotropy behavior of quasilinear p-forms under scalar extension to function fields of quasilinear p-hypersurfaces. More specifically, let ψ be a quasilinear p-form over F which is not split, and let $a_0, \dots, a_n \in F$ be such that $\psi \simeq \langle a_0, \dots, a_n\rangle$, with $a_0, a_1 \neq 0$. Then, by Remarks 2.6(1), we have an F-isomorphism of fields

$$F(\psi) \simeq F(T)\Big(\sqrt[p]{a_1^{-1}\big(a_0 + a_2T_2^p + \cdots + a_nT_n^p\big)}\Big),$$

where $T = (T_2, \dots, T_n)$ is an $(n-1)$-tuple of algebraically independent variables over F. Thus, putting Lemmas 2.26 and 2.27 and together, we obtain:

Lemma 2.31 (see [Hoffmann 2004, §§7.3, 7.4]). *Let ϕ be an anisotropic quasilinear p-form over F, and let ψ be as above. Then*:

(1) $\dim(\phi_{F(\psi)})_{\mathrm{an}} \ge \frac{1}{p}\dim\phi$.

(2) *Equality holds in* (1) *if and only if* $a_1^{-1}\big(a_0 + a_2T_2^p + \cdots + a_nT_n^p\big) \in G(\phi_{F(T)})$.

(3) $\mathrm{ndeg}(\phi_{F(\psi)}) \ge \frac{1}{p}\,\mathrm{ndeg}(\phi)$.

(4) *Equality holds in* (3) *if and only if* $a_1^{-1}\big(a_0 + a_2T_2^p + \cdots + a_nT_n^p\big) \in N(\phi_{F(T)})$.

(5) *The equivalent conditions of* (4) *are satisfied if $\phi_{F(\psi)}$ is isotropic.*

As a basic application, we have:

Corollary 2.32 (see [Hoffmann 2004, §§7.3, 7.4]). *Let ϕ and ψ be quasilinear p-forms over F such that ϕ is anisotropic and ψ is not split. Then*:

(1) $\dim(\phi_{F(\psi)})_{\mathrm{an}} \ge \frac{1}{p}\dim\phi$, *with equality holding if and only if $N(\psi) \subseteq G(\phi)$.*

(2) *If $\phi_{F(\psi)}$ is isotropic, then $N(\psi) \subseteq N(\phi)$. In particular*, $\mathrm{ndeg}(\psi) \le \mathrm{ndeg}(\phi)$.

Proof. We may assume that ψ is as in Lemma 2.31. In this case, we have $N(\psi) = F^p\big(\frac{a_0}{a_1}, \dots, \frac{a_n}{a_1}\big)$ (see Remark 2.9), and so (1) follows from the first two parts of the former lemma and Corollary 2.24. Similarly, since $N(\phi_L) = D((\phi_{\mathrm{nor}})_L)$ for any field extension L of F, (2) follows from Lemma 2.31(4,5) and Theorem 2.23. □

Finally, it will be useful to record in this section another basic application of the Cassels–Pfister theorem. To state it, let $T = (T_1, \dots, T_m)$ be a tuple of algebraically independent variables over F, let $g \in F[T]$ be an irreducible polynomial and let $F[g]$ denote the field $\mathrm{Frac}(F[T]/(g))$ (i.e., the function field of the integral hypersurface $\{g = 0\} \subset \mathbb{A}^m_F$). Given $f \in F[T]$, we write $\mathrm{mult}_g(f)$ for the multiplicity of g in f, i.e., the largest nonnegative integer s such that $f = g^s h$ for some $h \in F[T]$.[11]

Proposition 2.33. *In the above situation, let ϕ be a quasilinear p-form over F and let $f \in F[T]$. Suppose that $f \in D(\phi_{F(T)})$ and that $\phi_{F[g]}$ is anisotropic. Then* $\mathrm{mult}_g(f) \equiv 0 \pmod{p}$.

Proof. Let $s = \mathrm{mult}_g(f)$. After replacing f by $f/g^{kp} \in D(\phi_{F(T)})$ for a suitable integer $k \geq 0$, we may assume that $s < p$. Our goal is then to prove that $s = 0$. To see this, note first that there exists a $v \in V_\phi \otimes_F F[T]$ such that $\phi_{F(T)}(v) = f$ by Theorem 2.23. If $s \neq 0$, then the image $\bar{v}$ of v in $V_\phi \otimes_F F[g]$ is an isotropic vector for $\phi_{F[g]}$. By hypothesis, it follows that $\bar{v} = 0$, which means that $v = gw$ for some $w \in V_\phi \otimes_F F[T]$. But this implies that $f = g^p \phi_{F(T)}(w)$, which contradicts the fact that $s < p$. We conclude that $s = 0$, and so the proposition is proved. □

2I. *The divisibility index and scalar extension.* Let ϕ be a quasilinear p-form over F. We make some brief remarks concerning the behavior of the divisibility index $\mathfrak{d}_0(\phi)$ (see Section 2C) under scalar extension.

Lemma 2.34. *Let ϕ be a quasilinear p-form over F and let L be a field extension of F. If $(\phi_{\mathrm{sim}})_L$ is anisotropic, then $\mathfrak{d}_0(\phi_L) \geq \mathfrak{d}_0(\phi)$.*

Proof. As an L^p-vector space, $D((\phi_{\mathrm{sim}})_L)$ is spanned by $D(\phi_{\mathrm{sim}}) = G(\phi)$. Since we evidently have $G(\phi) \subseteq G(\phi_L) = D((\phi_L)_{\mathrm{sim}})$, and since $(\phi_{\mathrm{sim}})_L$ is anisotropic by hypothesis, Proposition 2.3 implies that $(\phi_{\mathrm{sim}})_L \subset (\phi_L)_{\mathrm{sim}}$. The desired assertion now follows from Corollary 2.19. □

In particular, this applies in the case where L is a separable extension of F (see Lemma 2.26). In the case where L is purely transcendental over F, we can say more:

Lemma 2.35. *Let ϕ be a quasilinear p-form over F and let L be a purely transcendental extension of F. Then $(\phi_L)_{\mathrm{sim}} \simeq (\phi_{\mathrm{sim}})_L$ and $\mathfrak{d}_0(\phi_L) = \mathfrak{d}_0(\phi)$.*

Proof. Continuing with the proof of Lemma 2.34, it is sufficient to show that in this case $G(\phi_L)$ is generated by $G(\phi)$ over L^p. If L is finitely generated over F, then this follows from Corollary 2.24. On the other hand, the general case reduces easily to the finitely generated case in view of Lemma 2.13, so the lemma is proved. □

[11] With the added convention that $\mathrm{mult}_g(0) = +\infty$.

2J. ***The Knebusch splitting pattern.*** Let ϕ be a quasilinear p-form over F. Following the construction outlined in Section 1 (see also [Hoffmann 2004, §7.5]), set $F_0 = F,\ \phi_0 = \phi_{\mathrm{an}}$, and recursively define

- $F_r = F_{r-1}(\phi_{r-1})$ (provided ϕ_{r-1} is not split), and
- $\phi_r = (\phi_{F_r})_{\mathrm{an}}$ (provided F_r is defined).

Note here that if ϕ_r is defined, then we have $\dim\phi_r < \dim\phi_{r-1}$ by Remarks 2.6(4). As such, the whole process is finite, terminating at the first nonnegative integer $h(\phi)$ for which $\dim\phi_{h(\phi)} \leq 1$. The integer $h(\phi)$ will be called the *height* of ϕ, and the tower of fields $F_0 \subset F_1 \subset \cdots \subset F_{h(\phi)}$ will be called the *Knebusch splitting tower* of ϕ. For each $0 \leq r \leq h(\phi)$, we set $\mathfrak{j}_r(\phi) = \mathfrak{i}_0(\phi_{F_r})$. If ϕ is not split and $r \geq 1$, then the difference $\mathfrak{j}_r(\phi) - \mathfrak{j}_{r-1}(\phi)$ will be called the *r-th higher isotropy index* of ϕ, and will be denoted by $\mathfrak{i}_r(\phi)$. In this case, the form ϕ_r will be called the *r-th higher anisotropic kernel* of ϕ. Finally, the sequence $\mathfrak{i}(\phi) = (\mathfrak{i}_1(\phi), \dots, \mathfrak{i}_{h(\phi)}(\phi))$ (understood to be empty if ϕ is split) will be called the *Knebusch splitting pattern* of ϕ.[12]

Remarks 2.36. Let ϕ be a quasilinear p-form over F.

(1) By the recursive nature of the above construction, we have $\mathfrak{i}_r(\phi) = \mathfrak{i}_1(\phi_{r-1})$ for every $1 \leq r \leq h(\phi)$.

(2) By Proposition 2.4, we have $\mathfrak{i}_r(\phi) = \dim\phi_{r-1} - \dim\phi_r$ for all $1 \leq r \leq h(\phi)$.

(3) Let L be a field extension of F. As already remarked in Section 1, it is not true in general that $\mathfrak{i}_0(\phi_L) = \mathfrak{j}_r(\phi)$ for some $0 \leq r \leq h(\phi)$; see Example 2.47 below.

Note that by Remarks 2.6(3) we have the following:

Lemma 2.37. *Let ϕ be a quasilinear p-form form of dimension ≥ 2 and let (F_r) denote its Knebusch splitting tower. Then $F_r \sim_F F(\phi^{\times r})$ for every $0 \leq r \leq h(\phi)$.*

In light of Lemma 2.37, we therefore have:

Corollary 2.38. *Let ϕ be a quasilinear p-form over F. Then, for every $0 \leq r \leq h(\phi)$, we have $\mathfrak{j}_r(\phi) = \mathfrak{i}_0\big(\phi_{F(\phi^{\times r})}\big)$.*

Given the results of Section 2H, we are now in a position to prove the following characterization of anisotropic quasi-Pfister p-forms:

Proposition 2.39 (see [Hoffmann and Laghribi 2004, Theorem 8.11]). *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F. Then $\dim\phi_1 \geq \frac{1}{p}\dim\phi$, and the following conditions are equivalent*:

(1) $\dim\phi_1 = \frac{1}{p}\dim\phi$.

(2) $\mathfrak{i}(\phi) = \big(p^{h(\phi)} - p^{h(\phi)-1}, p^{h(\phi)-1} - p^{h(\phi)-2}, \dots, p^2 - p, p - 1\big)$.

[12] See Footnote 3. We omit the term $\mathfrak{i}_0(\phi)$ from the sequence because we are ultimately interested in the case where ϕ is anisotropic (i.e., where $\mathfrak{i}_0(\phi) = 0$).

(3) *ϕ is similar to a quasi-Pfister p-form.*

Proof. The inequality $\dim \phi_1 \geq \frac{1}{p} \dim \phi$ holds by Corollary 2.32. The same result shows that equality holds if and only if $N(\phi) \subseteq G(\phi)$. By Corollary 2.20, the latter condition holds if and only if ϕ is divisible by ϕ_{nor}. In view of Lemma 2.11, this proves the equivalence of (1) and (3), as well as the implication (2) $\Rightarrow$ (3). On the other hand, if ϕ is similar to a quasi-Pfister p-form of dimension p^n, then ϕ_1 is similar to a quasi-Pfister p-form of dimension p^{n-1} by Corollary 2.8 and (1). Since $\mathfrak{i}_1(\phi) = \dim \phi - \dim \phi_1$ (see Remarks 2.36(2)), an easy induction on $h(\phi)$ then shows that (3) implies (2). □

Finally, a repeated application of Lemma 2.31(3–5) (with $\psi = \phi$) yields the following computation of the height $h(\phi)$:

Corollary 2.40 (see [Hoffmann 2004, Theorem 7.25(ii)]). *Let ϕ be a quasilinear p-form over F. Then $h(\phi) = \log_p(\mathrm{ndeg}(\phi))$.*

Together with Corollary 2.32(2), this implies the following useful result:

Corollary 2.41 (see [Scully 2016, Proposition 4.12]). *Let ϕ and ψ be quasilinear p-forms over F such that ϕ is anisotropic and ψ is not split. If $\phi_{F(\psi)}$ is isotropic, then $h(\psi) \leq h(\phi)$.*

2K. ***The quasi-Pfister height and higher divisibility indices.*** Let ϕ be a quasilinear p-form over F. As in [Scully 2016, §4.2], we define the *quasi-Pfister height* of ϕ, denoted $h_{\mathrm{qp}}(\phi)$, to be the smallest nonnegative integer l such that ϕ_l is similar to a quasi-Pfister p-form (this is well defined, since $\phi_{h(\phi)}$, being of dimension 1, is similar to a 0-fold quasi-Pfister p-form). We have:

Lemma 2.42. *Let ϕ be a quasilinear p-form over F and let $d = h(\phi) - h_{\mathrm{qp}}(\phi)$. Then $\mathfrak{i}(\phi) = \big(\mathfrak{i}_1(\phi), \dots, \mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi), p^d - p^{d-1}, p^{d-1} - p^{d-2}, \dots, p^2 - p, p - 1\big)$ and $\mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi) = \dim(\phi) - \mathfrak{j}_{h_{\mathrm{qp}}(\phi)-1}(\phi) - p^d < p^{d+1} - p^d$.*

Proof. The first statement is an immediate consequence of Proposition 2.39. The point here is that $\phi_{h_{\mathrm{qp}}(\phi)}$ is similar to a quasi-Pfister p-form of dimension p^d. By Remarks 2.36(2), we therefore have $\mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi) = \dim \phi_{h_{\mathrm{qp}}(\phi)-1} - \dim \phi_{h_{\mathrm{qp}}(\phi)} = \dim(\phi) - \mathfrak{j}_{h_{\mathrm{qp}}(\phi)-1}(\phi) - p^d$. Finally, since $\phi_{h_{\mathrm{qp}}(\phi)-1}$ is (by the definition of $h_{\mathrm{qp}}(\phi)$) not similar to a quasi-Pfister p-form, it must have dimension $< p^{d+1}$, again by Proposition 2.39. This proves the inequality in the second statement, and hence the lemma. □

The Knebusch splitting pattern of a quasilinear p-form ϕ is therefore determined by $h(\phi)$, $h_{\mathrm{qp}}(\phi)$ and the truncated sequence $\big(\mathfrak{i}_1(\phi), \dots, \mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi)\big)$. With a view to studying the latter invariant, we now introduce new invariants of ϕ which will be of central interest in the sequel. More specifically, for each $1 \leq r \leq h(\phi)$, we define the *r-th higher divisibility index* of ϕ, denoted $\mathfrak{d}_r(\phi)$, as the integer $\mathfrak{d}_0(\phi_r)$ (see

Section 2C). In other words, $\mathfrak{d}_r(\phi)$ is the largest integer s such that ϕ_r is divisible by an s-fold quasi-Pfister p-form (over the corresponding field of the Knebusch splitting tower of ϕ). The sequence of integers $(\mathfrak{d}_0(\phi), \dots, \mathfrak{d}_{h(\phi)}(\phi))$ will be denoted by $\mathfrak{d}(\phi)$. As per Lemma 2.42 (and the proof of Proposition 2.39), we have:

Lemma 2.43. *Let ϕ be a quasilinear p-form over F and let $d = h(\phi) - h_{\mathrm{qp}}(\phi)$. Then $\mathfrak{d}(\phi) = \big(\mathfrak{d}_0(\phi), \dots, \mathfrak{d}_{h_{\mathrm{qp}}(\phi)-1}(\phi), d, d-1, \dots, 1, 0\big)$.*

We also, however, have the following information concerning the "nontrivial part" of the sequence $\mathfrak{d}(\phi)$:

Lemma 2.44. *Let ϕ be a quasilinear p-form over F. Then $\mathfrak{d}_0(\phi) \leq \cdots \leq \mathfrak{d}_{h_{\mathrm{qp}}(\phi)}$.*

Proof. We may assume that ϕ is anisotropic and not split. We need to show that if ϕ is not similar to a quasi-Pfister p-form, then $\mathfrak{d}_1(\phi) \geq \mathfrak{d}_0(\phi)$. By Lemma 2.34 it will be sufficient to check that ϕ_{sim} remains anisotropic over $F(\phi)$. Suppose otherwise. Then, by Corollary 2.32(1), we have $N(\phi) \subseteq G(\phi)$. By Corollary 2.15 it follows that $N(\phi) = G(\phi)$, or, equivalently, that $\phi_{\mathrm{nor}} \simeq \phi_{\mathrm{sim}}$ (see Lemma 2.2). But, in view of Lemma 2.11 and Corollary 2.18, this implies that ϕ is similar to a quasi-Pfister p-form, thus contradicting our assumption. The lemma follows. $\square$

In particular, since $\mathfrak{i}_r(\phi) = \dim \phi_r - \dim \phi_{r-1}$ for all $1 \leq r \leq h(\phi)$ (see Remarks 2.36(2)), we obtain the following result concerning the integers $\mathfrak{i}_r(\phi)$:

Corollary 2.45. *Let ϕ be a quasilinear p-form over F. Then we have $\mathfrak{i}_r(\phi) \equiv 0 \pmod{\mathfrak{d}_{r-1}(\phi)}$ for all $1 \leq r \leq h_{\mathrm{qp}}(\phi)$.*

2L. *Some examples.* We now conclude this section with two basic computations which will be needed in the sequel, beginning with:

Lemma 2.46. *Let ϕ be an anisotropic quasilinear p-form over F and let $\psi = \phi_{F(T)} \perp \langle T \rangle$, where T is an algebraically independent variable over F. Then $\mathfrak{i}(\psi) = (1, \mathfrak{i}_1(\phi), \mathfrak{i}_2(\phi), \dots, \mathfrak{i}_{h(\phi)}(\phi))$ and $\mathfrak{d}(\psi) = (0, \mathfrak{d}_0(\phi), \mathfrak{d}_1(\phi), \dots, \mathfrak{d}_{h(\phi)}(\phi))$.*

Proof. The form ψ is clearly anisotropic. Now, the field $F(T)(\psi)$ is F-isomorphic to a purely transcendental extension of F (see the presentation of Remarks 2.6(1), for example). In particular, $\phi_{F(T)(\psi)}$ is anisotropic (Lemma 2.26), and so $\mathfrak{i}_1(\psi) = 1$ and $\psi_1 \simeq \phi_{F(T)(\psi)}$ (see Remark 2.5). Since $F(T)(\psi)$ is purely transcendental over F, the first statement now follows immediately from Lemma 2.26. In a similar way, Lemma 2.35 implies that $\mathfrak{d}_r(\psi) = \mathfrak{d}_{r-1}(\phi)$ for all $1 \leq r \leq h(\psi)$. Thus, to prove the second statement, it only remains to check that $\mathfrak{d}_0(\psi) = 0$. But, since $\mathfrak{i}_1(\psi) = 1$, this is an immediate consequence of Corollary 2.45. $\square$

Given this result, we can give an example of a quasilinear p-form whose (full) splitting pattern is *not* determined by its Knebusch splitting pattern:

Example 2.47 (see [Hoffmann and Laghribi 2004, Example 8.15]). Let $T = (T_1, \dots, T_{n+1})$ be a tuple of algebraically independent variables over a field F_0 of characteristic p, and let $F = F_0(T)$. Consider the form $\phi = \langle\langle T_1, \dots, T_n\rangle\rangle \perp \langle T_{n+1}\rangle$ over F. By Lemma 2.46 and Proposition 2.39, we have

$$\mathfrak{i}(\phi) = \bigl(1, p^n - p^{n-1}, p^{n-1} - p^{n-2}, \dots, p^2 - p, p - 1\bigr),$$

so that $\mathfrak{j}_r(\phi) = p^n - p^{n-r+1} + 1$ for all $1 \le r \le n+1$. On the other hand, the full splitting pattern of ϕ also contains all the integers $p^n - p^{n-r+1}$ $(1 \le r \le n+1)$. Indeed, if (L_s) denotes the Knebusch splitting tower of $\langle\langle a_1, \dots, a_n\rangle\rangle_L$, then we clearly have $\mathfrak{i}_0(\phi_{L_{r-1}}) = p^n - p^{n-r+1}$ for all $1 \le r \le n+1$ (again, we are using Lemma 2.26 and Proposition 2.39 here).

Our second computation is the following:

Lemma 2.48. *Let ϕ be a quasilinear p-form over F. Let $\psi = \langle\langle T_1, \dots, T_n\rangle\rangle \otimes \phi_{F(T)}$, where $T = (T_1, \dots, T_n)$ is a tuple of algebraically independent variables over F. Then $\mathfrak{i}(\psi) = \bigl(p^n \mathfrak{i}_1(\phi), \dots, p^n \mathfrak{i}_{h(\phi)}(\phi), p^n - p^{n-1}, p^{n-1} - p^{n-2}, \dots, p^2 - p, p-1\bigr)$ and $\mathfrak{d}(\psi) = \bigl(\mathfrak{d}_0(\phi) + n, \mathfrak{d}_1(\phi) + n, \dots, \mathfrak{d}_{h(\phi)-1}(\phi) + n, n, n-1, \dots, 1, 0\bigr)$.*

Proof. It is enough to treat the case where $n = 1$. To simplify the notation, let us write T for the variable T_1 and L for the rational function field $F(T)$. Now, by construction, we have $\mathrm{ndeg}(\psi) = p(\mathrm{ndeg}(\phi))$ (see Remark 2.10). In view of Corollary 2.40, it follows that $h(\psi) = h(\phi) + 1$. Let (L_r) and (F_r) denote the Knebusch splitting towers of ψ and ϕ respectively. We claim that, for every $0 \le r \le h(\phi)$, $(L_r)_T$ is F-isomorphic to a purely transcendental extension of F_r. The case where $r = 0$ is evident. In general, we have $(L_r)_T = (L_T)_r$, where $((L_T)_r)$ denotes the Knebusch splitting tower of ψ_{L_T}. But, since $\langle\langle T\rangle\rangle_{L_T} \sim \langle 1\rangle$, and since L_T is purely transcendental over F (the $r = 0$ case), we have $(\psi_{L_T})_{\mathrm{an}} \simeq \phi_{L_T}$. By Remark 2.5, it follows that $(L_r)_T$ is L-isomorphic to a purely transcendental extension of the free composite $F_r \cdot L_T$. Again, since L_T is purely transcendental over F, the claim follows. Given this, Proposition 2.30 and Lemma 2.26 together imply that

$$\mathfrak{j}_r(\psi) = \mathfrak{i}_0(\psi_{L_r}) = p\mathfrak{i}_0(\phi_{(L_r)_T}) = p\mathfrak{i}_0(\phi_{F_r}) = p\mathfrak{j}_r(\phi)$$

for all $0 \le r \le h(\phi)$, which proves the first statement of the lemma. Similarly, our claim, Proposition 2.30 and Lemma 2.35 together imply that

$$\mathfrak{d}_r(\psi) = \mathfrak{d}_0(\psi_{L_r}) = \mathfrak{d}_0(\phi_{(L_r)_T}) + 1 = \mathfrak{d}_0(\phi_{F_r}) + 1 = \mathfrak{d}_r(\phi) + 1$$

for all $0 \le r \le h(\phi)$, and so the second statement also holds. □

3. An incompressibility theorem and related results

In this section, we collect some of the farther-reaching results on the isotropy behavior of quasilinear p-forms over function fields of quasilinear p-hypersurfaces which have been obtained in recent years. These results will have an essential role to play in the sequel. We do not provide full details here, but the interested reader is referred to the original articles [Hoffmann and Laghribi 2004; Totaro 2008; Scully 2016] for further information.

3A. ***The incompressibility theorem.*** Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F with associated quasilinear p-hypersurface X_ϕ. As in [Scully 2013, §5], we define the *Izhboldin dimension* of X_ϕ, denoted $\dim_{\mathrm{Izh}}(X_\phi)$, to be the integer $\dim X_\phi - \mathfrak{i}_1(\phi) + 1$. The following result was proved in [loc. cit.]:

Theorem 3.1 [Scully 2013, Theorem 5.12]. *Let X be an anisotropic quasilinear p-hypersurface over F. Let Y be an algebraic variety over F such that $Y(F_{\mathrm{sep}}) = \varnothing$. If $\dim Y < \dim_{\mathrm{Izh}}(X)$, then there cannot exist a rational map $X \dashrightarrow Y$.*

Remarks 3.2. (1) In the case where $p = 2$, Theorem 3.1 is due to Totaro [2008, Theorem 5.1]. In fact, if $X_\phi = \{\phi = 0\}$ is an anisotropic projective quadric over a field k of any characteristic, and if Y is any complete k-variety possessing no closed points of odd degree, then it is known that the existence of a rational map $X \dashrightarrow Y$ necessarily implies that $\dim Y \geq \dim_{\mathrm{Izh}}(X_\phi)$, where $\dim_{\mathrm{Izh}}(X_\phi) = \dim X_\phi - \mathfrak{i}_1(\phi) + 1$. This result was first proved by Karpenko and Merkurjev [2003, Theorem 4.1] (see also [Elman et al. 2008, Theorem 76.5]) in the case where X_ϕ is smooth, and was later extended by Totaro [loc. cit.] to the singular case.

(2) The special case where Y is a closed subvariety of X shows that the canonical dimension of X (see Section 1) is equal to $\dim_{\mathrm{Izh}}(X)$ (the inequality $\mathrm{cdim}(X) \leq \dim_{\mathrm{Izh}}(X)$ is trivial; see [Scully 2013, Corollary 5.14]).

In the remainder of this section, we will recall some of the main applications of Theorem 3.1 (and its proof). Here, we mention the following:

Corollary 3.3 (see [Scully 2016, Corollary 5.4]). *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F, and let $\sigma \subset \phi$ be a subform of dimension $\leq \dim \psi - \mathfrak{i}_1(\psi)$. Then $\sigma_{F(\psi)} \subset (\phi_{F(\psi)})_{\mathrm{an}}$. In particular, $\sigma_{F(\psi)}$ is anisotropic.*

Proof. We trivially have $D(\sigma_{F(\psi)}) \subseteq D(\phi_{F(\psi)})$. In light of Proposition 2.3, it therefore suffices to check that $\sigma_{F(\psi)}$ is anisotropic, or, equivalently, that there does not exist a rational map $X_\psi \dashrightarrow X_\sigma$. But, since $X_\sigma(F_{\mathrm{sep}}) = \varnothing$ (see Lemma 2.26), this is an immediate consequence of Theorem 3.1. □

In particular, we have the following fundamental observation:

Corollary 3.4. *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and let $\psi \subset \phi$ be a subform of codimension $\mathfrak{i}_1(\phi)$. Then $\phi_1 \simeq \psi_{F(\phi)}$.*

Proof. By Corollary 3.3, we have $\psi_{F(\phi)} \subset \phi_1$. Since both forms have the same dimension by hypothesis, the result follows. □

3B. ***Neighbors and near neighbors, I.*** Let ψ and ϕ be anisotropic quasilinear p-forms of dimension ≥ 2 over F. We will say that ψ is a *neighbor* (resp. *near neighbor*) of ϕ if ψ is similar to a subform of codimension $< \mathfrak{i}_1(\phi)$ (resp. codimension $\mathfrak{i}_1(\phi)$) of ϕ. Our motivation here is the following extension of Corollary 3.4:

Lemma 3.5. *Let ϕ and ψ be anisotropic quasilinear p-forms over F such that ϕ is anisotropic of dimension ≥ 2 and ψ is similar to a subform of ϕ. Then $(\psi_{F(\phi)})_{\mathrm{an}}$ is similar to ϕ_1 if and only if ψ is a neighbor or near neighbor of ϕ.*

Proof. Without loss of generality, we may assume that $\psi \subset \phi$. Again, we trivially have $D(\psi_{F(\phi)}) \subseteq D(\phi_{F(\phi)}) = D(\phi_1)$. By Proposition 2.3, it therefore suffices to check that $\dim(\psi_{F(\phi)})_{\mathrm{an}} = \dim\phi - \mathfrak{i}_1(\phi)$ if and only if ψ has codimension $\leq \mathfrak{i}_1(\phi)$ in ϕ. The left-to-right implication here is trivial. Conversely, if ψ has codimension $\leq \mathfrak{i}_1(\phi)$ in ϕ, then we have $\dim(\psi_{F(\phi)})_{\mathrm{an}} \geq \dim\phi - \mathfrak{i}_1(\phi)$ by Theorem 3.1. Since the reverse inequality holds here by obvious dimension reasons, the lemma is proved. $\square$

Note here that while neighbors of ϕ become anisotropic over $F(\phi)$, its *near* neighbors do not (Corollary 3.3). This enables us to compute:

Proposition 3.6 (see [Scully 2013, Proposition 6.1]). *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F such that ψ is a codimension-d neighbor of ϕ. Then $\mathfrak{i}_1(\psi) = \mathfrak{i}_1(\phi) - d$.*

3C. ***The ruledness theorem.*** Another key application of Theorem 3.1 is the following extension of Proposition 3.6, which shows in a precise way that anisotropic quasilinear p-hypersurfaces having first higher isotropy index larger than 1 are ruled.

Theorem 3.7 (see [Scully 2013, Theorem 7.6]). *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F such that ψ is a codimension-d neighbor of ϕ. Then X_ϕ is birationally isomorphic to $X_\psi \times_F \mathbb{P}^d$.*

Proof. By Proposition 3.6, we have $\mathfrak{i}_1(\psi) = \mathfrak{i}_1(\phi) - d$. It is therefore enough to treat the case where $d = \mathfrak{i}_1(\psi) - 1$, and this is covered by [Scully 2013, Theorem 7.6]. $\square$

Remark 3.8. Again, in the case where $p = 2$, this result is due to Totaro [2008, Theorem 6.4]. Unlike Theorem 3.1, however, the analogous assertion remains open for generically smooth quadrics (in any characteristic; see [Totaro 2008; 2009]).

It is worth mentioning the following explicitly:

Corollary 3.9. *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F such that $\mathfrak{i}_1(\phi) > 1$ and ψ is similar to a codimension-1 subform of ϕ. Then we have an F-isomorphism of fields $F(\phi) \simeq F[\psi]$.*

Proof. By Theorem 3.7, $F(\phi)$ is F-isomorphic to a degree-one purely transcendental extension of $F(\psi)$. In view of Remarks 2.6(3), the result follows. $\square$

3D. ***Neighbors and near neighbors, II.*** Given Theorem 3.7, we extend Proposition 3.6 as follows:

Proposition 3.10 (see [Scully 2016, Proposition 6.2]). *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F such that ψ is a codimension-d neighbor of ϕ. Then we have $\mathfrak{i}(\psi) = (\mathfrak{i}_1(\phi) - d, \mathfrak{i}_2(\phi), \dots, \mathfrak{i}_{h(\phi)}(\phi))$ and $\mathfrak{d}(\psi) = (\mathfrak{d}_0(\psi), \mathfrak{d}_1(\phi), \dots, \mathfrak{d}_{h(\phi)}(\phi))$.*

Proof. By Theorem 3.7, $F(\phi)$ is F-isomorphic to a purely transcendental extension of $F(\psi)$. Thus, if (F_r) and $(F(\phi)_r)$ denote the Knebusch splitting towers of ψ and $\psi_{F(\phi)}$, respectively, then $F(\phi)_r$ is F_r-isomorphic to a purely transcendental extension of F_{r+1} for every $0 \leq r \leq h(\psi_{F(\phi)})$. In particular, we have $\mathfrak{i}_r(\psi_{F(\phi)}) = \mathfrak{i}_{r+1}(\psi)$ and $\mathfrak{d}_r(\psi_{F(\phi)}) = \mathfrak{d}_{r+1}(\psi)$ for any such r by Lemmas 2.26 and 2.35, respectively. On the other hand, Lemma 3.5 shows that $(\psi_{F(\phi)})_{\mathrm{an}}$ is similar to ϕ_1. Since $\mathfrak{i}_r(\phi) = \mathfrak{i}_{r-1}(\phi_1)$ and $\mathfrak{d}_r(\phi) = \mathfrak{d}_{r-1}(\phi_1)$ for all $1 \leq r \leq h(\phi)$, the proposition follows immediately. □

Let ϕ be a quasilinear p-form over F. We say that ϕ is a *quasi-Pfister p-neighbor* (of π) if there exists a quasi-Pfister p-form π over F such that ϕ is similar to a subform of π and $\dim \phi > \frac{1}{p} \dim \pi$. If ϕ is anisotropic, then it follows from Proposition 2.39 that ϕ is a quasi-Pfister p-neighbor if and only if it is a neighbor of some quasi-Pfister p-form in the sense of Section 3B. Forms of this type are of special importance in the general theory of quasilinear p-forms. Putting Lemma 2.11, Proposition 2.39, Corollary 2.40, Lemma 3.5 and Proposition 3.10 together, we obtain the following classification of anisotropic quasi-Pfister p-neighbors:

Corollary 3.11 (see [Scully 2016, Theorem 6.4]). *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and let n be the smallest nonnegative integer such that $p^{n+1} \geq \dim \phi$. Then the following are equivalent*:

(1) *ϕ is a quasi-Pfister p-neighbor.*

(2) *ϕ is a neighbor of ϕ_{nor}.*

(3) *$\phi_{F(\phi_{\mathrm{nor}})}$ is isotropic.*

(4) $\mathrm{ndeg}(\phi) = p^{n+1}$.

(5) $h(\phi) = n + 1$.

(6) $\mathfrak{i}_1(\phi) = \dim \phi - p^n$ *and* $\mathfrak{i}_2(\phi) = p^n - p^{n-1}$.

(7) $\mathfrak{i}(\phi) = \big(\dim \phi - p^n, p^n - p^{n-1}, p^{n-1} - p^{n-2}, \dots, p^2 - p, p - 1\big)$.

(8) *ϕ_1 is similar to a quasi-Pfister p-form* (*i.e.*, $h_{\mathrm{qp}}(\phi) \leq 1$).

Remark 3.12. In the case where $p = 2$, this result was proved earlier by Hoffmann and Laghribi [2004, Theorem 8.1] using different methods.

For arbitrary subforms, the situation is naturally more complicated, but we can nevertheless appeal to the following general result which was proved in [Scully 2016] (and whose proof again makes essential use of Theorems 3.1 and 3.7):

Proposition 3.13 (see [Scully 2016, Proposition 8.6]). *Let ϕ and ψ be anisotropic quasilinear p-forms of dimension ≥ 2 over F such that $\phi_{F(\psi)}$ is isotropic. Then either*

(1) *$(\psi_r)_{F_r(\phi)}$ is anisotropic for all $0 \leq r < h(\psi)$ and $\mathfrak{i}(\psi_{F(\phi)}) = \mathfrak{i}(\psi)$, or*

(2) *$\mathfrak{i}(\psi_{F(\phi)}) = \big(\mathfrak{i}_1(\psi), \dots, \mathfrak{i}_{s-1}(\psi), \mathfrak{i}_s(\psi) + \mathfrak{i}_{s+1}(\psi), \mathfrak{i}_{s+2}(\psi), \dots, \mathfrak{i}_{h(\psi)}(\psi)\big)$, where $s < h(\psi)$ is the smallest nonnegative integer such that $(\psi_s)_{F_s(\phi)}$ is isotropic.*

Note here that in the special situation where ψ is a neighbor of ϕ, we are necessarily in case (2) with s being equal to 0. Since $(\psi_{F(\phi)})_{\mathrm{an}} \simeq \phi_1$ in this instance (Lemma 3.5), we recover the computation of Proposition 3.10. By contrast, if ψ is a *near* neighbor of ϕ, then we can be in either of cases (1) and (2) (see Remark 3.15 below). Nevertheless, we still have $\psi_{F(\phi)} \simeq \phi_1$ (Corollary 3.4), and so we get:

Corollary 3.14 (see [Scully 2016, Corollary 6.10]). *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and let ψ be a near neighbor of ϕ. Then either*

(1) *$(\psi_r)_{F_r(\phi)}$ is anisotropic for all $0 \leq r < h(\psi)$ and $\mathfrak{i}(\psi) = \mathfrak{i}(\phi_1)$, or*

(2) *$\mathfrak{i}(\psi) = \big(\mathfrak{i}_2(\phi), \mathfrak{i}_3(\phi), \dots, \mathfrak{i}_s(\phi), \mathfrak{i}_s(\psi), \mathfrak{i}_{s+1}(\phi) - \mathfrak{i}_s(\psi), \mathfrak{i}_{s+2}(\phi), \dots, \mathfrak{i}_{h(\phi)}(\phi)\big)$, where $s < h(\psi)$ is the smallest positive integer such that $(\psi_s)_{F_s(\phi)}$ is isotropic.*

Remark 3.15. As per the comments above, neither of cases (1) and (2) can be ruled out here. Indeed, case (1) describes the situation where $h(\psi) = h(\phi) - 1$, while case (2) describes the situation where $h(\psi) = h(\phi)$. As the reader will easily verify using Corollary 2.40, both these situations can arise in practice.

3E. ***A comparison result.*** We now conclude this section by recalling the following comparison result for isotropy indices of quasilinear quadratic forms which was obtained in [Scully 2016] with the help of Theorem 3.1:

Proposition 3.16 (see [Scully 2016, Theorem 7.13, Remark 9.1]). *Assume that $p = 2$. Let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F and let L be a field extension of F such that ϕ_L is not split. Then*

$$\mathfrak{i}_0(\phi_{L(\phi)}) - \mathfrak{i}_1(\phi) \geq \min\big\{\mathfrak{i}_0(\phi_L), \big[\tfrac{1}{2}(\dim\phi - \mathfrak{i}_1(\phi) + 1)\big]\big\}.$$

Remark 3.17. A similar statement also holds for $p > 2$ (see [Scully 2016, Theorem 7.13]), but this will not be needed below.

Finally, it will be convenient to record here the following application of this result:

Theorem 3.18 [Scully 2016, Theorem 9.2]. *Let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F and write $\dim\phi = 2^n + m$ for uniquely determined integers $n \geq 0$ and $1 \leq m \leq 2^n$. Then, for any field extension L of F, we either have $\mathfrak{i}_0(\phi_L) \geq m$ or $\mathfrak{i}_0(\phi_L) \leq m - \mathfrak{i}_1(\phi)$.*

Remark 3.19. Taking $L = F(\phi)$ here, we see that if $\mathfrak{i}_1(\phi) < m$, then $\mathfrak{i}_1(\phi) \leq m/2$. This result will now be subsumed in our Theorem 1.3.

4. A motivational example

As a warm-up for the proof of our main result, we will now prove Proposition 1.4. Throughout this section, we assume that $p = 2$. By a *quasi-Pfister form*, we will mean a quasi-Pfister 2-form. Our assumption on the prime p is imposed for simplicity, but also because we will make use of the following fact already mentioned in the introduction, the analogue of which is unknown when $p > 3$ (see [Scully 2016, §4.1]):

Lemma 4.1 (see [Hoffmann 2004, Corollary 7.22]). *Let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F and let L be a field extension of F such that ϕ_L is isotropic. Then $\mathfrak{i}_0(\phi_L) \geq \mathfrak{i}_1(\phi)$.*

Now, let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F and let $\psi \subset \phi$ be a subform of codimension $\mathfrak{i}_1(\phi)$. By Corollary 3.4, $\psi_{F(\phi)}$ is isomorphic to the first higher anisotropic kernel of ϕ. In the next section, we will prove Theorem 1.5, which asserts that the latter form is divisible by a quasi-Pfister form of dimension $\geq \mathfrak{i}_1(\phi)$. Here, we will consider some special situations in which this divisibility property is already visible over the base field F. To this end, we will be interested in the following technical condition on the pair (ϕ, ψ):

($\star$) There exist elements $a \in D(\phi)\setminus\{0\}$ and $b \in D(\psi)\setminus\{0\}$ such that $a \neq c(bd+ef)$ for any $c, d, e, f \in D(\psi)$.

Example 4.2. If, in the above situation, we have $\mathrm{ndeg}(\psi) < \mathrm{ndeg}(\phi)$ (equivalently, if $h(\psi) < h(\phi)$; see Corollary 2.40), then ($\star$) holds for the pair (ϕ, ψ). Indeed, in this case, $D(\phi)$ is (evidently) not contained in $cN(\psi)$ for any $c \in F$. Since $bd + ef \in N(\psi)$ for every $b, d, e, f \in D(\psi)$, the validity of ($\star$) is immediately verified.

In light of Example 4.2, Proposition 1.4 is subsumed in the following result:

Proposition 4.3. *Let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F and let $\psi \subset \phi$ be subform of codimension $\mathfrak{i}_1(\phi)$ such that the pair (ϕ, ψ) satisfies ($\star$). Then there exist a quasi-Pfister form π, a subform $\sigma \subset \pi$, an element $\lambda \in D(\phi)$ and a form τ over F such that $\psi \simeq \pi \otimes \tau$ and $\phi \simeq \psi \perp \lambda\sigma$.*

Proof. Let $b \in D(\psi) \setminus \{0\}$ be as in ($\star$). Then ($\star$) also holds for the pair $(b\phi, b\psi)$. If this were not the case, then, for every $a \in D(\phi)$, we could find $c, d, e, f \in D(\psi)$ such that $ba = bc(b^3d + bebf)$, or, equivalently, such that $a = b^2c(bd + ef)$. But, since $D(\psi)$ is an F^2-linear subspace of F, we have $b^2c \in D(\psi)$, and so this would contradict the fact that ($\star$) holds for the original pair (ϕ, ψ). Since the exchange $(\phi, \psi) \to (b\phi, b\psi)$ does not affect the statement of the proposition, we can therefore assume that $b = 1$. In other words, we can assume that $1 \in D(\psi)$ and that:

($\star'$) There exists an element $a \in D(\phi) \setminus \{0\}$ such that $a \neq c(d + ef)$ for any $c, d, e, f \in D(\psi)$.

Let us fix $a \in D(\phi) \setminus \{0\}$ as in ($\star'$). We will now prove that the statement of the proposition holds with $\lambda = a$. First, we note that ($\star'$) implies:

(1) $a \notin D(\psi)$.

(2) $\psi_{F_{au}}$ is anisotropic for every $u \in D(\psi) \setminus \{0\}$.

Indeed, if a were in $D(\psi)$, then we could contradict ($\star'$) by taking $c = 1$, $d = a$ and $e = f = 0$. Similarly, if $\psi_{F_{au}}$ were isotropic in (2), then (since ψ is anisotropic) we could find nonzero elements $x, y \in D(\psi)$ such that $au = xy$ (see Lemma 2.27(2)); taking $c = x^{-1}$, $d = 0$, $e = y$ and $f = u$, this would again contradict ($\star'$).

Now, (1) implies that we have $\psi \perp \langle a \rangle \subset \phi$. If $\mathfrak{i}_1(\phi) = 1$, then the latter inclusion is an isomorphism, and the statement of the proposition holds with $\pi = \sigma = \langle 1 \rangle$ and $\tau = \psi$. Assume now that $\mathfrak{i}_1(\phi) > 1$. Since $1, a \in D(\phi)$, ϕ_{F_a} is isotropic by Lemma 2.27(2). By Lemma 4.1, it then follows that $\mathfrak{i}_0(\phi_{F_a}) \geq \mathfrak{i}_1(\phi)$. On the other hand, (2) (with $u = 1 \in D(\psi)$) shows that ψ_{F_a} is anisotropic, and, since $\dim \psi = \dim \phi - \mathfrak{i}_1(\phi)$, we conclude that $(\phi_{F_a})_{\mathrm{an}} \simeq \psi_{F_a}$. In other words, we have $D(\phi_{F_a}) = D(\psi_{F_a}) = D(\psi) + aD(\psi)$ (where the latter equality holds by Lemma 2.27(1)). By Lemma 2.2, it follows that we can write $\phi \simeq \psi \perp \langle a \rangle \perp a\sigma'$ for some form $\sigma' \subset \psi$. Let $\sigma = \langle 1 \rangle \perp \sigma'$, so that $\phi \simeq \psi \perp a\sigma$. As $1 \in D(\psi)$, we have $\sigma \subset \psi$. Since σ is anisotropic and represents 1, Lemma 2.11 shows that $\sigma \subset \sigma_{\mathrm{nor}}$. Thus, in order to complete the proof, it will be enough to prove that ψ is divisible by $\pi = \sigma_{\mathrm{nor}}$. By Corollary 2.20, this amounts to showing that $D(\sigma) \subset G(\psi)$. Let $x \in D(\sigma) \setminus \{0\}$. In order to show that $x \in G(\psi)$, we must check that $xy \in D(\psi)$ for all $y \in D(\psi)$ (see Lemma 2.13). If $y = 0$, then there is nothing to prove. Suppose now that $y \neq 0$. Then, by Lemma 2.27(2), $\phi_{F_{ay}}$ is isotropic. On the other hand, $\psi_{F_{ay}}$ is anisotropic by (2). Using Lemma 4.1 in the same way as before, we see that $(\phi_{F_{ay}})_{\mathrm{an}} \simeq \psi_{F_{ay}}$, or, equivalently, that $D(\phi_{F_{ay}}) = D(\psi_{F_{ay}}) = D(\psi) + ayD(\psi)$. But, letting $u = \sqrt{ay}$, we have $xy = (a^{-1}u)^2ax \in D(\phi_{F_{ay}})$. In particular, we can find $u, v \in D(\psi)$ such that $xy = u + ayv$. To complete the proof, it now only remains to show that $v = 0$. But, if $v \neq 0$, then we have $a = v^{-1}(x + uy^{-1})$, and we obtain a contradiction to ($\star'$) by taking $c = v^{-1}$, $d = x$, $e = u$ and $f = y^{-1}$. The result follows. □

In general, it is *not* always possible to find a subform $\psi \subset \phi$ of codimension $\mathfrak{i}_1(\phi)$ such that the pair (ϕ, ψ) satisfies condition $(\star)$. Indeed, note that if $\mathfrak{i}_1(\phi) = 2$ in the situation of Proposition 4.3, then ϕ is divisible by the binary (i.e., 2-dimensional) form σ. The following example, which is directly analogous to an example of Vishik from the characteristic $\neq 2$ theory of quadratic forms, shows that there exist anisotropic quasilinear quadratic forms which have first higher isotropy index equal to 2, but which are not divisible by a binary form: let a, b, c, d, e be algebraically independent variables over a field F_0 of characteristic 2, let $F = F_0(a, b, c, d, e)$ and consider the anisotropic F-form $\phi = \langle\langle ab, ac, ad\rangle\rangle \perp bcd\langle 1, ab, ac, ad\rangle \perp e\langle a, b, c, d\rangle$.

Lemma 4.4 (Vishik; see [Totaro 2009, Lemma 7.1]). *In the above situation, we have* $\mathfrak{i}_1(\phi) = 2$, *but* ϕ *is not divisible by a binary form.*

Proof. We will freely use some basic facts from the theory of symmetric bilinear forms over fields of characteristic 2. If B is such a form, then ϕ_B will denote the totally singular quadratic form $v \mapsto B(v, v)$. For all other notation and terminology, the reader is referred to [Elman et al. 2008, Chapter I].

Let $\mathfrak{b}$ be the bilinear form $\langle\langle ab, ac, ad\rangle\rangle_b \perp bcd\langle 1, ab, ac, ad\rangle_b \perp e\langle a, b, c, d\rangle_b$ over F, so that $\phi = \phi_{\mathfrak{b}}$. As the reader will immediately verify, we have

$$\mathfrak{b} \sim \mathfrak{c} := \langle\langle a, b, c, d\rangle\rangle_b \perp \langle\langle e\rangle\rangle_b \otimes \langle a, b, c, d\rangle_b,$$

where the symbol $\sim$ denotes Witt equivalence. Let $\pi = \langle\langle a, b, c, d\rangle\rangle_b$, and let π' denote the pure subform of π. Since $\operatorname{ndeg}(\phi_\pi) = 16 < 32 = \operatorname{ndeg}(\phi)$, Corollary 2.32(2) implies that $\pi_{F(\phi)}$ is anisotropic. At the same time, it follows from Remarks 2.6(4) and the definition of ϕ that $e \in D(\pi'_{F(\phi)})$. Thus, by [Elman et al. 2008, Lemma 6.1], there exists a 3-fold bilinear Pfister form η over $F(\phi)$ such that $\pi_{F(\phi)} \simeq \langle\langle e\rangle\rangle_b \otimes \eta$. In particular, $\mathfrak{c}_{F(\phi)}$ is divisible by $\langle\langle e\rangle\rangle_b$, and so $\mathfrak{i}_W(\mathfrak{c}_{F(\phi)})$ is even (see [Elman et al. 2008, Proposition 6.22]; $\mathfrak{i}_W$ denotes here the Witt index). Since $\dim \mathfrak{c} - \dim \mathfrak{b} = 8 \equiv 0 \pmod 4$, it follows that $\mathfrak{i}_W(\mathfrak{b}_{F(\phi)})$ is also even. In particular, this shows that $\mathfrak{i}_1(\phi) \geq 2$ (see [Laghribi 2007, Proposition 5.15]), and to prove that $\mathfrak{i}_1(\phi) = 2$, it suffices to find a field extension L of F such that ϕ_L is isotropic but $\mathfrak{i}_0(\phi_L) \leq 2$ (see Lemma 4.1). We claim that $L = F_{cde}$ is such an extension. First, note that since $cde = (bcd)(b^{-1}e)$ is a product of two nonzero elements of $D(\phi)$, ϕ_L is isotropic by Lemma 2.27(2). On the other hand, it is easy to see that the codimension-2 subform $\psi = \langle\langle ab, ac, ad\rangle\rangle \perp bcd\langle 1, ab, ac, ad\rangle \perp e\langle c, d\rangle \subset \phi$ remains anisotropic over L. Indeed, since

$$\psi_L \simeq \langle\langle ab, ac, ad\rangle\rangle \perp bcd\langle 1, ab, ac, ad\rangle \perp \langle d, c\rangle \subset \langle\langle a, b, c, d\rangle\rangle_L,$$

it suffices to check that $\langle\langle a, b, c, d\rangle\rangle$ remains anisotropic over L. But, since $cde \notin F^2(a, b, c, d) = N(\langle\langle a, b, c, d\rangle\rangle)$, this follows from Lemma 2.27(4). Since the anisotropy of ψ_L readily implies that $\mathfrak{i}_0(\phi_L) \leq 2$, we have proved our claim.

It now remains to check that ϕ is not divisible by a binary form. For the sake of contradiction, suppose instead that ϕ is divisible by $\langle\langle u\rangle\rangle$ for some $u \in F \setminus F^2$. We claim that $u \in F^2(ab, ac, ad)$. Again, let us assume that this is not the case. Then the quasi-Pfister form $\tau = \langle\langle ab, ac, ad\rangle\rangle$ remains anisotropic over F_u by Lemma 2.27(4). Let $\sigma = \tau \perp \langle bcd\rangle \subset \phi$ and $\eta = \tau \perp \langle ae\rangle \subset \phi$. Since $\mathfrak{i}_0(\phi_{F_u}) = \frac{1}{2}\dim\phi = 8$ (Lemma 2.27(6)), and since $\dim\sigma = \dim\eta = 9$, both σ_{F_u} and η_{F_u} are necessarily isotropic. Since τ_{F_u} is anisotropic, this means that $bcd, ae \in D(\tau_{F_u}) = F^2(ab, ac, ad, u)$. But this implies that $F^2(a, b, c, d, e) \subseteq F^2(ab, ac, ad, u)$, thus contradicting the fact that the elements a, b, c, d, e are algebraically independent over F_0. This proves our claim, and so we can write $u = v + w$ for some $v \in D(\langle 1, ab, ac, ad\rangle)$ and $w \in D(\langle bc, bd, cd, abcd\rangle)$. Now, Lemma 2.27(6) implies that $u \in G(\phi)$. In particular, applying Lemma 2.13 to the elements $ae, acd, abd \in D(\phi)$, we see that

(1) $aeu \in D(\phi)$,

(2) $acdu \in D(\phi)$, and

(3) $abdu \in D(\phi)$.

We can now complete the proof: first, note that we have $aev \in D(\phi)$, since $ae\langle 1, ab, ac, ad\rangle \simeq e\langle a, b, c, d\rangle \subset \phi$. By (1), this implies that $aew \in D(\phi)$. Note however that $ae\langle bc, bd, cd, abcd\rangle \simeq \langle abce, abde, acde, bcde\rangle$, and the latter form does not represent any nonzero element of $D(\phi)$. It follows that $w = 0$, and so $u \in D(\langle 1, ab, ac, ad\rangle)$. Next, consider the form $\rho = acd\langle 1, ab, ac, ad\rangle \simeq \langle acd, bcd, c, d\rangle$. By (2), we have $acdu \in D(\rho) \cap D(\phi)$. Since the elements a, b, c, d, e are algebraically independent over F_0, direct inspection shows that the former intersection is equal to $acd D(\langle 1, ab\rangle)$, and so $u \in D(\langle 1, ab\rangle)$. Finally, we can use (3) in a similar way to show that $u \in F^2$, thus providing us with the needed contradiction. The lemma is proved. □

In fact, Vishik's example shows more: it is generally not possible to find a subform $\psi \subset \phi$ of codimension $\mathfrak{i}_1(\phi)$ such that the pair (ϕ, ψ) satisfies condition $(\star)$, *even after making arbitrary rational extensions of* F — this stronger assertion follows from Lemma 2.35. This leads us to consider the possibility that decompositions of the kind suggested by Proposition 4.3 may be found by passing to suitable transcendental extensions of the base field, and, ultimately, to our main result. Nevertheless, it is still interesting to ask for conditions on ϕ (or, more specifically, on the Knebusch splitting pattern of ϕ) which automatically ensure the existence of the needed subform ψ. To this end, it is natural to look to the extremities where ϕ is "far from generic" (e.g., where $\mathfrak{i}_1(\phi)$ is "large"). The basic example is provided here by Proposition 2.39, which characterizes (scalar multiples of) anisotropic quasi-Pfister forms in terms of $\mathfrak{i}_1$, and one can hope that similar characterizations exist for "sufficiently simple" forms (see also Theorem 6.8 below). A better understanding of all these problems would have important implications for the study of symmetric bilinear forms of "low complexity" (e.g., of "small" height) in characteristic 2.

5. Main theorem

We are now ready to give the proof of Theorem 1.5. In order to treat the case where $p > 2$, the statement needs to be modified as follows:

Theorem 5.1. *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and let s be the smallest nonnegative integer such that $p^s \geq \mathfrak{i}_1(\phi)$. If ϕ is not a quasi-Pfister p-neighbor, then ϕ_1 is divisible by an s-fold quasi-Pfister p-form.*

Remark 5.2. Nothing is lost here by assuming that ϕ is not a quasi-Pfister p-neighbor. Indeed, if ϕ is a quasi-Pfister p-neighbor, and n denotes the smallest nonnegative integer such that $p^{n+1} \geq \dim \phi$, then ϕ_1 is similar to an n-fold quasi-Pfister p-form by Corollary 3.11. If $p = 2$, then we have $\mathfrak{i}_1(\phi) \leq \frac{1}{2} \dim \phi \leq 2^n$ by Proposition 2.39, so that $n \geq s$, where s is the integer defined in the statement of the theorem. Note, however, that if $p > 2$, then n may be strictly smaller than s (again, see Corollary 3.11). This explains why the additional hypothesis is needed here, but not in the statement of Theorem 1.5 (i.e., the case where $p = 2$).

Proof. To simplify the notation, we will write $\mathfrak{i}_1$ instead of $\mathfrak{i}_1(\phi)$ in what follows. If $\mathfrak{i}_1 = 1$, then the statement of the theorem holds trivially. We therefore assume henceforth that $\mathfrak{i}_1 > 1$. After multiplying ϕ by a nonzero scalar if necessary, we may also assume that $1 \in D(\phi)$. In particular, we can find $a_1, \ldots, a_n \in F$ such that $\phi \simeq \langle 1, a_1, \ldots, a_n\rangle$. For the remainder of the proof, we let $\phi' = \langle a_1, \ldots, a_n\rangle$, and we write $\phi'(T)$ for the "generic value" of ϕ', i.e., $\phi'(T) = \sum_{i=1}^n a_i T_i^p \in F[T]$, where $T = (T_1, \ldots, T_n)$ is a tuple of algebraically independent variables over F. By Corollary 3.9, the function field $F(\phi)$ is F-isomorphic to $F[\phi']$, and may therefore be identified with $\mathrm{Frac}\big(F[T]/(\phi'(T))\big)$ (see Remarks 2.6(1)). Fixing this identification henceforth, we will write $\bar{f}$ for the image of a polynomial $f \in F[T]$ under the canonical F-algebra homomorphism $F[T] \to F(\phi)$. We will also write $m(f)$ for the multiplicity $\mathrm{mult}_{\phi'(T)}(f)$ of $\phi'(T)$ in f, i.e., the largest integer k such that $f = \phi'(T)^k h$ for some $h \in F[T]$. Note here that we have $\bar{f} \neq 0$ if and only if $m(f) = 0$.

Now, let $\psi \subset \phi$ be any subform of codimension $\mathfrak{i}_1(\phi)$ such that $1 \in D(\psi)$. Then:

Lemma 5.3. *In the above situation, we can find elements $g_{i,j} \in D(\psi_{F[T]})$ $(1 \leq i < \mathfrak{i}_1$, $1 \leq j < p)$ such that*

$$\phi_{F(T)} \simeq \psi_{F(T)} \oplus \phi'(T)\langle 1, f_1, \ldots, f_{\mathfrak{i}_1 - 1}\rangle,$$

where $f_i = \sum_{j=1}^{p-1} g_{i,j}\phi'(T)^{j-1}$ for each $1 \leq i < \mathfrak{i}_1$.

Proof. First, let us note that $\phi'(T) \notin D(\psi_{F(T)})$. Indeed, if $\psi_{F(T)}$ were to represent $\phi'(T)$, then it would follow from Theorem 2.23 that $a_1, \dots, a_n \in D(\psi)$. Since $1 \in D(\psi)$ by hypothesis, this would imply that $D(\phi) \subseteq D(\psi)$, or, equivalently, that $\phi \subset \psi$ (see Proposition 2.3), which is impossible for dimension reasons (recall here that $\mathfrak{i}_1 > 1$ by assumption). It follows that $\psi_{F(T)} \oplus \langle \phi'(T) \rangle$ is anisotropic, and so $\psi_{F(T)} \oplus \langle \phi'(T) \rangle \subset \phi_{F(T)}$ by Proposition 2.3. Now, the affine function field $F[\phi]$ may be identified (over F) with the field $K = F(T)_{\phi'(T)}$ (see Remarks 2.6(1)). Since $F[\phi]$ is F-isomorphic to a purely transcendental extension of $F(\phi)$ (Remarks 2.6(2)), Corollary 3.4 and Lemma 2.26 together imply that $\psi_K \simeq (\phi_K)_{\mathrm{an}}$. In particular, we have $D(\phi_{F(T)}) \subset D(\phi_K) = D(\psi_K) = \sum_{j=0}^{p-1} D(\psi_{F(T)})\phi'(T)^j$ (where the last equality holds by Lemma 2.27(1)). Thus, by Lemma 2.2, we can complete the subform inclusion $\psi_{F(T)} \oplus \langle \phi'(T) \rangle \subset \phi_{F(T)}$ to an isomorphism $\phi_{F(T)} \simeq \psi_{F(T)} \oplus \langle \phi'(T) \rangle \oplus \langle f'_1, \dots, f'_{\mathfrak{i}_1 - 1} \rangle$, where, for each i, we have $f'_i = \sum_{j=0}^{p-1} g_{i,j} \phi'(T)^j$ for some $g_{i,j} \in D(\psi_{F(T)})$. Note, however, that every element of $D(\psi_{F(T)})$ is (trivially) the ratio of an element of $D(\psi_{F[T]})$ and a p-th power in $F[T]$. Since multiplying the f'_i by p-th powers in $F[T]$ does not change the $F(T)$-form $\langle f'_1, \dots, f'_{\mathfrak{i}_1 - 1} \rangle$ up to isomorphism (see Lemma 2.2), we can arrange it so that the $g_{i,j}$ belong to $D(\psi_{F[T]})$. Similarly, since subtracting elements of $D(\psi_{F(T)})$ from the f'_i does not change the isomorphism class of $\psi_{F(T)} \oplus \langle f'_1, \dots, f'_{\mathfrak{i}_1 - 1} \rangle$ (again, see Lemma 2.2), we can also arrange it so that $g_{i,0} = 0$ for all i. The remaining $g_{i,j}$ then satisfy the statement of the lemma. □

Let us now fix elements $g_{i,j} \in D(\psi_{F[T]})$ (and the associated polynomials f_i) satisfying the statement of Lemma 5.3. We are searching here for a sufficiently large quasi-Pfister divisor of ϕ_1, and we would like to try to build this quasi-Pfister p-form from the elements $g_{i,1}$. The basic point here is the following:

Lemma 5.4. *In the above situation, we have $\overline{g_{i,1}} b \in D(\phi_1)$ for all $b \in D(\psi)$ and all $1 \le i < \mathfrak{i}_1$.*

Proof. If $b = 0$, then the statement is trivial. Let us now fix $b \in D(\psi) \setminus \{0\}$. We will need another lemma:

Lemma 5.5. *In the above situation, there exist elements $s_{i,j} \in D(\psi_{F[T]})$ and $t_{i,j} \in F[T] \setminus \{0\}$ ($1 \le i < \mathfrak{i}_1$, $0 \le j < p$) such that, for every $1 \le i < \mathfrak{i}_1$, we have:*

(1) *$bf_i = \sum_{j=0}^{p-1} (s_{i,j}/t_{i,j}^p)(\phi'(T)^j/b^j)$ in $F(T)$.*

(2) *For each $0 \le j < p$, at least one of $\overline{s_{i,j}}$ and $\overline{t_{i,j}}$ is nonzero.*

Proof. Let $1 \le i < \mathfrak{i}_1$, and consider the field $L = F(T)_u$, where $u = \phi'(T)/b$. In view of Remarks 2.6(1), L is F-isomorphic to the affine function field $F[\eta]$, where η denotes the F-form $\langle b \rangle \perp \phi'$. Now, since $b \in D(\psi)$, and since $\psi \subset \phi$, we have $D(\phi') \subseteq D(\eta) \subseteq D(\phi)$. In particular, if $\eta \not\simeq \phi$, then it follows from Lemma 2.2 that

$\eta_{\mathrm{an}} \simeq \phi'$. Either way, we see that L is F-isomorphic to a degree-1 purely transcendental extension of $F(\phi)$ — in the first case, see Remarks 2.6(2); in the second, see Remarks 2.6(3) and Corollary 3.9. By Corollary 3.4 and Lemma 2.26, it follows that $\psi_L \simeq (\phi_L)_{\mathrm{an}}$. In other words, we have $D(\phi_L) = D(\psi_L) = \sum_{j=0}^{p-1} D(\psi_{F(T)})u^j$ (again, see Lemma 2.27(1) for the final equality). Now, since u is a p-th power in L, we have $bf_i = \phi'(T) f_i/u \in D(\phi_L)$. We can therefore write $bf_i = \sum_{i=0}^{p-1} q_j u^j$ for some $q_j \in D(\psi_{F(T)})$. Since every element of $D(\psi_{F(T)})$ is the quotient of an element of $D(\psi_{F[T]})$ and a p-th power in $F[T]$, and since $u = \phi'(T)/b$, this shows that we can find elements $s_{i,j} \in D(\psi_{F[T]})$ and $t_{i,j} \in F[T] \setminus \{0\}$ such that (1) holds. Finally, since $\psi_{F(\phi)}$ is anisotropic, Proposition 2.33 implies that $m(s_{i,j}) \equiv 0 \pmod p$ for all $0 \leq j < p$. For each such j, let $m_j = \min(m(s_{i,j}), pm(t_{i,j}))$, and put $s'_{i,j} = s_{i,j}/\phi'(T)^{m_j}$ and $t'_{i,j} = t_{i,j}/\phi'(T)^{m_j/p}$. Then, by Theorem 2.23, we again have $s'_{i,j} \in D(\psi_{F[T]})$. Thus, replacing $s_{i,j}$ by $s'_{i,j}$ and $t_{i,j}$ by $t'_{i,j}$ (for each j), we arrive at the situation where, for any j, either $m(s_{i,j}) = 0$ or $m(t_{i,j}) = 0$. In other words, at least one of $\overline{s_{i,j}}$ and $\overline{t_{i,j}}$ is nonzero, as we wanted. □

Returning now to the proof of Lemma 5.4, let $s_{i,j} \in D(\psi_{F[T]})$ and $t_{i,j} \in F[T]\setminus\{0\}$ be as in Lemma 5.5. In particular, we have the equation

$$\sum_{l=1}^{p-1} bg_{i,l}\phi'(T)^{l-1} = bf_i = \sum_{j=0}^{p-1} \frac{s_{i,j}}{t_{i,j}^p}\frac{\phi'(T)^j}{b^j}$$

in $F(T)$. Clearing denominators, we obtain

$$\prod_k t_{i,k}^p \sum_{l=1}^{p-1} bg_{i,l}\phi'(T)^{l-1} = \sum_{j=0}^{p-1}\prod_{k\neq j} t_{i,k}^p s_{i,j}\frac{\phi'(T)^j}{b^j}. \tag{5-1}$$

Now, we claim that, for all $0 \leq j < p$, we have $\overline{t_{i,j}} \neq 0$, or, equivalently, $m(t_{i,j}) = 0$. To see this, let $m = \min\bigl\{\sum_{k\neq j} m(t_{i,k}) \mid 0 \leq j < p\bigr\}$. Then our claim amounts to the assertion that $m = \sum_{k=0}^{p-1} m(t_{i,k})$. Suppose that this is not the case, and let $0 \leq j < p$ be minimal so that $\sum_{k\neq j} m(t_{i,k}) = m$. Then, reducing both sides of (5-1) modulo $\phi'(T)^{pm+j+1}$, we see that $s_{i,j} \equiv 0 \pmod{\phi'(T)}$. In other words, we have $\overline{s_{i,j}} = 0$. By the choice of the $s_{i,j}$ and $t_{i,j}$, this implies that $\overline{t_{i,j}} \neq 0$, or, equivalently, that $m(t_{i,j}) = 0$. But then $m = \sum_{k\neq j} m(t_{i,k}) = \sum_{k=0}^{p-1} m(t_{i,k})$, which contradicts our assumption. The claim is therefore proved, and so, reducing (5-1) modulo $\phi'(T)$ and dividing through by $\prod_k \overline{t_{i,k}}^p$, we obtain the equality $\overline{g_{i,1}}b = \overline{s_{i,0}}/\overline{t_{i,0}}^p$ in $F(\phi)$. As $s_{i,0} \in D(\psi_{F[T]})$, this shows that $\overline{g_{i,1}}b \in D(\psi_{F(\phi)})$. But since $\psi_{F(\phi)} \simeq \phi_1$, we have $D(\psi_{F(\phi)}) = D(\phi_1)$, and the lemma is therefore proved. □

Continuing with the proof of Theorem 5.1, let us now choose elements $g_{i,j}$ in $D(\psi_{F[T]})$ as in the statement of Lemma 5.3 so that the integer $\sum_{i=1}^{i_1-1} \deg_{T_1}(g_{i,1})$ is minimal, where $\deg_{T_1}(g)$ denotes the degree of any $g \in F[T]$ viewed as an element of the ring $F(T_2, \dots, T_n)[T_1]$, i.e., as a polynomial in the single variable T_1 (with the added convention that $\deg_{T_1}(0) = 0$). Consider the form $\sigma = \langle 1, \overline{g_{1,1}}, \dots, \overline{g_{\mathfrak{i}_1-1,1}} \rangle$ over $F(\phi)$. The final step in the proof of the theorem will be to prove the following statement:

Lemma 5.6. *In the above situation, σ is anisotropic.*

Before proving the lemma, let us explain how this concludes the proof of Theorem 5.1. First, we claim that ϕ_1 is divisible by the quasi-Pfister p-form σ_{nor}. By Corollary 2.20, this amounts to checking that $D(\sigma) \subseteq G(\phi_1)$. Since $G(\phi_1)$ is a subfield of F containing F^p (Corollary 2.15), it suffices to show here that $\overline{g_{i,1}} \in G(\phi_1)$ for all $1 \leq i < \mathfrak{i}_1$. But by Lemma 2.13, this is equivalent to showing that, for all such i, we have $\overline{g_{i,1}}\, D(\phi_1) \subseteq D(\phi_1)$. Since $D(\phi_1) = D(\psi_{F(\phi)})$ is spanned as an $F(\phi)^p$-vector space by $D(\psi)$, this follows immediately from Lemma 5.4. The claim is therefore proved, and to finish the proof of the theorem, it only remains to check that $\dim \sigma_{\mathrm{nor}} \geq p^s$. But, σ is anisotropic by Lemma 5.6, and so $\sigma \subset \sigma_{\mathrm{nor}}$ by Lemma 2.11. In particular, we have $\dim \sigma_{\mathrm{nor}} \geq \dim \sigma = \mathfrak{i}_1$, which is precisely the assertion that $\dim \sigma_{\mathrm{nor}} \geq p^s$ (because $\dim \sigma_{\mathrm{nor}}$ is necessarily a power of p). Now, in order to prove Lemma 5.6, we need another auxiliary statement:

Lemma 5.7. *If σ is isotropic, then $p > 2$, and there exist polynomials $g_j \in D(\psi_{F[T]})$ $(1 \leq j < p)$ and an integer $2 \leq k < p$ such that:*

(1) $\sum_{j=1}^{p-1} g_j \phi'(T)^j \in D(\phi_{F[T]})$.

(2) $m(g_l) > 0$ *for all* $1 \leq l < k$.

(3) $m(g_k) = 0$.

Proof. If σ is isotropic, then since $\phi'(T)$ is a Fermat-type polynomial of degree p, we can find an integer $1 \leq m \leq p$ and polynomials $h_0, \dots, h_{\mathfrak{i}_1-1}, h \in F[T]$ such that:

(i) $h_0^p + g_{1,1}h_1^p + \dots + g_{\mathfrak{i}_1-1,1}h_{\mathfrak{i}_1-1}^p = \phi'(T)^m h$ in $F[T]$.

(ii) $\deg_{T_1}(h_i) < p$ for all $0 \leq i < \mathfrak{i}_1$.

(iii) $\bar{h_i} \neq 0$ for some $1 \leq i < \mathfrak{i}_1$.

First, let us note that we have $\phi'(T)^m h \in D(\psi_{F[T]})$ by (i) and the definition of the elements $g_{i,1}$. Since $\psi_{F(\phi)}$ is anisotropic, it follows from Proposition 2.33 that $h = 0$ or $m = p$. Either way, we can assume henceforth that $m = p$. Now, by (iii), there exists an $1 \leq l < \mathfrak{i}_1$ such that $h_l \neq 0$. Among all such integers l, let us fix one so that $\deg_{T_1}(g_{l,1}h_l^p)$ is maximal. Consider now the polynomial $f = h_0^p + \sum_{k=1}^{\mathfrak{i}_1-1} f_k h_k^p \in F[T]$, where the f_k are as in the statement of Lemma 5.3. Since $h_l \neq 0$, Lemma 2.2 implies that $\langle 1, f_1, \dots, f_{\mathfrak{i}_1-1} \rangle \simeq \langle 1, f_1, \dots, f_{l-1}, f, f_{l+1}, \dots, f_{\mathfrak{i}_1-1} \rangle$ as $F(T)$-forms. In particular, we have

$$\phi_{F(T)} \simeq \psi_{F(T)} \oplus \phi'(T)\langle 1, f_1, \dots, f_{l-1}, f, f_{l+1}, \dots, f_{\mathfrak{i}_1-1} \rangle. \tag{5-2}$$

Now, by definition, $f=\sum_{j=1}^{p-1} g_j'\phi'(T)^{j-1}$, where $g_1'=\phi'(T)^m h\in D(\psi_{F[T]})$ and $g_j'=\sum_{k=1}^{\mathrm{i}_1-1} g_{k,j}h_k^p\in D(\psi_{F[T]})$ for all $2\le i<p$. Let $r=\min\{m(g_j')\mid 1\le j<p\}$. Since $g_j'\in D(\psi_{F[T]})$ for all j, and since $\psi_{F(\phi)}$ is anisotropic, another application of Proposition 2.33 shows that $r\equiv 0 \pmod p$. In particular, for each $j\ge 1$, we have $g_j:=g_j'/\phi'(T)^r\in D(\psi_{F[T]})$. In view of (5-2), it follows that the exchange $g_{l,j}\to g_j$ does not alter the statement of Lemma 5.3. By our choice of the $g_{i,j}$, we therefore have

$$\deg_{T_1}(\phi'(T)^{p-r}h)=\deg_{T_1}(g_1)\ge\deg_{T_1}(g_{l,1}). \tag{5-3}$$

Now, we claim that the elements g_j (together with an appropriate integer k) satisfy the conditions of the lemma. We have already seen here the validity of (1). At the same time, we have $m(g_j)=0$ for some $j\ge 1$ by construction. Thus, in order to prove the existence of an integer k such that (2) and (3) are satisfied, we just need to check that $m(g_1)>0$. Recall again that we have $g_1=\phi'(T)^{p-r}h$. If $h=0$, then there is nothing to prove. Suppose now that $h\neq 0$. By (i) and the choice of the integer l, we have $\deg_{T_1}(g_{l,1}h_l^p)\ge\deg_{T_1}(\phi'(T)^p h)\ge p^2+\deg_{T_1}(h)$. Since $\deg_{T_1}(h_l)<p$ (by (ii)), it follows that $\deg_{T_1}(g_{l,1})>\deg_{T_1}(h)$. In view of (5-3), we see that $r=0$ in this case. In particular, we have $g_1=\phi'(T)^p h$, and so $m(g_1)\ge p>0$, as we wanted. □

We are now ready to prove Lemma 5.6 and thus complete the proof of Theorem 5.1. If $p=2$, then the statement was already proved in Lemma 5.7. Suppose now that $p>2$, and assume for the sake of contradiction that σ is isotropic. Let g_j $(1\le j<p)$ and k be as in the statement of Lemma 5.7. By condition (2) of the lemma, we can, for each $l<k$, write $g_l=\phi'(T)^{m_l}h_l$ for some positive integer m_l and some polynomial h_l. By a now familiar application of Proposition 2.33, we have $m_l\equiv 0 \pmod p$ for every such l. In particular, the m_l are all strictly larger than k. Now, by condition (1) of the lemma, the element

$$\phi'(T)^k\big(g_k+g_{k-1}\phi'(T)+\cdots+g_{\mathrm{i}_1-1}\phi'(T)^{p-k-1}+h_1\phi'(T)^{m_1-k}+\cdots+h_{k-1}\phi'(T)^{m_{k-1}-k}\big)$$

lies in $D(\phi_{F[T]})$. Using the very same argument as that used to prove Lemma 5.4 above (and the fact that the integers m_l-k $(l<k)$ are all positive), one readily shows that $b^k\overline{g_k}\in D(\phi_1)$ for every $b\in D(\psi)$. Note, however, that $\overline{g_k}\neq 0$ by condition (3) of Lemma 5.7. Since $2\le k<p$, and since $\phi_1\simeq\psi_{F(\phi)}$, it follows from Lemma 2.22 that ϕ_1 is a quasi-Pfister p-form. But, by Corollary 3.11, this in turn implies that ϕ is a quasi-Pfister p-neighbor, thus contradicting our original hypothesis. The lemma and theorem are therefore proved. □

6. First applications of the main theorem

We now give the basic applications of Theorem 1.5.

6A. ***Possible values of the Knebusch splitting pattern.*** Let ϕ be a quasilinear p-form of dimension ≥ 2 over F. In the previous section we have shown that the first higher anisotropic kernel ϕ_1 of ϕ is divisible by a quasi-Pfister p-form of dimension $\geq \mathfrak{i}_1(\phi)$, provided that ϕ_{an} is not a quasi-Pfister p-neighbor. In the terminology of Section 2K, this amounts to the assertion that if $h_{\mathrm{qp}}(\phi) \geq 2$, then $\mathfrak{d}_1(\phi) \geq \log_p(\mathfrak{i}_1(\phi))$ (here we are also making use Corollary 3.11). By virtue of the inductive nature of the Knebusch splitting tower construction, we also obtain analogous restrictions on the higher isotropy indices $\mathfrak{i}_r(\phi)$ $(2 \leq r < h_{\mathrm{qp}}(\phi))$ in terms of the corresponding higher divisibility indices $\mathfrak{d}_r(\phi)$. Taking the observations of Section 2K into account, our results may be summarized as follows:

Theorem 6.1. *Let ϕ be a quasilinear p-form over F and let $d = h(\phi) - h_{\mathrm{qp}}(\phi)$. Then*:

(1) $\mathfrak{i}(\phi) = \big(\mathfrak{i}_1(\phi), \dots, \mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi), p^d - p^{d-1}, p^{d-1} - p^{d-2}, \dots, p^2 - p, p-1\big)$.

(2) $\mathfrak{i}_{h_{\mathrm{qp}}}(\phi) = \dim \phi - \mathfrak{j}_{h_{\mathrm{qp}}(\phi)-1}(\phi) - p^d < p^{d+1} - p^d$.

(3) $\mathfrak{d}(\phi) = \big(\mathfrak{d}_0(\phi), \dots, \mathfrak{d}_{h_{\mathrm{qp}}(\phi)-1}(\phi), d, d-1, \dots, 1, 0\big)$.

(4) $\mathfrak{d}_0(\phi) \leq \mathfrak{d}_1(\phi) \leq \cdots \leq \mathfrak{d}_{h_{\mathrm{qp}}(\phi)} = d$.

(5) $\mathfrak{i}_r(\phi) \equiv 0 \pmod{p^{\mathfrak{d}_{r-1}(\phi)}}$ *for all* $1 \leq r < h_{\mathrm{qp}}(\phi)$.

(6) $\mathfrak{d}_r(\phi) \geq \log_p(\mathfrak{i}_r(\phi))$ *for all* $1 \leq r < h_{\mathrm{qp}}(\phi)$.

(7) *For every* $1 \leq r < h_{\mathrm{qp}}(\phi)$, $\mathfrak{i}_r(\phi) - 1$ *is the remainder of* $\dim \phi - \mathfrak{j}_{r-1}(\phi) - 1$ *modulo* $p^{\mathfrak{d}_r(\phi)}$.

Proof. Parts (1), (2), (3), (4) and (5) are the statements comprising Lemmas 2.42, 2.43, 2.44 and Corollary 2.45. Since $\mathfrak{i}_r(\phi) = \mathfrak{i}_1(\phi_{r-1})$, $\mathfrak{d}_r(\phi) = \mathfrak{d}_1(\phi_{r-1})$ and $\dim \phi_{r-1} = \dim \phi - \mathfrak{j}_{r-1}(\phi)$ for all $1 \leq r \leq h(\phi)$, parts (6) and (7) follow immediately from Theorem 5.1 and Corollary 3.11. □

In order to highlight the general shape of the Knebusch splitting pattern exposed by Theorem 6.1, it is worth writing down the following result explicitly (here the notation $a \mid b$ means that a divides b):

Corollary 6.2. *Let ϕ be a quasilinear p-form over F. Then*

$$\mathfrak{i}_1(\phi) \leq p^{\mathfrak{d}_1(\phi)} \mid \mathfrak{i}_2(\phi) \leq p^{\mathfrak{d}_2(\phi)} \mid \cdots \mid \mathfrak{i}_{h_{\mathrm{qp}}(\phi)-1}(\phi) \leq p^{\mathfrak{d}_{h_{\mathrm{qp}}(\phi)-1}(\phi)} \mid \mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi).$$

Remark 6.3. In the special case where $p = 2$, the chain of inequalities

$$\mathfrak{i}_1(\phi) \leq \mathfrak{i}_2(\phi) \leq \cdots \leq \mathfrak{i}_{h_{\mathrm{qp}}(\phi)}(\phi)$$

was previously obtained in [Scully 2016, Theorem 9.5] using Proposition 3.16. Here, we have given a more precise and natural explanation of this phenomenon.

We now show that, as far as the Knebusch splitting pattern is concerned, one cannot do any better than Theorem 6.1 in general:

Proposition 6.4. *Let n be any positive integer. Suppose that we are given a nonnegative integer $k \le n$ and two sequences $(\mathfrak{d}_0, \mathfrak{d}_1, \dots, \mathfrak{d}_k = d)$ and $(\mathfrak{i}_1, \dots, \mathfrak{i}_k)$ of $k+1$ and k nonnegative integers, respectively, such that the following conditions hold*:

(i) $\mathfrak{i}_k = n - \sum_{j=1}^{k-1} \mathfrak{i}_j - p^d < p^{d+1} - p^d$.

(ii) $\mathfrak{d}_0 \le \mathfrak{d}_1 \le \cdots \le \mathfrak{d}_k = d$.

(iii) $1 \le \mathfrak{i}_r \equiv 0 \pmod{p^{\mathfrak{d}_{r-1}}}$ *for all* $1 \le r < k$.

(iv) $\mathfrak{d}_r \ge \log_p(\mathfrak{i}_r)$ *for all* $1 \le r < k$.

(v) *For every* $1 \le r < k$, $\mathfrak{i}_r - 1$ *is the remainder of* $n - \left(\sum_{j=1}^{r-1} \mathfrak{i}_j\right) - 1$ *modulo* $p^{\mathfrak{d}_r}$.

Then there exists a (purely transcendental) field extension L of F and an anisotropic quasilinear p-form ϕ of dimension n over L such that:

(1) $h_{\mathrm{qp}}(\phi) = k$.

(2) $h(\phi) = k + d$.

(3) $\mathfrak{d}(\phi) = \left(\mathfrak{d}_0(\phi), \dots, \mathfrak{d}_{k-1}(\phi), d, d-1, \dots, 1, 0\right)$.

(4) $\mathfrak{d}_r(\phi) \ge \mathfrak{d}_r$ *for all* $0 \le r < k$.

(5) $\mathfrak{i}(\phi) = \left(\mathfrak{i}_1, \dots, \mathfrak{i}_k, p^d - p^{d-1}, p^{d-1} - p^{d-2}, \dots, p^2 - p, p-1\right)$.

Proof. We argue by induction on k. If $k = 0$, then Proposition 2.39 shows that we can take $L = F(T)$ and $\phi = \langle\langle T_1, \dots, T_d \rangle\rangle$, where $T = (T_1, \dots, T_d)$ is a d-tuple of algebraically independent variables over F. Suppose now that $k > 0$, and let $n' = (n - \mathfrak{i}_1)/p^{\mathfrak{d}_1}$ and $\mathfrak{i}'_r = \mathfrak{i}_{r+1}/p^{\mathfrak{d}_1}$ for all $1 \le r < k$. By our hypotheses, these ratios are, in fact, positive integers. Setting $\mathfrak{d}'_r = \mathfrak{d}_{r+1} - \mathfrak{d}_1$ for all $0 \le r < k$, and putting $d' = \mathfrak{d}'_{k-1}$, conditions (i)–(v) then imply the following:

(i′) $\mathfrak{i}'_{k-1} = n' - \sum_{j=1}^{k-2} \mathfrak{i}'_j - p^{d'} < p^{d'+1} - p^{d'}$.

(ii′) $\mathfrak{d}'_0 \le \mathfrak{d}'_1 \le \cdots \le \mathfrak{d}'_{k-1} = d'$.

(iii′) $1 \le \mathfrak{i}'_r \equiv 0 \pmod{p^{\mathfrak{d}'_{r-1}}}$ for all $1 \le r < k-1$.

(iv′) $\mathfrak{d}'_r \ge \log_p(\mathfrak{i}'_r)$ for all $1 \le r < k-1$.

(v′) For every $1 \le r < k-1$, $\mathfrak{i}'_r - 1$ is the remainder of $n' - \left(\sum_{j=1}^{r-1} \mathfrak{i}'_j\right) - 1$ modulo $p^{\mathfrak{d}'_r}$.

By the induction hypothesis, there exists a (purely transcendental) field extension L_0 of F and an anisotropic quasilinear p-form ψ of dimension n' over L_0 such that:

(1′) $h_{\mathrm{qp}}(\psi) = k - 1$.

(2′) $h(\psi) = k - 1 + d'$.

(3′) $\mathfrak{d}(\psi) = \left(\mathfrak{d}_0(\psi), \mathfrak{d}_1(\psi), \dots, \mathfrak{d}_{k-2}(\psi), d', d'-1, \dots, 1, 0\right)$.

(4′) $\mathfrak{d}_r(\psi) \ge \mathfrak{d}'_r$ for all $0 \le r < k-1$.

(5′) $\mathfrak{i}(\psi) = \left(\mathfrak{i}'_1, \dots, \mathfrak{i}'_{k-1}, p^{d'} - p^{d'-1}, p^{d'-1} - p^{d'-2}, \dots, p^2 - p, p-1\right)$.

Consider now the form $\sigma = \psi_{L_1} \perp \langle T_0 \rangle$ over the rational function field $L_1 = L_0(T_0)$. By (3′), (5′) and Lemma 2.46, we have:

(a) $\mathfrak{d}(\sigma) = \big(0, \mathfrak{d}_0(\psi), \mathfrak{d}_1(\psi), \dots, \mathfrak{d}_{k-2}(\psi), d', d'-1, \dots, 1, 0\big)$.

(b) $\mathfrak{i}(\sigma) = \big(1, \mathfrak{i}'_1, \dots, \mathfrak{i}'_{k-1}, p^{d'} - p^{d'-1}, p^{d'-1} - p^{d'-2}, \dots, p^2 - p, p-1\big)$.

We would like to modify this further. Consider the product $\tau = \langle\langle T_1, \dots, T_{\mathfrak{d}_1} \rangle\rangle \otimes \sigma_L$ over $L = L_1(T)$, where $T = (T_1, \dots, T_{\mathfrak{d}_1})$ is a $\mathfrak{d}_1$-tuple of algebraically independent variables over L_1. Then, by (a), (b) and Lemma 2.48, we have:

(c) $\mathfrak{d}(\tau) = \big(\mathfrak{d}_1, \mathfrak{d}_1(\psi) + \mathfrak{d}_1, \mathfrak{d}_2(\psi) + \mathfrak{d}_1, \dots, \mathfrak{d}_{k-2}(\psi) + d_1, d, d-1, \dots, 1, 0\big)$.

(d) $\mathfrak{i}(\tau) = \big(p^{\mathfrak{d}_1}, \mathfrak{i}_2, \dots, \mathfrak{i}_k, p^d - p^{d-1}, p^{d-1} - p^{d-2}, \dots, p^2 - p, p-1\big)$.

Now, by (iv), we have $\mathfrak{i}_1 = p^{\mathfrak{d}_1} - s$ for some $0 \leq s < p^{\mathfrak{d}_1}$. By (ii) and (iii), s is divisible by $p^{\mathfrak{d}_0}$. Let ϕ be any codimension-s subform of τ which is divisible by $\langle\langle T_1, \dots, T_{\mathfrak{d}_0} \rangle\rangle$. Clearly ϕ is anisotropic, and by (c), (d) and Proposition 3.10, we have:

(e) $\mathfrak{d}(\phi) = \big(\mathfrak{d}_0(\phi), \mathfrak{d}_1(\psi) + \mathfrak{d}_1, \dots, \mathfrak{d}_{k-2}(\psi) + \mathfrak{d}_1, d, d-1, \dots, 1, 0\big)$.

(f) $\mathfrak{i}(\phi) = \big(\mathfrak{i}_1, \mathfrak{i}_2, \dots, \mathfrak{i}_k, p^d - p^{d-1}, p^{d-1} - p^{d-2}, \dots, p^2 - p, p-1\big)$.

The second statement shows that ϕ satisfies conditions (2) and (5). At the same time, since $\mathfrak{d}_0(\phi) \geq \mathfrak{d}_0$ by construction, and since $\mathfrak{d}_2(\psi) + \mathfrak{d}_1 \geq \mathfrak{d}'_r + \mathfrak{d}_1 = \mathfrak{d}_{r+1}$ for all $0 \leq r < k-1$ by (4′), (e) shows that (3) and (4) are also satisfied. Finally, since $\mathfrak{i}_k < p^{d+1} - p^d$, Proposition 2.39 shows that $h_{\mathrm{qp}}(\phi) = k$, i.e., that (1) holds for ϕ. The pair (L, ϕ) therefore has all the desired properties. □

Remark 6.5. In general, it is not possible to arrange it so that $\mathfrak{d}_r(\phi) = \mathfrak{d}_r$ for all $0 \leq r < k$ in the statement of Proposition 6.4. For example, suppose that $p = 2$, and take $n = 2^{s+1} - 2$ for some $s \geq 3$, $k = 1$, $\mathfrak{d}_0 = 0$, $\mathfrak{d}_1 = s$, $\mathfrak{i}_0 = 0$, $\mathfrak{i}_1 = 2^s - 2$. As the reader will readily verify, these integers satisfy conditions (i)–(v) of the proposition. On the other hand, let (L, ϕ) be any pair consisting of a field extension L of F and an anisotropic form ϕ of dimension $2^{s+1} - 2$ over L such that $\mathfrak{i}_1(\phi) = 2^s - 2$. By Theorem 6.8 below (see also [Scully 2016, Theorem 9.6]), ϕ is necessarily a quasi-Pfister 2-neighbor, and therefore satisfies conditions (1)–(5) of the proposition (see Corollary 3.11). We claim, however, that $\mathfrak{d}_0(\phi) > 0$, i.e., that ϕ is divisible by a binary form. To see this, note that there exists an anisotropic $(s+1)$-fold bilinear Pfister form π over L and a subform $\mathfrak{b} \subset \pi$ such that ϕ is given by the assignment $v \mapsto \mathfrak{b}(v, v)$. Since $\mathfrak{b}$ (being a codimension-2 subform of π) becomes split over the extension $K = L_{\det(\mathfrak{b})}$ (where det denotes the determinant), we have $\mathfrak{i}_0(\phi_K) \geq \frac{1}{2} \dim \phi$ (see [Laghribi 2007, Proposition 5.15]), whence ϕ is divisible by $\langle\langle \det(\mathfrak{b}) \rangle\rangle$ (Lemma 2.27(6)).

Theorem 6.1 and Proposition 6.4 thus give a complete solution to the problem of determining the possible values of the Knebusch splitting pattern for quasilinear

p-forms. In particular, we have an answer to Question 1.1 in the totally singular case. As noted in Example 2.47, the Knebusch and full splitting patterns need not agree in general for quasilinear p-forms. In Section 7 below, we will consider the problem of determining the possible values of the full splitting pattern in the case where $p = 2$.

6B. ***Canonical dimensions of quasilinear p-hypersurfaces.*** Since the canonical dimension of an anisotropic quasilinear p-hypersurface is determined by the Knebusch splitting pattern of its underlying form (this is the basic fact underlying our proof of Theorem 5.1; see Theorem 3.1 and Remarks 3.2(2)), we also obtain a list of all restrictions on the possible values of the former invariant:

Theorem 6.6. *Let X be an anisotropic quasilinear p-hypersurface over F. Then there exists a nonnegative integer s such that*:

(1) $p^s - 1 \leq \dim X < p^s + \mathrm{cdim}(X)$.

(2) $\mathrm{cdim}(X) \equiv -1 \pmod{p^s}$.

There are no further restrictions on $\mathrm{cdim}(X)$.

Proof. As per the above discussion, this follows readily from Theorems 3.1 and 6.1 and Proposition 6.4 (see also Remarks 3.2(2)). $\square$

Remarks 6.7. (1) We emphasize again that $\mathrm{cdim}(X)$ is to be understood here as the minimum dimension of the image of a rational self-map $X \dashrightarrow X$. For $p > 3$, we do not know if this agrees with the definition given in [Elman et al. 2008, §90] (although this is certainly expected; see [Scully 2016, Question 4.4, Proposition 4.7]).

(2) The problem of determining all possible values of the canonical dimension for smooth (projective) hypersurfaces X of prime degree $p > 2$ in characteristic $\neq p$ remains open in general. Using a degree formula, Merkurjev [2003, §7.3] showed that if $\dim X \geq p^n - 1$ for some nonnegative integer n, then we also have $\mathrm{cdim}(X) \geq p^n - 1$ provided that X has no points of degree *prime to* p. Little else is known. By way of specialization, our Theorem 6.6 may be of some use when considering specific examples over certain fields of characteristic zero.

6C. ***Quasilinear p-forms with maximal splitting.*** Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and write $\dim\phi = p^n + m$ for uniquely determined integers $n \geq 0$ and $1 \leq m \leq p^{n+1} - p^n$. By Theorem 6.1 (see also [Scully 2013, Corollary 6.8]), we have $\mathfrak{i}_1(\phi) \leq m$. If equality holds here, then we say that ϕ has *maximal splitting*. The basic examples of forms having this property are given by anisotropic quasi-Pfister p-neighbors (see Corollary 3.11). It is interesting to ask here to what extent this property characterizes quasi-Pfister p-neighbors. Given Theorem 5.1, we can now prove the following general result:

Theorem 6.8. *Let ϕ be an anisotropic quasilinear p-form of dimension ≥ 2 over F and let n be the smallest nonnegative integer such that $p^{n+1} \geq \dim \phi$. If ϕ has maximal splitting, and if either*

(1) $p > 2$ *and* $\dim \phi > p^n + p^{n-1}$, *or*

(2) $p = 2$ *and* $\dim \phi > 2^n + 2^{n-2}$,

then ϕ is a quasi-Pfister p-neighbor.

Proof. By Theorem 5.1, we may assume that $\mathfrak{d}_1(\phi) \geq \mathfrak{i}_1(\phi)$. Suppose first that $p > 2$. Since ϕ has maximal splitting, we have $\mathfrak{i}_1(\phi) > p^{n-1}$ by (1), and so $\mathfrak{d}_1(\phi) \geq p^n$. On the other hand, we have $\dim \phi_1 = \dim \phi - \mathfrak{i}_1(\phi) = p^n$ (see Remarks 2.36(2)). It follows that ϕ_1 is similar to an n-fold quasi-Pfister p-form, and so ϕ is a quasi-Pfister p-neighbor by Corollary 3.11. If $p = 2$, the same argument (and (2)) shows that ϕ_1 is a form of dimension 2^n which is divisible by an $(n-1)$-fold quasi-Pfister 2-form. Since every binary form is similar to a quasi-Pfister 2-form in this case, ϕ_1 is, in fact, similar to an n-fold quasi-Pfister 2-form, and we now conclude as before. □

Remark 6.9. The statement of Theorem 6.8 was originally conjectured in [Hoffmann 2004, Remark 7.32] (see also [Scully 2016, Question 7.6]). The result is the best possible, in the sense that one has examples of anisotropic quasilinear p-forms with maximal splitting which are not quasi-Pfister p-neighbors in every dimension omitted in the statement of the theorem (see [Hoffmann 2004, Example 7.31]). In the special case where $p = 2$, Theorem 6.8 was previously established in [Scully 2016, Theorem 9.6] using Proposition 3.16. Here, we obtain a more natural explanation of this phenomenon by way of Theorem 5.1. Note that the $p = 2$ case of the theorem is a direct analogue of a conjecture of Hoffmann in the theory of nonsingular quadratic forms which remains open, even over fields of characteristic different from 2 (see [Hoffmann 1995, §4; Izhboldin and Vishik 2000, Conjecture 1.6].

7. Further remarks on the splitting of quasilinear quadratic forms

Having determined all possible *standard* splitting patterns of quasilinear p-forms (Theorem 6.1, Proposition 6.4), we now turn our attention towards the problem of obtaining a similar result for the *full* splitting pattern. Here, we restrict our considerations to the case of quasilinear *quadratic* forms, where we can take direct inspiration from the following theorem of Vishik (which may be deduced from the existence of "excellent connections" in the integral Chow motives of anisotropic quadrics over fields of characteristic $\neq 2$; see [Vishik 2011, Theorem 1.3]):

Theorem 7.1 [Vishik 2011]. *Let ϕ be an anisotropic quadratic form of dimension ≥ 2 over a field k of characteristic $\neq 2$. Let* $\dim \phi - \mathfrak{i}_1(\phi) = 2^{r_1} - 2^{r_2} + \cdots + (-1)^{s-1} 2^{r_s}$ *for uniquely determined integers $r_1 > r_2 > \cdots > r_{s-1} > r_s + 1 \geq 1$. Let $1 \leq l \leq s$, and put $D_l = \sum_{i=1}^{l-1} (-1)^{i-1} 2^{r_i - 1} + \epsilon(l) \sum_{j=l}^{s} (-1)^{j-1} 2^{r_j}$, where $\epsilon(l) = 1$ (resp. $\epsilon(l) = 0$)*

if l is even (resp. odd). Then, for any field extension L of k, we either have $\mathrm{i}_0(\phi_L) \geq D_l + \mathrm{i}_1(\phi)$ *or* $\mathrm{i}_0(\phi_L) \leq D_l$.

Proof. This is nothing else but a restatement of [Vishik 2011, Proposition 2.6] in terms of Witt indices. To see how it may be derived from [Vishik 2011, Theorem 2.1] in further detail, we refer the reader to [Scully 2016, Proof of Theorem 1.2]. □

Examples 7.2. Let ϕ be an anisotropic quadratic form of dimension ≥ 2 over a field k of characteristic $\neq 2$, and write $\dim \phi = 2^n + m$ for uniquely determined integers $n \geq 0$ and $1 \leq m \leq 2^n$.

(1) For $l = 1$, Theorem 7.1 asserts that $\mathrm{i}_0(\phi_L) \geq \mathrm{i}_1(\phi)$ whenever ϕ_L is isotropic.

(2) For $l = 2$, Theorem 7.1 asserts that, for any field extension L of k, we either have $\mathrm{i}_0(\phi_L) \geq m$ or $\mathrm{i}_0(\phi_L) \leq m - \mathrm{i}_1(\phi_L)$.

(3) For $l = s$, Theorem 7.1 asserts that, for any field extension L of k, we either have $\mathrm{i}_0(\phi_L) \geq \frac{1}{2}(\dim \phi + \mathrm{i}_1(\phi) - 2^{r_s})$ or $\mathrm{i}_0(\phi_L) \leq \frac{1}{2}(\dim \phi - \mathrm{i}_1(\phi) - 2^{r_s})$. Taking $L = \bar{k}$ $\left(\text{so that } \mathrm{i}_0(\phi_L) = \left[\frac{1}{2} \dim \phi\right]\right)$, we see that $\mathrm{i}_1(\phi) \leq 2^{r_s}$, which gives a proof of Hoffmann's Conjecture 1.2 in this setting (see [Vishik 2011, Theorem 2.5.])

We expect, in fact, that Theorem 7.1 extends verbatim to our setting. Henceforth, let us assume that $p = 2$. We state the following:

Conjecture 7.3. Let ϕ be an anisotropic quasilinear quadratic form of dimension ≥ 2 over F, and write $\dim \phi - \mathrm{i}_1(\phi) = 2^{r_1} - 2^{r_2} + \cdots + (-1)^{s-1} 2^{r_s}$ for uniquely determined integers $r_1 > r_2 > \cdots > r_{s-1} > r_s + 1 \geq 1$. Let $1 \leq l \leq s$, and put $D_l = \sum_{i=1}^{l-1} (-1)^{i-1} 2^{r_i - 1} + \epsilon(l) \sum_{j=l}^{s} (-1)^{j-1} 2^{r_j}$, where $\epsilon(l) = 1$ (resp. $\epsilon(l) = 0$) if l is even (resp. odd). If L is any field extension of F, then we either have $\mathrm{i}_0(\phi_L) \geq D_l + \mathrm{i}_1(\phi)$ or $\mathrm{i}_0(\phi_L) \leq D_l$.

This expectation is partly justified by:

Proposition 7.4. *Conjecture 7.3 holds for $l \leq 2$.*

Proof. As in Examples 7.2(1) (resp. (2)), the $l = 1$ (resp. $l = 2$) case is nothing else but Lemma 4.1 (resp. Theorem 3.18). □

At present, we do not have a general approach to the $l > 2$ case of Conjecture 7.3. Using Theorem 1.3, however, we can provide further evidence for the $l = s$ case. First, it is worth stating here the following lemma:

Lemma 7.5. *In order to prove Conjecture 7.3, we may assume that* $\mathrm{i}_1(\phi_L) \geq \mathrm{i}_1(\phi)$.

Proof. Suppose that the statement of the conjecture fails to hold. In other words, suppose that $\mathrm{i}_1(\phi) > 1$ and $\mathrm{i}_0(\phi_L) = D_l + t$ for some $1 \leq t < \mathrm{i}_1(\phi)$. Note that we necessarily have $r_s > 0$ by Theorem 1.3. Now, let $\sigma = (\phi_L)_{\mathrm{an}}$. Since $D(\sigma)$ is spanned by elements of $D(\phi)$, there exists a subform $\rho \subset \phi$ such that $\sigma \simeq \rho_L$

(see Lemma 2.2). Let ψ be any codimension-$(t-1)$ subform of ϕ containing ρ. By construction, we have $D(\psi_L) \subseteq D(\phi_L) = D(\sigma) = D(\rho_L) \subseteq D(\psi_L)$, whence $D(\psi_L) = D(\sigma)$, or, equivalently, $(\psi_L)_{\mathrm{an}} \simeq \sigma$. In particular, we have $\mathfrak{i}_0(\psi_L) = D_l + 1$. On the other hand, since $t < \mathfrak{i}_1(\phi)$, ψ is a neighbor of ϕ, and so $\dim\psi - \mathfrak{i}_1(\psi) = \dim\phi - \mathfrak{i}_1(\phi) = 2^{r_1} - 2^{r_2} + \cdots + (-1)^{s-1}2^{r_s}$ (Proposition 3.10). We therefore conclude that the statement of the conjecture also fails for the triple (ψ, l, L). Since we are looking to produce a contradiction, we can replace ϕ by ψ in order to arrive at the case where $t = 1$. In this case, we claim that $\mathfrak{i}_1(\phi_L) \geq \mathfrak{i}_1(\phi)$. To see this, note first that $L(\phi)$ is L-isomorphic to a purely transcendental extension of $L(\sigma)$ (Remarks 2.6(3)). In view of Lemma 2.26, it follows that $\mathfrak{i}_1(\phi_L) = \mathfrak{i}_1(\sigma_{L(\phi)})$. In particular, we have $\mathfrak{i}_0(\phi_{L(\phi)}) = \mathfrak{i}_0(\phi_L) + \mathfrak{i}_0(\sigma_{L(\phi)}) = \mathfrak{i}_0(\phi_L) + \mathfrak{i}_1(\phi_L)$. The statement of Proposition 3.16 may therefore be rewritten here as

$$\mathfrak{i}_1(\phi_L) + \mathfrak{i}_0(\phi_L) - \mathfrak{i}_1(\phi) \geq \min\bigl\{\mathfrak{i}_0(\phi_L), \bigl[\tfrac{1}{2}(\dim\phi - \mathfrak{i}_1(\phi) + 1)\bigr]\bigr\}. \tag{7-1}$$

But, since $r_s > 0$, we have $\mathfrak{i}_0(\phi_L) = D_l + 1 \leq 2^{r_1-1} - 2^{r_2-1} + \cdots + (-1)^{s-1}2^{r_s-1} = \frac{1}{2}(\dim\phi - \mathfrak{i}_1(\phi))$. Inequality (7-1) therefore yields the desired assertion, and so the lemma is proved. □

We can now prove:

Proposition 7.6. *Conjecture 7.3 holds when $l = s$ and $L = F(Q)$ is the function field of any integral* (*affine or projective*) *quadric Q over F.*

Proof. In view of Lemma 2.26, the statement of the conjecture is stable under replacing F by any separable extension of itself. We may therefore assume that L is a purely inseparable quadratic extension of F. By Lemma 7.5, we may also assume that $\mathfrak{i}_1(\phi_L) \geq \mathfrak{i}_1(\phi)$. Suppose now that the statement fails to hold, so that $\mathfrak{i}_1(\phi) > 1$ and $\mathfrak{i}_0(\phi_L) = D_s + t$ for some $1 \leq t < \mathfrak{i}_1(\phi)$. We then have

$$\dim\,(\phi_L)_{\mathrm{an}} = \begin{cases} 2^{r_1-1} - 2^{r_2-1} + \cdots - 2^{r_{s-2}-1} + 2^{r_{s-1}-1} + \mathfrak{i}_1(\phi) - t & \text{if } s \text{ is even,} \\ 2^{r_1-1} - 2^{r_2-1} + \cdots - 2^{r_{s-1}-1} + 2^{r_s} + \mathfrak{i}_1(\phi) - t & \text{if } s \text{ is odd.} \end{cases}$$

Since $\mathfrak{i}_1(\phi_L) \geq \mathfrak{i}_1(\phi)$, and since $1 \leq \mathfrak{i}_1(\phi) - t < \mathfrak{i}_1(\phi)$, Theorem 1.3 implies that

$$\mathfrak{i}_1(\phi_L) \geq \begin{cases} 2^{r_{s-1}-1} + \mathfrak{i}_1(\phi) - t & \text{if } s \text{ is even,} \\ 2^{r_s} + \mathfrak{i}_1(\phi) - t & \text{if } s \text{ is odd,} \end{cases}$$

from which we conclude that

$$\mathfrak{i}_0(\phi_L) + \mathfrak{i}_1(\phi_L) \geq \begin{cases} 2^{r_1-1} - 2^{r_2-1} + \cdots + 2^{r_{s-1}-1} + \mathfrak{i}_1(\phi) & \text{if } s \text{ is even,} \\ 2^{r_1-1} - 2^{r_2-1} + \cdots + 2^{r_{s-2}-1} + \mathfrak{i}_1(\phi) & \text{if } s \text{ is odd.} \end{cases}$$

But $\mathfrak{i}_0(\phi_L)+\mathfrak{i}_1(\phi_L)=\mathfrak{i}_0(\phi_{L(\phi)})=\mathfrak{i}_1(\phi)+\mathfrak{i}_0((\phi_1)_{L(\phi)})$ (see the proof of Lemma 7.5), and so we have

$$\mathfrak{i}_0((\phi_1)_{L(\phi)}) \geq \begin{cases} 2^{r_1-1}-2^{r_2-1}+\cdots+2^{r_{s-1}-1} & \text{if } s \text{ is even,} \\ 2^{r_1-1}-2^{r_2-1}+\cdots+2^{r_{s-2}-1} & \text{if } s \text{ is odd.} \end{cases}$$

Either way, we see that $\mathfrak{i}_0((\phi_1)_{L(\phi)}) > 2^{r_1-1}-2^{r_2-1}+\cdots+(-1)^{s-1}2^{r_s-1} = \frac{1}{2}(\dim\phi-\mathfrak{i}_1(\phi))=\frac{1}{2}\dim\phi_1$. Since $L(\phi)$ is a purely inseparable quadratic extension of $F(\phi)$, this is impossible by Lemma 2.27(5), and so the result follows. □

In general, it suffices to prove Conjecture 7.3 in the case where L is a finite purely inseparable extension of F. Proposition 7.6 shows that the $l=s$ case of the conjecture holds for degree-2 purely inseparable extensions. More generally, the proposition covers the case where $\operatorname{ndeg}(\phi_L)=\frac{1}{2}\operatorname{ndeg}(\phi_L)$. Indeed, using Lemmas 2.26 and 2.27(3), one can easily reduce this case to that of a quadratic extension.

Acknowledgements

This work was carried out while I was a visiting postdoctoral fellow at Max-Planck-Institut für Mathematik in Bonn. I would like to thank this institution for its support and hospitality throughout the duration of my stay.

References

[EGA II 1961] A. Grothendieck, "Eléments de géométrie algébrique, II: Étude globale élémentaire de quelques classes de morphismes", *Inst. Hautes Études Sci. Publ. Math.* **8** (1961), 5–222. MR 0163909 Zbl 0118.36206

[Elman et al. 2008] R. Elman, N. Karpenko, and A. Merkurjev, *The algebraic and geometric theory of quadratic forms*, American Mathematical Society Colloquium Publications **56**, American Mathematical Society, Providence, RI, 2008. MR 2427530 Zbl 1165.11042

[Haution 2013] O. Haution, "On the first Steenrod square for Chow groups", *Amer. J. Math.* **135**:1 (2013), 53–63. MR 3022956 Zbl 1267.14012

[Haution 2015] O. Haution, "Detection by regular schemes in degree two", *Algebr. Geom.* **2**:1 (2015), 44–61. MR 3322197 Zbl 1322.14036

[Hoffmann 1995] D. W. Hoffmann, "Isotropy of quadratic forms over the function field of a quadric", *Math. Z.* **220**:3 (1995), 461–476. MR 1362256 Zbl 0840.11017

[Hoffmann 2004] D. W. Hoffmann, "Diagonal forms of degree p in characteristic p", pp. 135–183 in *Algebraic and arithmetic theory of quadratic forms*, edited by R. Baeza et al., Contemp. Math. **344**, American Mathematical Society, Providence, RI, 2004. MR 2058673

[Hoffmann and Laghribi 2004] D. W. Hoffmann and A. Laghribi, "Quadratic forms and Pfister neighbors in characteristic 2", *Trans. Amer. Math. Soc.* **356**:10 (2004), 4019–4053. MR 2058517 Zbl 1116.11020

[Hoffmann and Laghribi 2006] D. W. Hoffmann and A. Laghribi, "Isotropy of quadratic forms over the function field of a quadric in characteristic 2", *J. Algebra* **295**:2 (2006), 362–386. MR 2194958 Zbl 1138.11012

[Izhboldin 2004] O. T. Izhboldin, "Virtual Pfister neighbors and first Witt index", pp. 131–142 in *Geometric methods in the algebraic theory of quadratic forms*, edited by J.-P. Tignol, Lecture Notes in Math. **1835**, Springer, Berlin, 2004. MR 2066517 Zbl 1053.11033

[Izhboldin and Vishik 2000] O. Izhboldin and A. Vishik, "Quadratic forms with absolutely maximal splitting", pp. 103–125 in *Quadratic forms and their applications* (Dublin, 1999), edited by E. Bayer-Fluckiger et al., Contemp. Math. **272**, American Mathematical Society, Providence, RI, 2000. MR 1803363 Zbl 0972.11017

[Karpenko 2003] N. A. Karpenko, "On the first Witt index of quadratic forms", *Invent. Math.* **153**:2 (2003), 455–462. MR 1992018 Zbl 1032.11016

[Karpenko and Merkurjev 2003] N. Karpenko and A. Merkurjev, "Essential dimension of quadrics", *Invent. Math.* **153**:2 (2003), 361–372. MR 1992016 Zbl 1032.11015

[Knebusch 1976] M. Knebusch, "Generic splitting of quadratic forms, I", *Proc. London Math. Soc.* (3) **33**:1 (1976), 65–93. MR 0412101 Zbl 0351.15016

[Laghribi 2004a] A. Laghribi, "On splitting of totally singular quadratic forms", *Rend. Circ. Mat. Palermo* (2) **53**:3 (2004), 325–336. MR 2165184 Zbl 1097.11017

[Laghribi 2004b] A. Laghribi, "Quasi-hyperbolicity of totally singular quadratic forms", pp. 237–248 in *Algebraic and arithmetic theory of quadratic forms*, edited by R. Baeza et al., Contemp. Math. **344**, American Mathematical Society, Providence, RI, 2004. MR 2060200 Zbl 1143.11313

[Laghribi 2006] A. Laghribi, "The norm theorem for totally singular quadratic forms", *Rocky Mountain J. Math.* **36**:2 (2006), 575–592. MR 2234821 Zbl 1142.11020

[Laghribi 2007] A. Laghribi, "Sur le déploiement des formes bilinéaires en caractéristique 2", *Pacific J. Math.* **232**:1 (2007), 207–232. MR 2358037 Zbl 1214.11044

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics **211**, Springer, New York, 2002. MR 1878556 Zbl 0984.00001

[Merkurjev 2003] A. Merkurjev, "Steenrod operations and degree formulas", *J. Reine Angew. Math.* **565** (2003), 13–26. MR 2024643 Zbl 1091.14006

[Scully 2013] S. Scully, "Rational maps between quasilinear hypersurfaces", *Compos. Math.* **149**:3 (2013), 333–355. MR 3040743 Zbl 1315.11025

[Scully 2016] S. Scully, "On the splitting of quasilinear p-forms", *J. Reine Angew. Math.* **713** (2016), 49–83.

[Totaro 2008] B. Totaro, "Birational geometry of quadrics in characteristic 2", *J. Algebraic Geom.* **17**:3 (2008), 577–597. MR 2395138 Zbl 1144.11031

[Totaro 2009] B. Totaro, "Birational geometry of quadrics", *Bull. Soc. Math. France* **137**:2 (2009), 253–276. MR 2543476 Zbl 1221.14014

[Vishik 1998] A. Vishik, "Integral motives of quadrics", preprint 1998-13, Max Planck Institute for Mathematics, 1998, available at https://www.mpim-bonn.mpg.de/preblob/286.

[Vishik 1999] A. Vishik, "Direct summands in the motives of quadrics", preprint, 1999, available at https://www.maths.nottingham.ac.uk/personal/av/Papers/dirsumE.pdf.

[Vishik 2004] A. Vishik, "Motives of quadrics with applications to the theory of quadratic forms", pp. 25–101 in *Geometric methods in the algebraic theory of quadratic forms*, edited by J.-P. Tignol, Lecture Notes in Math. **1835**, Springer, Berlin, 2004. MR 2066515 Zbl 1047.11033

[Vishik 2011] A. Vishik, "Excellent connections in the motives of quadrics", *Ann. Sci. Éc. Norm. Supér.* (4) **44**:1 (2011), 183–195. MR 2760197 Zbl 1223.14005

Communicated by Raman Parimala
Received 2015-08-07 Revised 2016-02-24 Accepted 2016-03-24

sscully@ualberta.ca *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton AB T6G 2G1, Canada*

ALGEBRA AND NUMBER THEORY 10:5 (2016)
dx.doi.org/10.2140/ant.2016.10.1133

K3 surfaces over finite fields with given L-function

Lenny Taelman

The zeta function of a K3 surface over a finite field satisfies a number of obvious (archimedean and ℓ-adic) and a number of less obvious (p-adic) constraints. We consider the converse question, in the style of Honda–Tate: given a function Z satisfying all these constraints, does there exist a K3 surface whose zeta-function equals Z? Assuming semistable reduction, we show that the answer is yes if we allow a finite extension of the finite field. An important ingredient in the proof is the construction of complex projective K3 surfaces with complex multiplication by a given CM field.

Introduction

Let X be a K3 surface over $\mathbb{F}_q$. The zeta function of X has the form

$$Z(X/\mathbb{F}_q, T) = \frac{1}{(1-T)L(X/\mathbb{F}_q, qT)(1-q^2T)}$$

where the polynomial $L(X/\mathbb{F}_q)$ is defined by

$$L(X/\mathbb{F}_q, T) := \det\bigl(1 - T\,\mathrm{Frob}, \mathrm{H}^2(X_{\bar{\mathbb{F}}_q}, \mathbb{Q}_\ell(1))\bigr) \in \mathbb{Q}[T].$$

We have $L(X/\mathbb{F}_q, T) = \prod_{i=1}^{22}(1-\gamma_i T)$ with the γ_i of complex absolute value 1. The polynomial $L(X/\mathbb{F}_q, T)$ factors in $\mathbb{Q}[T]$ as $L = L_{\mathrm{alg}} L_{\mathrm{trc}}$ with

$$L_{\mathrm{alg}}(X/\mathbb{F}_q, T) = \prod_{\gamma_i \in \mu_\infty} (1 - T\gamma_i), \quad L_{\mathrm{trc}}(X/\mathbb{F}_q, T) = \prod_{\gamma_i \notin \mu_\infty} (1 - T\gamma_i),$$

where μ_∞ is the group of complex roots of unity.

Theorem 1. *Let X be a K3 surface over $\mathbb{F}_q$ with $q = p^a$. Assume that X is not supersingular. Then*

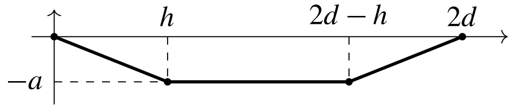
(1) *all complex roots of $L_{\mathrm{trc}}(X/\mathbb{F}_q, T)$ have absolute value* 1;

(2) *no root of $L_{\mathrm{trc}}(X/\mathbb{F}_q, T)$ is a root of unity*;

MSC2010: primary 14J28; secondary 14G15, 14K22, 11G25.
Keywords: K3 surfaces, zeta functions, finite fields.

(3) $L_{\mathrm{trc}}(X/\mathbb{F}_q, T) \in \mathbb{Z}_\ell[T]$ *for all* $\ell \neq p$;

(4) *the Newton polygon of* $L_{\mathrm{trc}}(X/\mathbb{F}_q, T)$ *at* p *is of the form*



with h *and* d *integers satisfying* $1 \leq h \leq d \leq 10$;

(5) $L_{\mathrm{trc}}(X/\mathbb{F}_q, T) = Q^e$ *for some* $e > 0$ *and some irreducible* $Q \in \mathbb{Q}[T]$, *and* Q *has a unique irreducible factor in* $\mathbb{Q}_p[T]$ *with negative slope.*

The above theorem collects results of Deligne, Artin, Mazur, Yu and Yui, and slightly expands on these, see Section 1 for the details. The integer h in the theorem is the height of X (which is finite by the assumption that X is not supersingular), and assuming the Tate conjecture (which is now known in almost all cases [Charles 2013; 2014; Madapusi Pera 2015]) the Picard rank of $X_{\bar{\mathbb{F}}_q}$ is $22 - 2d$.

Definition 2 (Property ($\star$)). A K3 surface X over a finite extension k of $\mathbb{Q}_p$ is said to satisfy ($\star$) if there exists a finite extension $k \subset \ell$ and a proper flat algebraic space $\mathfrak{X} \to \operatorname{Spec} \mathcal{O}_\ell$ such that

(1) $\mathfrak{X} \times_{\operatorname{Spec} \mathcal{O}_\ell} \operatorname{Spec} \ell \cong X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell$,

(2) $\mathfrak{X}$ is regular,

(3) the special fiber of $\mathfrak{X}$ is a reduced normal crossings divisor with smooth components,

(4) $\omega_{\mathfrak{X}/\mathcal{O}_\ell} \cong \mathcal{O}_{\mathfrak{X}}$.

Property ($\star$) is a strong form of potential semistability. It is expected that every X satisfies ($\star$), but this is presently only known for special classes of K3 surfaces, see [Maulik 2014, §4] and [Liedtke and Matsumoto 2015, §2]. Our main result is the following partial converse to Theorem 1.

Theorem 3. *Assume every K3 surface* X *over a* p*-adic field satisfies* ($\star$). *Let*

$$L = \prod_{i=1}^{2d} (1 - \gamma_i T) \in 1 + T\mathbb{Q}[T]$$

be a polynomial which satisfies properties (1)–(5) *of Theorem 1. Then there exists a positive integer* n *and a K3 surface* X *over* $\mathbb{F}_{q^n}$ *such that*

$$L_{\mathrm{trc}}(X/\mathbb{F}_{q^n}, T) = \prod_{i=1}^{2d} (1 - \gamma_i^n T).$$

The proof of Theorem 3 follows the same strategy as the proof of the Honda–Tate theorem [Tate 1971]: given L_{trc}, one constructs a K3 surface over a finite field by first producing a complex projective K3 surface with CM by a suitably chosen CM field, then descending it to a number field, and finally reducing it to the residue field at a suitably chosen prime above p. In the final step a criterion of good reduction is needed, which has been obtained recently by Matsumoto [2015] and Liedtke and Matsumoto [2015], under the assumption ($\star$).

A crucial intermediate result, that may be of independent interest, is the following theorem.

Theorem 4. *Let E be a CM field with $[E : \mathbb{Q}] \le 20$. Then there exists a K3 surface over $\mathbb{C}$ with CM by E.*

See Section 2 for the definition of "CM by E", and see Section 3 for the proof of this theorem.

Remark 5. I do not know if one can take $n = 1$ in Theorem 3. Finite extensions are used in several parts of the proof, both in constructing a K3 surface X over some finite field, and in verifying that the action of Frobenius on H^2 is the prescribed one.

Recently Kedlaya and Sutherland [2015] obtained some computational evidence suggesting that the theorem might hold with $n = 1$. They enumerated all polynomials L satisfying (1)–(5) with $q = 2$, $\deg L = \deg Q = 20$ and with $L(1) = 2$ and $L(-1) \ne 2$. There are 1995 such polynomials. If $L = L_{\mathrm{trc}}(X/\mathbb{F}_2, T)$ for a K3 surface over $\mathbb{F}_2$, then the Artin–Tate formula [Milne 1975; Elsenhans and Jahnel 2015] puts strong restrictions on the Néron–Severi lattice of X. These restrictions suggest that X should be realizable as a smooth quartic, and indeed for each of the 1995 polynomials Kedlaya and Sutherland manage to identify a smooth quartic X defined over $\mathbb{F}_2$ with $L = L_{\mathrm{trc}}(X/\mathbb{F}_2, T)$.

If one can take $n = 1$ in Theorem 3, then new ideas will be needed to prove this. Indeed, there is no reason at all that the X constructed in the current proof is defined over $\mathbb{F}_q$. A similar problem occurs in the proof of the Honda-Tate theorem [Tate 1971]: given a q-Weil number one first constructs an abelian variety over a finite extension of $\mathbb{F}_q$, and then identifies the desired abelian variety as a simple factor of the Weil restriction to $\mathbb{F}_q$. Perhaps a variation of this argument in the context of hyperkähler varieties can be made to work in our setting?

Remark 6. By the work of Madapusi Pera [2015], for every d there is an étale map $M_{2d} \to \mathrm{Sh}_{2d}$ from the moduli space of quasipolarized K3 surfaces of degree $2d$ to a an integral model of a certain Shimura variety, over $\mathbb{Z}[1/2]$. It is surjective over $\mathbb{C}$, and assuming ($\star$), one can deduce from the criterion of Liedtke and Matsumoto that it is surjective on $\bar{\mathbb{F}}_p$-points. In odd characteristic, Kottwitz [1990] and Kisin [2013] have given a group-theoretic description of the isogeny classes in $\mathrm{Sh}_{2d}(\bar{\mathbb{F}}_p)$, for every d. With arguments similar to those in Section 3, it should be possible to deduce Theorem 1 and Theorem 3 from the above results.

1. p-adic properties of zeta functions of K3 surfaces

1.1. ***Recap on the formal Brauer group of a K3 surface.*** Let X be a K3 surface over a field k. Artin and Mazur [1977] have shown that the functor

$$R \mapsto \ker(\operatorname{Br} X_R \to \operatorname{Br} X)$$

on Artinian k-algebras is prorepresentable by a (one-dimensional) formal group $\hat{\operatorname{Br}}\, X$ over k. This formal group is called the *formal Brauer group* of X.

Assume now that k is a perfect field of characteristic $p > 0$ and that X is not supersingular. Then $\hat{\operatorname{Br}}\, X$ has finite height h satisfying $1 \leq h \leq 10$. We denote by $\mathbb{D}(\hat{\operatorname{Br}}\, X)$ the (covariant) Dieudonné module of $\hat{\operatorname{Br}}\, X$. This has the structure of an F-crystal over k. It is free of rank h over the ring W of Witt vectors of k.

We denote by $\mathrm{H}^2_{\mathrm{crys}}(X/W)_{<1}$ the maximal sub-F-crystal of $\mathrm{H}^2_{\mathrm{crys}}(X/W)$ that has all slopes < 1.

Proposition 7. *If X is not supersingular, then there is a canonical isomorphism*

$$\mathrm{H}^2_{\mathrm{crys}}(X/W)_{<1} \cong \mathbb{D}(\hat{\operatorname{Br}}\, X)$$

of F-crystals over k.

Proof. By [Illusie 1979, §7.2] there is a canonical isomorphism of F-crystals

$$\mathrm{H}^2_{\mathrm{crys}}(X/W) = \mathrm{H}^2(X, \mathrm{W}\mathcal{O}_X) \oplus \mathrm{H}^1(X, \mathrm{W}\Omega^1_{X/k}) \oplus \mathrm{H}^0(X, \mathrm{W}\Omega^2_{X/k}), \tag{1}$$

coming from the de Rham–Witt complex, and by [Artin and Mazur 1977, Corollary 4.3] we have an isomorphism of F-crystals

$$\mathrm{H}^2(X, \mathrm{W}\mathcal{O}_X) = \mathbb{D}(\hat{\operatorname{Br}}\, X).$$

Since $\hat{\operatorname{Br}}\, X$ is a formal group, the slopes of $\mathrm{H}^2(X, \mathrm{W}\mathcal{O}_X) = \mathbb{D}(\hat{\operatorname{Br}}\, X)$ are < 1. On the other hand, since F is divisible by p^i on $\mathrm{H}^{2-i}(X, \mathrm{W}\Omega^i_{X/k})$, the slopes of the other summands in (1) are ≥ 1. This proves the theorem. □

1.2. ***Proof of Theorem 1.***

Proof. Property (2) holds by definition, (3) is a formal consequence of the trace formula in ℓ-adic cohomology (see, e.g., [Deligne 1974, §1]), and (1) is part of the Weil conjectures [Deligne 1972; 1974].

The other properties make use of crystalline cohomology. Property (4) is well, known. It follows for example from Mazur's proof of "Newton above Hodge" [Mazur 1972; 1973] for liftable varieties with torsion-free cohomology, see [Mazur

1972, §2]. Property (5) is a sharpening of a result of Yu and Yui [2008, Proposition 3.2]. The argument is essentially the same as in [loc. cit.], we repeat it for completeness.

For a polynomial $Q = \prod(1-\gamma_i T) \in \mathbb{Q}_p[T]$ we denote by $Q_{<0}$ the product

$$Q_{<0} = \prod_{v_p(\gamma_i)<0} (1-\gamma_i T) \in \mathbb{Q}_p[T].$$

Let K be the field of fractions of W. If $q = p^a$, then by Proposition 7 we have

$$L_{\mathrm{trc},<0} := L_{\mathrm{trc},<0}(X/\mathbb{F}_q, T) = \det_K\bigl(1 - F^a T,\ K \otimes_W \mathbb{D}(\hat{\mathrm{B}}\mathrm{r}\, X)(1)\bigr)$$

in $\mathbb{Q}_p[T] \subset K[T]$. Since $\hat{\mathrm{B}}\mathrm{r}\, X$ is a one-dimensional formal group of finite height, the crystal $\mathbb{D}(\hat{\mathrm{B}}\mathrm{r}\, X)$ is indecomposable. It follows that the endomorphism F^a of $\mathbb{D}(\hat{\mathrm{B}}\mathrm{r}\, X)$ has an irreducible minimum polynomial over K, and hence $L_{\mathrm{trc},<0} = P^e_{<0}$ for some irreducible $P_{<0} \in \mathbb{Q}_p(T)$. Let Q be an irreducible factor of L_{trc}. Then Q has a reciprocal root γ with $v_p(\gamma) < 0$, for otherwise the roots of Q would be algebraic integers and hence roots of unity. In particular $Q_{<0} = P_{<0}$. Apparently any two irreducible factors of L_{trc} share a common root, hence $L_{\mathrm{trc}} = Q^e$. This proves (5). □

2. CM theory of K3 surfaces

This section collects results of Zarhin, Shafarevich and Rizov.

2.1. ***Hodge theoretic aspects.*** For a projective K3 surface X over $\mathbb{C}$ we denote by $\mathrm{NS}(X)$ its Néron–Severi group and by $T(X) \subset \mathrm{H}^2(X, \mathbb{Z}(1))$ the transcendental lattice, i.e., $T(X)$ is the orthogonal complement of $\mathrm{NS}(X)$. We have a decomposition

$$\mathrm{H}^2(X, \mathbb{Q}(1)) = \mathrm{NS}(X)_{\mathbb{Q}} \oplus T(X)_{\mathbb{Q}}.$$

The Hodge structure $T(X)_{\mathbb{Q}}$ is irreducible [Zarhin 1983, Theorem 1.4.1]. The cup product pairing defines even symmetric bilinear forms on $\mathrm{NS}(X)$ and $T(X)$ of signature $(1, \rho-1)$ and $(2, 20-\rho)$, with $\rho = \mathrm{rk}\,\mathrm{NS}(X)$.

Proposition 8 [Zarhin 1983, §2]. *Let X be a projective $K3$ surface over $\mathbb{C}$. Then the following are equivalent*:

(1) *The Hodge group of $T(X)_{\mathbb{Q}}$ is commutative.*

(2) $E := \mathrm{End}_{\mathrm{HS}}\, T(X)_{\mathbb{Q}}$ *is a CM field and* $\dim_E T(X)_{\mathbb{Q}} = 1$. □

Definition 9. If X satisfies the equivalent conditions (1) and (2) of Proposition 8, then we say that X is a K3 surface *with CM* (by E).

Remark 10. Another equivalent condition is that $T(X)_{\mathbb{Q}}$ is contained in the Tannakian category of Hodge structures generated by the H^1 of CM abelian varieties.

If E is a CM field, then we denote the canonical complex conjugation of E by $z \mapsto \bar{z}$, and its fixed field by E_0. We have $[E : E_0] = 2$, and E_0 is a totally real number field.

Proposition 11. *Let X be a K3 surface with CM by E. Then*

(1) *$ax \cdot y = x \cdot \bar{a}y$ for all $a \in E$ and $x, y \in T(X)_{\mathbb{Q}}$;*

(2) *the group of Hodge isometries of $T(X)_{\mathbb{Q}}$ is* $\ker(\mathrm{Nm} : E^\times \to E_0^\times)$.

Proof. The cup product pairing induces an isomorphism

$$T(X)_{\mathbb{Q}} \xrightarrow{\sim} \mathrm{Hom}(T(X)_{\mathbb{Q}}, \mathbb{Q})$$

of Hodge structures, and hence the action of E on $T(X)$ induces an "adjoint" homomorphism $\varphi : E \to E$ such that $ax \cdot y = x \cdot \varphi(a)y$. Considering the induced action on $\mathrm{H}^{0,2}(X)$ one sees that $\varphi(a) = \bar{a}$, which proves the first assertion. The second is an immediate consequence of the first. □

2.2. *Arithmetic aspects: the Main Theorem of CM.* Let X be a K3 surface over $\mathbb{C}$ with CM by E. Consider the algebraic torus G over $\mathbb{Q}$ which is the kernel of the norm map $E^\times \to E_0^\times$ (seen as map of tori over $\mathbb{Q}$). Then $G(\mathbb{Q})$ is the group of E-linear isometries of $T(X)_{\mathbb{Q}}$.

If X is defined over a subfield $k \subset \mathbb{C}$, then we have canonical isomorphisms

$$\mathrm{H}^2_{\mathrm{et}}(X_{\bar{k}}, \mathbb{Q}_\ell(1)) = \mathrm{H}^2(X(\mathbb{C}), \mathbb{Q}(1)) \otimes_{\mathbb{Q}} \mathbb{Q}_\ell.$$

Since the Galois action on the left-hand side respects the intersection pairing and the subgroup $\mathrm{NS}(X_{\bar{k}}) = \mathrm{NS}(X_{\mathbb{C}})$, we see that both Gal_k and $G(\mathbb{Q}_\ell)$ act on $T(X)_{\mathbb{Q}_\ell}$. If we denote by $\mathbb{A}_f$ the finite adèles of $\mathbb{Q}$, i.e., $\mathbb{A}_f = \mathbb{Q} \otimes \hat{\mathbb{Z}}$, then we obtain actions of Gal_k and $G(\mathbb{A}_f)$ on $T(X)_{\mathbb{A}_f}$.

Theorem 12 (Main Theorem of CM for K3 surfaces [Rizov 2010]). *There exists a number field $k \subset \mathbb{C}$ containing E such that*

(1) *X is defined over k,*

(2) *the Galois action on $T(X)_{\mathbb{A}_f}$ factors over a map $\rho : \mathrm{Gal}_k \to G(\mathbb{A}_f)$*

(3) *the diagram*

$$\begin{tikzcd} \mathrm{Gal}_k \arrow[r, tail] \arrow[d, "\rho"] & \mathrm{Gal}_E \arrow[r, two heads, "CFT"] & \mathbb{A}_{E,f}^\times / E^\times \arrow[d, "z \mapsto \bar{z}/z"] \\ G(\mathbb{A}_f) \arrow[rr] & & G(\mathbb{A}_f)/G(\mathbb{Q}) \end{tikzcd}$$

commutes.

Proof. This is a reformulation of [Rizov 2010, Corollary 3.9.2]. Note however that the stated definition of complex multiplication [Rizov 2010, Definition 1.4.3] needs to be corrected (the condition $\dim_E T_{\mathbb{Q}} = 1$ is missing) for the proof and statement to be correct. □

Remark 13. A priori, the moduli space of polarized complex K3 surfaces has two natural models over $\mathbb{Q}$: the "canonical model" of the theory of Shimura varieties [Deligne 1971, §3], which is defined in terms of the Galois action on special points, and the model coming from the moduli interpretation. The essential content of Theorem 12 is that these two models coincide. (See also [Madapusi Pera 2015, §3]).

3. Existence of K3 surface with CM by a given CM field

In this section we prove Theorem 4. By the surjectivity of the period map for K3 surfaces, this reduces to a problem about quadratic forms over $\mathbb{Q}$.

3.1. ***Invariants of quadratic forms over*** $\mathbb{Q}$. We quickly recall some basic facts about quadratic forms over $\mathbb{Q}$. We refer to [Cassels 1978; Scharlau 1985; Serre 1970] for details and proofs. Let k be a field of characteristic different from 2. A *quadratic space* over k is a pair $V = (V, q)$ consisting of a finite-dimensional vector space over k and a nondegenerate symmetric bilinear form $q : V \times V \to k$. To such a space one associates the following invariants:

(1) the *dimension* $\dim(V)$;

(2) the *determinant* $\det(V) \in k^\times/k^{\times 2}$;

(3) the *Hasse invariant* $w(V) \in \mathrm{Br}(k)[2]$.

Any form V over k is isomorphic to a diagonal form $\langle \alpha_1, \ldots, \alpha_n \rangle$ with $n = \dim V$, and for such a form the invariants are

$$\det(V) = \prod_i \alpha_i \qquad \in k^\times/k^{\times 2},$$

$$w(V) = \sum_{i<j} (\alpha_i, \alpha_j)_k \quad \in \mathrm{Br}(k)[2],$$

where $(\alpha, \beta)_k$ denotes the class of the quaternion algebra generated by i and j with $i^2 = \alpha$, $j^2 = \beta$, $ij = -ji$.

We denote the orthogonal sum of two quadratic spaces by $V \oplus W$.

Lemma 14. *Let V and W be quadratic spaces over k. Then*

(1) $\det(V \oplus W) = \det(V)\det(W)$;

(2) $w(V \oplus W) = w(V) + w(W) + (\det(V), \det(W))_k$.

Proof. This follows from the above formulas for the determinant and Hasse invariant of a diagonal quadratic form, and the bilinearity of $(\alpha, \beta)_k$. □

Theorem 15. *Two forms over $\mathbb{Q}_p$ are isomorphic if and only if they have the same dimension, determinant and Hasse invariant. For every $d \geq 3$, $\delta \in \mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2}$ and $w \in \mathrm{Br}(\mathbb{Q}_p)[2]$ there exists a form of dimension d, determinant δ and Hasse invariant w.* □

If $k = \mathbb{Q}$ then a fourth invariant is given by the signature of the form $V_\mathbb{R}$.

Theorem 16. *Two forms over $\mathbb{Q}$ are isomorphic if and only if they have the same signature, determinant, and Hasse invariant. All forms V over $\mathbb{Q}$ of signature (r, s) satisfy*

(1) *the sign of $\delta(V)$ is $(-1)^s$;*

(2) *the image of $w(V)$ in $\mathrm{Br}(\mathbb{R})[2] = \mathbb{Z}/2\mathbb{Z}$ is $s(s-1)/2 \bmod 2$.*

If $r + s \geq 3$, and if δ and w satisfy (1) and (2) above, then there exists a quadratic space over $\mathbb{Q}$ with signature (r, s), determinant δ and Hasse invariant w. □

Lemma 17. *For $\Lambda_{K3,\mathbb{Q}} = \mathbb{Q} \otimes_\mathbb{Z} \Lambda_{K3}$, the Hasse invariant $w(\Lambda_{K3,\mathbb{Q}}) \in \mathrm{Br}(\mathbb{Q})[2]$ is the class of the quaternion algebra $(-1,-1)_\mathbb{Q}$, and $\det(\Lambda_{K3,\mathbb{Q}}) = -1$.*

Proof. We have $\Lambda_{K3} \cong (-E_8) \oplus (-E_8) \oplus U \oplus U \oplus U$ where U is the standard hyperbolic plane. Using this explicit description, one computes (over $\mathbb{Q}$) an orthogonal basis, and computes the invariants using the formula for diagonal forms. □

3.2. ***The form q_λ.*** Let E be a CM field with maximal totally real subfield E_0. Put $d := [E_0 : \mathbb{Q}]$. Denote by $z \mapsto \bar{z}$ the complex conjugation on E. For $\lambda \in E_0^\times$ the map

$$q_\lambda : E \times E \mapsto \mathbb{Q}, \quad (x, y) \mapsto \mathrm{tr}_{E_0/\mathbb{Q}}(\lambda x \bar{y})$$

is a nondegenerate symmetric bilinear form over $\mathbb{Q}$.

We denote the discriminant of the number field E by $\Delta(E/\mathbb{Q})$.

Lemma 18 [Bayer-Fluckiger 2014, Lemma 1.3.2].

$$\det(q_\lambda) = (-1)^d\, \Delta(E/\mathbb{Q}) \quad \text{in } \mathbb{Q}^\times/\mathbb{Q}^{\times 2}.$$

Lemma 19. *If $\lambda \in E_0^\times$ has signature (r, s), then q_λ has signature $(2r, 2s)$.* □

3.3. ***Construction of a K3 surface with CM by E.*** A key ingredient in the proof of Theorem 4 is the following proposition on rational quadratic forms. I am grateful to Eva Bayer for pointing me to her work on maximal tori in orthogonal groups [Bayer-Fluckiger 2014], and for explaining how it simplifies an earlier version of the proof below.

Proposition 20. *Let E be a CM field with maximal totally real subfield E_0, and assume $d := [E_0 : \mathbb{Q}] \leq 10$. Then there exists a $\lambda \in E_0^\times$ of signature $(1, d-1)$ and a quadratic space V such that, as quadratic spaces over $\mathbb{Q}$,*

$$(E, q_\lambda) \oplus V \cong \Lambda_{K3,\mathbb{Q}}$$ □

Proof. If $d < 10$ then we claim that for every choice of λ a complement V exists. Indeed, given a choice of λ, then the dimension, signature, determinant and Hasse invariant of V are determined by Lemma 14. These invariants satisfy conditions (1) and (2) of Theorem 16 because they are satisfied by the invariants of (E, q_λ) and $\Lambda_{K3,\mathbb{Q}}$. Since $\dim(V) > 2$, the theorem then guarantees the existence of a form V with $(E, q_\lambda) \oplus V \cong \Lambda_{K3,\mathbb{Q}}$.

So we assume $d = 10$. Let $\delta = \Delta(E/\mathbb{Q}) \in \mathbb{Q}^\times/\mathbb{Q}^{\times 2}$. Note that $\delta > 0$ (since d is even). Consider the diagonal quadratic space $V = \langle -1, \delta\rangle$. By the same reasoning as above, there exists a unique quadratic space W of dimension 20 such that

$$W \oplus \langle -1, \delta\rangle \cong \Lambda_{K3,\mathbb{Q}}.$$

We will show that W can be realized as (E, q_λ) for a suitable choice of $\lambda \in E_0^\times$. Note that W has signature $(2, 18)$, so by Lemma 19 the scalar λ will automatically have signature $(1, 9)$.

By Corollary 4.0.3 and Proposition 1.3.1 of [Bayer-Fluckiger 2014], there exists a λ with $(E, q_\lambda) \cong W$ if and only if the following three conditions hold:

(1) the signature of W is even;

(2) $\mathrm{disc}(W) = \delta$;

(3) for every prime p such that all places of E_0 above p split in E, we have that $W_{\mathbb{Q}_p}$ is isomorphic to an orthogonal sum of 10 hyperbolic planes.

Our W clearly satisfies the first two conditions. For the third, consider a prime p such that all places of E_0 above p split in E. Then the image of δ in $\mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2}$ is 1. Together with Lemma 14 and Lemma 17 this allows us to compute the invariants of $W_{\mathbb{Q}_p}$, and we find $\det(W_{\mathbb{Q}_p}) = 1$ and $w(W_{\mathbb{Q}_p}) = (-1, -1)_{\mathbb{Q}_p}$. These are the same as the invariants for 10 copies of the hyperbolic plane, so with Theorem 15 we see that W satisfies the third condition, which finishes the proof of the proposition. □

Finally, we show that for every CM field E of degree at most 20 there exists a projective K3 surface X with CM by E.

Proof of Theorem 4. Choose $\lambda \in E_0$ and V as in Proposition 20. This guarantees that there exists an integral lattice

$$\Lambda \subset (E, q_\lambda) \oplus V$$

with $\Lambda \cong \Lambda_{K3}$. Choose such a Λ, and choose an embedding $\epsilon : E \hookrightarrow \mathbb{C}$ with $\epsilon(\lambda) > 0$. Then we have a splitting

$$\Lambda_{\mathbb{C}} = \mathbb{C}_\epsilon \oplus \mathbb{C}_{\bar{\epsilon}} \oplus (\oplus_{\sigma \neq \epsilon, \bar{\epsilon}} \mathbb{C}_\sigma) \oplus V_{\mathbb{C}}.$$

We make Λ into a pure $\mathbb{Z}$-Hodge structure of weight 0 by declaring $\mathbb{C}_\epsilon$ to be of type $(1, -1)$, its conjugate $\mathbb{C}_{\bar{\epsilon}}$ of type $(-1, 1)$, and all the other terms of type $(0, 0)$.

By construction, the bilinear form $\Lambda \otimes \Lambda \to \mathbb{Z}$ is a morphism of Hodge structures. Note that E acts on $E \subset \Lambda_{\mathbb{Q}}$ via Hodge structure endomorphisms, so that E is irreducible and hence

$$\Lambda^{0,0} \cap \Lambda_{\mathbb{Q}} = V.$$

For every nonzero $z \in \mathrm{H}^{2,0}$ we have $z \cdot \bar{z} \in \mathbb{R}_{>0}$ since $\epsilon(\lambda) > 0$, so that the surjectivity of the period map [Todorov 1980] gives the existence of a complex analytic K3 surface X and a Hodge isometry $\Lambda \cong \mathrm{H}^2(X, \mathbb{Z}(1))$. A priori, it may not be clear that X is algebraic. However, as $\mathrm{Pic}(X)_{\mathbb{Q}} \cong V$ has signature $(1, 21-2d)$, there exists an $h \in \mathrm{Pic}(X)$ with $h \cdot h > 0$. By [Barth et al. 2004, Theorem IV.6.2] this implies that the surface X is projective. By construction, X is a K3 surface with CM by E. □

Remark 21. A similar construction has been used by Piatetski-Shapiro and Shafarevich [1973, §3] in showing the existence of some K3 surfaces with CM. The new ingredients that allow us to obtain a stronger result are the use of rational (as opposed to integral) quadratic forms, the results of Bayer on quadratic forms q_λ, and the use of the algebraicity criterion from [Barth et al. 2004], which avoids the delicate question of identifying an ample $h \in \mathrm{Pic}(X)$.

4. Existence of K3 surface with given L_{trc}

In this section we will prove Theorem 3. So let

$$L = \prod_{i=1}^{2d} (1 - \gamma_i T) \in 1 + T\mathbb{Q}[T]$$

be a polynomial satisfying properties (1)–(5) of Theorem 1. Consider the number field $F := \mathbb{Q}(\gamma_1)$.

Lemma 22. *F is a CM field and $\bar{\gamma}_1 \gamma_1 = 1$.*

Proof. The image γ of γ_1 under any homomorphism $F \to \mathbb{C}$ satisfies $|\gamma| = 1$, hence $\bar{\gamma} = \gamma^{-1}$. Moreover γ cannot be real, since then $\gamma = \pm 1$, contradicting the fact that γ_1 is not a root of unity. It follows that F is a CM field with complex conjugation $\gamma_1 \mapsto \gamma_1^{-1}$. □

By property (5), the number field F has a unique valuation v above the prime p such that $v(\gamma_1) < 0$.

Lemma 23. *There exists an extension E of F with $[E : \mathbb{Q}] = 2d$, and such that*

(1) *E is a CM field*;

(2) *the valuation v has a unique extension to E.*

Proof. Let F_0 be the maximal totally real subfield of F. Let v_0 be the place of F_0 under v. Now choose a polynomial $P(X) \in F_0[X]$ such that

(1) $\deg P = e$;

(2) P has e real roots for every embedding $F_0 \hookrightarrow \mathbb{R}$;

(3) P is irreducible in $(F_0)_{v_0}[X]$.

Note that v_0 splits in F, since by the preceding lemma $\bar{v}(\gamma_1) > 0$ and hence $\bar{v} \neq v$. In particular $P(X)$ is irreducible in $F_v[X]$, and it follows that $E := F[X]/P(X)$ is a field satisfying the desired conditions. □

We fix an E satisfying the conditions of the lemma. Abusing notation, we will denote the unique extension of v to E by the same symbol v.

Lemma 24. $[E_v : \mathbb{Q}_p] = h$.

Proof. Since $L = Q^e$, and since v is the unique place with $v(\gamma_1) < 0$, we see from properties (4) and (5) in Theorem 1 that $[F_v : \mathbb{Q}_p] = h/e$. But $[E : F] = e$ and v has a unique extension to E, hence $[E_v : \mathbb{Q}_p] = h$. □

Let X be a K3 surface over $\mathbb{C}$ with CM by E. By the Main Theorem of CM (Theorem 12) this surface is defined over a number field k containing E. Let w be a place of k lying above v. We extend the commutative diagram of Theorem 12 to include the local-global compatibility of class field theory:

$$\begin{tikzcd}
W_{k_w} \arrow[r, tail] \arrow[d, tail] & W_{E_v} \arrow[r, two heads, "\mathrm{LCFT}"] \arrow[d, tail] & E_v^\times \arrow[d] \\
\mathrm{Gal}_k \arrow[r, tail] \arrow[d, "\rho"] & \mathrm{Gal}_E \arrow[r, two heads, "\mathrm{GCFT}"] & \mathbb{A}_{E,f}^\times/E^\times \arrow[d, "z \mapsto z/\bar{z}"] \\
G(\mathbb{A}_f) \arrow[rr] & & G(\mathbb{A}_f)/G(\mathbb{Q})
\end{tikzcd} \tag{2}$$

Here $W_{k_w} \subset \mathrm{Gal}_{k_w}$ denotes the Weil group of the local field k_w. Extending k if necessary, we may assume that the residue field $\mathbb{F}_w$ is an extension of $\mathbb{F}_q$.

Choose a prime $\ell \neq p$. Then the image of inertia I_{k_w} in $G(\mathbb{Z}_\ell)$ is finite, hence after replacing k by a finite extension, we may assume that the action of Gal_{k_w} on $\mathrm{H}^2_{\mathrm{et}}(X_{\bar{k}}, \mathbb{Q}_\ell(1))$ is unramified.

Now assume X_{k_w} satisfies ($\star$). Then, replacing k once more by a finite extension, we may assume by the criterion of Liedtke and Matsumoto [2015, Theorem 2.5] that X has good reduction at w. Let $\bar{X}/\mathbb{F}_w$ be the reduction of X/k at w.

Let $\sigma \in W_{k_w}$ be a Frobenius element. Note that γ_1 lies in $G(\mathbb{Q}) \subset E^\times$.

Proposition 25. *There is an* $m > 0$ *such that for all* $\ell \neq p$ *we have in* $G(\mathbb{Q}_\ell)$

$$\rho(\sigma^m)_\ell = \gamma_1^{m[\mathbb{F}_w:\mathbb{F}_q]}.$$

Proof. Let $\pi \in E_v^\times$ be the image of σ under the CFT map. Then

$$v(\pi) = \frac{e(k_w : E_v)}{f(E_v : \mathbb{Q}_p)}.$$

The image of π in $G(\mathbb{A}_f)/G(\mathbb{Q})$ is the class of the idèle

$$(1, \ldots, 1, \pi, \bar{\pi}^{-1}, 1, \ldots, 1) \in \mathbb{A}_{E,f}^\times,$$

where $\bar{\pi} \in E_{\bar{v}}$ denotes the image of π under the isomorphism $E_v \to E_{\bar{v}}$ induced by complex conjugation on E.

We have $v(\gamma_1) = -[\mathbb{F}_q : \mathbb{F}_p]/h$ from which we compute

$$v(\gamma_1^{[\mathbb{F}_w:\mathbb{F}_q]}) = -v(\pi),$$

and hence $\bar{v}(\gamma_1^{[\mathbb{F}_w:\mathbb{F}_q]}) = v(\pi)$. Moreover, γ_1 is a unit at all places of E different from v and $\bar{v}$. It follows that the idèle

$$\alpha := \gamma_1^{[\mathbb{F}_w:\mathbb{F}_q]} \cdot (1, \ldots, 1, \pi, \bar{\pi}, 1, \ldots, 1) \in \mathbb{A}_{E,f}^\times.$$

lies in the maximal compact subgroup

$$\mathcal{K} = \{g \in (\mathcal{O}_E \otimes \hat{\mathbb{Z}})^\times \mid g\bar{g} = 1\} \subset G(\mathbb{A}_f).$$

Since Gal_k is compact also, $\rho(\sigma)$ lies in $\mathcal{K}$. From the commutativity of the diagram (2) we conclude that $\rho(\sigma)/\alpha$ lies in the kernel of the map

$$\mathcal{K} \to G(\mathbb{A}_f)/G(\mathbb{Q}).$$

This kernel equals $\{g \in \mathcal{O}_E^\times \mid g\bar{g} = 1\}$, which is finite by the Dirichlet unit theorem. We conclude that $\rho(\sigma^m) = \alpha^m$ for some m, and hence

$$\rho(\sigma^m)_\ell = \gamma_1^m$$

in $G(\mathbb{Q}_\ell)$ for all $\ell \neq p$. □

We have

$$L(\bar{X}/\mathbb{F}_w) = \det_{\mathbb{Q}_\ell}(1 - \sigma T,\ \mathrm{H}^2_{\mathrm{et}}(X_{\bar{\mathbb{F}}_w}, \mathbb{Q}_\ell(1))).$$

Since none of the conjugates of γ_1 are roots of unity, we conclude from the preceding proposition that there is a finite extension $\mathbb{F}_w \subset \mathbb{F}'_w$ such that

$$L_{\mathrm{trc}}(\bar{X}_{\mathbb{F}'_w}/\mathbb{F}'_w) = \det_{\mathbb{Q}}(1 - \gamma_1^{[\mathbb{F}'_w:\mathbb{F}_q]}T,\ E),$$

or in other words

$$L_{\mathrm{trc}}(\bar{X}_{\mathbb{F}'_w}/\mathbb{F}'_w) = \prod_i (1 - \gamma_i^{[\mathbb{F}'_w:\mathbb{F}_q]}T),$$

which finishes the proof of Theorem 3.

Acknowledgements

This paper has benefited significantly from enlightening discussions with Eva Bayer, Christophe Cornut, Johan de Jong, Nick Katz, Mark Kisin, Christian Liedtke, Ronald van Luijk, Yuya Matsumoto, Davesh Maulik, Ben Moonen and Raman Parimala — I am most grateful for their questions, answers, comments and suggestions. Most of the research leading to this paper was performed while I was a member at the IAS. My gratitude to the Institute and its staff for maintaining the most stimulating research environment cannot be overstated. Special thanks go to Momota Ganguli who promptly and enthusiastically fulfilled even my most obscure bibliographical requests. Finally, I wish to thank Lotte Meijer and Michel Reymond for the necessary diversions outside geometry.

The author acknowledges financial support of the National Science Foundation (NSF) and the Netherlands Organization for Scientific Research (NWO).

References

[Artin and Mazur 1977] M. Artin and B. Mazur, "Formal groups arising from algebraic varieties", *Ann. Sci. École Norm. Sup.* (4) **10**:1 (1977), 87–131. MR 56 #15663 Zbl 0351.14023

[Barth et al. 2004] W. P. Barth, K. Hulek, C. A. M. Peters, and A. Van de Ven, *Compact complex surfaces*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **4**, Springer, Berlin, 2004. MR 2004m:14070 Zbl 1036.14016

[Bayer-Fluckiger 2014] E. Bayer-Fluckiger, "Embeddings of maximal tori in orthogonal groups", *Ann. Inst. Fourier* (*Grenoble*) **64**:1 (2014), 113–125. MR 3330542 Zbl 1318.11054

[Cassels 1978] J. W. S. Cassels, *Rational quadratic forms*, London Mathematical Society Monographs **13**, Academic Press, London, 1978. Reprinted by Dover, Mineola, NY, 2008. MR 80m:10019 Zbl 0395.10029

[Charles 2013] F. Charles, "The Tate conjecture for K3 surfaces over finite fields", *Invent. Math.* **194**:1 (2013), 119–145. MR 3103257 Zbl 1282.14014

[Charles 2014] F. Charles, "Birational boundedness for holomorphic symplectic varieties, Zarhin's trick for K3 surfaces, and the Tate conjecture", preprint, 2014. arXiv 1407.0592

[Deligne 1971] P. Deligne, "Travaux de Shimura", pp. 123–165 in *Séminaire Bourbaki* 1970/1971 (Exposé 389), Lecture Notes in Mathematics **244**, Springer, Berlin, 1971. MR 58 #16675 Zbl 0225.14007

[Deligne 1972] P. Deligne, "La conjecture de Weil pour les surfaces K3", *Invent. Math.* **15** (1972), 206–226. MR 45 #5137 Zbl 0219.14022

[Deligne 1974] P. Deligne, "La conjecture de Weil, I", *Inst. Hautes Études Sci. Publ. Math.* **43** (1974), 273–307. MR 49 #5013 Zbl 0287.14001

[Elsenhans and Jahnel 2015] A.-S. Elsenhans and J. Jahnel, "On the characteristic polynomial of the Frobenius on étale cohomology", *Duke Math. J.* **164**:11 (2015), 2161–2184. MR 3385131 Zbl 06486373

[Illusie 1979] L. Illusie, "Complexe de de Rham–Witt et cohomologie cristalline", *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR 82d:14013 Zbl 0436.14007

[Kedlaya and Sutherland 2015] K. S. Kedlaya and A. V. Sutherland, "A census of zeta functions of quartic K3 surfaces over F_2", preprint, 2015. arXiv 1511.06945

[Kisin 2013] M. Kisin, "Mod p points on Shimura varieties of abelian type", preprint, Harvard University, 2013, Available at http://www.math.harvard.edu/~kisin/dvifiles/lr.pdf.

[Kottwitz 1990] R. E. Kottwitz, "Shimura varieties and λ-adic representations", pp. 161–209 in *Automorphic forms, Shimura varieties, and L-functions* (Ann Arbor, MI, 1988), vol. 1, edited by L. Clozel and J. S. Milne, Perspectives in Mathematics **10**, Academic Press, Boston, MA, 1990. MR 92b:11038 Zbl 0743.14019

[Liedtke and Matsumoto 2015] C. Liedtke and Y. Matsumoto, "Good reduction of K3 surfaces", preprint, 2015. arXiv 1411.4797

[Madapusi Pera 2015] K. Madapusi Pera, "The Tate conjecture for K3 surfaces in odd characteristic", *Invent. Math.* **201**:2 (2015), 625–668. MR 3370622 Zbl 06480962

[Matsumoto 2015] Y. Matsumoto, "Good reduction criterion for K3 surfaces", *Math. Z.* **279**:1-2 (2015), 241–266. MR 3299851 Zbl 1317.14089

[Maulik 2014] D. Maulik, "Supersingular K3 surfaces for large primes", *Duke Math. J.* **163**:13 (2014), 2357–2425. MR 3265555 Zbl 1308.14043

[Mazur 1972] B. Mazur, "Frobenius and the Hodge filtration", *Bull. Amer. Math. Soc.* **78** (1972), 653–667. MR 48 #8507 Zbl 0258.14006

[Mazur 1973] B. Mazur, "Frobenius and the Hodge filtration (estimates)", *Ann. of Math.* (2) **98** (1973), 58–95. MR 48 #297 Zbl 0261.14005

[Milne 1975] J. S. Milne, "On a conjecture of Artin and Tate", *Ann. of Math.* (2) **102**:3 (1975), 517–533. MR 54 #2659 Zbl 0343.14005

[Piatetski-Shapiro and Shafarevich 1973] I. I. Piatetski-Shapiro and I. R. Shafarevich, "Арифметика поверхностей типа K3", *Trudy Mat. Inst. Steklov.* **132** (1973), 44–54. Translated as "The arithmetic of surfaces of type K3" in *Proc. Steklov Inst. Math.* **132** (1973), 45–57. MR 49 #302 Zbl 0293.14010

[Rizov 2010] J. Rizov, "Kuga–Satake abelian varieties of K3 surfaces in mixed characteristic", *J. Reine Angew. Math.* **648** (2010), 13–67. MR 2012e:14089 Zbl 1208.14031

[Scharlau 1985] W. Scharlau, *Quadratic and Hermitian forms*, Grundlehren der Mathematischen Wissenschaften **270**, Springer, Berlin, 1985. MR 86k:11022 Zbl 0584.10010

[Serre 1970] J.-P. Serre, *Cours d'arithmétique*, Collection SUP: "Le Mathématicien" **2**, Presses Universitaires de France, Paris, 1970. Translated as *A course in arithmetic*, Graduate Texts in Mathematics **7**, Springer, New York, 1973. MR 41 #138 Zbl 0225.12002

[Tate 1971] J. Tate, "Classes d'isogénie des variétés abéliennes sur un corps fini (d'après T. Honda)", pp. 95–110 in *Séminaire Bourbaki* 1968/1969 (Exposé 352), Lecture Notes in Mathematics **175**, Springer, Berlin, 1971. MR 3077121 Zbl 0212.25702

[Todorov 1980] A. N. Todorov, "Applications of the Kähler–Einstein–Calabi–Yau metric to moduli of K3 surfaces", *Invent. Math.* **61**:3 (1980), 251–265. MR 82k:32065 Zbl 0472.14006

[Yu and Yui 2008] J.-D. Yu and N. Yui, "K3 surfaces of finite height over finite fields", *J. Math. Kyoto Univ.* **48**:3 (2008), 499–519. MR 2010j:11099 Zbl 1174.14034

[Zarhin 1983] Y. G. Zarhin, "Hodge groups of K3 surfaces", *J. Reine Angew. Math.* **341** (1983), 193–220. MR 84g:14009 Zbl 0506.14034

Communicated by Kiran S. Kedlaya
Received 2015-08-17 Revised 2015-11-27 Accepted 2015-12-27

l.d.j.taelman@uva.nl *Korteweg-de Vries Instituut, Universiteit van Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, Netherlands*

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. 5 2016